Boston College
Lynch School of Education

Department of
Teaching, Curriculum, and Society

Curriculum and Instruction

SUPPORTING SCIENCE TEACHER LEARNING IN CURRICULUM-BASED PROFESSIONAL DEVELOPMENT

Dissertation
By

BENJAMIN R. LOWELL

submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

May 2022

Abstract

Supporting science teacher learning in curriculum-based professional development

Benjamin R. Lowell, Author

Katherine L. McNeill, Chair

Science education is shifting from a vision of students memorizing facts towards engaging in figuring out the natural world as students build ideas from their own experiences and backgrounds. This shift is hard for teachers. One way to support teachers is curriculum-based professional development, which pairs high-quality instructional materials with professional development to help teachers understand the philosophy of those materials and what that looks like in practice. This three-paper dissertation uses the OpenSciEd middle school field test, a curriculum-based professional development program, as a context to investigate how to support teachers with this shift.

The first paper is a quantitative look at teacher surveys taken across the first two years of the OpenSciEd field test. I tracked changes in teachers' beliefs about science instruction and confidence in implementing OpenSciEd. I used Hierarchical Linear Modeling to identify teacher characteristics associated with differences in those changes. Beliefs and confidence changed initially and leveled out over time, but confidence took longer to change than beliefs. Teachers who had more experience and found the PD more valuable were less likely to hold traditional beliefs and more likely to have higher confidence.

The second paper is a conceptual look at practice-based professional development activities focused on one new one: the student hat. Student hat is when teachers engage in science activities while considering ideas and experiences their students might bring to

them and sharing those ideas using students' language. Student hat uniquely helps teachers to consider students' relationship to the science ideas under discussion and their cognitive and affective responses to reform science instruction.

The third paper is a qualitative look at the use of the student hat in one professional development workshop. I engaged in thematic analysis of interviews and video to determine what student hat helped teachers to learn and how. Student hat provided safety for teacher confusion, allowing teachers to learn science ideas. It also helped teachers develop their epistemic empathy for students, helping them to learn about their students and the OpenSciEd instructional approach.

Acknowledgements

To my family, thank you for always being my strongest supporters. To my parents, Cliff and Kerry, thank you for the healthy skepticism of the PhD degree you instilled in me from an early age but the unconditional support you provided when I decided to pursue one. To my sister, Katie, thank you for constantly reminding me that work may be important but it is not life. I feel confident that me receiving my doctoral degree will not pressure you into feeling like you need one. To Percy, during my many days writing at home alone, thanks for keeping me company in the way only a puppy can.

To Kate, my advisor and committee chair, I do not have the words to share my gratitude. That could be because I used up all of my words on this dissertation. You have helped me grow as a researcher and teacher in more ways than I can name, and I will always be grateful for your generosity and support. I feel so lucky to have you as a mentor and friend. After five years learning from you, I certainly owe you a nice bottle of scotch.

I am extraordinarily thankful to my committee, who improved my abilities as a reader, thinker, and writer: Drs. Marilyn Cochran-Smith, Laura O'Dwyer, and Brian Reiser. You helped me to sharpen my thinking, clarify my methods, and connect to theory in ways that I found both frustrating and ultimately useful. As a result of your feedback, this dissertation has become significantly stronger.

My time in graduate school has been shaped by the research teams I have been on, and I would like to thank members of both the ILSP and OpenSciEd teams. Becca, thank you for having an office that I always knew housed a welcome ear, a comfortable couch, and some nice chocolate. Renee, you are the professional development facilitator I dream to be someday. Thank you for your amazing example, your sassiness, and those car pickles. Kevin, thank you for being a graduate student I could look up to and partner with. Casandra, Sam, and Julie, working with you was a pleasure, and we did a pretty good job of sharing that tiny little cube, too.

In addition to co-workers, my time at BC was deeply improved by the friendship of my cohort. Will Peters, Melita, and Reid especially, thank you for mid-afternoon coffee dates, the ridiculous gossip, and the shared pain and joy of going through this process together.

Finally, my deepest thanks to my husband, Will, who was my primary supporter throughout the entire doctoral program. You were there to push me to write when I didn't want to and help me to calm down when I got hysterical. Because of you, we travelled widely and you helped me enjoy my time in graduate school outside of graduate school. And during that one week I spent every waking hour at my dissertation station at the dining room table, you helped make sure that I stayed fed and watered. I love you.

Table of Contents

List of Tables

List of Figures

Section I: **Introduction**

The last thirty years in American K-12 education has been marked by a strong policy focus on standards, testing, and accountability as a supposed path towards improved educational outcomes for children (Cochran-Smith, 2003; Taubman, 2009). Multiple standards-based reform efforts have grown out of this era of accountability, including the Common Core State Standards for English and mathematics (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) and the Next Generation Science Standards for science (NGSS Lead States, 2013). While these policies and the national discourses around them are influential and important, ultimately, they make little difference if they do not impact the actual day-to-day work that takes place in American classrooms.

Overall, research has found that various policy initiatives have had visible but mixed and unexpected impacts on teacher practice (Coburn et al., 2016). Despite the mixed results, a common focus of much of these studies was on the teacher as the primary implementer of policy change (Coburn et al., 2016). The ways teachers ultimately implement new policies are influenced by the ideas and experiences they bring to those new demands (Coburn, 2001). Because of their role as implementers of these policies and the ways their ideas and understandings influence that implementation, teachers are central figures in understanding how educational change comes to take place in American classrooms (Cuban, 1990; National Academies of Sciences, Engineering, and Medicine, 2015).

This is not to say that teachers are the *only* factor that determines the success of major instructional reforms. Many factors influence what takes place inside a classroom, including poverty, hunger, or housing segregation, and the focus on teacher practice has been used to place the blame for these societal problems on the school system (Labaree, 2008). In other words, many who argue that the teacher is important do so from a “thin equity” lens that implies that teacher quality is ultimately responsible for educational inequity or failure of reform efforts (Cochran-Smith et al., 2016). I am not making such an argument, but rather that within the diverse array of factors that influence what takes place in schools, how teachers come to understand and implement instructional reform is an important one.

Despite this focus on supporting teacher practice, the pace of educational change is slow, and shifting actual teaching practice can be particularly challenging (Cuban, 1990; Tyack & Cuban, 1995). For example, teachers might take up the language of new reforms, but actually be using new words to describe the same instructional methods they were using pre-reform (Cohen, 1990; McNeill et al., 2017; Tyack & Cuban, 1995). As teacher educators and educational researchers, therefore, it is incumbent upon us to understand how to support teachers as they seek to change and improve their own instructional practices.

One of the most common approaches to facilitating teacher learning as they implement instructional reform is professional development (Borko, 2004; Darling-Hammond et al., 2017; Desimone et al., 2002; Wilson, 2013). Professional development (PD) can be particularly powerful because it can help teachers to change their beliefs around what strong teaching and learning can be (Maeng et al., 2020; Nespor, 1987; Voet & De Wever, 2019). At the same time, when grounded in problems of practice teachers encounter in the course of teaching and aligned with teachers' accountability structures, PD can also support concrete changes in the ways teachers teach (Allen & Penuel, 2015; Ball & Cohen, 1999).

Science education is a particularly useful context to study teacher learning and PD because of large-scale reform efforts currently underway in the discipline (Bybee, 2014). These efforts are characterized by a vision of learning in which students actively engage in the processes of science to figure out important ideas about the natural world, and has been formalized by the release of the Next Generation Science Standards (National Research Council, 2012; NGSS Lead States, 2013). This approach stands in contrast to traditional methods of teaching science education, which emphasized students knowing “final form” science knowledge, meaning that they could explain what we know about the natural world without explaining how or why we know it (Duschl & Osborne, 2002; McNeill & Berland, 2017).

Given the scope of these changes, therefore, science teachers need substantial support in understanding and implementing these changes (National Academies of Sciences, Engineering, and Medicine, 2015). Professional development can play a key role in that support, but we need to know more about how to plan professional development targeted towards the particular needs of science teachers in this moment of reform. This three-paper dissertation seeks to address that need by studying one recent science teacher professional development program to offer insights about what changes it supported in teacher thinking and how. In the rest of this section, I outline the science education reforms that motivated this particular PD program, provide context of the program, and discuss the theory of teacher learning in which it was grounded. I then briefly outline the approaches of each paper to show how they contribute to our understanding of designing science teacher professional development in support of helping students to figure out the natural world.

