Designing and Implementing Cognitively Demanding Science Tasks for Fostering Productive Disciplinary Engagement

Miray Tekkumru-Kisa and Ozlem Akcil-Okan
mtekkumrukisa@fsu.edu, oakcilokan@fsu.edu
Florida State University

Abstract: Recent ambitious instructional reforms in science education emphasize developing classroom learning environments that promote students' productive disciplinary engagement (PDE). Many teachers struggle with how to design classroom cultures that support PDE. This study proposes that PDE can be fostered through the design and effective implementation of cognitively demanding science tasks. In our analysis of the implementation of two contrasting lessons by the same teacher with the same group of students, we focused on how the teacher, students, and the task work together in a classroom activity system to support students' PDE. Going beyond what PDE looks like, the study findings shed light on how it is fostered and the pedagogical factors that derive it.

Introduction

One of the major questions in current science education research is how to productively engage students in the discourses and practices of science in ways that are supportive of their "grasp of practice" (Ford, 2008) and effective for their understanding of core science ideas. Consistently, recent science education reforms in the United States since the release of the Framework for K-12 Science Education (NRC, 2012) embrace a vision that integrates science content with scientific practices. One of the key principles emerging from the NRC (2015) report to guide the implementation of this vision is engaging students in classroom cultures that support productive disciplinary engagement (Anderson et al., 2018; Engle, 2012). However, many teachers struggle with how to engage students in science in authentic ways and simultaneously move the class as a whole toward the development of core ideas. As an attempt to understand what it takes for teachers to promote these rich learning environments that can provide the opportunity for students' deep involvement in and facility with scientific concepts and practices, we will focus on a chemistry teacher's class by situating our exploration in a theoretical frame that addresses how productive disciplinary engagement (PDE) can be fostered in a science classroom (Engle and Conant, 2002; Engle, 2012). Through an in-depth analysis of the implementation of two contrasting lessons by the same teacher with the same group of students, we focus on how the teacher, students, and the task work together in a classroom activity system (Greeno, 2006) to support students' PDE.

PDE has been used to study student engagement and disciplinary learning (e.g., Engle, 2012; Mortimer & de Araújo, 2014). How teachers can create learning environments that support PDE has received much less attention (Forman et al., 2014). As the recent instructional reforms in science education are calling for PDE in science classrooms, our study addresses this need by exploring how it can be fostered in science classrooms and how to engineer the learning environment to support PDE. We posit the idea that the design and effective implementation of cognitively demanding science tasks can support facilitating PDE in science classrooms, an area that was overlooked in the field of science education. For us, cognitive demand is about the degree to which opportunities provided for students with the tasks they are assigned create and maintain intellectual challenge that is conducive to students' high-level thinking and sensemaking through meaningful engagement in scientific practices (Tekkumru-Kisa, Stein, & Schunn, 2015). As the field seeks ways to create opportunities for students' PDE in science classrooms, we argue that introducing cognitively demanding tasks into the design and enactment of science instruction can provide opportunities for PDE in science classrooms.

Theoretical framework

A fundamental aspect of the recent reforms in the United States is that science instruction has to provide opportunities for students to use scientific ideas and practices in service of sensemaking, which occurs as students try to figure out and offer possible mechanisms for why a phenomenon occurs (Achieve, 2019; Odden & Russ, 2019). This sensemaking perspective positions students as capable of asking questions, constructing and critiquing knowledge claims in the classroom, which are critical for PDE, a comprehensive approach to characterize students' engagement in disciplinary ideas and practices (Engle, 2012; Haverly et al., 2018; Penuel et al., 2019).

Engle and Conant (2002) defined PDE as "students' deep involvement in, and progress on, concepts and/or practices characteristic of the discipline they were learning about" (p. 403). They presented four principles of fostering PDE. First is *problematizing*, which requires engendering uncertainty to motivate students to take on the explanation of phenomena. Problematizing is shown to be essential to students' inquiry in science classrooms (Phillips, Watkins, & Hammer, 2018). Second, learning environments need to give students the *authority* to define and address problems and reposition them as knowledge producers in the classroom. Third, students should be held *accountable* to others and disciplinary norms. Finally, it is necessary to provide relevant *resources* to support students' sensemaking. As Suarez (2017) stated, "the realization of PDE relies on the dynamic balance between all four principles, with some easily linked due to their contributions to students' sensemaking." (p. 57)

