Science
Education

Rigor in Elementary Science Students' Discourse: The Role of Responsiveness and Supportive Conditions for Talk

CAROLYN COLLEY, MARK WINDSCHITL
Department of Curriculum and Instruction, College of Education, University of Washington, Seattle, WA 98195, USA

Received 14 September 2015; accepted 2 May 2016
DOI 10.1002/sce.21243
Published online 25 August 2016 in Wiley Online Library (wileyonlinelibrary.com).

ABSTRACT: Teaching that is responsive to students' ideas can create opportunities for rigorous sense-making talk by young learners. Yet we have few accounts of how thoughtful attempts at responsive teaching unfold across units of instruction in elementary science classrooms and have only begun to understand how responsiveness encourages rigor in conversations. In this study, the first author taught an electric circuits unit to four upper elementary science classes, exercising a responsive teaching stance. We found that rigorous episodes of whole-class talk were associated with the teacher's use of open-ended questions, follow-up prompts, references to activity or representations, prediscussion tasks, and asking students to comment on their peers' ideas. Overall, higher rigor talk co-occurred with these conditions when used in combination. Despite being responsive to students' emerging ideas, all four classes addressed the science ideas for the unit—an outcome we attribute to the use of an anchoring phenomenon and the teacher's awareness of the concepts required to construct evidence-based explanations for it. Finally, concerted attempts to teach in responsive ways—while also attending to rigor—surfaced pedagogical tensions that problematize efforts to create such discourse-rich environments and inform how this type of instruction might be enacted by others. © 2016 Wiley Periodicals, Inc. *Sci Ed* **100:**1009–1038, 2016

Correspondence to: Carolyn Colley, e-mail: cdawson1@uw.edu

INTRODUCTION

The call to engage students in productive talk is now at the heart of current science education reforms—students are expected to participate in disciplinary discourse at more rigorous levels than ever before (National Research Council, 2007, 2012). These aspirations, however, contrast with commonplace norms in American science classrooms. For example, current elementary science instruction is frequently based on activities that are only loosely linked to science ideas via conversations. Often, whole-class "discussions" are times when teachers simply explain science ideas to students (Corcoran & Gerry, 2011; Davis & Krajcik, 2005; Roth et al., 2006). These pedagogies are not sufficient in light of the *Next Generation Science Standards* (NGSS Lead States, 2013), which demand that students take a more central role in the discourse-intensive work of developing and revising models, creating explanations, and arguing with evidence—while simultaneously learning and using conceptual content in these practices. To accomplish these goals through talk requires that the teacher uses specialized discourse routines, talk norms, and a strategic balance of dialogic and authoritative communication with students (Engle, 2006; Leinhardt & Steele, 2005; Mercer, 2008; Minstrell & Kraus, 2005; Mortimer & Scott, 2003; Sfard & McClain, 2002; Van Zee, Hammer, Bell, Roy, & Peter, 2005).

While prior research has described what productive academic talk sounds like in classrooms (Chin, 2006; Engle & Conant, 2002; Michaels, O'Connor, & Resnick, 2008), few studies have examined whole-class conversations across entire units of instruction to systematically assess the kinds of classroom conditions that both reflect teacher responsiveness and afford rigorous intellectual work to be accomplished by students through talk. In this study, we document talk opportunities for students during a 3-week unit on electrical circuits, taught by the same teacher in four different elementary classrooms at two school sites. We seek answers to the following research questions:

1. Can a teacher's use of responsive discourse with students consistently support rigorous whole-class conversations across units of instruction?
2. Under what teacher-mediated conditions (TMCs) are students more likely to engage in rigorous science talk during whole-class discussions?
3. Does the teacher's attention to students' science ideas shift the trajectories of units in ways that preclude engagement with some important curricular topics?

We intend to make contributions to the literature on responsive discourse and sense making in two ways. First, we focus on the explanatory rigor of students' talk and the teacher moves and conditions that co-occur with this talk in whole-class settings. Research focusing on "responsive instruction" often lacks close attention to the development of substantive disciplinary ideas which can include students' attempts at causal accounts of phenomena (Coffey, Hammer, Levin, & Grant, 2011). Yet, the literature describing teachers learning to notice, assess, and respond to students' ideas is at least partly based on the assumption that these skills will assist young learners in developing or analyzing explanations and to cultivate an understanding of science concepts as part of this effort (Levin, Hammer, & Coffey, 2009; Michaels et al., 2008; Rosebery & Warren, 2008; Sherin & van Es, 2009). To ground our notion of rigor, we draw upon Ford and Wargo's (2011) assertion about disciplinary discourse—that the dialogic understanding of an idea includes: being able to use it to explain a phenomenon to others, being aware that any given explanation is one among many alternatives, and being able to assess the credibility of one's ideas or explanations based on evidence. Engaging in talk that reflects these understandings and abilities is important not only because of its instrumental value in developing conceptual knowledge, but also because it embodies a powerful epistemic stance toward knowledge building that can be used in any domain of science (Framework; NRC, 2012). The second contribution of this study is the documentation of all whole-class discussions in four different classrooms during the same unit of instruction with the same basic attempts at discourse moves and scaffolding. We examine, in particular, whether *combinations* of these moves/scaffolding co-occur with rigorous talk by students. Most close analyses of discourse focus on a defined set of conversations within the context of a single classroom or lesson. While these have yielded valuable theoretical insights about what is possible for students and teachers, they do not directly test hypotheses that correlate rigorous talk by students with teacher moves and/or forms of scaffolding over the duration of a unit, where the circumstances of whole-class conversations each day can vary significantly. The design used in this study further allows us to look at four different classrooms so that, in aggregate, trends are less likely to be explained by the presence of a few students who dominate the talk environment, or by unique conditions (topic of the day, student involvement in an exciting activity, etc.). Using the same teacher across multiple contexts allows us to assess the influence of consistently applied moves or uses of scaffolding that is hard to achieve when working with multiple teachers.

BACKGROUND

What Does it Mean to be Responsive to Students' Ideas?

Recent research in the areas of student learning, expert teaching, and knowledge construction in the disciplines, has converged on the vision of classrooms as communities in which the careful orchestration of talk by teachers mediates increasingly productive forms of reasoning and activity by students (Engle, 2006; Leinhardt & Steele, 2005; Minstrell & Kraus, 2005; Mortimer & Scott, 2003; Sfard & McClain, 2002). Sense making is a form of reasoning that requires students to "act on" ideas—generated from both authoritative academic sources and from one's own base of knowledge and experience—by elaborating on them, disambiguating them from related ideas, questioning their coherence, comparing and contrasting them with other ideas, considering how to test them, assessing their credibility against evidence, representing them in various forms, reconstructing them, and applying them to explain different situations (Michaels et al., 2008). From a sociocultural perspective, much of this work is best conducted on the social plane of the classroom. Students can participate in intellectual work with peers, access the reasoning of others, experience how a community deals with uncertainty, and observe how "what we know" changes over time (Brown & Campione, 1996; Engle & Conant, 2002; Scardamalia & Bereiter, 2006). By necessity, teaching that accommodates this kind of collective sense making in effective ways uses students' ideas, questions, and everyday experiences to fuel sense-making conversations.

Being responsive under these conditions also requires an ongoing press by teachers and students for creating mutual understandings about the science ideas in play. Pierson (2008) characterizes responsiveness as the "attempts to understand what another is thinking, displayed in how a conversational partner builds, questions, probes, clarifies, or takes up that which another has said" (p. 25). From a teacher's perspective, responsiveness includes deciding from a repertoire of moves how one might create an opportunity for students' meaning making given the conversational context (i.e., by posing a question based on what was said, making a connection back to a prior contribution, asking students to compare two of their recent claims, etc.; see Ball, Lubienski, & Mewborn, 2001). To further facilitate talk, teachers can engage students in tasks that help them represent and inspect each other's ideas (in writing or drawing) over the course of the unit (Danish & Enyedy, 2006; Linn & Hsi, 2000; Radinsky, Oliva, & Alamar, 2010). In such classrooms, thinking is made visible and public, so that it can be acted upon by the group as well as by individuals. The scientific practice of modeling, for example, is designed to allow communities (of scientists or young learners) to make their current thinking explicit so that ideas represented in the models can be discussed and tested. Collaborative sense making is valuable but challenging in classrooms because it is socially negotiated, contingent upon shared knowledge and intersubjectivity, and always a work in progress.

There is evidence that students can take up discourse norms characterized by responsiveness to peers (see Thompson et al., 2016; Stroupe, 2014; Jadallah et al., 2011). For example, in whole-group discourse, the teacher can provide examples of what probing an idea sounds like. Students themselves may take up such probes in whole-group conversation or in small group and pair interactions. Responsive science teaching can demonstrate for students how to respond to a peer's idea by asking for more information to clarify a word or phrase, adding onto the idea by agreeing, posing an alternative idea or disagreeing with the reasoning, etc. At first, students may appropriate these superficially, enacting talk moves without attention to the content of their peer's ideas; however, with practice and reflection, both students and teachers improve at being responsive to the thinking of others in productive ways.

On a larger timescale, responsiveness can be reflected in changes teachers make to lesson or unit trajectories. When students express particular interest in or confusion about some aspect of a phenomenon, the teacher may decide to shift the makeup of lessons so those science ideas can be explored in more depth. In one study, for example, seventh-grade students working on a unit about batteries and the transformation of energy became interested in the different ways that they could charge their cell phones (Windschitl, 2013). In response, the teacher recast the next few lessons so that students could use their everyday experiences with cell phones as the basis to help them understand the relationship between chemical and electrical energy. Adaptations to lessons and resequencing unit activities based on students' interests, puzzlement, and confusion are becoming more frequent in the literature as examples of skilled pedagogical practice and equitable instruction (Calabrese-Barton & Tan, 2009; Rosebery, Warren, & Conant, 1992; Stroupe, 2014).

Students' Ideas as Resources

A fundamental principle underlying both responsive teaching and equitable teaching in general is that students' science ideas are leverage-able intellectual resources that can be used by educators to support reasoning and, in turn, learning. This stance acknowledges a broad range of assets that students work with in developing their own understandings and rejects the notion that soliciting and privileging "correct answers" from students engenders meaningful forms of understanding. One class of resources relevant to science is concrete, phenomenon-specific intuitions and experiences that can serve as referents to inform more coherent and generalizable scientific theories (diSessa, 1993; Maskiewicz & Winters, 2012). Maskiewicz and Winters (2012) use the term "resources" (rather than expertise, knowledge, beliefs, skills, or conceptions) to emphasize that students' contributions are often composed of small-sized, disjoint, context-sensitive ideas that can, with instructional guidance, serve as building blocks for productive theorizing. Other student resources are epistemic (e.g., ideas that knowledge about the natural world can be constructed rather than received from authority figures) and are hypothesized to support the ability to participate in activities related to the generation of knowledge (e.g., analogy work, argumentation, modeling) and guide the direction of inquiry activity (Hammer & Elby, 2002; Louca, Elby, Hammer,

& Kagey, 2004; May, Hammer, & Roy, 2006). Students' ideas are resources not just for teachers but for their peers as well. To be used as such, thinking has to be made public (Michaels et al., 2008) and teachers have to help everyone in the classroom develop the habits of listening to and critiquing the partial understandings of others.