Current Reforms in Science Education

In the United States, the two major policy documents guiding current shifts in science education are the *Framework for K-12 Science Education* (National Research Council, 2012) and the Next Generation Science Standards (NGSS) based off of the *Framework* (NGSS Lead States, 2013). The theory of learning undergirding these documents is that students are always engaging with the world around them to figure out how things work and to build explanations of the phenomena they experience (National Research Council, 2007, 2012). Many in science education have called this student learning theory “sensemaking,” meaning “a dynamic process of building or revising an explanation to ‘figure something out’—to ascertain the mechanism underlying a phenomenon to resolve a gap or inconsistency in one's understanding” (Odden & Russ, 2019, pp. 191–192). Students are natural sensemakers, and are always trying to develop their understanding of the world around them, even if teachers do not recognize those sensemaking practices (Rosebery et al., 2010). The goal of the NGSS is that teachers both honor and value the sensemaking work that students do and provide students with structured experiences to develop more rich, complex, and well supported explanations of natural phenomena than they might have on their own (Manz, 2015a).

This shift in science teaching and learning has been described as a move from “learning about” to “figuring out” (Schwarz et al., 2017), which is a helpful way to summarize the differences between traditional and reform approaches to science instruction. In traditional approaches to instruction, students are often “learning about” science facts or how scientists come to understand those facts, which might involve memorization of facts and terminology in ways that are disconnected from the natural world or reading about the things scientists do to learn about the natural world without meaningfully engaging in any of those practices (National Research Council, 2015; Schwarz et al., 2017). In classrooms dedicated to students “figuring out” the natural world, they are exposed to intriguing or unexpected natural phenomena and engage in learning activities in order to figure out how to describe, explain, or predict those phenomena (Schwarz et al., 2017).

In order to support students in engaging in science as a process of figuring out, teachers must also change their roles and responsibilities in the classroom (Bybee, 2014). This change in role includes a shift away from the teacher as a source of information towards the teacher as a facilitator of investigations and sensemaking conversations (National Academies of Sciences, Engineering, and Medicine, 2019). In this view, the teacher is responsible for selecting appropriate phenomena for students to make sense of, facilitating learning activities and conversations to help students build their explanations of those phenomena, and ensuring that students develop their understanding of how science is done through engaging in the practices of science. This shift is difficult. Therefore, teachers need an array of supports to understand why and how to engage students in figuring out the natural world (National Academies of Sciences, Engineering, and Medicine, 2015).

Supporting Science Teachers' Learning

Two major strategies that have emerged to support instructional change at the classroom level are the design and release of reform-based curricular materials and the development of teacher PD programs to teach educators about desired reforms (Lynch et al., 2019). Both of these approaches have complementary benefits and drawbacks. For example, high quality curricular materials can help teachers see how to design learning environments to support students' sensemaking, but without training they may implement those materials in traditional ways (Alozie et al., 2010). Similarly, PD can help change teachers' beliefs and vision of science instruction, but may not provide sufficient supports to help teachers truly change their instructional practice (Mills et al., 2019).

Some work has shown that teachers who use reform-based curriculum materials can successfully improve their students' understanding of the subject (Geier et al., 2008; Harris et al., 2015). Educative curriculum materials are curriculum materials with supports designed into the materials to help teachers understand how and why the materials are designed the way they are to explicitly support teacher learning (Davis & Krajcik, 2005). These supports can take many forms, including written callout boxes beside lesson plans or multimedia supports such as video of teachers engaging in the target lesson (Loper et al., 2017). There is some evidence that educative curriculum materials on their own can help teachers support their students in learning conceptual science ideas (Schuchardt et al., 2017).

While the value of strong curriculum materials, particularly those with educative features, is well accepted in the science education literature, there is also a plethora of work arguing that instructional materials on their own may be insufficient to support successful instructional change. Many studies have found that teachers may implement reform curricula in traditional ways, thereby undercutting the goals for student learning built into those materials (Alozie et al., 2010; McNeill et al., 2017; Tekkumru-Kisa et al., 2019). This may be because teachers bring their own backgrounds and experiences to interpreting curriculum materials, interpreting reforms through the lens of their current ideas, which might lead them to lose sight of the key differences inherent in the reforms (Spillane et al., 2002). If teachers are not supported to notice how reform curricula are different from traditional ones, they may not process how these materials support fundamentally different relationships between teacher, student, and content (Amador et al., 2017; Dietiker et al., 2018).

Professional development is a popular and important strategy for supporting teacher learning (Kennedy, 2016; Wilson, 2013). Unfortunately, PD also has a mixed record in the literature, with multiple studies showing that it has little to no effect on teacher practice (Kennedy, 2016) while others arguing that it has positive impacts depending on the design (Borko, 2004). In order to address these concerns about the impact of PD on teacher learning, there are a plethora of studies examining what makes it effective. Some consensus has been reached on five critical features for professional development: content focus, use of active learning strategies, sustained duration in terms of both hours of time and span of time over which the PD is spread, and collective participation of teachers in the same school or department (Desimone, 2009; Garet et al., 2001). This general framework has also been applied successfully to examining content-specific professional learning opportunities, such as in science, which have shown that these same features by and large are valuable when supporting teachers' ability to teach science specifically (Penuel et al., 2007).

Recently, more attention has been paid to combining reform-based curriculum materials with high-quality professional development, a strategy called *curriculum-based professional development* (McNeill et al., in press; Short & Hirsh, 2020). In fact, a recent meta-analysis of studies on PD for science, technology, engineering, and mathematics teachers has found that pairing curriculum and professional development is associated with statistically significant increases in teacher learning outcomes (Lynch et al., 2019). Because of the recent successes of some curriculum-based PD, it represents a fruitful format to investigate further as we build our understanding of how PD can help teachers to learn about and implement instructional reforms.

The OpenSciEd Context

The three papers in this dissertation draw specifically from the OpenSciEd curriculum-based professional development. OpenSciEd is a nonprofit organization that develops high quality instructional materials and accompanying professional development workshops for science and releases them as freely available open educational resources (OpenSciEd, 2020b). The organization began its work in January of 2018 with middle school and finished releasing its full middle school curriculum in February of 2022. It has since moved on to developing high school curriculum and plans to expand into elementary school as well.

Funded by a coalition of private foundations led by the Carnegie Corporation of New York, OpenSciEd has a somewhat unusual institutional structure. The organization itself is relatively small, consisting of a handful of staff members who coordinate the planning, outreach, and development work. In order to develop the actual curricula, the organization hires consortia of developers based out of universities and larger non-profit education research organizations. The middle school curriculum was developed by a consortium consisting of teams from BSCS Science Learning, Northwestern University, Boston College, the Dana Center at the University of Texas Austin, and the University of Colorado Boulder.

The curriculum and PD materials went through an extensive design, field test, and revision process before being released publicly (Edelson et al., 2021). At the beginning of development, representatives of every state's chief science education officer were offered the opportunity to participate in the field test; ten states agreed to participate in the field test, and they formed a steering committee that supervised the development of the materials. For the middle school curriculum, the ten field test states were: California, Iowa, Louisiana, Massachusetts, Michigan, New Jersey, New Mexico, Oklahoma, Rhode Island, and Washington. Each unit and PD was initially developed, and then teachers from the field test states participated in the PD, taught the unit, and provided data and feedback to the developers. Based on this feedback, the units and PD were revised and the units were evaluated by the NGSS Peer Review Panel using the EQuIP rubric (Achieve, 2016) as an external mark of quality. Once approved, the units and PD were released publicly. This occurred in rounds with three new units being field tested every six months from fall of 2018 through the spring of 2021.

The focus of the OpenSciEd organization is on the development of curriculum materials and associated curriculum-based professional development, but a growing group of academics are also developing research projects around the materials and their implementation. For example, the nonprofit research organization Digital Promise, with support from the Carnegie Corporation, has published an OpenSciEd research agenda and is providing seed grants to support initial research into the curricular materials and their use (McElhaney et al., 2021).