Although there are accounts of PDE in prior research (e.g., Engle, 2012; Kumpulainen, 2014), what remains missing in the literature is how to the design learning environments that can support PDE and what resources can allow opportunities for students to engage in PDE (Forman et al., 2014; Kelly, 2014). Engle and Conant (2002) emphasize the importance of materials and intellectual resources to support students in their sensemaking. We conjecture that using cognitively demanding science tasks as a resource and implementing them by drawing on students' intellectual resources can help to foster PDE in science classrooms. Prior research provided evidence for the role of attention to students' ideas in facilitating students' engagement in cognitively demanding tasks (Tekkumru-Kisa et al., 2019). Therefore, selecting and effectively implementing cognitively demanding tasks can provide tangible practices for teachers to promote PDE.

Science tasks can be considered as the context within which students think about and intellectually engage in disciplinary ideas and practices (Doyle, 1988). A science task could be a project, an investigation, or a question to explain—any academic work that students are given to engage in the classroom. Not all tasks provide similar opportunities for students to engage in scientific ideas and practices (Tekkumru-Kisa et al., 2015). In our earlier work, we developed Task Analysis Guide in Science (TAGS) to analyze the level and kind of student thinking demanded by a science task (Tekkumru-Kisa et al., 2015). Science tasks that are categorized into the higher-level categories of the TAGS (i.e., levels 4 and 5) require students' high-level thinking in core ideas, crosscutting concepts, and meaningful engagement in the scientific practices as they make sense of phenomena. In other words, these high-level tasks require students to make sense of how and why a phenomenon happens. When these tasks are effectively enacted in the classroom, students would be able to experience their intellectual work as figuring something out. As Berland and colleagues (2015) stated, "…if a student were asked: 'Why are you doing this activity?' they would say, 'To help us figure out how and why [a particular phenomenon] happens,' rather than, 'Because the teacher (or worksheet) asked us" (p. 1086).

While using cognitively demanding tasks offers a strong promise for engaging students' in high-level thinking and sensemaking, research has consistently shown that it is not easy to implement these tasks in the classrooms effectively. Cognitively demanding tasks are often transformed once they are placed into real classroom settings; they have a history of being proceduralized in science classrooms (e.g., Kang et al., 2016; Tekkumru-Kisa, Schunn, Stein, & Reynolds, 2019). These transformations result in students' superficial engagement in disciplinary ideas and practices (e.g., pseudoargumentation, Berland & Hammer, 2012). Based on the consistent patterns observed in science and mathematics classrooms, the Task Framework in Figure 1 was developed. It helps to identify and describe the changes observed in students' thinking across the phases of a task that is (1) the task as designed (e.g., as it appears in written materials) characterizes the potential level and kind of thinking in which students are invited to engage; (2) the task as announced by the teacher characterizes the framing of the intellectual work for the students; (3) the task as perceived by each student and as enacted by the teacher and the students, that is the actual intellectual work in which students engage (Doyle, 1988; Stein & Smith, 1998; Tekkumru-Kisa et al., 2019).

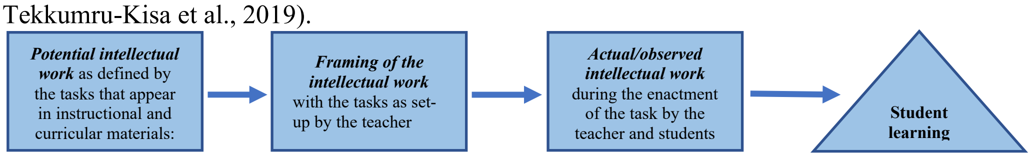

Figure 1. The Task Framework.

Therefore, while selecting cognitively demanding tasks is essential, it is not sufficient to engage students in high levels of thinking; the interaction between the teacher and students during the launch and enactment phases of these tasks maintains or reduces the demand on students' thinking (Tekkumru-Kisa et al., 2019). Thus, we argue that not only the design of cognitively demanding tasks as a resource that can be used in science classrooms, but also their effective enactment by drawing on students' intellectual resources in the classroom would support fostering PDE. To uncover this relationship between cognitive demand and PDE, we conducted an in-depth analysis of two contrasting science lessons implemented by the same teacher with the same group of students. Uncovering this relationship is important because while PDE principles helped to explain students' engagement during the rich educational innovations, this work provides little guidance to teachers and teacher educators in terms of how to integrate and actualize these ideas in the classroom (Williams-Candek, 2015). This study offers one possible means to put PDE into practice through teachers' creating learning environments with the design and effective enactment of cognitively demanding science tasks.