Elementary teachers typically do not have difficulty surfacing students' resources; however, helping them develop these into connected and accurate conceptions over time is more challenging (Harris, Phillips, & Penuel, 2012; Richards & Robertson, 2015). The treatment of questions and answers when teaching responsively differs from that of traditional science pedagogy. In traditional science teaching, the teacher responds to students' answers by acknowledging them and/or evaluating them; students' questions are treated as requests for information, which the teacher typically provides, acting as the science authority. In these cases, the teacher may be *responding* to students when providing requested information but is not being *responsive* to students' ideas. To be responsive to students' ideas, relevant resources of all kinds must first be activated and then collectively refined and revised over time through purposeful activity and discourse.

How Scaffolding Supports Classroom Discourse

Scaffolding is necessary for rich classroom discourse, in part because students are asked to do unfamiliar interactive work with the ideas of others and to engage with abstract science ideas. Successful scaffolding strategies have taken the form of providing structure as students attempt challenging performances (Hmelo-Silver, Duncan, & Chinn, 2007; Sandoval & Reiser, 2004), using specialized intellectual roles for students in group work (Walqui & van Lier, 2010; Cohen, 1994), making features of scientific work and thinking explicit (Bell, 2004, Edelson, Gordin, & Pea, 1999, Sandoval & Reiser, 2004; Lee & Songer, 2004), and prompts for students to use particular reasoning strategies (e.g., Derry, Hmelo-Silver, Nagarajan, Chernobilsky, & Beitzel, 2006; White & Frederiksen, 1998). A small body of studies has demonstrated potential links between scaffolding and students' contributions in discussions. Explicit requests by teachers to make sense of complex phenomena by referencing recent learning activities have encouraged sense-making discussions (Brown & Campione, 1996; Forman & Ansell, 2001). Modeling environments, and in particular, opportunities for students to generate pictorial representations of natural events, have coincided with more productive whole-class argumentation (Buckland, 2008; Lehrer & Schauble, 2012; Passmore & Stewart, 2002). A study investigating middle school students' responses to teacher prompts for discussion found that students were more likely to share their thinking and describe a diversity of ideas in writing as opposed to whole-class talk (Ruiz-Primo & Furtak, 2006; Pimentel & McNeill, 2013). This suggests that there may be some benefit to preconversation writing tasks to help students make their ideas public and prime them to comment on the ideas of others. These forms of scaffolding however are likely ineffective for generating talk unless norms for discussion and activity structures are in place (see, e.g., Herrenkohl, Palincsar, DeWater, & Kawasaki, 1999).

The Role of Rigor in Responsive Teaching

In this study, most opportunities to engage in disciplinary thinking came about through students' involvement in the scientific practices of explanation, modeling, and argumentation (Lehrer & Schauble, 2005; McNeill & Krajcik, 2008; Reiser, 2004). These practices involve the gradual construction of "knowledge objects" and along the way engage students in two fundamental epistemic conversations: What do we think we know? and Why are we convinced we know it? (Duschl, 2008). Inherent in this vision of teaching is a commitment to merging ideas about rigor and responsiveness under the general umbrella of reasoning about phenomena and constructing evidence-based explanations in a way that values the *progress of students' ideas* as a disciplinary norm (Bereiter, 1994).

Within this context, rigor is not a characteristic of curriculum. Rather, it is assessed by the levels of intellectual work students engage in during class discussions that are broadly associated with the following: clarifying or challenging conceptions that are central to an explanation, attempting to link observations to unobservable events and processes, comparing and contrasting explanations, and using evidence to support or refute an idea. These are not unlike the higher cognitive demand tasks used in the mathematics education community (Stein & Lane, 1996) and similar to how Ford and Wargo (2011) describe the aims of disciplinary discourse–to use ideas to create an explanation, select from multiple explanations, and assess the credibility of explanations using evidence.

No conversation of rigor is complete without addressing the perennial debate about coverage of material (see Osborne & Dillon, 2008; Sykes, Bird, & Kennedy, 2010). Responsive and rigorous teaching spends more time on fewer (but more central) science ideas than does traditional teaching. Concerns about engaging students with the full range of science ideas that have been expressed in the new standards are real. The *Framework* document for the *Next Generation Science Standards* (NRC, 2012), for example, argues for a balance between depth and breadth but does not describe what that would look like in a classroom. There have been only a few studies that demonstrate how both depth and breadth are possible when teaching responsively (see, e.g., Hammer, 1997), and none have intentionally compared the variations in topics students addressed in multiple classrooms where similar units of responsive teaching have been enacted. Our study tests the assumption that working on and with students' ideas over the course of a unit does not preclude the exploration of a wide range of related ideas that are described in a curriculum.

A DISCOURSE FRAMEWORK FOR RESPONSIVENESS

Our framework for responsive and rigorous science teaching first recognizes that any talk happening in classrooms is influenced by particular conversational norms, interactional routines, and the availability of tools. These can constrain or enable different types of discursive exchanges. We propose two possible pathways through which discursive exchanges can unfold in whole-class discussion, which are associated with increasing or decreasing the likelihood of engaging students in rigorous discussion. Both pathways begin with an initiation which is typically a question or prompt posed by the teacher to which students respond (Figure 1).

Either pathway could also be initiated by a spontaneous student contribution (represented in box labeled "Student Response") to which the teacher must decide how to respond. In the moment, the teacher decides which responses by students should be incorporated into the discussion, how they will be treated, and how the teacher's response might encourage others to then act upon the idea (Edwards & Mercer, 1987).

In the *recognition/evaluation* pathway, the teacher poses a question or prompt that intends a discrete and easily evaluable response from students. When one is offered, the teacher recognizes or evaluates it (e.g., "Thank you for that interesting idea ... " or "Not quite, anyone else?"). This can happen respectfully; however, the teacher's recognition/evaluation terminates that particular line of thinking by the student—at least in terms of public discourse. The exchange is typically followed by the teacher calling on another student who introduces a new response, perhaps to a different question, obviating the need for discussion about the original question and response. This dialogue pattern has been labeled as I–R–E (initiation–response–evaluation) or triadic discourse (see Cazden, 1986; Mehan,

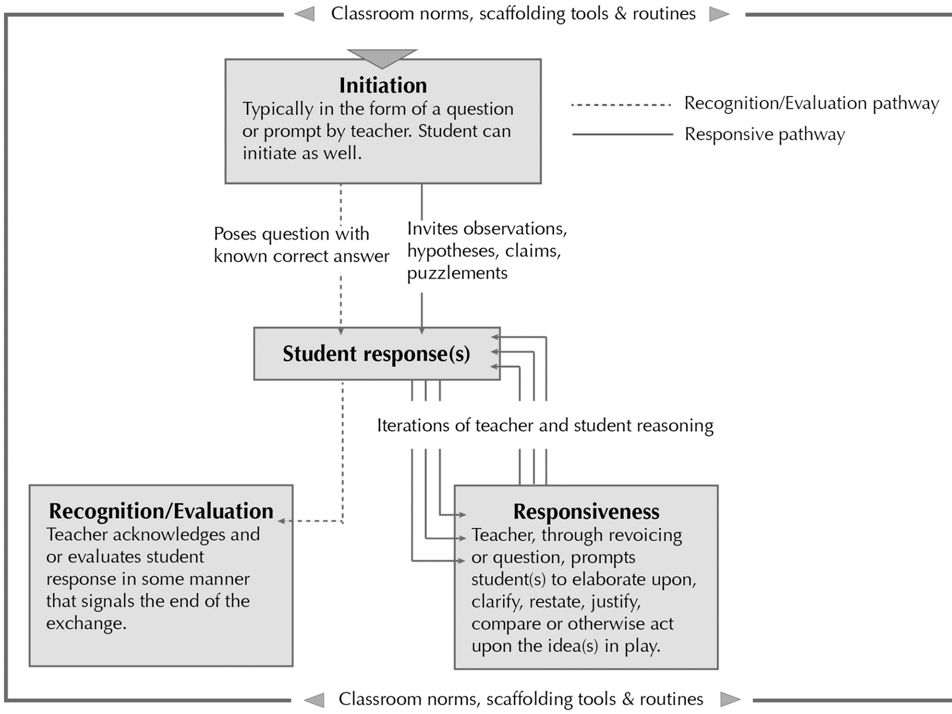

Figure 1. Pathways for whole-class talk.

1979). The result of overusing the recognition/evaluation pathway is little intellectual work done by students other than what is needed to provide a series of answers. Through this routine, most students learn that they will not be held accountable for listening to their peers or comparing different ideas.

In the *responsive* pathway, the teacher begins with an open-ended question. The nature and quality of the questions teachers pose matter for the nature and quality of student thinking they reveal and promote (Black & Wiliam, 1998). Substantial research supports the value of open-ended questions that require and afford more thinking than short phrase responses. Such questions elicit more information from students (Nystrand, Wu, Gamoran, Zeiser, & Long, 2003) which then provides teachers with more data and prompts new questions that can spark deeper student thinking (Minstrell & van Zee, 2003; Pierson, 2008). Following the students' response to the open-ended questions (or perhaps multiple student responses without evaluation), the teacher then enlists the responding student and his/her peers to engage further with the ideas, observations, or experiences that have been made public by asking others to do something more with these statements (e.g., "So I'm hearing you say X, how is that like Inara's idea about Y?" or "Can someone else elaborate on Jose's comment about X?").

At this point the pathway becomes cyclical, but more importantly the pathway requires the teacher to listen to the entire student utterance, then select a particular segment of it that they should respond to while simultaneously selecting a discourse move that will provoke further student reasoning—choosing, for example, who in the classroom should be addressed, with what request (to revoice a peer's statement, respond to a what–if scenario, turn and talk to a partner), and for what purpose (clarification, purposeful puzzlement, comparison, contrast, rerepresentation of ideas, etc.). The teacher must be responsive to the *substance* of student thinking (Coffey et al., 2011) and then ask other students to engage with the original comment, deepening the conversation around a specific and important idea. The responsive pathway allows for students to engage in multiple cycles of commentary around one idea, connecting ideas, clarifying language, and making tacit knowledge more explicit – all of which contribute to deepening the level of rigor emerging in the conversation.

Which pathway the teacher takes up depends on the purpose of the conversation—both kinds of talk are useful in science classrooms. The recognition/evaluation pathway can be appropriate when reprising what we, as a class, know so far; however, this pathway can limit further thinking. More sustained, iterative talk is required for meaningful learning. We note here that even though students can take on these responding roles to their peers, for this study we are focused on the teacher's initiations and responses within a whole-class discussion setting.