The work in this dissertation comes from studies I performed as a member of the middle school developer's consortium. Boston College was responsible for designing the professional development materials, and as part of that work we also collected data on the implementation of the professional development and field test units. I drew from that corpus of data to look specifically at the outcomes of the field test PD sessions and the experiences of teachers during PD. In order to understand this work, it is important to understand how OpenSciEd conceptualized science learning for students and learning for teachers. In the following sections, therefore, I briefly discuss the theory of student learning that undergirded the development of the OpenSciEd curricular materials and the theory of teacher learning that informed the development of the PD materials.

OpenSciEd Theory of Student Learning: Supporting Student Sensemaking

The OpenSciEd curriculum materials use an approach called "storylines" that is designed to support students to incrementally and collaboratively build science ideas based on their own experiences and structured investigations into an anchoring phenomenon (Edelson et al., 2021; Reiser et al., 2021). These materials take the same approach that the NGSS as a whole does to student learning: that students are making sense of natural phenomena around them in order to develop explanations of the world (Odden & Russ, 2019). Storyline units are particularly concerned with ensuring that learning activities are designed in a way that makes sense to students as they are developing complex understandings of natural phenomena, which means that investigations and learning activities emerge from students' need to understand something rather than from the teacher telling students what is happening next (Manz, 2015b). In this way, students build more complex understandings of the science ideas over time starting with their own questions and ideas. The instructional model supports this student learning through a series of routines: the *anchoring phenomenon routine* in which the class experiences and develops questions about a phenomenon; the *navigation routine* in which they work together to determine the next steps to investigate the anchoring phenomenon; the *putting the pieces together routine* in which they synthesize what they have learned from investigations in order to explain the phenomenon; and the *problematizing routine* in which the teacher foregrounds gaps or inconsistencies in students' explanations in order to motivate further investigation (Reiser et al., 2021). These routines are all designed to help students be meaningful contributors to the knowledge building work of the classroom rather than passive receivers of information (Reiser et al., 2021).

A core tenet of storyline curricula that the routines are designed to support is that the learning is *coherent from the student perspective* (Reiser et al., 2021; Zivic et al., 2018). This is the idea that each new investigation or activity should make sense to students who are building their understanding of the phenomenon. Traditional science curricula are often designed to be coherent from the disciplinary perspective, meaning the order of activities makes sense to someone who has a strong grasp of the discipline, like a scientist, but not necessarily a novice (National Academies of Sciences, Engineering, and Medicine, 2019; Reiser, Novak, et al., 2017). For example, a standard approach to chemical reactions might include telling students that mass is always conserved during a reaction and then having them measure the mass of a reaction in a closed system before and after to show that to be true. But if students have not established a reason why they would do this investigation or a question that it helps answer, then this approach becomes more about students knowing something because their teacher told them to rather than answering their own questions about the natural world (Berland et al., 2016). In contrast, the OpenSciEd unit focusing on chemical reaction begins with students observing bath bombs fizz and disappear when placed into water, which leads students to wonder if gas was trapped inside the solid bath bomb. In order to test that question, students measure the mass of a bath bomb reacting with water in an open and closed system to show that matter is conserved during reactions in closed systems (OpenSciEd, 2020a). Although these two examples ultimately have students doing the same investigation—measuring the mass of a reaction in a closed system—the OpenSciEd approach is based in the theory that work that makes sense to students will better support them to understand the key science ideas addressed in the investigation (Reiser et al., 2021; Zivic et al., 2018).

Theory of Teacher Learning: Situated in Practice

The OpenSciEd professional development was designed on the theory that teachers learn when work is situated in their practice and aligned with their implementation context (Ball & Cohen, 1999; Lampert, 2010; McNeill et al., in press). This means that teachers learn when they engage with complex theories like coherence from the student perspective in the context of the work that will be done in the classroom and how that might play out. This focus on practice does not conceive of teaching as a series of disconnected moves for teachers to master, but rather that teachers develop their understanding of strong instruction as they consider how that might be implemented in practice and what makes it different from traditional approaches (Forzani, 2014).

In order to create PD sessions aligned with this practice-based theory of teacher learning, the OpenSciEd PD was designed with four principles in mind: support teachers to take the student perspective, analyze images of classroom instruction, examine contrasting curricular cases, and engage in cycles of planning, enactment, and reflection (McNeill et al., in press). Encouraging teachers to take the student perspective was designed to help teachers understand what it is like as a student to experience OpenSciEd instruction and how that is different from traditional approaches to science teaching. Images of classroom instruction took the form of written vignettes and videos to allow teachers to get a sense of what it looks like to facilitate OpenSciEd curricular materials. Contrasting curricular cases, such as vignettes of the same idea taught in a traditional and OpenSciEd approach, were used to help teachers problematize current practice and identify key features of the OpenSciEd approach that better support student sensemaking. Finally, engaging teachers in cycles of planning, enactment, and reflection helped teachers to ground their learning in their own context in order to support change in their practice. Together, the goal of these features was to help teachers develop a robust theoretical understanding of the goals of OpenSciEd in the context of the application to their own classroom (McNeill et al., in press).

Overview of the Dissertation

The three papers in this dissertation take distinct methodological approaches to the fundamental questions of *what* teachers learned during the OpenSciEd field test professional development and *how* the design of the PD session supported that learning. The first paper takes a quantitative approach to understanding outcomes from the field test PD over time. The second paper is a conceptual exploration of one particular activity common in the OpenSciEd PD, and the third paper takes an empirical qualitative approach to studying that activity. In the following sections, I describe the purpose, research questions (where applicable), approach, and target audience for each paper.

Paper 1: Changes in Teachers' Beliefs and Confidence: A Longitudinal Study of Science Teachers Engaging in Storyline Curriculum-Based Professional Development

The purpose of my first paper is to investigate what teachers learned from multiple rounds of curriculum-based professional development with materials designed to support teachers' shifts toward teaching science as a process of figuring out the natural world. I use OpenSciEd as the context and investigate shifts in beliefs and confidence in implementing the teaching practices supporting learning as figuring out. Based on surveys the field test teachers from all 10 field test states took before the first PD and after each round of PD, I track their beliefs and confidence over the first two years of the OpenSciEd field test. I use hierarchical linear modeling (Raudenbush & Bryk, 2002) to create a model that partially explains differences in patterns of beliefs and confidence between teachers. My research questions for this paper include two questions focused on teachers' beliefs and two questions focused on teachers' confidence: 1a) how do *teachers' beliefs about teaching* change over the course of learning about and implementing new science curricula? and 1b) what teacher characteristics predict differences in initial *beliefs* and change over time? 2a) how do teachers' *confidence* change over the course of learning about and implementing new science curricula and 2b) what teacher characteristics predict differences in both initial *confidence* and their change over time? The target audience for this paper is science education researchers and professional development providers, so the paper will be sent to a science education journal.

Paper 2: The Student Hat: A New Tool in Practice-Based Professional Development

The purpose of paper 2 is to take a conceptual look at an activity that was common in the OpenSciEd professional development sessions: the "student hat." Broadly, student hat activities are ones in which teachers engage in science investigations or discussions that are part of the curriculum while being asked to actively consider the ideas and experiences students might bring to that work and use the language students might use while speaking (McNeill et al., in press; Reiser, Michaels, et al., 2017). Using the instructional triangle (Cohen et al., 2003), I look at five common activities in practice-based PD to describe what they help teachers to think about in terms of science learning. I then argue that the student hat is uniquely positioned to help teachers consider students' cognitive and affective experiences of engaging in NGSS-aligned science instruction. Finally, I illustrate how student hat might look in a PD session and how it can be used in concert with other more common activities such as watching classroom video or analyzing student work. My target audience for this paper is science teacher educators and professional development providers, and it will also be submitted to a science education specific journal.

Paper 3: The Student Hat in Professional Development: Building Epistemic Empathy to Support Teacher Learning

The third paper takes an empirical, qualitative approach to studying the student hat activity structure I define in paper 2. I focus on one 2.5-day PD session in one of the field test states. The analysis includes interviews with 12 teachers who participated in the PD to understand what the student hat helped them to learn. I then analyze video from the PD session to make claims about how the student hat helped teachers to learn. In order to do this analysis, I use productive struggle theory (Hiebert & Grouws, 2007) to look for moments that teachers struggled to use student language when being asked to talk in student hat. The theory here is that that struggle was productive in supporting teacher learning, and looking at moments of struggle will help us to understand what and how teachers were learning. My research questions for this study are: What does the student hat activity structure help teachers learn during curriculum-based professional development? How does it support that learning? Unlike the other papers in this dissertation, the target audience for this paper is not science specific, but rather the learning sciences community; therefore, the paper is focused more on how we can design the student hat into teacher learning environments rather than targeting science teacher educators specifically.