Methods

This study emerged from a genuine puzzling teaching phenomenon, which was about the difference between two lessons enacted by the same teacher in the same classroom in close proximity in time: Bending Water Lesson and Equilibrium Lesson. While both of the lessons were designed by the same teacher within the context of a professional development (PD) that aims to promote productive science discourse, they were experienced differently both by the teacher and the students. In the interviews, neither the teacher nor the students found the Equilibrium Lesson as thought-provoking as the Bending Water Lesson. Mr. Smith thought that among the four lessons that he co-designed in the PD, "the worst being the one about equilibrium." About the Bending Water Lesson, he said, "I think the one that I enjoyed the most that was the one where we were bending water with the household things." His students brought attention to a similar contrast among the lessons. One of the students said, "This lab [Equilibrium Lesson] probably did not help me to understand equilibrium, but the notes that we went through were more useful". Similarly, another student said, "I do not remember this lab...I do remember that we add something that makes the color change, but I do not remember what we tried to solve...I could not explain this one." However, regarding the Bending Water lesson, the same student said that "This [bending water] made a lot of more sense after the lab for sure. I was like I could explain that; it made sense."

Contrasting cases can help to make the distinctive features between the cases more salient (e.g., Bransford & Schwartz, 1999). We thought that this contrast that was expressed consistently by the teacher and students between the two lessons could help to uncover the differences in teaching and learning that was experienced and fostered in this classroom. Consistently, this case study (Creswell, 2003) was guided by two research questions: (1) How did the level and kind of student thinking change across the phases of a science task (i.e., tasks as designed, launched, and enacted) throughout the trajectory of the focal contrasting lessons in a chemistry classroom? (2) How did PDE differ between the focal contrasting lessons in a chemistry classroom?

Context

The study was conducted as part of an NSF-funded project that focuses on supporting teachers' learning to facilitate productive talk in science classrooms. The project involves the design of a PD structured around science teachers' co-designing, teaching, and reflecting on science lessons to learn to facilitate productive science talk in their classrooms. PD started in summer 2018; four of the teachers from summer agreed to participate in a year-long PD during the 2018-2019 academic year, which consisted of four cycles of Design-Teach-Analyze sessions. In the Design sessions, teachers co-designed lessons to promote productive discourse in their classrooms. Then, they taught these lessons, which were video recorded by the research team. Next, in the Analyze sessions, teachers analyzed 5-10 min video clips depicting their implementation of the co-designed lessons that were identified by the research team. Our focal teacher in this study, Mr. Smith, was a chemistry teacher with five years of teaching experience. As part of the PD, he partnered to co-design his lessons with the second author, who has nine years of chemistry teaching experience. In their collaboration, Mr. Smith took the primary role in the design of lessons and made the final decision about their features. He chose to implement these lessons in his 11th grade Advanced Placement Chemistry, which consisted of 22 students (14 White, 4 African American, and 4 Asian).

Contrasting lessons

The *Bending Water Lesson* focused on intermolecular forces and was guided by the question: "Why and how does the water bend towards the charged PVC rod?" Throughout this 3-day lesson, the main task was designing and conducting an investigation to explain a phenomenon. The lesson was launched with a demonstration that showed the bending of a water stream by a charged rod. The students were asked to explain how and why it happened. After briefly brainstorming in their small groups, they had a whole group discussion about their claims for why and how they thought the water stream bent, and how they would go about testing their claims. After this discussion in the launch phase of the lesson, students started to work in small groups to design and conduct an investigation to test their claims. The *Equilibrium Lesson* was related to chemical equilibrium and focused on identifying various stresses that can cause a shift in a chemical equilibrium system. The guiding question for this lesson was, "What is the equilibrium position for each given chemical equilibrium system?" The lesson was launched with a video of the Briggs-Rauscher reaction mechanism, followed by a discussion about why the reaction oscillates between its different colors. Next, Mr. Smith introduced three chemical equilibrium reaction systems which were *not* related to the reactions in the video. Students were asked to provide various stresses that would lead to shifts in the equilibrium systems by using their prior knowledge and exposure to this topic. After this brief discussion, Mr. Smith wrote a list of possible stresses on the board. Then, the students were tasked with deciding the specific tests they would conduct to determine the equilibrium position of each system by drawing on their lecture notes. The enactment of the tasks throughout each of these three-day lessons was separated into four parts based on the nature of the work in which students engaged, as summarized in Table 1. This allowed us to conduct a fine-grained analysis that we will describe next.