Despite the impact of particular talk moves in sustaining conversations, other supports, in the form of scaffolding, are needed. For the current study, we identified several kinds of scaffolding that could be used to support whole-class discussion. One of these is identifying a referent for discussions (Mercer, 2008). The referent functions as the object of conversation and the focus of attention as students engage in the joint work of discussion. Referents used for this purpose could include a recently completed activity, a public record of ideas generated by students, or a model that students create of a phenomenon. Another scaffold strategy is a prediscussion task in which students either write individually or converse in small groups to prepare them for whole-class talk (Walqui & van Lier, 2010). This can serve to activate students' prior knowledge and allow them to try out "rough draft" explanations before offering them in a whole-class setting. Another form of scaffolding relevant to discussions is the teacher making their thinking explicit, or demonstrating disciplinary ways of expressing ideas and addressing explanations or evidence offered by others (Engle & Conant, 2002; Herrenkohl et al., 1999; Sandoval & Reiser, 2004). These strategies often work well in combination to support substantive conversation by a wide variety of students.

METHOD

Participants and Context

Participants in this study were students in two fourth-grade classes and two fifth-grade classes at schools located in the Pacific northwest. Classes were selected using criterion-based sampling (Merriam, 2009) where teachers were interested in deepening their understanding of students' science discourse, open to having science lessons led by an experienced guest teacher, and were planning to teach units on electrical circuits to fourth and fifth graders. Two classroom teachers, Isabel and Mary, and their Grade 4 classes participated as part of a K–8 school, we call Glenloch Elementary, which reported 11% free and reduced lunch, 78% White students, and 19% special education identified students. One additional classroom teacher, James, and his Grade 5 morning and afternoon classes were part of a K–5 school, we refer to as Southern Hills Elementary, which reported 40% free and reduced lunch, 57% White students, and 11% special education identified students. Class size ranged between 27 and 29 students in both schools.

Although the three classroom teachers collaborated with the primary author in planning lessons, the first author served as a guest teacher for all lessons during these units in four classrooms and facilitated all whole-class discussions. The first author is an experienced elementary classroom teacher and instructional coach with a background in responsive instruction. She spent time before each unit getting to know the students and the culture of the participating classrooms (e.g., going out to recess with students, volunteering

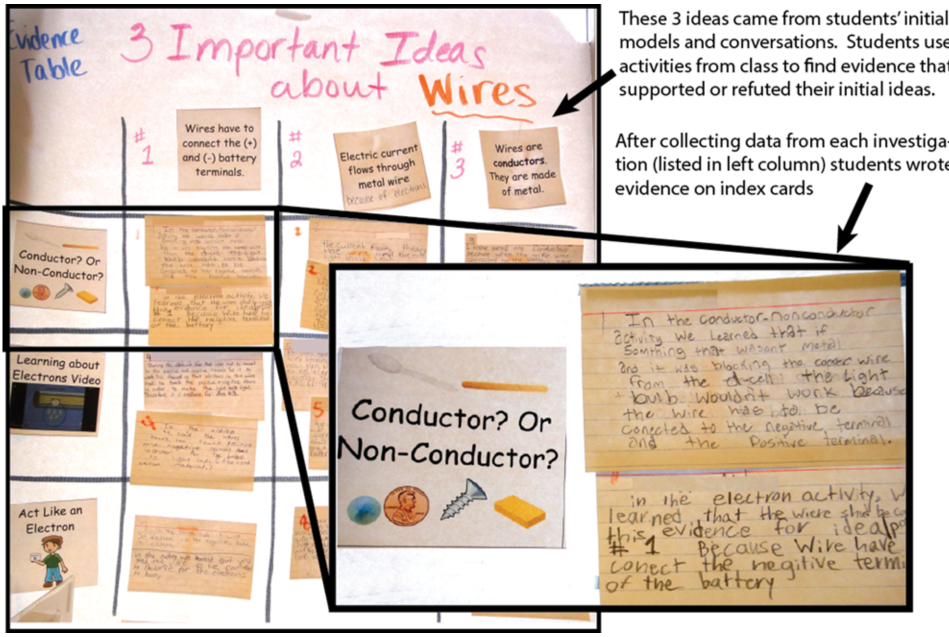

Figure 2. Example of on-going public record of students' collective thinking; columns based on students' initial theories. Enlargement shows students' writing about conductors and electrons.

in the classroom prior to the unit). By serving as a guest teacher for these units, the first author took a similar responsive teaching stance with all four classes. This ensured that particular conditions and a repertoire of discourse moves were similarly enacted during each unit. By teaching an entire unit, rather than stand-alone lessons, the guest teacher developed rapport with students over time and gained a day-to-day working understanding of the development of students' science ideas across a unit. Being "inside" the work allowed first-person insights into the kinds of dilemmas and ambiguities attendant to responsive instruction.

The 3-week unit about electric circuits was anchored in the challenge of understanding a complex yet accessible phenomenon for which students gradually developed models and evidence-based explanations. The phenomenon was that of a "dying" flashlight—the device was accidentally left on for a period of days, during which time it grew steadily dimmer and finally stopped working. Students gathered support for their initial ideas through a series of investigations (modified from curriculum and selected based on students' shared ideas) and were engaged in discourse on a daily basis around how particular pieces of evidence connected to and helped explain the dying flashlight phenomenon. Discourse occurred in pairs, in small groups, and whole class. The guest teacher and classroom teachers used students' models to continuously assess students' thinking. During whole-class discourse, the students created public records of their thinking based on their science ideas, recorded evidence from activities, and also modeled circuit pathways (Figure 2). These tools facilitated student communication, and they referred to these often during the unit to elaborate on and justify their thinking.

The selection of this phenomenon was influenced by the district's learning objectives and state standards such that the majority of component ideas necessary to explain it were covered within the grade-level science standards (13 of 20 concepts). Several additional ideas were added to the lessons as they were needed for students to fully explain how a

TABLE 1
Component Science Ideas for Students to Marshal in Explaining the Flashlight Circuit Phenomenon

Theme	Component Science Ideas
Matter (M)	M.1** Matter is composed of particles (atoms, particles, electrons).
	M.2** Particles have charges which makes some particles attracted to other particles.
	M.3** Particle arrangement in materials affects how electrons behave.
Conductors (C)	C.1* Some materials conduct (metals, minerals); while others are non-conductors or insulators of electrical energy (glass, plastic).
	C.2* Conductors are a pathway for electrical energy.
	C.3** Electron behavior affects the conductivity of a material.
	C.4** The “bumping” of electrons is the energy moving through matter
Energy Source (B)	B.1* There is energy/power stored in a battery.
	B.2** There are particular materials (chemicals) inside the battery that react.
	B.3* Batteries can “go out,” be “used up,” or “die” when chemicals inside the battery change.
	B.4* There is less energy available in the battery over time.
Energy Story (E)	E.1* There are different forms/kinds/types of energy (i.e., chemical, electrical, light, heat).
	E.2* Energy can be moved (transferred) through matter.
	E.3* Energy can be changed (transformed) by interactions with matter.
	E.4** Energy cannot be created or disappear, only transformed or transferred.
System (S)	S.1* Parts of a circuit system must be connected in a particular way for energy to be transferred.
	S.2* Individual parts of the circuit system must each be working properly.
	S.3* Parts each have a particular function.
	S.4* There can be more than one of a part in the system (i.e., multiple batteries, multiple bulbs).
	S.5* One part of the system can affect another (i.e., more batteries increases bulb brightness; dead battery affects bulb output)

*Denotes science standard and/or district grade level expectation (13 of 20).
**Denotes ideas that were added to fully explain flashlight phenomenon (7 of 20).

flashlight works and why it flashlight “dies.” As part of the planning practice, the guest teacher and classroom teachers worked together to compose a gapless scientific explanation of the event resulting in a list of 20 component science ideas described in Table 1. These elements were then prioritized for inclusion in the unit and relative emphasis based on their utility in helping students understand the scientific phenomenon. These planning considerations were a necessary prerequisite for responsive teaching that guided, in part, when and how the guest teacher made decisions about how to utilize students’ evolving ideas.

During the unit, students conducted investigations, engaged in activities, and collected evidence about each component idea, and then made sense of the evidence in relation to other science ideas, making connections and seeing relationships that would ultimately allow them the opportunity to stitch together their own evidence-based scientific explanations. These explanations could take various legitimate forms and were not reproductions of any textbook account.

Sources of Data

We video-recorded whole-class discourse throughout each of the units. For technical reasons, a small number of lessons could not be recorded in their entirety. These few gaps for the whole-class talk data were not systematic, but randomly occurred across all four classes. Following each class, the guest teacher and the classroom teacher debriefed the lesson, commenting on students' thinking and how that might influence the instructional choices during the following days' activities. Other data sources included samples of student work and public representations of science ideas created during class.

Data Analysis

We followed video analysis protocols recommended by Derry et al. (2010). After teaching each of the units, the videos were transcribed in full. In broad strokes, these were examined for (1) the full breadth of students' ideas, experiences, and responses to one another; (2) the intellectual rigor present in each discourse episode; (3) teacher discourse that preceded or co-occurred with student responses for students to build on ideas in productive and rigorous ways, as well as teacher responses that may have attenuated or redirected conversations; and (4) other TMCs that characterized students' opportunities to participate. Talk was analyzed at the level of utterance—meaning an uninterrupted stretch of speaking by an individual (Rowe, 2011). Each transcript was segmented based on the science content discussed. This resulted in a total of 90 separate whole-group discussion episodes for analysis. This allowed us to see if and when students' ideas were acted upon, how, by whom, and how long a particular idea was entertained. Each of these episodes received codes for rigor as well as marked for which components of science ideas were present. The initial coding categories are described below.

To enhance interrater reliability, the first and second authors began by collaboratively developing clear code definitions. The number of codes was kept limited as was the granularity of data the code was applied to (Miles & Huberman, 1994). The two coders independently rated a subset of the observations; these transcripts included both typical examples of whole-class talk and samples that were thought to be exemplary in terms of rigor. The coders then compared their analyses for both the level of rigor in student conversation and the presence of particular TMCs. Disagreements were highlighted, and code definitions were further clarified. Successive iterations of ratings, comparisons between rater's analyses, and code clarification resulted in approximately 80% joint probability of agreement for rigor and for TMCs.

Components of Big Science Ideas. Transcriptions of each lesson were coded for the presence of components of science concepts using student talk (see Table 1). Students typically voiced these by one of three means within whole-group discourse: (1) spontaneously proposing an idea or hypothesis, (2) responding to a teacher's question with evidence or reasoning, or (3) responding to another student's idea by building onto it, typically by agreeing or disagreeing and then providing additional reasoning. Each idea represented a concept that was "in play" (or remained "in play").