Conclusion

In the final section of the dissertation, I draw connections between the three papers in order to paint the picture of what we know about practice-based teacher professional development to support science teacher learning. I end this section by proposing future directions for this work in terms of research, practice, and policy at two levels: professional learning programs and specific PD session activities. At the program level, I argue that we need to better understand how teacher learning associated with the NGSS changes over time as teachers gain more experience with well-designed curricular materials. At the PD session level, I point to the potential value of the student hat and how it requires further work to understand qualities of effective facilitation and support of student hat activities. Finally, I point out that all of this teacher learning work must take place in the context of accountability and policy systems that ultimately support teachers to learn about and engage their students in deep sensemaking about natural phenomena.

References

Achieve. (2016). *EQuIP rubric for lessons and units: Science*. https://www.nextgenscience.org/sites/default/files/EQuIPRubricforSciencev3.pdf

Allen, C. D., & Penuel, W. R. (2015). Studying teachers' sensemaking to investigate teachers' responses to professional development focused on new standards. *Journal of Teacher Education*, *66*(2), 136–149. https://doi.org/10.1177/0022487114560646

Alozie, N. M., Moje, E. B., & Krajcik, J. S. (2010). An analysis of the supports and constraints for scientific discussion in high school project-based science. *Science Education*, *94*(3), 395–427. https://doi.org/10.1002/sce.20365

Amador, J. M., Males, L. M., Earnest, D., & Dietiker, L. (2017). Curricular noticing: Theory on and practice of teachers' curricular use. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Eds.), *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks* (pp. 427–443). Springer International Publishing. https://doi.org/10.1007/978-3-319-46753-5_25

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (Vol. 1, pp. 3–22). Jossey-Bass.

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*, *53*(7), 1082–1112. https://doi.org/10.1002/tea.21257

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3–15. https://doi.org/10.3102/0013189X033008003

Bybee, R. W. (2014). NGSS and the next generation of science teachers. *Journal of Science Teacher Education*, *25*(2), 211–221. https://doi.org/10.1007/s10972-014-9381-4

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, *23*(2), 145–170. https://doi.org/10.3102/01623737023002145

Coburn, C. E., Hill, H. C., & Spillane, J. P. (2016). Alignment and accountability in policy design and implementation: The Common Core State Standards and implementation research. *Educational Researcher*, *45*(4), 243–251. https://doi.org/10.3102/0013189X16651080

Cochran-Smith, M. (2003). The unforgiving complexity of teaching: Avoiding simplicity in the age of accountability. *Journal of Teacher Education*, *54*(1), 3–5. https://doi.org/10.1177/0022487102238653

Cochran-Smith, M., Stern, R., Sánchez, J. G., Miller, A., Stringer Keefe, E., Fernández, M. B., Chang, W.-C., Cummings Carney, M., Burton, S., & Baker, M. (2016). *Holding teacher preparation accountable: A review of claims and evidence*. National Education Policy Center. http://nepc.colorado.edu/publication/teacher-prep

Cohen, D. K. (1990). A Revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, *12*(3), 311–329. https://doi.org/10.3102/01623737012003311

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*(2), 119–142. https://doi.org/10.3102/01623737025002119

Cuban, L. (1990). A fundamental puzzle of school reform. In A. Lieberman (Ed.), *Schools as collaborative cultures: Creating the future now* (pp. 71–77). The Falmer Press.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, *34*(3), 3–14. https://doi.org/10.3102/0013189X034003003

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81–112. https://doi.org/10.3102/01623737024002081

Dietiker, L., Males, L. M., Amador, J. M., & Earnest, D. (2018). Curricular noticing: A framework to describe teachers' interactions with curriculum materials. *Journal for Research in Mathematics Education*, *49*(5), 521–532.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, *38*, 39–72. https://doi.org/10.1080/03057260208560187

Edelson, D. C., Reiser, B. J., McNeill, K. L., Mohan, A., Novak, M., Mohan, L., Affolter, R., McGill, T. A. W., Buck Bracey, Z. E., Deutch Noll, J., Kowalski, S. M., Novak, D., Lo, A. S., Landel, C., Krumm, A., Penuel, W. R., Van Horne, K., González-Howard, M., &

Suárez, E. (2021). Developing research-based instructional materials to support large-scale transformation of science teaching and learning: The approach of the OpenSciEd middle school program. *Journal of Science Teacher Education*, *32*(7), 780–804. https://doi.org/10.1080/1046560X.2021.1877457

Forzani, F. M. (2014). Understanding "core practices" and "practice-based" teacher education: Learning from the past. *Journal of Teacher Education*, *65*(4), 357–368. https://doi.org/10.1177/0022487114533800

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945. https://doi.org/10.3102/00028312038004915

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., & Clay-Chambers, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *Journal of Research in Science Teaching*, *45*(8), 922–939. https://doi.org/10.1002/tea.20248

Harris, C. J., Penuel, W. R., D'Angelo, C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., Cheng, B. H., & Krajcik, J. S. (2015). Impact of project-based curriculum materials on student learning in science: Results of a randomized controlled trial. *Journal of Research in Science Teaching*, *52*(10), 1362–1385. https://doi.org/10.1002/tea.21263

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age Publishing.

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Labaree, D. F. (2008). The winning ways of a losing strategy: Educationalizing social problems in the United States. *Educational Theory*, *58*(4), 447–460. https://doi.org/10.1111/j.1741-5446.2008.00299.x

Lampert, M. (2010). Learning teaching in, from, and for practice: What do we mean? *Journal of Teacher Education*, *61*(1–2), 21–34. https://doi.org/10.1177/0022487109347321

Loper, S., McNeill, K. L., & González-Howard, M. (2017). Multimedia educative curriculum materials (MECMs): Teachers' choices in using MECMs designed to support scientific argumentation. *Journal of Science Teacher Education*, *28*(1), 36–56. https://doi.org/10.1080/1046560X.2016.1277600

Lynch, K., Hill, H. C., Gonzalez, K. E., & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis*, *41*(3), 260–293. https://doi.org/10.3102/0162373719849044

Maeng, J. L., Whitworth, B. A., Bell, R. L., & Sterling, D. R. (2020). The effect of professional development on elementary science teachers' understanding, confidence, and classroom implementation of reform-based science instruction. *Science Education*, *104*(2), 326–353. https://doi.org/10.1002/sce.21562

Manz, E. (2015a). Resistance and the development of scientific practice: Designing the mangle into science instruction. *Cognition and Instruction*, *33*(2), 89–124. https://doi.org/10.1080/07370008.2014.1000490

Manz, E. (2015b). Representing student argumentation as functionally emergent from scientific activity. *Review of Educational Research*, *85*(4), 553–590. https://doi.org/10.3102/0034654314558490

McElhaney, K., Baker, A., Kasad, Z., Roschelle, J., & Chillmon, C. (2021). *A field-driven, equity-centered research agenda for OpenSciEd* [White paper]. Digital Promise. https://digitalpromise.org/wp-content/uploads/2021/10/A-Field-Driven-Equity-Centered-Research-Agenda-for-OpenSciEd-Oct-2021.pdf

McNeill, K. L., Affolter, R., & Reiser, B. J. (in press). Anchoring science professional learning in curriculum materials enactment: Illustrating theories in practice to support teacher learning. In A. Castro Superfine, S. R. Goldman, & M.-L. M. Ko (Eds.), *Teacher learning in changing contexts: Perspectives from the learning sciences*. Routledge.