Table 1: Details about the enactment phases of the focal contrasting chemistry lessons

	Enactment	Time segments	Brief description of activities
Bending Water Lesson	Enact #1	Day 1: ~5 min Small groups ~21 min Whole group	Discussion of students' claims, explanations for the phenomenon and plans for how to test them.
	Enact #2	Day 1: ~7 min Small groups Day 2: ~11 min Whole group	Designing investigations to test students' claims regarding the phenomenon.
	Enact #3	Day 2: ~22 min Small groups Day 3: ~13 min Small groups	Conducting investigations to test their claims; developing explanations and arguments.
	Enact #4	Day 3: ~15 min Whole group	Presenting posters with their arguments and developing a consensus for the mechanism of the phenomenon.
Equilibrium Lesson	Enact #1	Day 1: ~10 min Small groups	Selecting the tests to apply for shifting the equilibrium positions of the given reaction systems.
	Enact #2	Day 2: ~35 min Small groups	Conducting experiments and data collection.
	Enact #3	Day 2: ~14 min Small groups Day 3: ~19 min Small groups	Interpreting the findings to determine the equilibrium position and preparing posters with the findings.
	Enact #4	Day 3: ~27 min Whole group	Presenting posters regarding the equilibrium position.

Data collection and analysis

The data sources that we draw on to answer the study questions were the lesson artifacts, which included planning documents and tasks assigned to students, video-records of the implementation in Mr. Smith's classroom, and the interviews with Mr. Smith and a subset of his voluntary students about these focal lessons.

In the analysis of the lesson artifacts and videos, we used several analytical lenses. First, we used three of the Instructional Quality Assessment in Science (IQA-Science) rubrics (Tekkumru-Kisa, Preston, Kisa, Oz, & Morgan, 2019), which were developed based on the TAGS and the Task Framework (see Figure 1) to provide a nuanced evidence about the level and type of thinking in each phase of a task. In IQA-Science, (1) Potential rubric assesses the level and kind of thinking (i.e., cognitive demand) in which students could potentially engage as they work on the assigned science task; (2) Launch rubric identifies the framing of the intellectual work for students in set-up of the lesson; (3) Implementation rubric assesses the level and type of thinking that students engage during the enactment of the task. In our analysis using these rubrics, we focused on the main task that students were assigned, how it was framed for the students at the beginning of the lesson (i.e., lesson launch), and how it was enacted in the rest of the lesson (see enactment parts in Table 1). These analyses addressed the first research question, which is about change in level and kind of student thinking throughout the trajectory of a lesson.

The next set of analyses addressed the second research question, which is about students' PDE during the implementation of the focal lessons. Analysis of the main science task that students were assigned in each lesson provided insights into the *resources* and how these *resources* provided opportunities to *problematize* phenomena regarding two essential principles of PDE. To provide further insight into PDE, we also focused on the *authority* and *accountability* given to students during the enactment phase of each lesson. We modified from the literature a rubric for *intellectual authority* (Stein & Kaufman, 2010) to assess the extent to which students were positioned to have the authority to reason with the scientific ideas and practices to solve the problems and produce knowledge as suggested by Engle and Conant (2002). Intellectual authority codes ranged from low (the teachers keep the authority on themselves for deciding what ideas are right or wrong; the judgments of correctness derived from teacher or text) to high (students are positioned as critical thinkers that have the authority to reason the scientific ideas and judgments of correctness derived from scientific reasoning). These codes are consistent with students' epistemic agency (e.g., Stroupe, 2014) and epistemic authority in relation to PDE (e.g., Forman & Ford, 2014; Sandoval et al., 2019). To assess the accountability to others in the classroom and to the disciplinary norms for knowledge building within the context of the classroom talk during the enactment phase of the lessons, we modified the accountable talk rubrics (Boston, 2012; Resnick et al., 2010). They were developed to assess accountability to the learning community, knowledge, and reasoning by considering the nature of the teacher's and students' contributions in the talk (i.e., Boston, 2012; Engle 2012; Resnick et al., 2010). Both of the researchers independently examined the lesson artifacts attending to the criteria described in the rubrics. They watched the video recordings of the lessons, took notes with detailed descriptions of what they noticed in the videos about PDE and cognitive demand on students' thinking, and rated the lessons based on the rubrics described above. Next, they discussed their coding by providing justification for their codes from their notes. This discussion led to the consensus coding with detailed justifications (Maxwell, 2012) that we summarize in the next section. Analysis of the interviews focused on Mr. Smith's and his students' comments about the focal lessons to further explain the patterns observed in videos. We read through the transcripts, highlighted the teacher's and students' comments about the focal lessons, and identified the common themes across these comments that explain their engagement during the two focal lessons.