Episode Rigor. Each sample of whole-group talk was broken into episodes by the component idea(s) voiced. A new episode was designated when the students or teacher brought up a different idea or otherwise shifted the direction of the conversation. In some episodes, multiple ideas were in play; however, because students were referencing them against each

TABLE 2
Coding for rigor in whole-group talk episodes

Code	Level	Definition
3	High	Connects claims about observations to unobservable causal mechanisms; building on others' ideas; justifying claims using evidence or logic; comparing or contrasting ideas; conjecturing about relevant "what–if" scenarios in principled ways.
2	Moderate	Includes partial sense making about a science idea, typically pairing an observation with an inference about an unobservable feature, or naming observable conditions under which a phenomenon is likely to occur.
1	Low	Relevant descriptions of activity (observations) or sharing a related personal story/experience; some agreement and/or description (one-word answers, short phrases) without elaboration.

other in one segment of time, these two or three intertwined ideas constituted a single episode. Once episodes were identified and coded for the components of science ideas (using Table 1), each was then assigned a level of rigor based on how students were interacting with science concepts in the episode (see Table 2).

During analysis, moderate and high codes for rigor were grouped together for clearer comparison to low rigor episodes. The differences between rigor levels 1 and 2 (see Table 2) in terms of intellectual work demanded of students turned out to be more substantial than the difference in intellectual work required between levels 2 and 3. Because the goal of whole-class talk was *collective* sense making, to count as moderate-or-high rigor the contributing student(s) must be acting upon an idea in ways that were public and accessible to other students.

Using a subset of the data, we conducted a first pass analysis to identify segments of the transcripts using components of science ideas and then coded each segment with a level of rigor. Then we more closely examined the conditions immediately preceding and within the episodes to create a list of conditions that co-occurred with higher rigor discussions (see Table 3).

Responsive Talk Moves. A finer grained turn-of-talk analysis was used for each episode to code for the teachers' (and peers') responsiveness to students' comments. Broadly stated, the teacher's responses to a student's utterance were initially coded as recognition or as a bid to elaborate. Example responses that would be coded under the recognition category are utterances such as "Thank you for sharing," "That's an interesting idea," or a more general comment about the norms of science discourse such as, "I like how you included evidence to support your idea," which typically ended the talk episode about a particular concept. Alternatively, the teacher could use requests to elaborate (such as asking to clarify or asking for evidence) in response to a student's idea, which could be directed at the idea's author or to the whole class, and typically resulted in students continuing to discuss the idea(s) in play. These responsive requests became part of the list of TMCs that we hypothesized might be associated with rigorous student talk.

Teacher-Mediated Conditions. Whole-group discourse episodes co-occurred with or were preceded by conditions that we theorized would influence student talk. The initial

TABLE 3
Teacher-Mediated Conditions

Condition	Description	General Examples
A. Open question	The teacher (or student) poses a question for which there could be multiple legitimate responses or asks for a level of explanation.	What are some reasons that . . . ? What are some differences between . . . ? How do you think X happens? Why do you think X happened in that way?
B. Individual prompt to elaborate	Teacher asks individual to clarify, elaborate, connect ideas that they expressed.	Say more about . . . What do you mean by . . . What evidence do you have for that idea?
C. Prompts others to elaborate	Teacher asks others to elaborate, connect, agree/disagree with a statement by a peer.	If you agree or disagree with what [student] just said about X, how come? How is your idea like [student]'s? What evidence can we use to support that claim?
D. Reference to recent activity	Talk episode initiated by the teacher that occurs directly after science activity.	Students share observations and claims just after an activity is completed. Sketch of experimental set-up and list of observations created as public record.
E. Prediscussion task	Individual, partner, or small group task to consider an [open-ended] question before bringing the ideas to the whole-class level.	Use a format such as think–pair–share or a quick write to compare two student-generated hypotheses, to decide which one they agree with, or to explain how or why they think process or event occurs.
F. Availability of a referent for the talk; either a recent activity or an inscribed representation of it on the classroom walls	The teacher or student gestures/refers to an inscription of an idea. Materials are at hand for reference or manipulation.	Written conclusions for activity on public chart (reminder of learning). Class-created model or diagram (helps students explain ideas).

list of TMCs was drawn from the literature as potentially supportive of talk (see Table 3). Conditions A–C represent talk moves, and conditions D–F are scaffolds in the form of rehearsing ideas (E: prediscussion task) or identifying a referent to focus attention and/or activate prior knowledge (D, recent activity itself; F, public record of ideas or evidence). These were used to code each episode within whole-group discourse. After an initial round of analysis, we collapsed conditions D ("reference to recent activity") and F ("reference to tools that represented recent activity and ideas discussed") because we determined through video analysis these were not functionally unique.

FINDINGS

In our analysis of whole-class science discussions, we looked at the co-occurrence of particular levels of rigor and presence of TMCs as well as how the guest teacher made instructional changes to be responsive to students' ideas, interests, and puzzlements while also attending to the same science concepts for each group of students. We begin by providing examples of what different levels of rigor sound like in whole-class discussion, followed by a description of which combinations of TMCs increased the likelihood of medium-or-high rigor emerging in student discussion. We conclude the findings by looking at how the guest teacher was responsive to student ideas by adapting lesson sequences and describing the frequency of science concepts discussed by students across the units.

Examples of Levels of Rigor in Whole-Class Discourse

We begin with examples of talk episodes to illustrate "what counts" as high and low rigor and show that both types could occur within the same whole-class conversation. The following transcript illustrates the guest teacher's use of three TMCs (open-ended questions, individual prompts, and the use of a referent) as co-occurring with students' rigorous talk. These students were engaged in thinking about how and why energy transformed from chemical to electrical energy to heat and light energy in a flashlight. They used their observations of a working flashlight, personal experiences with friction, and what they learned in this unit about electrons and chemicals inside a battery to try to make sense of the energy transformations. The teacher sets up the episode using a combination of TMCs (see Figure 3) and then allows time for multiple students to contribute to the main question about what makes energy transform.

By the end of this excerpt, there are two hypotheses on the table: (1) Daisy's idea about friction in the filament caused by energy "rubbing" that then changes the electrical energy to heat and (2) Francine's idea about different metals in the circuit that help change the energy at different points along the pathway. The level of reasoning required to articulate these explanations is more demanding than simply naming parts of a circuit or labeling types of energy on a circuit diagram. The teacher entered this lesson attempting to engage students in deeper levels of reasoning and planned to prompt students to say more about their thinking, moving from observable features to unobservable causes. In this example, the combination of three TMCs seemed to help the class sustain a conversation about energy transformations, about where they happen and why they happen.

Within the same lesson (approximately 1 minute after the prior episode ended), there was an example of a low rigor episode (Figure 4). Here a student shares a personal experience (unanticipated) and the guest teacher must respond spontaneously to help other students make sense of the story and make connections between the student's personal story and the lesson objective about energy transformations. In this example, the guest teacher was unable to use talk moves in a way that would help other students make sense of one individual's personal experience to deepen an understanding of energy transformation. The teacher used only one TMC during this episode—providing individual follow-up prompts (in this case to help clarify and visualize the story for other students). Because Gary did not make a clear connection between his story and any prior idea (and the instructor was unable to help Gary do this work), the inferred idea about energy transformations was not made publicly apparent to others and therefore this excerpt is coded as low rigor–Gary is making some

Line	*Lesson 8 (Energy Story) from Isabel's Class (Duration of excerpt: 4 minutes, 18 seconds)*	*Teacher-mediated condition (TMC)*
1	**Teacher: Alright, so I think the energy, I think I can agree with Francine that**	
2	**electrical energy is in the wire, we may still consider kinetic, but electrical is**	
3	**more likely** (writing on poster)**. But wait, hold on, to go from here to here**	(F) Use of Referent
4	(pointing at poster) **what made the energy change in the wire to heat and light**	
5	**in the light bulb? What made the energy transform?**	
6	Zeke: THE WIRE!	
7	**Teacher: What wire? What made it transform?...In our flashlight system what makes**	
8	**the energy change from inside the wire being electrical then it comes out as**	(A) Open-ended question
9	**light and heat? What makes it change? What do we think?...Amanda?**	
10	Amanda: Well, there might be like something in the light bulb that once electricity comes	
11	in that it changes or something, to change the energy.	
12	**Teacher: So then the question that I thought of when you said that, is at what point**	(F) Use of Referent
13	**does it change? Does it change to light energy here? here? here?** [pointing at	
14	different places in light bulb diagram] **At what place does that change**	
15	**happen? Omar?**	
16	Omar: In the filament.	(B) Individual Prompt
17	**Teacher: In the filament, that's that squirrely piece of wire in the middle** [pointing at	
18	filament on chart]**. It's the thin piece, the curly Q piece of wire. He says that**	
19	**that's where it turns into light....How do you know that, Omar?**	
20	Omar: I just guessed.	(B) Individual Prompt
21	**Teacher: Have you seen that part light up before?**	
22	Omar: Mm-hmmm… It generates it usually it just is two wires on either side. I could see	
23	there's two wires and then the light is in the middle.	
24	**Teacher: Okay. Daisy, then Francine.**	
25	Daisy: So I think that I agree that it changes in the filament thing and I think that it changes	
26	there because it's so curly Q'ed so tightly that the energy rubs against each other	
27	which makes a spark and makes light.	
28	**Teacher: Oh! Can everybody rub their hands together?**	
29	Students shouting out: Friction! Spark! Spark! Hot! Hot!	
30	**Teacher: So you feel what kind of energy?**	(B) Individual prompt
31	Omar: It's heat energy	
32	**Teacher: Wait a minute, so when the electrons are rubbing their way through that**	
33	**thin wire, what's happening?**	
34	Multiple students: HEAT!	
35	**Teacher: That gets to Gary's idea that not only does it produce, not produce, we don't**	(F) Use of Referent
36	**make energy, we just change energy. It doesn't ...it goes through that thin**	
37	**wire [points to chart showing circuit pathway] and there's a little friction**	
38	**perhaps so it changes, the electrical energy in the wire into heat energy like**	
39	**Gary said and also light energy like we've seen. That was one of your first**	
40	**ideas from like two weeks ago, wasn't it?**	
41	Several students collectively: Yeah	
42	**Teacher: So Francine, then Gary. So... go...**	
43	Francine: Um I agree with both about how the heat, how it creates the heat but when it	
44	creates the light, not creates but changes…	
45	**Teacher: Right. We don't create energy.**	
46	Francine: I agree with both about how the heat, how it creates the heat but when it creates	
47	the light, not creates, changes… in the filament … because they (electrons) are so	
48	tightly packed but I think it's a difference in metal which is why it changes when	
49	it goes from a battery to the wire it's because it's a change… I mean we don't	
50	know that, but it does change…	

Figure 3. High rigor episode example where students are considering how energy is transformed in the light bulb and circuit system.

sense of his own experience, but not in a way that is accessible, taken up, or responded to by others.