McNeill, K. L., & Berland, L. (2017). What is (or should be) scientific evidence use in k-12 classrooms? *Journal of Research in Science Teaching*, *54*(5), 672–689. https://doi.org/10.1002/tea.21381

McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2017). Moving beyond pseudoargumentation: Teachers' enactments of an educative science curriculum focused on argumentation. *Science Education*, *101*(3), 426–457. https://doi.org/10.1002/sce.21274

Mills, K., Jass Ketelhut, D., & Gong, X. (2019). Change of teacher beliefs, but not practices, following integration of immersive virtual environment in the classroom. *Journal of Educational Computing Research*, *57*(7), 1786–1811. https://doi.org/10.1177/0735633119854034

National Academies of Sciences, Engineering, and Medicine. (2015). *Science teachers' learning: Enhancing opportunities, creating supportive contexts*. The National Academies Press. https://doi.org/10.17226/21836

National Academies of Sciences, Engineering, and Medicine. (2019). *Science and Engineering for Grades 6-12: Investigation and Design at the Center* (B. Moulding, N. Songer, & K. Brenner, Eds.). National Academies Press. https://doi.org/10.17226/25216

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for English language arts and literacy in history/social studies, science, and technical subjects*. http://www.corestandards.org/ELA-Literacy/

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Academies Press. https://doi.org/10.17226/11625

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press. https://doi.org/10.17226/13165

National Research Council. (2015). *Guide to Implementing the Next Generation Science Standards*. National Academies Press.

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, *19*(4), 317–328. https://doi.org/10.1080/0022027870190403

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. National Academies Press. https://doi.org/10.17226/18290

Odden, T. O. B., & Russ, R. S. (2019). Defining sensemaking: Bringing clarity to a fragmented theoretical construct. *Science Education*, *103*(1), 187–205. https://doi.org/10.1002/sce.21452

OpenSciEd. (2020a). *How can we make something new that was not there before?* https://www.openscied.org/instructional-materials/7-1-chemical-reactions-matter/

OpenSciEd. (2020b). *OpenSciEd Teacher Handbook* (Version 3.0). https://www.openscied.org/teacher-handbook/

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, *44*(4), 921–958. https://doi.org/10.3102/0002831207308221

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed). Sage Publications.

Reiser, B. J., Michaels, S., Moon, J., Bell, T., Dyer, E., Edwards, K. D., McGill, T. A. W., Novak, M., & Park, A. (2017). Scaling up three-dimensional science learning through teacher-led study groups across a state. *Journal of Teacher Education*, *68*(3), 280–298. https://doi.org/10.1177/0022487117699598

Reiser, B. J., Novak, M., & McGill, T. A. W. (2017). *Coherence from the students' perspective: Why the vision of the framework for K-12 science requires more than simply "combining" three dimensions of science learning*. The National Academies of Sciences, Engineering and Medicine, Board on Science Education. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_180270.pdf

Reiser, B. J., Novak, M., McGill, T. A. W., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal of Science Teacher Education*, *32*(7), 805–829. https://doi.org/10.1080/1046560X.2021.1884784

Rosebery, A. S., Ogonowski, M., Dischino, M., & Warren, B. (2010). "The coat traps all your body heat": Heterogeneity as fundamental to learning. *Journal of the Learning Sciences*, *19*(3), 322–357. https://doi.org/10.1080/105084062010491752

Schuchardt, A. M., Tekkumru-Kisa, M., Schunn, C. D., Stein, M. K., & Reynolds, B. (2017). How much professional development is needed with educative curriculum materials? It depends upon the intended student learning outcomes. *Science Education*, *101*(6), 1015–1033. https://doi.org/10.1002/sce.21302

Schwarz, C. V., Passmore, C., & Reiser, B. J. (2017). Moving beyond "knowing about" science to making sense of the natural world. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 3–21). NSTA Press.

Short, J., & Hirsh, S. (2020). *The elements: Transforming teaching through curriculum-based professional learning*. Carnegie Corporation of New York. https://www.carnegie.org/publications/elements-transforming-teaching-through-curriculum-based-professional-learning/

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, *72*(3), 387–431. https://doi.org/10.3102/00346543072003387

Taubman, P. M. (2009). *Teaching by numbers: Deconstructing the discourse of standards and accountability in education*. Routledge.

Tekkumru-Kisa, M., Schunn, C., Stein, M. K., & Reynolds, B. (2019). Change in thinking demands for students across the phases of a science task: An exploratory study. *Research in Science Education*, *49*(3), 859–883. https://doi.org/10.1007/s11165-017-9645-z

Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia*. Harvard University Press.

Voet, M., & De Wever, B. (2019). Teachers' adoption of inquiry-based learning activities: The importance of beliefs about education, the self, and the context. *Journal of Teacher Education*, *70*(5), 423–440. https://doi.org/10.1177/0022487117751399

Wilson, S. M. (2013). Professional development for science teachers. *Science*, *340*(6130), 310–313. https://doi.org/10.1126/science.1230725

Zivic, A., Smith, J. F., Reiser, B. J., Edwards, K. D., Novak, M., & McGill, T. A. W. (2018). *Negotiating epistemic agency and target learning goals: Supporting coherence from the students' perspective*. International Conference of the Learning Sciences, London, U.K.

Section II: Paper #1 - Changes in Teachers' Beliefs and Confidence: A Longitudinal Study of Science Teachers Engaging in Storyline Curriculum-Based Professional Development

The current era of science education reform is focused on moving away from asking students to learn information about the world and the scientific process towards figuring out how natural phenomena work using similar practices that professional scientists do while building from their own experiences and ideas (Bang et al., 2012; Schwarz et al., 2017). In the United States, these ideas are most clearly laid out by the *Framework for K-12 Science Education* and the Next Generation Science Standards (NGSS) that came out of it (National Research Council, 2012; NGSS Lead States, 2013). Since their publication, the NGSS have been adopted or adapted by 44 states representing 71% of US students, demonstrating the influence of these reforms and their importance in the national conversation around science education (National Science Teaching Association, n.d.).

Changes in the NGSS can be characterized with four key features: that science curriculum and instruction be more phenomenon-based, three-dimensional, supportive of student epistemic agency, and coherent from the student perspective (Lowell et al., 2021). These features represent a fundamental shift in the relationship between teacher, student, and curriculum materials because they ask teachers to engage students in the social and cultural practices of science rather than having teachers primarily explain final-form knowledge to students (J. Osborne, 2014; Pruitt, 2014). For example, rather than using natural phenomena as "discrepant events" to highlight and correct student misconceptions, the NGSS asks teachers to use them as opportunities to provide all students with experiences around a particular event and draw out their own ideas related to it (Furtak & Penuel, 2019). Similarly, in order to truly support students' epistemic agency, teachers must attend and respond to the ideas students bring with them to the classroom and the ways they make sense of the learning experiences they engage with during class rather than focusing on asking students to memorize canonical information without regard to their existing funds of knowledge (Miller et al., 2018; Robertson et al., 2015; Smith et al., 1994). This work cannot be completely scripted ahead of time and requires that teachers have strong understanding of the science content, the ways students might respond to that content, and the pedagogical moves that can help students develop their understandings of the natural world (Crawford & Capps, 2018).

Given the degree of change represented by these shifts, it is not surprising, therefore, that teachers would find them hard to understand and implement (McNeill & Knight, 2013; National Research Council, 2015; Ricketts, 2014). Some have argued that this difficulty could be related to teachers' lack of knowledge of what these reforms actually mean, which could lead to relabeling traditional instructional practice as reformed rather than truly changing instruction (Capps et al., 2016; McNeill et al., 2017). In addition to teachers understanding *what* these reforms are and might look like in classrooms, however, it is also important that they *believe* that this approach to science education truly can help all students better understand the natural world because teachers who do not understand and believe in instructional reforms are unlikely to implement them as intended (National Academies of Sciences, Engineering, and Medicine, 2015).

Two main strategies have been used to support changes in teachers' beliefs: creating strong curriculum materials that include educative features to support teachers' learning as they implement the materials (Davis & Krajcik, 2005; Davis et al., 2017) and professional development (PD) to introduce teachers to and help them learn about the NGSS (Wilson, 2013). Recent evidence has shown that these strategies are particularly effective when combined into *curriculum-based professional development* in which PD sessions are paired with reform-based curricular materials and the focus of the sessions is on understanding the reforms inherent in the materials and developing teachers' ability to implement the target curriculum (Lynch et al., 2019; Short & Hirsh, 2020; Taylor et al., 2015).