Findings

Consistent with the differences that Mr. Smith and students consistently noted in their reflections on the focal lessons, our analysis revealed a clear disparity in how PDE was experienced and fostered in these contrasting lessons. This contrast in PDE patterns is also consistent with the cognitive demand of the tasks designed by Mr. Smith for these lessons, and the level and kind of thinking that students engaged in during their enactment.

Addressing the first research question, demand on students' thinking differed throughout the trajectory of the Bending Water Lesson, and Equilibrium Lesson. The level and kind of student thinking was maintained at a high-level throughout the trajectory of the Bending Water Lesson (see Figure 2). The cognitive demand of the task was rated as level-5 based on the potential demand of a task rubric of IQA-Science. According to this rubric, tasks at level-5 have the potential to engage students in sensible versions of the actual intellectual work in which scientists engage; to develop a sense of and experience how scientific knowledge is produced. By maintaining the potential demand of the task on students' thinking, Mr. Smith launched it with a demonstration of a puzzling phenomenon that is bending of a water stream by a charged rod and asking students to explain how and why that happened. In this launch phase, which was also rated high based on the IQA-Science rubric for task launch, the lesson was framed as figuring out the bending water phenomenon, and this phenomenon guided students' exploration in the rest of the lesson. During all four parts of the enactment phase, students were positioned as sense-makers as they tried to explain, with the data that they gathered, the mechanisms behind bending of the water stream by/towards the charged rod. Thus, the level and kind of student thinking in all parts of the enactment was rated as level-5 based on the IQA-Science implementation rubric, which indicates that students continued to think and act like scientists while experiencing how scientific knowledge is generated. These patterns also revealed that problematizing of a phenomenon was integrated into the design of the Bending Water Lesson, it was used to engender uncertainty to motivate students to take on explaining this phenomenon both in the launch and the enactment phases throughout the entire lesson.

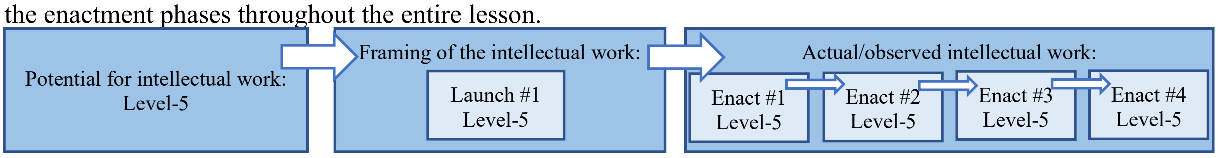

Figure 2. The level and kind of student thinking throughout the trajectory of the Bending Water Lesson.

As opposed to the maintenance of level and kind of students thinking throughout the trajectory of the Bending Water Lesson, the Equilibrium Lesson provided limited opportunities for students' thinking (see Figure 3). The cognitive demand of the task was rated as level-3 based on the potential demand of a task rubric of IQA-Science. This means, even though the task had the potential to engage students in high-level thinking processes such as finding relations or analyzing information, understanding the science content was foregrounded in the task without providing a robust opportunity for doing science. Level-3 tasks do not position students to understand both the pursuit (i.e., seeking an explanation for a phenomenon) and the body of knowledge that results from that pursuit. Consistently, the task of the Equilibrium Lesson emphasized learning about chemical equilibrium rather than using this idea to figure out a phenomenon. While launching the lesson, Mr. Smith tried to increase the cognitive demand of the task by creating uncertainty through a puzzling phenomenon. Specifically, after showing the Briggs-Rauscher reaction video, he asked students why the reaction oscillated between its different colors. Thus, we coded this first part of the task launch as level-4 based on the IQA-Science launch rubric. Students could mostly explain the color change in the reaction shown in the video. Without getting into the mechanisms in play, Mr. Smith told students that it was a complex phenomenon, so they would not be able to explain it entirely, which resulted in shifting the intellectual work for the students. In the next part of the launch phase, Mr. Smith declined the demand on students' thinking to level-3 by shifting the conversation to the introduction of a verification lab about shifting chemical equilibrium that they covered in the previous lesson. He framed the work for the students as finding the equilibrium position of the given reaction systems by applying several stresses to shift the equilibrium position. Therefore, the assigned science task, combined with its launch, did not promote problematizing of phenomena because students did not appear to experience an uncertainty around a phenomenon that they were puzzled with and tried to explain. During all four parts of the enactment phase, the level and kind of student thinking were maintained at level-3; students conducted experiments to verify theoretical expectations informed by the science content with which they were presented before. The enactment phase helped students to make use of their lecture notes on chemical equilibrium, Le Chatelier's principle, stresses on the dynamic equilibrium to verify how different stresses shift the equilibrium systems in chemical equations provided.