These two examples illustrate how a range of rigor can emerge within one whole-class discussions and how rigor can depend, in part, on how the teacher sets up and sustains the discussion utilizing talk moves and scaffolding to facilitate student interactions. The next

Transcript Lesson 9 from Isabel's Class	*TMC*
Gary [hand raised]: It's something about making heat energy. It relates.	
Teacher: So changing to heat energy?	
Gary: But it relates, um, my dad just got a um when we were in Portland my Dad got this special new [inaudible] kind of flint and steel. And technically the old one is since it's using iron and flint it's um the electrons are hitting...	
Teacher: Flint's a kind of rock, right? I just want to make sure we're all visualizing this	(B) Individual prompt
Gary: It's a big sharp, it's a big black sharp rock.	
Teacher: What do you use it for?	(B) Individual prompt
Gary: You use it for, there's a special steel thing like this you put it in your hand, the old one and you go like that and you do it as hard as you can and it will make a big spark.	
Other student [shouts out]: Yeah, my dad has one of those.	
Teacher: So you're turning the mechanical, the friction, energy where you rub into like a heat energy with a spark?	(B) Individual prompt
Gary: And then there's a special one…	
Brady [interrupting]: You can do the same thing with two rocks	
Teacher: Hold on, let him finish. Gary, can you finish your thought then Brady can add on.	
Gary: Um and then they also have these special ones where there's a scraper, a wood handle, a piece of I don't know what it's called. It's this light really, really flammable metal .. um and it has then it has a big bump of flint on it.	
Teacher: Why do you think it would have a wood piece on that for?	(B) Individual prompt
Gary: You can scrape it off or it's a really good handle.	
Teacher: Oh a good handle. I was…	
Gary: You can scrape off some of the flint or you can scrape off some of the metal and then scrape off some of the flint and it will make a spark and it will make and the the uh metal can go up over five thousand degrees.	
Teacher: That's super hot. I liked how you were thinking about energy transformations in a different system than our flashlight but it's kind of the same thing because we're talking about energy and how it can change.	
Gary: Uh huh. Yup.	

Figure 4. Example of low rigor talk episode.

section describes how many episodes were coded for low rigor compared to moderate-to-high rigor and which TMCs were present in these episodes.

Combinations of Conditions That Mediated Rigorous Talk

Within the 90 episodes of whole-group talk across the four classes during this unit, a set of six TMCs emerged from the data (see Table 3). Three of these conditions, occurring in various combinations were more frequently associated with moderate-to-high levels of rigorous student talk than other conditions. Before considering the efficacy of *combinations* of conditions, in terms of what they afforded students for talk, we first examine which levels of rigor co-occurred with *single* conditions present (compared to combinations of TMCs). These data are presented in Table 4.

Overall, there is a strong pattern between combinations of TMCs with higher rigor. Few examples of rigorous whole-class talk happened when only one TMC was present (4% across all classes). This suggests that combinations of TMCs seemed to be helpful to students during these sense-making conversations—65% of higher rigor episodes featured such combinations. Perhaps as noteworthy was that conditions widely thought to help generate talk, such as open-ended questioning, were uniformly unproductive for all four classrooms when left unsupported by other strategies.

TABLE 4
Percentage of All Whole-Group Talk Episodes by the Number of TMCs Present

	Low Rigor (%)		Moderate-to-High Rigor (%)		
Class	Single TMC	Two or more TMC	Single TMC	Two or more TMC	Total Whole-Class Episodes
James' AM	18	9	9	64	=100
James' PM	8	33	0	58	= 100
Isabel	7	11	4	78	= 100
Mary	17	24	4	55	= 100
Overall	13	18	4	65	= 100

TABLE 5
Percent of Moderate-and-High Rigor Episodes Co-Occurring with Combinations of the Three Most Frequently Occurring TMCs

Combinations of TMCs Present with Higher Rigor			Percent of All Moderate-and-High Rigor Episodes by Class			
A	B	D/F	Glenloch Elementary		Southern Hills Elementary	
Open-Ended Question	Individual Follow-Up Prompt	Reference to Activity or Material Representation	James AM	James PM	Isabel	Mary
X	X		6	0	27	30
X		X	25	43	9	6
	X	X	25	14	9	24
X	X	X	12	43	27	12

Column totals do not add up to 100% because this table shows the three most prevalent conditions (A, B, and D/F). Other combinations also co-occurred with higher rigor episodes but not as frequently as combinations of these three; these conditions make up the remaining percentage.

We found one particular set of combinations of TMCs that were more frequently associated with moderate-to-high rigor talk episodes than others. These TMCs were (1) use of open-ended questions, (2) follow-up prompts for the respondent or other students to elaborate on an idea, and (3) direct reference by the teacher or student to a recent activity or to a public inscription representing that activity. Of the total whole-class talk episodes, 69% were rated at moderate-to-high rigor, and 21% of these had all three of these TMCs, and 53% of these had two of the three present. Table 5 shows a breakdown by class and school for combinations of these three conditions.

Across the classes, each combination of these three conditions varied in their correlation with rigorous talk, though some classes featured the conditions above more frequently than others during episodes to support rigorous talk. For example, James' classes show a higher percentage of rigorous talk when a reference to activity or material was made by the teacher during whole-class talk in combination with one or two other conditions (43% and 25% each); without such referents the numbers drop to 0% and 16% in each of the two classes. Mary's class, on the other hand, did not seem to require the direct referent condition but responded readily to open-ended questioning and follow-up prompts (30% of episodes of moderate-or-high rigor in her class contained open-ended questions and follow-up prompts).

Two other conditions, which occurred less frequently, were strongly associated with higher rigor conversations. One was using a prediscussion task such as talking with a partner about a focal question before bringing ideas out to the whole class. The other was prompting someone to comment on the idea of one of their peers during whole-class discussion. Twenty-three percent of the total number of episodes contained a prediscussion task—in addition to other conditions like open-ended questioning— and 86% of these resulted in moderate-or-high rigor talk. The guest teacher invited other students to comment on his/her peer's ideas in 20% of the total episodes, and 89% of those resulted in moderate-or-high rigor talk. There were not enough occurrences of these conditions to make a strong claim that they influenced opportunities for sense making. There is, however, enough evidence to suggest a positive relationship, one that warrants further study.

In summary, we are not claiming that any one particular combination of TMCs is a silver bullet to engage students in rigorous science discourse but there is a strong pattern that utilizing particular TMCs in combination seem to increase the likelihood of students deepening their engagement with the science concepts in whole-group settings and having rigorous discourse with peers about science ideas.

Adapting Lesson Sequences To Be Responsive to Students' Ideas, Interests, and Puzzlements

Here we describe the relationship between a larger "grain size" of responsiveness by the teacher and students' cumulative engagement with key science ideas at both the lesson and unit levels. This section describes conditions under which major shifts in unit trajectories occurred, and how all classes addressed the same science concepts by the end of the unit.

Adaptations to lesson plans created variations in the frequency and timing of when during the unit students were engaged with target science concepts during whole-group discourse. We illustrate this with an example from Isabel's class partway through the unit. This shift came about because of how students answered an assessment question from the curriculum guide. The question asked students to trace the current flow through a battery and bulb on a circuit diagram. In James' and Mary's classes, all students traced the current going one direction in a circle. In Isabel's class, however, it became clear by looking across students' notebooks that there were two distinct hypotheses about the direction of the current flow. Isabel's students drew and explained the path in different ways within the class, falling into two "current direction" hypotheses. The images in Figure 5 show contrasting examples taken from student notebooks.

The student work on the right describes the current flowing in one direction around the circuit, out of one battery terminal, around the circuit, and into the other battery terminal – a representation of what the class called the "one-way" hypothesis (shown by arrows drawn on wires in the circuit and described in words). The student work on the left describes a pathway where "all the electrons go up each side" of the support wires and "rub together making themselves hot" in the filament then "burst into nothing" (as shown in the drawing with dots rising up from the filament). This is what the class named the "two-way" hypothesis. Several students were invested in this alternative theory.

In response to this situation, the teacher decided to pursue this current flow debate rather than proceed with the planned unit lessons about fuses and switches—to which the other three classes had moved on. Isabel and the guest teacher decided during a planning session

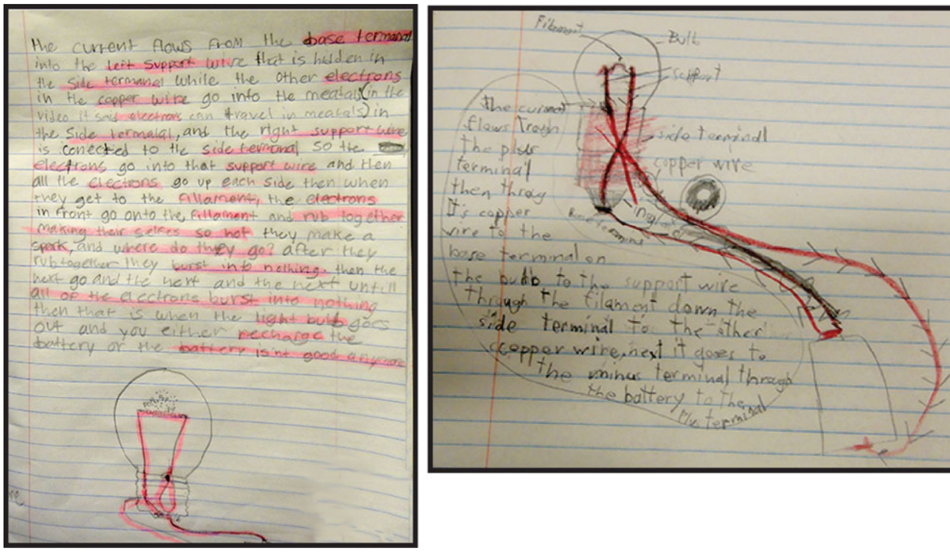

Figure 5. Contrasting examples, taken from student notebooks, showing different "current direction" hypotheses.

that the directional flow of electricity through the flashlight's batteries, wires, and bulb was a crucial idea for students. It was a science concept that could apply to and partially explain the functioning of a wide range of everyday devices students were familiar with. Without a grasp of this idea, the operation of switches and fuses could not be fully understood. The following day, students from each hypothesis group came up to the front of the class to explain their thinking, calling on peers in the audience to agree, disagree, and/or ask for evidence for each hypothesis.

This current flow debate represents one major alteration in the unit trajectory for Isabel's class. This opportunity allowed students to explain their hypotheses, draw upon previously gathered evidence, and debate the direction in which current flows in the circuit pathway. This instructional choice of responding to student thinking by replacing lessons from the curriculum guide represents of how a responsive stance toward science instruction can result in a significant change to the unit trajectory, and how the principled professional reasoning that guides these decisions. We note that students had *fully appropriated* responsibility for this conversation. One student, Colin, began by calling on a peer to respond. His classmate, Amanda, signaled a civil disagreement and referred to a drawing to support her assertions. Later, four different students unpacked whether a glass separator (bead) at the base of the bulb filament conducts or insulates against how the electric current flows. There were as many as 17 turns of student talk without an utterance by the guest teacher or classroom teacher. Resulting from this initial debate, the guest teacher and Isabel added additional lessons about current flow and electron behavior which helped students come to a consensus about current flow.