One of the biggest weaknesses in the design of PD is that it often occurs over short durations, mostly one-off workshops that do not provide teachers enough time to make meaningful changes to their own beliefs, ideas, and practices (Desimone, 2009). This is also true of research on PD, which often reports out the design or results of individual workshops over short periods of time (Borko, 2004). These designs conflict with well-accepted ideas that in order to truly change their beliefs and knowledge, teachers require significant time as well as opportunities for implementation, metacognition, and reflection (Clarke & Hollingsworth, 2002; Crawford & Capps, 2018). Similarly, not all teachers change their ideas at the same rate (Allen & Penuel, 2015), and therefore it might be helpful to understand characteristics of teachers that require more or less support in making sense of the reforms inherent in the NGSS. Therefore, we need to better understand how teachers' beliefs change over time as they engage with curriculum-based PD as well as teacher characteristics associated with variation in this rate of change so that we can better design interventions that support teachers in this change.

Background

In this section, I begin by articulating what I mean by teacher beliefs and how they have been defined and studied in the past. In addition, I briefly review our understanding of PD design to highlight what we currently know about effective PD and what gaps still remain in the research literature.

Teachers' Educational Beliefs

Beliefs are a well-studied psychological construct that can encompass a range of individual thoughts and opinions (Jones & Leagon, 2014; Pajares, 1992). Beliefs and knowledge are inherently, and some argue inseparably, connected (Jones & Leagon, 2014; Pajares, 1992). The key feature that many have used to distinguish the two is the influence of affective and evaluative components (e.g. Nespor, 1987). Whereas both knowledge and beliefs are connected to an individual's cognition, beliefs are connected to a person's emotions and therefore can be harder to change because they involve a strong affective component (Crawford, 2007; Jones & Leagon, 2014; Nespor, 1987; Pajares, 1992). Each person has many different kinds of beliefs that exist together in a system, and within this system, teachers' *educational beliefs* are the set of beliefs teachers hold about education (Pajares, 1992; Rokeach, 1968). For the purposes of this study, I use Luft and colleagues' definition of teachers' educational beliefs as "personal constructs that are important to a teacher's practice, as they guide instructional decisions, influence classroom management, and impact the representation of content" (Luft et al., 2011, p. 1202).

Even the system of teachers' educational beliefs is broad, however, and can encompass multiple different kinds of beliefs about education (Pajares, 1992). Jones and Leagon (2014) defined five different kinds of science teachers' educational beliefs: beliefs about knowledge, about science, about self, about teaching, and about students. All of these kinds of educational beliefs can be related to changes in teacher practice (Jones & Leagon, 2014), but for this study, I am particularly interested in two of these sub-constructs: teachers' beliefs about teaching and their beliefs about themselves.

Beliefs about Teaching

Past work has defined teachers' beliefs about teaching as "judgements of appropriate goals and purposes of instruction" (Voet & De Wever, 2019, p. 425). These beliefs are inherently connected to what a teacher believes about *learning* because how a teacher believes students learn best is connected to how they choose to teach (Crawford, 2007; Jones & Leagon, 2014). In the context of this study, the most relevant beliefs about teaching are those connected to the reforms inherent in the NGSS. Primarily, the NGSS posits that students should engage in sensemaking around a natural phenomenon rather than listening to or reading, processing, and then confirming existing scientific knowledge (Schwarz et al., 2017). By sensemaking, I mean that students should be "building or revising an explanation in order to 'figure something out'—to ascertain the mechanism underlying a phenomenon in order to resolve a gap or inconsistency in [their] understanding" (Odden & Russ, 2019, pp. 191–192). This stands in contrast to past approaches to science education that have focused on teaching students scientific vocabulary and facts and then asking them to use that information to describe the natural world (Russ & Berland, 2019). In order to effectively implement NGSS-aligned curricular materials, teachers must believe that this approach to teaching will ultimately support students' science learning.

Beliefs about the Self: Confidence

Self-efficacy is beliefs about the self, which can incorporate general ideas about one's own confidence and more contextualized forms such as science, teaching, or personal self-efficacy (Jones & Leagon, 2014). The concept of self-efficacy was defined by Bandura as "beliefs in one's capabilities to organize and execute the course of action" (Bandura, 1997, p. 3). This is a broader construct than the one addressed by this study. Translating this definition into the context of teachers implementing the NGSS, therefore, I consider self-efficacy to be teachers' confidence in their ability to implement the key features of NGSS instruction as designed in the particular curricular materials they are using.

Self-efficacy and confidence have a long history of study in the literature, with a plethora of evidence connecting teacher self-efficacy with both effective classroom practices and positive psychological indicators such as job satisfaction and commitment (Zee & Koomen, 2016). Empirical work has also shown that teachers' confidence in their own abilities can be improved through experiencing PD and that change in self-efficacy can be connected to changing teacher practice (Maeng et al., 2020; Sandholtz & Ringstaff, 2014). Around the NGSS specifically, work has focused on how to improve teachers' confidence in their ability to implement various aspects of the new standards such as integrating the science ideas, practices, and crosscutting concepts. For example, Kang and colleagues (2019) found that after engaging in a professional development program consisting of workshops, coaching, and team meetings, 2nd grade teachers reported higher levels of confidence for including the science practices in their teaching. Similarly, Lo and colleagues (2014) found that a course dedicated to NGSS instructional design helped teachers become more confident engaging their students in the scientific practice of modeling.

Studies of Professional Development

One of the most popular approaches to studying PD is to attempt to define one or more externally observable features of PD associated with increased teacher and/or student learning (Kennedy, 2016). Perhaps the most well-known entry in this research tradition is the following set of five features: content-focus, active learning, coherence, duration, and collective participation (Darling-Hammond et al., 2017; Desimone, 2009; Garet et al., 2001; Penuel et al., 2007). As outlined by Desimone (2009), in this framework content-focused means the PD is discipline-specific and focused on teachers learning both the content they will teach and how students learn that content. Active learning means that teachers build and make sense of the key ideas in the PD rather than simply listening to presentations about those ideas. Coherence means the messages sent by the PD align with the values and accountability structures in place in teachers' schools. Duration refers to both the total number of contact hours of the PD and the time over which those contact hours are distributed, with both needing to be "substantial." Finally, collective participation requires that multiple stakeholders from the same team or building participate together in the PD.

These features were constructed based on large-scale surveys of teachers reporting what features of PD they believed were helpful for their learning (Garet et al., 2001). In the intervening years, many studies, often written by the PD providers themselves, have used these features to establish that their PD is "well designed" and then reported some other or more specific measure to justify the program's positive impact (e.g. Greenleaf et al., 2011). While these kinds of studies are effective in providing rich portraits of particular methods of PD, they leave open questions about the variations possible in these features and how those variations might support (or not support) teacher and student learning (Borko, 2004; National Academies of Sciences, Engineering, and Medicine, 2015).

Another concern with focusing on defining, describing, and quantifying "features" of effective PD is that can mask more complex questions about how teachers learn (Kennedy, 2016). For example, many interested in exploring the feature of "duration" ask how many hours of PD and/or over what period of time makes that PD "effective" for teacher learning (e.g. Yoon et al., 2007). The problem here is that while hours are quantifiable, they ignore the idea that the rate of learning is highly dependent on *what* is being learned and how it does (or does not) cohere with teachers' existing systems or structures (Allen & Penuel, 2015).

It is rare to have access to longitudinal data on teachers' learning over extended periods of time, and therefore it can be particularly powerful to use that data to understand the shape of change over time and what profiles of teacher characteristics relate to variation in this shape. For example, some may think that teachers with more teaching experience will change their beliefs more slowly because they have been inducted into a uniform "teaching culture" that values traditional approaches (Brousseau et al., 1988). On the other hand, others may argue that experience brings stronger understanding of the nuances of teaching and ability to change their thinking more quickly. By looking at how teachers' beliefs about teaching and their own self-efficacy change over time as they engage in curriculum-based PD, we can begin to answer these questions, providing a more nuanced picture of teacher learning than trying to determine how many hours is "sufficient."