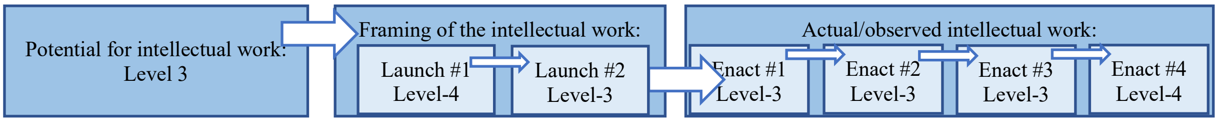

Figure 3. The level and kind of student thinking throughout the trajectory of the Equilibrium Lesson.

All in all, addressing the first research question, our analysis revealed a clear difference in the level and kind of student thinking in the potential, launch, and enactment of the main science tasks used in these two lessons. The cognitive demand of these tasks designed by Mr. Smith afforded different opportunities for students to engage in the core disciplinary ideas and science practices. The demand of these tasks was maintained during their launch and enactment by sustaining the contrast between these two lessons (Figures 2 and 3).

Analysis addressing the second research question about the PDE revealed similar differences between these contrasting lessons. The patterns that we summarized above started to reveal the differences in how PDE was fostered in these two lessons. As these patterns indicate, designing tasks at different cognitive demand levels provided different opportunities for students' engagement in scientific ideas and practices. While the design of the science task and how it was launched through a puzzling phenomenon engendered uncertainty that students tried to explain throughout the trajectory of the Bending Water Lesson, this was not the case in the Equilibrium Lesson. The chemical reaction task and how it was launched did not offer sufficient opportunities to problematize a phenomenon for students to figure out throughout the Equilibrium Lesson. These differences between the lessons in how PDE was fostered was supported by the authority and accountability ratings presented in Table 2.

Table 2. Summary of the intellectual authority and accountability codes for enactment of focal chemistry lessons

	Bending Water Lesson				Equilibrium Lesson			
	Enact #1	Enact #2	Enact #3	Enact #4	Enact #1	Enact #2	Enact #3	Enact #4
Students' intellectual authority	4	4	4	4	2	3	3	3
Accountability to community:								
Teacher Linking	4	1	3	4	1	1	2	3
Student Linking	4	4	4	4	3	3	3	4
Accountability to knowledge & reasoning:								
Teacher Press	4	4	4	4	1	1	3	4
Student Contribution	4	4	4	4	2	3	3	4

** The ratings are on a 4-point scale: Low(1)- Medium(2)- Moderately High(3)- High(4).

Throughout the enactment of the Bending Water Lesson, students were given opportunities to achieve authority for producing knowledge by developing explanations of a puzzling phenomenon, as reflected in the intellectual authority ratings presented in Table 2. Students were given the agency to design an investigation to test their claims and develop arguments and critique each other's claims. They were positioned as critical thinkers and active participants that had the authority to reason with the scientific ideas; judgments of correctness derived from scientific and collective reasoning. Students consistently appeared to hold the authority to construct knowledge claims; one group, for instance, claimed, "The greater electrostatic force, the greater the bending is, and the more polar a substance is, the more likely it is going to bend." In addition, accountability ratings were mostly coded high both for the teacher and students during all the parts of the enactment phase of the Bending Water Lesson. Except in Enact-2 (see Table 2), Mr. Smith explicitly linked students' contributions to each other; he consistently asked students for their reasoning and to justify their claims with evidence. For example, he mostly asked the following questions during the discussions: "Why do you think it is?","What is that reason again?", "Are there any other ideas for why we are seeing this?", "Does anyone else have another idea that is different?" Similarly, his students carefully listened to each other, discussed alternative arguments, and built on each other's ideas while making sense of the observed scientific phenomenon. These patterns provided further evidence for the opportunities for PDE that was fostered and engendered throughout the Bending Water Lesson.