Another responsive shift occurred at the beginning of the unit in James' morning class. In the first part of the initial lesson, students were tasked with creating different configurations of simple circuits to light a bulb. James' morning class talked about their circuits in terms of where to connect the parts, whereas his afternoon class had already begun to suggest hypotheses about why and how their circuit worked and posing "what if" scenarios based on what they thought was happening (i.e., What would happen to the bulb if we used all the batteries at our table?). The next day, James' students in the morning class were given an

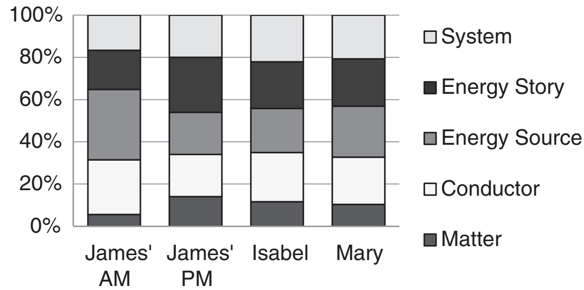

Figure 6. Categories of science ideas as percentage of whole-group talk across unit.

extended opportunity to continue experimenting with circuit materials before developing initial explanatory models. Though this change is on a smaller timescale (a matter of 20 minutes), compared to the aforementioned current flow debate (3 days), the decision was driven by the teacher's attention to the substance of students' ideas and desire to provide this additional opportunity which she viewed would support students' future reasoning. This choice meant a trade-off in how much time students would have to develop their initial models. However, this was not a significant concern, given the unit-long focus on revisiting and revising student-created models in light of evidence from activities meaning that all students would engage with core ideas more than once.

Responsive shifts in lessons or lesson sequences happened as the teacher interpreted and responded to students' current reasoning about key science concepts. Next, we describe how these key science concepts unfolded in each of the four classes.

Addressing Science Concepts

Returning now to a broader unit-level analysis of the frequency of science concepts in whole-group discourse, all classes generally engaged with the same core ideas by the end of the unit; however, the times when these concepts presented themselves within whole-group discourse differed, even as students were engaged in similar lessons. Each class addressed concepts in all five categories of concepts (matter, conductors, energy source, energy story, and systems) through whole-group discourse during the unit. Classes varied by when, how, and to what degree each of these categories were represented in whole-group discussions. The following displays characterize these units by how often classes surfaced particular categories of science concepts by (1) how often the five categories of science ideas appeared within the unit by class and (2) by comparing frequency of each of the 20 subcomponent science ideas by class. Figure 6 shows the percent of whole-group talk episodes for each class addressing the five categories. This representation shows that, even as the guest teacher was being responsive to the unique ideas and experiences students brought into discussions in each class, all four classes addressed all five categories of science ideas in whole-class talk during the unit. Figure 7 shows a finer grained display of the 20 subcomponents students discussed whole group. Most of the grade-level standards (denoted with *) were surfaced in whole-class discourse across the unit. Additional concepts were also used to explain the flashlight phenomenon (not asterisked); however, these were not part of grade-level standards. Furthermore, the order and frequency of ideas students voiced during whole-class discussion varied across the unit, yet all key ideas were part of a whole-class discussion at some point during the unit. Classes surfaced all, but a handful of the 20 subcomponent ideas needed to fully explain the dying flashlight; many of these

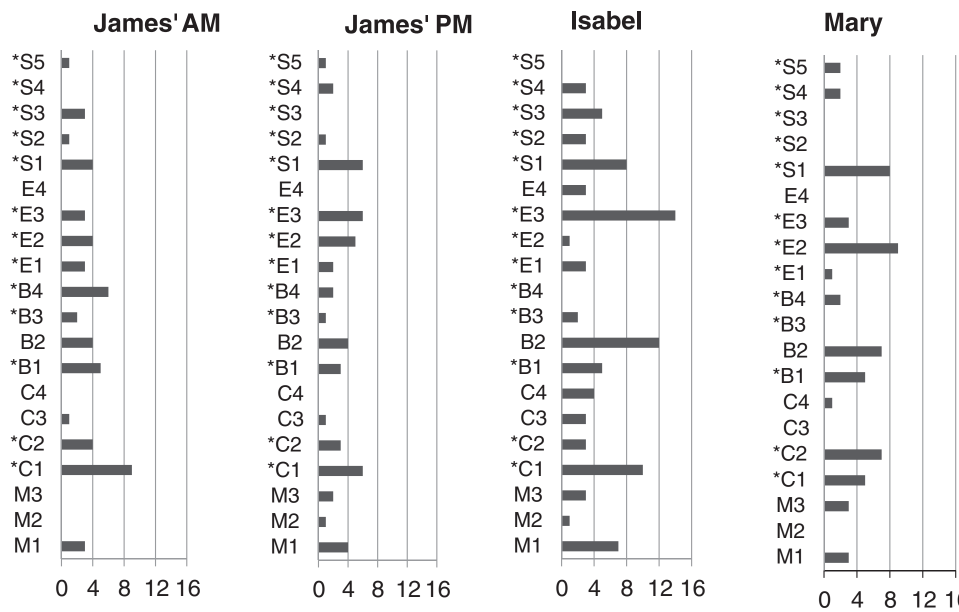

Figure 7. Aggregate number of appearances of each science idea during whole-class discussion across the entire unit, compared by class.

same ideas surfaced in other settings (i.e., small group talk, student writing, during model revisions).

DISCUSSION

Teachers' Responsive Discourse and Supporting Rigorous Student Conversations

We tested the hypothesis that efforts to infuse responsive talk in four different classrooms would support regular opportunities for students to engage in rigorous science conversations. Two patterns stand out in the data. The first is that more than half of all discussions in each classroom (ranging from 58% to 82%) were rated moderate or high in rigor. These discussions pressed students to unpack, compare, and test their ideas as a group and to use observations, evidence, and everyday experiences to regularly refine their thinking about the flashlight phenomenon. These outcomes contrast with talk in more common classroom situations, where students experience activities but never have the chance to distinguish important concepts related to these activities from trivia or to integrate their existing knowledge with new science ideas. Across several studies of typical classrooms, less than one third of discussions showed students engaged in any type of higher order thinking (Abrahams & Reiss, 2012, Corcoran & Gerry, 2011; Weiss, Pasley, Smith, Banilower, & Heck, 2003). In the present study, the teacher solicited and regularly challenged students' ideas. For example, the teacher takes Francine's hypothesis (Figure 3, lines 1–2) that electrical energy is in the wire (in the flashlight) and then turns to the rest of the class, asking "But wait ... what made the energy change in the wire to heat and light in the light bulb? What made the energy transform?" (lines 3–5). Later in the same conversation, she asks a student about how he knows the energy is transformed in the filament of the bulb (lines 18–19). These types of interaction with students' ideas were both responsive and intellectually probing. Occasionally, but not always, these moves were generative of new ideas. The 4-minute conversation produced several responses by students that were three words or less, but it also produced an extension of one student's filament idea by another and a theory about how electrons behave differently when passing through various metals in the circuit. Being responsive appeared to increase the chances of rigorous conversations by students but did not guarantee it.

Increasing the Likelihood of Productive Talk Through TMCs

Across the four classes, the teacher's use of open-ended questions and follow-up prompts for students to elaborate on their ideas or those of their peers were associated with rigorous whole-class conversations. Additionally, the use of a referent, as a scaffold, seemed particularly useful in the classrooms where students were less accustomed to extended turns at talk (i.e., most responses early in the unit were one word or short phrases) and not as fluent in discourse norms (i.e., not addressing each other but talking to/through the teacher).

Importantly, *none of these conditions used in isolation* were effective in encouraging sense-making talk. From a theoretical perspective, the open-ended question is necessary to frame what will be talked about but not sufficient because the actual object of collective inquiry by students are their own observations, hypotheses, and interpretations. The teacher facilitates this with a second move—a follow-up—that asks members of the class to comment on an idea of their peer, to link that idea with another idea or experience, challenge it, or otherwise act on that line of thinking. These are unnatural ways for students to respond to the statement of another person; few nine-year olds are asked in out-of-school settings to "Explain why you think that" or "Tell me how your idea is different from the one we just heard." The responsibility then initially falls to the teacher to coordinate talk with a series of strategic follow-ups that stimulates sense making and insight. If the teacher can demonstrate this explicitly, students may be able to take up this role themselves (Jadallah et al., 2011). Indeed this was observed in Isabel's classroom when her students carried on a constructive dialogue to resolve a disagreement about their electricity theories—a conversation that carried on for several minutes without the teacher.

The third condition associated with high rigor conversations was the availability of a recently completed activity or a public record of thinking from that activity (see Figure 5) to use as a referent in the conversations. A number of studies with elementary age students clearly indicate that they use knowledge products such as evidence tables (Figure 2) or models (Figure 5) to support assertions to peers (Danish & Enyedy, 2006; Linn & Hsi, 2000; Radinsky et al., 2010). While these referents served as a scaffold and appeared to add value to the rigor of the conversations we are unsure if it was their availability that mattered or that the referent was always talked about in terms of a central phenomenon—in this case the dying flashlight.

Many whole-class conversations had two levels of sense-making occurring, one was aimed at understanding a new idea, for example, conductors versus insulators, and the other was aimed at using these same ideas to deepen their explanation of the flashlight example. It seems plausible that intentionally revisiting the central and increasingly familiar phenomenon with new ideas and information could set the stage for more students to participate in the science conversations. Anchoring learning in a complex and puzzling event is not part of common curricula; however, research studies that report engagement, productive talk, and deep learning for elementary students are increasingly designed this way (see Magnusson & Palincsar, 2005; Rosebery et al., 1992; Roth et al., 2009). Two other conditions were associated with moderate or high rigor conversations— a prediscussion task and follow-ups by the teacher asking a student to comment on a peer's idea. The small number of instances of these conditions/moves (21 and 18, respectively) moderate our claims about their impact on the level of intellectual dialogue by students. Nonetheless, these features of the talk environment deserve further study.

What Responsive Teaching Means for Addressing Science Content Objectives

Teaching in a way that was responsive to students' thinking did not come at the expense of addressing content objectives. Although the lesson and unit trajectories varied, each class had opportunities to engage with curricular standards and lesson objectives by the conclusion of the unit. The amount of whole-class discussion around each key idea varied and the order in which these ideas surfaced was not identical; however, the teacher attempted to be responsive to the ideas that students were most intrigued by but also were within the bounds of the unit learning goals. There are few if any other studies that have analyzed the variability in which key science ideas emerge across units in responsive science classrooms (see Hammer, 1997, for an exception).

Responsive teaching may be viewed by some as an unwieldy approach to instruction. According to this view, indulgence in students' ideas can divert attention from exploring canonical ideas and allow alternative conceptions to flourish. We argue, based on our data, that responsive teaching is a pedagogy characterized by careful design and structure. From the design perspective, the unit planning began with identifying core science ideas related to the topic to be taught. The anchoring phenomenon was developed based on these ideas, and the activities in the unit were mapped onto the explanation for the phenomenon—What would students need to know, and when? From the structuring perspective, throughout the unit whole-class discussions were started by the teacher with a goal in mind. At times the goal was to elicit initial hypotheses, on other occasions it was to make sense of an activity, and still others it was designed to use observations from an activity as evidence to support part of an explanation for the anchoring event. During the whole-class conversations, the teacher constantly monitored the students' contributions for their potential to illuminate or challenge certain ideas in play and to move students toward the goal of the discussion. Such purposeful talk opportunities are examples of blending formative assessment with instruction (Black & Wiliam, 1998; Ruiz-Primo & Furtak, 2006) and represent the teacher's attention to the *substance* of science as she/he interacts with students (Coffey et al., 2011). The teacher can use what students say to comment on the quality of their reasoning, but student talk is also feedback to the teacher and allows more informed in-the-moment decisions during conversations. On a larger timescale, the content of conversations helps the teacher determine what lessons need to be extended, adapted, added, or discarded. Both the conversational moves and the adaptation of lessons are examples of responsiveness, but on different scales of action.