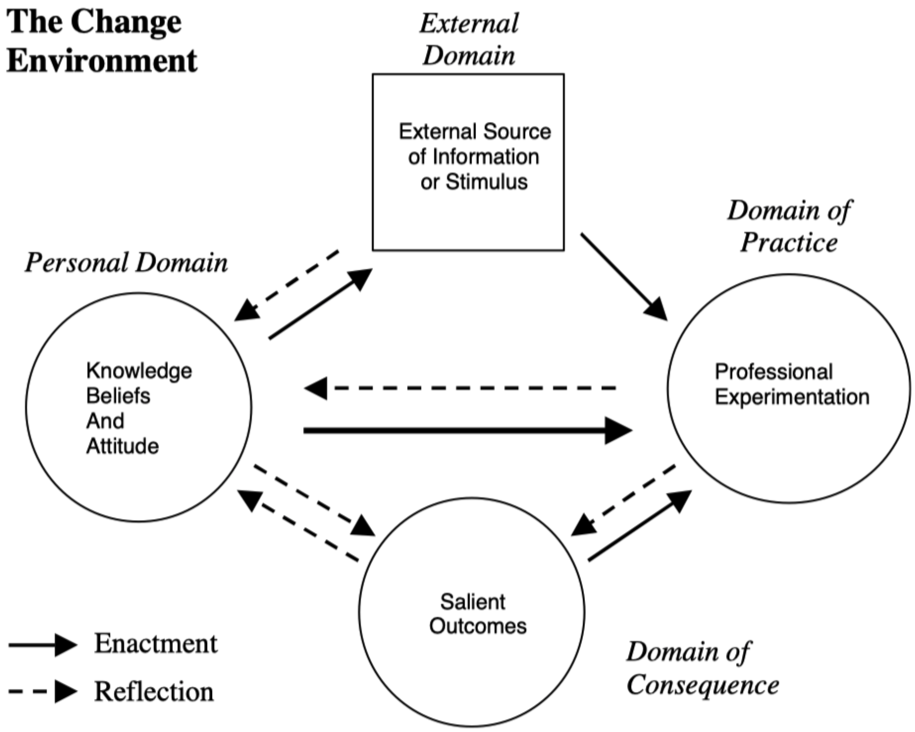
Theoretical Framework

Clarke and Hollingsworth's (2002) Integrated Model of Professional Growth (IMPG) helps to conceptualize the relationship between teachers' experience in PD, their own beliefs, and their instruction. For many early designers and researchers of PD, the theory of change undergirding the work was that PD could initiate changes in teachers' thinking and beliefs, which would then lead to change in their practice, and ultimately changes in student learning (Clarke & Hollingsworth, 2002). This theory, which was relatively implicit in initial PD studies, was directly challenged by Guskey (1986), who posited that the order of outcomes was flawed. Instead, he suggested that PD leads to a change in teacher practices, which results in improved learning outcomes, and ultimately a change in teacher's beliefs about learning. What both

Guskey and the earlier theory of change he was critiquing had in common, therefore, was a view of teacher learning as a linear series of changes resulting from engagement in PD.

The IMPG came as a growth from these theories as well as developing understandings of teacher learning as both the development of multiple kinds of knowledge (Shulman, 1987) and as situated in an ecosystem of interactions between individuals and their contexts (Greeno, 1997). Drawing on these ideas, the IMPG posits that change takes place in four different domains: the external domain, meaning the external stimuli or sources of information provided to a teacher; the personal domain, meaning the teacher's knowledge, beliefs, and attitudes; the domain of practice, meaning what the teacher actually does in the classroom; and the domain of consequence, meaning the outcomes the teacher values, such as student performance (Clarke & Hollingsworth, 2002). Figure II.1 below outlines this model. The two key ideas from this theory of professional learning are that changes in these four domains are not linear but rather cyclically build across each other and that the processes that mediate these changes are teachers' reflection and enactment.

Figure II.1: The Integrated Model of Professional Growth



Note. From Clarke & Hollingsworth, 2002, figure 3

Research Questions

Given the role of teachers' beliefs and confidence in influencing both what they learn from PD and how they ultimately implement those ideas in the classroom, I investigated how these constructs changed over time as a group of science teachers engaged in curriculum-based PD. Specifically, I asked parallel questions about teachers' beliefs and confidence: 1a) how do *teachers' beliefs about teaching* change over the course of learning about and implementing new science curricula? and 1b) what teacher characteristics predict differences in initial *beliefs* and change over time? 2a) how do teachers' *confidence* change over the course of learning about and implementing new science curricula and 2b) what teacher characteristics predict differences in both initial *confidence* and their change over time?

Methods

To answer these research questions, I analyzed survey results from a group of teachers who participated in four rounds of curriculum-based PD from mid-2018 through early 2020. This curriculum-based PD took place in the context of the field test for the OpenSciEd middle school science curriculum, so I begin by describing the curriculum and PD context before discussing the participants, measures, and analyses.

Context

OpenSciEd is a non-profit organization that designs reform-oriented science curricula and releases the curricular materials and curriculum-based PD as open educational resources (OpenSciEd, 2020). It began in 2018 with middle school (6th, 7th, and 8th grade) curriculum development. Each unit went through a multistep development process; after being written, units were field tested in classrooms in the 10 OpenSciEd partner states, externally reviewed by the NextGenScience Peer Review Panel, and revised before public release (Edelson et al., 2021).

The OpenSciEd curricular materials are designed to take a storyline approach, which means they are centered around students figuring out a complex anchoring phenomenon and each lesson is motivated by questions students co-construct with the teacher about that phenomenon (Reiser et al., 2021; Zivic et al., 2018). For example, the eighth-grade unit on sound begins with students watching a video of music playing from a car causing windows on the other end of a parking lot to move, which leads to questions about how sound is generated and travels. As a result of these questions, the class then investigates various instruments and sound makers to determine how sound is related to vibration before investigating how sound moves and transfers energy through air (OpenSciEd, 2019). The idea for this approach is that even though the teacher has an idea of where the class might go next, the steps are motivated by student questions about the phenomenon rather than a list of topics or sections stipulated by the textbook (Reiser et al., 2021). This curricular approach is relatively novel, and therefore most teachers need support in effectively implementing it. To provide that support, OpenSciEd also has designed curriculum-based PD to align with the curricular materials it releases (Edelson et al., 2021; McNeill & Reiser, 2018).

Curriculum-Based Professional Development Context

This study focuses on surveys taken after the curriculum-based PD given during the field test of each OpenSciEd unit. PD sessions were designed and implemented as new units were field tested, which happened every six months beginning in June 2018 and lasting until January 2021, totaling six rounds of PD and curriculum in three years. Because the last two rounds were substantially disrupted by the COVID pandemic, I focus on the first four rounds of PD, which were field tested in June 2018, January 2019, June 2019, and January 2020. The first round was a four-day introduction to OpenSciEd, the storyline approach, and the particular curricular unit teachers would be implementing (which was different for 6th, 7th, and 8th grade teachers). Subsequent rounds were either 2 or 2.5 days long and included a focus on a central topic and an introduction to the new curricular unit. These PD topics were determined based on feedback from teachers after past rounds and included identifying and reflecting on key elements of OpenSciEd (round 2), supporting equitable sensemaking through class discussion (round 3), and using assessment to support sensemaking (round 4).

The PD sessions for each round had a similar structure: each day began with participants from across all three grade levels working together to learn about the problem of practice, they then split into grade-level groups to concentrate on their particular unit, and finally ended the day coming back together to synthesize their learning across grade levels. A set of design principles guided the choice of the PD structure and the particular activities selected for each round (Edelson et al., 2021). Those principles were supporting teachers in: taking the student perspective, analyzing images of classroom instruction, examining contrasting cases of curriculum or instruction, and engaging in cycles of enactment and reflection (McNeill et al., in press). The first principle informed the decision to ask teachers to experience lessons from the curriculum as a student. The second informed the decision to use classroom video to illustrate key moments from the curriculum. The third informed the decision to read vignettes that contrasted traditional and OpenSciEd approaches to instruction. The fourth principle informed the overall structure of multiple rounds of PD taking place over time while teachers enacted new units between each round. Table I1.1 gives a summary of the focus and activities of the first four rounds of PD.