The Equilibrium Lesson, on the other hand, included limited opportunities for students to engage in PDE, as is evident in the patterns presented in Table 2. Students' intellectual authority was mostly scored as moderately high, indicating that although students were positioned as critical thinkers and had the authority to judge the correctness and solve the problems, the assessment of correctness sometimes derived by the teacher or the text.

Consistently, Mr. Smith's contributions during the talk, as evident in the teacher press and linking codes in Table 2, were mostly rated as low or medium for his accountability to the community, knowledge, and reasoning. These indicate that Mr. Smith did not regularly ask students to explain their arguments by presenting their reasoning, evidence. Also, he did not do much of connecting students' ideas to promote discussion of ideas and common understanding. These patterns in the accountability ratings were also consistent for students' contributions to the talk. Relative to Mr. Smith, students did more work to link their ideas with their peers, and to contribute with evidence and reasoning during their talk throughout the Equilibrium Lesson. However, their accountability ratings were not as high as in the Bending Water Lesson except the last part of enactment.

Analysis of the interviews provided confirming evidence for the patterns observed in the videos. About the Equilibrium Lesson, for example, Mr. Smith discussed limited opportunities for problematizing; he stated, "I think it was what they were expected… They did a thing; they expected to see a change; they saw that change. When they did a thing, they did not see a change; they were like 'Why did I not see that change?' There was not a whole lot of unexpected. There were a whole lot of expected results happening, and they were just kind of confirming what they saw already…" Similar patterns emerged in the student interviews. One student, for example, stated, "The bending water, the one I understood the best at the end. I would probably explain it best; the equilibrium was the least. I could not understand the video of the lab..." The same student also stated that, "I honestly remember in that lab [Equilibrium Lesson] just not understanding anything and it's really frustrating. … Here's the difference between the polarity one [Bending Water Lesson] and this one [Equilibrium Lesson] was with the polarity one; he [Mr. Smith] gave you freedom, but you understood what was going on and it felt like you were figuring something out. It felt like you were learning by yourself. With the equilibrium one, he gave you freedom but wouldn't, but we never got that satisfaction in the end of being able to fully understand it."

All in all, PDE differed between the focal lessons in Mr. Smith's classroom, and these patterns were consistent with the cognitive demand of the task selected, how it was launched to frame students' intellectual work, and enacted to engage students in high levels of thinking and reasoning. These patterns support our conjecture that using cognitively demanding science tasks as a context for student thinking and effectively implementing them by drawing on students' intellectual resources appears to have the potential to foster PDE.

Discussion

As the ambition in science education demands new ways of thinking and acting in science classrooms, many teachers struggle with how to engage their students in authentic ways of doing science while deepening their understanding of core disciplinary ideas and practices. Engle and Conant (2002), argued that "For engagement to be 'disciplinary,' there must be 'some contact between what students are doing and the issues and practices of a discipline's discourse" (p.402). We argue that providing students opportunities to work on cognitively demanding science tasks while maintaining cognitive demand, which can engage students in doing science, can provide this "contact" for the students. Many teachers today try to find the happy medium between holding students' hands during well-defined scientific investigations, as they typically do, and having the challenges so large that students are lost at sea as they try to ensure that students make progress in their understanding of content and practices. We propose that supporting teachers learning to design and effectively implement cognitively demanding tasks can support *productive* disciplinary engagement in science classrooms in this happy medium where students engage in productive struggle as they make sense of phenomena. Therefore, by focusing on the roles of the task, teacher, and students for fostering PDE in science classrooms, the study findings support the means for how to design these learning environments that support PDE. There has been a growing emphasis on supporting teachers' learning to design tasks and lessons aligned with the reform vision (e.g., Penuel et al., 2019; Tekkumru-Kisa et al., 2018). Expanding these efforts by focusing on the design and effective implementation of cognitively demanding tasks, the study offers means for fostering PDE in science classrooms.

References

Anderson, C. W., de los Santos, E. X., Bodbyl, S., Covitt, B. A., Edwards, K. D., Hancock, J. B., ... & Welch, M. M. (2018). Designing educational systems to support enactment of the Next Generation Science Standards. *Journal of Research in Science Teaching*, *55*(7), 1026-1052.

Berland, L. K., & Hammer, D. (2012). Framing for scientific argumentation. *Journal of research in science Teaching*, *49*(1), 68-94.

Berland, K., Cooper, V. R., Lee, K., Schröder, E., Thonhauser, T., Hyldgaard, P., & Lundqvist, B. I. (2015). van der Waals forces in density functional theory: a review of the vdW-DF method. *Reports on Progress in Physics*, *78*(6).