The "First-Person" Dilemmas of Responsive Teaching: Managing Endemic Uncertainty

First-hand experiences in four different classrooms highlighted a number of tensions about responsive and rigorous teaching that can inform how we might prepare others to take up these practices. Such tensions have been written about before (Hammer, 1997; Lampert, 1990; Michaels et al., 2008), and they generally arise from the uncertainties about "what counts" as knowing and how one responds in real time to ideas that may be partial, self-contradictory, based on experiences not shared by others in the classroom, or on the verge of insight. We describe three such tensions and then lay out the kinds of skills, tools, and stances toward knowledge building we found to be critical to work with young learners' ideas.

Tension 1: Making Decisions in Real Time About Students' Ideas and Experiences. When teaching responsively, students ask questions, tell stories, or pose hypotheses that go beyond the boundaries of the unit's curriculum goals. Therefore, the teacher needs to be able to listen to their thinking, compare their contributions to science concept targets, and decide how much, if any, whole-group discourse should be devoted to pursuing this line of thinking. Teachers cannot always assume, however, that they understand the students' intent. We cannot emphasize enough how often students offered remarks that seemed unrelated to the topic at hand, but when the teacher simply said "How do you think that helps you understand the flashlight?" the connection with lesson goals became clear. Even the generic prompt "Can you tell me more about that?" gave time for students to rephrase, elaborate on, or contextualize their initial utterances in a way that allowed their peers and the teacher to see the relevance of the idea and comment on it.

Responsiveness has been described as "attempts to understand what another is thinking, displayed in how a conversational partner builds, questions, probes, clarifies, or takes up that which another has said" (Pierson, 2008, p. 25). This characterizes what one might *see happening* in the classroom. Based on the extended time the first author had to attempt responsive talk with a wide variety of students in several classrooms, we can offer a characterization of the intellectual work that is much *less visible*. In real time, the teacher has to process what has been said, consider how to respond, and then reenter the conversation. The list below describes the complex intellectual work required on the part of the teacher to be responsive to students' ideas. This logic assumes that the student is directing a comment to the teacher (we acknowledge however that student-to-student conversations are valuable, if not preferable in many situations).

Within few moments, the teacher must

- comprehend what the student is saying. [if not, "Tell me in a different way."];
- intuit what the intention of the remark is, since student talk can be aimed at any previous comment and not necessarily be what the teacher most recently asked for;
- weigh out the potential for the student remark to provide insight to others. The teacher could coordinate a follow-up with the student's original remark to move the conversation forward (teacher creates an analogy using students' example, prompts students to recall ideas related to student comments from previous units). Alternatively, the teacher could provide a follow-up that would directly prompt other students to act upon their peer's idea, thereby improving the chance that collective reasoning would occur; and
- decide to whom the remark should be addressed (to original speaker or the whole class) or simply wait for other student comments.

If the teacher selects a student:

- Weigh out the challenge level of the question or prompt in light of the individuals' comfort or frustration level with the topic and whether they have already contributed recently to the point of excluding others.
- Compose the response, if there is one, and deliver it.

All this makes responsive teaching an intimidating challenge and suggests that deliberate practice of these interactions is necessary to achieve even a modest level of proficiency.

Tension 2: Developing Ideas With Students and Dealing With Their Fear of Being Wrong. Because a responsive teacher shares authority for classroom discourse by positioning students as intellectual partners, the dilemma arises about what to do if the student says something that is incongruent with the current scientific understandings. What should a responsive teacher do if a student expresses a "wrong" idea? In James' class, for example, one student stated that "electrons inside want to get to electrons outside." Though not technically correct, the general idea the student was describing about attraction between particles of opposite charges is actually part of the scientific explanation of current flow in circuits. When students' ideas are being used as intellectual resources in the classroom, it is unfair to judge them as correct or incorrect. Even the use of language like "correct" signals to students that the aim of talk is to provide the minimal information necessary to receive the approving nod from the teacher—at which point the responsibility to reason about ideas temporarily ceases.

But teachers can press students to reconcile their idea with available evidence if relevant experiences have already been provided, or provide students with new experiences that allow students to modify their claims. Reconciling can also happen for two divergent student ideas. For example, in Isabel's class the current flow debate sparked a need to add lessons to help students critique two opposing hypotheses, one of which was scientifically accurate and the other not, through exposure to new experiences. In all four classrooms, students were attempting to develop and refine explanations over the course of the unit, not simply after one lesson. This allowed opportunities for all types of ideas to be aired out and revisited over time with new information and new logic to evaluate them. This is the benefit of using puzzling and complex phenomena to anchor units of instruction, something advocated in the *Framework* for the *Next Generation Science Standards* (NRC, 2012). In this context, asking whether someone is "wrong" is not as productive as the question "How coherent is your reasoning when we take into account this evidence, this information, or these new ideas?"

Tension 3: Too Much and Too Little Participation. Too much participation creates its own dilemmas as does too little. In the two classes at Southern Hills, students comfortably used discussion stems, addressed each other, used pieces of other's ideas to continue conversations to the point where students would talk over each other, interrupting, seemingly excited to make connections and share out—these talk norms were in place long before the unit began. The two classes at Glenloch were less accustomed to answering questions in complete sentences and orienting toward one another. Often seemingly open-ended questions were met with awkward silence or a one word or phrase answer. This does not mean that Glenloch students had impoverished ideas, rather more probing and different discourse strategies were needed to access their thinking. As we know from literatures on equity, when the teacher is explicit about the "rules of the game" (Delpit, 1988; Nasir, Rosebery, Warren, & Lee, 2006) and when these are made clear and acted out for students, it opens the door to participation.

Alternatively, when the majority of students has their hands raised wanting to share, the challenge becomes the length of time it would take for each student to share individually. In this situation, the contributions can become a laundry list of disjointed ideas rather than a focused conversation building toward a collective understanding. By choosing to hear from some students and not others in these situations it creates yet another equity concern; some voices are simply not heard. In some of these situations, the teacher used a pair-share first before returning to a whole-class discussion to allow all students to communicate their thinking to at least one other person and listen to their partner in return. Then pairs were asked to volunteer to share whole group. Interestingly, the pair-share was also used successfully when students were reluctant to talk in whole group before conferring with a peer. Thus this partner talk strategy was used to deal with issues of both too much and too little initial participation.

A related tension that arose for the teacher, and that is unique to whole-group discourse, is deciding how long to press one student about his/her idea while others are waiting to contribute. Since responsive teaching needs ideas to be "in play," pressing students to elaborate beyond simply sharing observations is important; but this requires that peers hold their own thoughts temporarily and follow the line of thinking by their classmate. Part of the issue would be addressed if class norms and routines encouraged student-to-student talk during discussions. Instead of the teacher pressing or probing, students would engage in these talk moves with each other. From our data, this level of student-driven responsive discourse was only observed in Isabel's class most prominently and during the debate about current flow. Other research describes how students actively take on this kind of responsive stance toward their peers, giving them more authority over content and understanding that they will be held accountable to address the problem at hand by their peers (Engle & Conant, 2002; Rogat, Witham, & Chinn, 2014).

CONCLUSION

Responsiveness in discourse plays a key role in the larger framework of ambitious and equitable science teaching (Windschitl & Calabrese-Barton, 2016). This study provided evidence that a collective attention to students' thinking in the classroom can support rigorous kinds of intellectual work that persist and build over multiple lessons. We note that our analysis looked only at whole-class conversations; however, important science concepts were also the focus of small group and partner tasks and individual assignments. The whole-group analysis then does not show the full picture of how these ideas were addressed in each class. Our findings and those from a previous study (Thompson et al., 2016) support the claim that authentic forms of rigor in classroom talk are more likely to occur when teachers are responsive to students' ideas. This investigation suggests that having a toolkit of discourse moves is necessary but not sufficient to support ongoing sense making. Norms for productive talk are important for students to break from the inherited routines of recitation-style participation in classroom talk. Having a scientifically rich phenomenon that requires the building of an explanation over time also seemed to be an equally important condition of instruction to support discourse. Preparing teachers to do such complex work will require them to develop a theory of practice around sense making and science ideas that can organize their thinking about working with students in new ways. It will require deeper content knowledge to anticipate student contributions and to respond in ways that prompt further reasoning. And finally, attempts at responsive teaching cannot be undertaken in isolation from other professionals. The work requires feedback from one's colleagues and the opportunity to observe others attempting this pedagogy. In this study, we benefitted from our researcher–practitioner partnerships—it allowed us to take risks together and we would encourage such collaborations as a way for theory and practice to be developed together.

REFERENCES

Abrahams, I., & Reiss, M. (2012). Practical work: It's effectiveness in primary and secondary schools in England. Journal of Research in Science Teaching, 49(8), 1035–1055.

Ball, D. L., Lubienski, S., & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), Handbook of research on teaching (4th ed., pp. 433–456). New York, NY: Macmillan

Bell, P. (2004). Promoting students' argument construction and collaborative debate in the science classroom. In M. C. Linn, E. A. Davis, & P. Bell (Eds.), Internet environments for science education (pp. 115–143). Mahwah, NJ: Erlbaum.

Bereiter, C. (1994). Implications of Postmodernism for science, or, science as progressive discourse. Educational Psychologist, 29(1), 3–12.

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

Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), Innovations in learning: New environments for education (pp. 289–325). Mahwah, NJ: Erlbaum.

Buckland, L. (2008). Seventh-graders' epistemic criteria for model-based reasoning. Paper presented at the annual meeting of the American Educational Research Association, March 2008, New York, NY.

Calabrese-Barton, A., & Tan, E. (2009). Funds of knowledge, discourses and hybrid space. Journal of Research in Science Teaching, 46(1), 50–73.

Cazden, C. B. (1986). Classroom discourse. In M. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 432–463). New York, NY: Macmillan.

Chin, C. (2006). Classroom interaction in science: Teacher questioning and feedback to students' responses. International Journal of Science Education, 28(11), 1315–1346.

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.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. Review of Educational Research, 64(1), 1–35.

Corcoran, T., & Gerry, G. (2011). Science instruction in Newark Public Schools.Consortium for Policy Research in Education. Research report # RR-71. Teachers College, Columbia University, New York, NY.

Danish, J. A., & Enyedy, N. (2006). Unpacking the mediation of invented representations. International Conference of the Learning Sciences, Proceedings, 1, 113–119.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. Educational Researcher, 34(3), 3–14.