Boston, M. (2012). Assessing instructional quality in mathematics. *The Elementary School Journal*, *113*(1).

Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: Rethinking transfer: A simple proposal with multiple

implications. *Review of research in education*, *24*(1), 61-100.
Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist*, *23*(2), 167-180.
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.
Engle, R. A. (2012). The productive disciplinary engagement framework: Origins, key concepts, and developments. In *Design research on learning and thinking in educational settings* (pp. 170-209).
Forman, E. A., Engle, R. A., Venturini, P., & Ford, M. J. (2014). Introduction to special issue: International examinations and extensions of the productive disciplinary engagement framework. *International Journal of Educational Research*, *64*, 149-155.
Forman, E. A., & Ford, M. J. (2014). Authority and accountability in light of disciplinary practices in science. *International Journal of Educational Research*, *64*, 199-210.
Haverly, C., Calabrese Barton, A., Schwarz, C. V., & Braaten, M. (2018). "Making Space": How Novice Teachers Create Opportunities for Equitable Sense-Making in Elementary Science. *Journal of Teacher Education*.
Kang, H., Windschitl, M., Stroupe, D., & Thompson, J. (2016). Designing, launching, and implementing high quality learning opportunities for students that advance scientific thinking. *Journal of Research in Science Teaching*, *53*(9), 1316-1340.
Kelly, G. J. (2014). The social bases of disciplinary knowledge and practice in productive disciplinary engagement. *International Journal of Educational Research*, *64*, 211-214.
Kumpulainen, K. (2014). The legacy of productive disciplinary engagement. *International Journal of Educational Research*, *64*, 215-220.
Maxwell, J. A. (2012). *Qualitative research design: An interactive approach* (Vol. 41). Sage publications.
Mortimer, E. F., & de Araújo, A. O. (2014). Using productive disciplinary engagement and epistemic practices to evaluate a traditional Brazilian high school chemistry classroom. *International Journal of Educational Research*, *64*, 156-169.
Odden, T. O. B., & Russ, R. S. (2019). Defining sensemaking: Bringing clarity to a fragmented theoretical construct. *Science Education*, *103*(1), 187-205.
Penuel, W. R., Turner, M. L., Jacobs, J. K., Van Horne, K., & Sumner, T. (2019). Developing tasks to assess phenomenon-based science learning: Challenges and lessons learned from building proximal transfer tasks. *Science Education*, *103*(6), 1367-1395.
Phillips, A. M., Watkins, J., & Hammer, D. (2018). Beyond "asking questions": Problematizing as a disciplinary activity. *Journal of Research in Science Teaching*, *55*(7), 982-998.
Resnick, L. B., Michaels, S., & O'Connor, C. (2010). How (well structured) talk builds the mind. *Innovations in educational psychology: Perspectives on learning, teaching and human development*, 163-194
Sandoval, W. A., Enyedy, N., Redman, E. H., & Xiao, S. (2019). Organizing a culture of argumentation in elementary science. *International Journal of Science Education*, *41*(13), 1848-1869.
Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics teaching in the middle school*, *3*(4), 268-275.
Stein, M. K., & Kaufman, J. H. (2010). Selecting and supporting the use of mathematics curricula at scale. *American Educational Research Journal*, *47*(3), 663-693.
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. Standards. *Journal of Research in Science Teaching*, *55*(7), 1026-1052.
Tekkumru-Kisa, M., & Stein, M. K. (2015). Learning to see teaching in new ways: A foundation for maintaining cognitive demand. *American Educational Research Journal*, *52(1)*, 105-136.
Tekkumru-Kisa, M., & Schunn C. (2019). Integrating a space for teacher interaction into an educative curriculum: design principles and teachers' use of the iPlan tool. *Technology, Pedagogy, and Education, 28*(2), 1-23.
Tekkumru-Kisa, M., Preston, C., Kisa, Z., Ozulku, E., & Morgan, J. (2019, April). Measuring instructional quality in science in the NGSS era. Paper presented at the Annual Meeting of the American Educational Research Association (AERA), Toronto, Canada.

Acknowledgments

Work on this paper was supported by a grant DRK-12 (#1720587) from the National Science Foundation to Florida State University. All opinions and conclusions in this paper are those of the authors and do not necessarily reflect the views of the funding agency. We would like to thank all members of the project team and our participants for their contributions to this study.