Delpit, L. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. Harvard Educational Review, 58(3), 280–298.

Derry, S. J., Hmelo-Silver, C. E., Nagarajan, A., Chernobilsky, E., & Beitzel, B. D. (2006). Cognitive transfer revisited: Can we exploit new media to solve old problems on a large scale?. Journal of Educational Computing Research, 35(2), 145–162.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., & Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. Journal of the Learning Sciences, 19(1), 3–53

diSessa, A. (1993). Towards an epistemology of physics. Cognition and Instruction, 10(2–3), 105–225.

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. Review of Research in Education, 32(1), 268–291.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. Journal of the Learning Sciences, 8(3-4), 391–450

Edwards, D., & Mercer, N. (1987). Common knowledge: The development of joint understanding in the classroom, London, England: Methuen.

Engle, R. A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. Journal of the Learning Sciences, 15(4), 451–498.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. Cognition and Instruction, 20(4), 399–483.

Ford, M., & Wargo, B. (2011). Dialogic framing of scientific content for conceptual and epistemic understanding. Science Education, 96(3), 369–391.

Forman, E., & Ansell, E. (2001). The multiple voices of a mathematics classroom community. Educational Studies in Mathematics, 46, 115–142. doi:10.1007/0-306-48085-9_4

Hammer, D. (1997). Discovery learning and discovery teaching. Cognition and Instruction, 15(4), 485–529.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (pp. 169–190). Mahwah, NJ: Erlbaum.

Harris, C. J., Phillips, R. S., & Penuel, W. R. (2012). Examining teachers' instructional moves aimed at developing students' ideas and questions in learner-centered science classrooms. Journal of Science Teacher Education, 23(7), 769–788.

Herrenkohl, L. R., Palincsar, A. S., DeWater, L. S., & Kawasaki, K. (1999). Developing scientific communities in classrooms: A sociocognitive approach. Journal of the Learning Sciences, 8(3 & 4), 451–493. doi:10.1080/10508406.1999.9672076

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). Educational Psychologist, 42, 99–107. doi:10.1080/00461520701263368

Jadallah, M., Anderson, R., Nguyen-Jahiel, K., Miller, B., Kim, I.-H., Kuo, L.-J., ... Xiaoying, W. (2011). Influence of a teacher's scaffolding moves during child-led small-group discussions. American Educational Research Journal, 48(1), 194–230.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. American Educational Research Journal, 27(1), 29–63.

Lee, H. S., & Songer, N. B. (2003). Making Authentic Science Accessible to Students. International Journal of Science Education, 25(1), 1–26.

Lehrer, R., & Schauble, L. (2005). Developing modeling and argument in elementary grades. In T. A. Romberg, T. P. Carpenter, & F. Dremock (Eds.) Understanding mathematics and science matters. (pp. 29–53). Mahwah, NJ: Erlbaum

Lehrer, R., & Schauble, L. (2012). Seeding evolutionary thinking by engaging children in modeling its foundations. Science Education, 96(4), 701–724.

Leinhardt, G., & Steele, M. (2005). Seeing the complexity to standing to the side: Instructional dialogues. Cognition and Instruction, 23(1), 87–163.

Levin, D. M., Hammer, D., & Coffey, J. (2009). Novice teachers' attention to student thinking. Journal of Teacher Education, 60(2), 142–154. doi:10.1177/0022487108330245

Linn, M. C., & Hsi, S. (2000). Computers, teachers, peers: Science learning partners. Mahwah, NJ: Erlbaum.

Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. Educational Psychologist, 39(1), 57–68.

Magnusson, S., & Palincsar, A. (2005). Teaching to promote the development of scientific knowledge and reasoning about light at the elementary school level. In M. S. Donovan & J. Bransford (Eds.) How students learn science in the classroom (pp. 421–474). Washington DC: National Academies Press.

Maskiewicz, A., & Winters, V. (2012). Understanding the co-construction of inquiry practices: A case study of a responsive teaching environment. Journal of Research in Science Teaching, 49 (4), 429–464.

May, D. B., Hammer, D., & Roy, P. (2006). Children's analogical reasoning in a 3rd-grade science discussion. Science Education, 90(2), 316–330.

McNeill, K. L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. Journal of Research in Science Teaching, 45(1), 53–78.

Mehan, H. (1979). Learning lessons: Social organization in the classroom. Cambridge, MA: Harvard University Press.

Mercer, N. (2008). The seeds of time: Why classroom dialogue needs a temporal analysis. Journal of the Learning Sciences, 17(1), 33–59.

Merriam, S. (2009). Qualitative research: A guide to design and implementation. San Francisco: Jossey-Bass.

Michaels, S., O'Connor, C., & Resnick, L. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. Studies in Philosophy and Education. 27(4), 283–297.

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis (2nd ed.). Thousand Oaks, CA; Sage.

Minstrell, M., & Kraus, P. (2005). Guided inquiry in science classrooms. In M. S. Donovan & J. Bransford (Eds.) How students learn science in the classroom, (pp. 475–514). Washington DC: National Academies Press.

Minstrell, J., & van Zee, E. H. (2003). Using questioning to assess and foster student thinking. In J. M. Atkin & J. E. Coffey Everyday assessment in the science classroom. Arlington, VA: National Science Teachers Association Press.

Nasir, N. S., Rosebery, A. S., Warren, B., & Lee, C. D. (2006). Learning as a cultural process: Achieving equity through diversity. In R. K. Sawyer (Ed.), The Cambridge handbook of the learning sciences (pp. 489–504). New York, NY: Cambridge University Press.

Mortimer, E., & Scott, P. (2003). Meaning making in secondary science classrooms. Maidenhead, England: Open University Press.

National Research Council (2007). Taking science to school: Learning and teaching science in grades K-8. Committee on Science Learning, Kindergarten Through Eighth Grade. Duschl R. A., Schweingruber H. A., & Shouse A. W., (Eds.). Board on Science Education, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: The National Academies Press.

NGSS Lead States (2013). Next Generation Science Standards: For states, by states. Washington, DC: The National Academies Press.

Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S., & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. Discourse Processes, 35(2), 135 – 198.

Osborne, J., & Dillon, J. (2008). Science education in Europe: Critical reflections.. Kings College London, England: A Report to the Nuffield Foundation.

Passmore, C., & Stewart, J. (2002). A modeling approach to teaching evolutionary biology in high schools. Journal of Research in Science Teaching, 39(3), 185 – 204.

Pierson, J. (2008). The relationship between patterns of classroom discourse and mathematics learning. Unpublished Doctoral Dissertation, The University of Texas at Austin, TX.

Pimentel, D. S., & McNeill, K. L. (2013). Conducting talk in science classrooms: Investigating instructional moves and teachers' beliefs. Science Education, 97(3), 367 – 394.

Radinsky, J., Oliva, S., & Alamar, K. (2010). Camila, the earth, and the sun: Constructing an idea as shared intellectual property. Journal of Research in Science Teaching, 47(6), 619 – 642.

Richards, J., & Robertson, A. D. (2015). A review of the research on responsive teaching in science and mathematics. In A. D. Robertson, R. E. Scherr, & D. Hammer (Eds.), Responsive teaching in science. New York, NY: Routledge.

Reiser, B. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. Journal of the Learning Sciences, 13(3), 273 – 304.

Rogat, T. K., Witham, S. A., & Chinn, C. (2014). Teachers' autonomy-relevant practices within an inquiry-based science curricular context: Extending the range of academically significant autonomy-supportive practices. Teachers College Record, 116(7), 1 – 46.

Rosebery, A., & Warren, B. (Eds.). (2008). Teaching science to English language learners: Building on students' strengths. Arlington, VA: NSTA Press.

Rosebery, A. S., Warren, B., & Conant, F. R. (1992). Appropriating scientific discourse: Findings from language minority classrooms. Journal of the Learning Sciences, 2(1), 61 – 94.

Roth, K. J., Druker, S. L., Garnier, H. E., Lemmens, M., Chen, C., Kawanaka, T., & Gallimore, R. (2006). Teaching science in five countries: Results from the TIMSS 1999 video study (NCES 2006-11). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Roth, K., Chen, C., Lemmens, M., Garnier, H., Wickler, N., Atkins, L., & Zembal-Saul, C. (2009). Coherence and science content storylines in science teaching: Evidence of neglect? Evidence of effect? Colloquium and paper presented at the annual meeting of the National Association for Research in Science Teaching (NARST). Garden Grove, CA.

Rowe, S. (2011). Discourse in activity and activity as discourse (revised). In R. Rogers (Ed). An introduction to critical discourse analysis in education (2nd ed., pp. 227 – 241). New York, NY: Routledge.

Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. Educational Assessment, 11(3&4), 205 – 235.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. Science Education, 88(3), 345 – 372.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, technology. In R. K. Sawyer (Ed.), The Cambridge handbook of the learning sciences (pp. 97 – 118). New York, NY: Cambridge University Press.

Sherin, M. G., & Van Es, E. (2009). Effects of video club participation on teachers' professional vision. Journal of Teacher Education, 60, 20 – 37.

Sfard, A., & McClain, K. (2002). Analyzing tools: Perspectives on the role of designed artifacts in mathematics learning. Journal of the Learning Sciences, 11(2 & 3), 153 – 161.

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. Educational Research and Evaluation, 52(5), 659 – 685.

Stroupe, D. (2014). Examining classroom science practice communities: How teachers and students negotiate epistemic agency and learn science-as-practice. Science Education, 98(3), 487 – 516.

Sykes, G., Bird, T., & Kennedy, M. (2010). Teacher education: Its problems and some prospects. Journal of Teacher Education, 61(5), 464 – 476.

Thompson, J., Hagenah, S., Kang, H., Stroupe, D., Braaten, M., Colley, C., & Windschitl, M. (2016). Rigor and responsiveness in classroom activity. Teachers College Record, 118(5). 1 – 58. Retrieved from http://www.tcrecord.org/Home.asp

Van Zee, E. H., Hammer, D., Bell, M., Roy, P., & Peter, J. (2005). Learning and teaching science as inquiry: A case study of elementary school teachers' investigations of light. Science Education, 89(6), 1007–1042.

Walqui, A., & Van Lier, L. (2010). Scaffolding the academic success of adolescent English language learners: A pedagogy of promise (pp. 1–41). San Francisco, CA: WestEd.

Weiss, I., Pasley, J., Smith, S., Banilower, E., & Heck, D. (2003). Looking Inside the Classroom: A study of K–12 mathematics and science education in the United States, Chapel Hill, NC: Horizon Research.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. Cognition and Instruction, 16(1), 3–118. doi:10.1207/s1532690xci1601_2

Windschitl, M. (2013). The beginner's repertoire: The design and testing of core instructional practices for teacher preparation. Presented at American Educational Research Conference, San Francisco, May 2013.

Windschitl, M., & Calabrese-Barton, A. (2016). Rigor and equity by design: Seeing a core of practices for the science education community. In D. Gitomer & C. Bell (Eds.) AERA handbook of research on teaching, 5th Edition (pp. 1099–1158). American Educational Research Association.