Table II.1: Summary of OpenSciEd PD Activities

	Day 1	**Day 2**	**Day 3**	**Day 4**
Round 1: Getting to Know OpenSciEd (June 2018)	Introduction to OpenSciEd, experience an OpenSciEd phenomenon as a student Experience anchoring phenomenon of target unit as a student	Watch classroom video & read vignettes to consider coherence between lessons Construct summary of unit storyline, experience key lessons as a student	Consider talk moves to support discussions, watch classroom video of discussions Plan and rehearse a classroom discussion, experience key lessons as a student	Discuss assessment in OpenSciEd, analyze a sample assessment Analyze assessment and student work from the unit, experience key lessons as a student

	Day 1	Day 2	Day 3	Day 4
Round 2: OpenSciEd Key Instructional Elements (January 2019)	Reflect on experience teaching OpenSciEd, analyze video for key elements of instruction Experience anchoring phenomenon of target unit as a student, construct summary of unit storyline	Analyze classroom video for teacher moves that support key elements of instruction Experience key lessons as a student	---	---
Round 3: Supporting Equitable Sensemaking (June 2019)	Analyze video & written scenarios for students' diverse ideas and resources Experience anchoring phenomenon of target unit as a student (half day)	Examine resources and watch classroom video considering how to support equitable classroom culture. Construct summary of unit storyline, experience key lessons as a student	Read discussion planning tool, watch video of a teacher planning & implementing a discussion. Plan and rehearse a classroom discussion, experience key lessons as a student	---
Round 4: Using Assessment to Support Sensemaking (January 2020)	Reflect on and analyze strong assessments Experience anchoring phenomenon of target unit as a student, analyze unit mid-point assessment, construct summary of unit storyline	Analyze student work and discuss how to use assessment data Experience key lessons as a student, review assessments in unit, analyze student work from unit	---	---

Participants

The OpenSciEd units and accompanying PD sessions were field tested in 10 partner states that represent both geographic and demographic diversity: California, Iowa, Louisiana, Massachusetts, Michigan, New Jersey, New Mexico, Oklahoma, Rhode Island, and Washington. In order to facilitate efficient delivery of each round of PD across all 10 states, a cadre of PD facilitators was trained centrally before each round of the field test.

Each state recruited a cohort of teachers to field test both the PD and curricular materials. The size of those cohorts ranged from 14 to 50 teachers, although each state experienced some attrition throughout the field test and some added new teachers part way through the field test. The sample for this study is the field test teachers who engaged in at least one PD workshop, which totaled 322 participants. In terms of gender, race, and teaching experience, the sample is roughly similar to the national population of middle school teachers. Among the participants, 72% identified as a woman, 77% identified as White, and the average years of teacher experience was 12.2 (s.d. = 8.26) with a range of 0 to 39. Nationally, 72.1% of middle school teachers identify as women, 79.2% identify as White, and the average years of experience of middle school teachers is 13.9 (Taie & Goldring, 2020).

Measures

Before the first PD, participants took a survey asking about their background and teaching context, their instructional beliefs, and their confidence in implementing the OpenSciEd curricular approach. After each round of PD, they took a survey asking about their instructional beliefs, implementation confidence, and their perception of the value of the various PD activities. This study draws from five surveys taken across the first two years of the project: the background, post-PD1, post-PD2, post-PD3, and post-PD4 surveys.

Outcome Variables: Traditional Beliefs and Implementation Confidence

To assess the teachers' instructional beliefs, the pre-survey, PD1, and PD3 surveys included a set of items initially reported by Reiser and colleagues (2017). Some of the items in the set were initially asked by Banilower and colleagues in the 2012 *National Survey of Science and Mathematics Education* (2013) while others were originally written by Reiser and colleagues (2017). Eight of these items (items 1-8 in Table II.2) were a set of items that Reiser and colleagues called "beliefs about traditional instruction" because they represent ideas that teachers might hold that run counter to intended instructional reforms, such as a focus on students memorizing vocabulary and engaging in "hands on" investigations solely in order to increase engagement (Furtak & Penuel, 2019). The other two (items 9 and 10) were from a set Reiser and colleagues (2017) called "beliefs about using student ideas in instruction." Participants were asked to rate their level of agreement with each belief item on a six-point Likert scale ranging from "strongly disagree" to "strongly agree."

The confidence items were based on confidence items used by Reiser and colleagues (2017). They were designed to determine the degree to which teachers believed they could implement the routines of storyline curricula, or their own self-efficacy (Bandura, 1997). The items in this survey were written to target the specific components of the OpenSciEd curricular approach, asking teachers to rate their confidence on a five-point Likert scale ranging from "very unsure" to "very confident." Table II.2 shows all of the belief and confidence items participants answered.

Validity refers to whether an instrument actually measures what it claims to measure (Field, 2013). Evidence for validity includes that the instrument has been used in past studies to measure the same construct being measured in this study. Reiser et al. (2017) designed this instrument to measure science teachers' traditional educational beliefs and confidence implementing storyline curricula. As part of that design, they used some items from Banilower and colleagues (2013), who also designed their instrument to measure science teachers' traditional educational beliefs. Therefore, based on the past use of this instrument by both Reiser and colleagues (2017) and Banilower and colleagues (2013), I concluded this was a valid instrument to measure teachers' traditional instructional beliefs and confidence for implementing storyline curricula.

Table II.2: Belief and Confidence Items

Belief Items[a]	
For each of the statements below, state the degree to which you agree or disagree.	
1	Teachers should have students do interesting hands-on activities, even if the activities do not relate closely to the concept being studied.
2	Hands-on/laboratory activities should be used primarily to reinforce a science idea that the students have already learned.
3	At the beginning of instruction on a science idea, students should be provided with definitions for new scientific vocabulary that will be used.
4	Teachers should provide students with the outcome of an activity in advance so students know they are on the right track as they do the activity.
5	Students should do hands-on or laboratory activities, even if they do not have opportunities to discuss them as a class.
6	Teachers should explain an idea to students before having them consider evidence that relates to the idea.
7	Students should know what the results of an experiment are supposed to be before they carry it out.
8	When students do a hands-on activity and the data don't come out right, teachers should tell students what they should have found.
9	Students' ideas about a science concept should be deliberately brought to the surface prior to a lesson or unit so that students are aware of their own thinking.
10	It is better for science instruction to focus on ideas in depth, even if that means covering fewer topics

Implementation Confidence Items[b]	
How confident are you that you can...	
11	Get students to ask questions at the beginning of a unit that guide the lessons that follow?
12	Work with students to motivate the next step in investigating a phenomenon, rather than just telling them what they will do next?
13	Help students use science practices to figure out pieces of core science ideas?
14	Push students to go deeper to revise their explanatory models of phenomena?
15	Help students put pieces together related to disciplinary core ideas and crosscutting concepts?

[a] Answer options:1 = strongly disagree, 2 = moderately disagree, 3 = slightly disagree, 4 = slightly agree, 5 = moderately agree, 6 = strongly agree.
[b] Answer options:1 = very unsure, 2 = somewhat unsure, 3 = in the middle, 4 = somewhat confident, 5 = very confident.

I performed an exploratory factor analysis on the belief items and confidence items separately at each time point to determine how those variables were behaving. For the belief items, there was one factor with an eigenvalue higher than 1 at each time point, which consisted of items 1-8 from Table II.2. The remaining two items (9 and 10) did not load onto this factor and were therefore excluded from analysis. All five confidence items loaded onto one factor at each time point. I named these two factors "implementation confidence" and "traditional beliefs." See the Appendix for detailed results of this factor analysis.

When doing factor analysis, reliability testing is an important step to ensure the identified factors can be interpreted consistently across many different situations; in other words, two people who are the same on the factor measured should receive the same score (Field, 2013). The most common way to test reliability is Cronbach's Alpha, which is a measure of the internal consistency, and therefore reliability, of a factor (Field, 2013). I tested the reliability of the two factors I identified by calculating Cronbach's alpha at each time point, as shown in Table II.3. For every time point, the reliability was above 0.7, which is generally considered acceptable, and in fact all but one reliability measure was above 0.8, which is generally considered a benchmark of good reliability (Field, 2013; Netemeyer et al., 2003). Therefore, I concluded that this was a reliable measure of traditional beliefs and implementation confidence.

Table II.3: Reliability Analysis of Outcome Variables (Cronbach's Alpha)

Factor	Pre-PD	Post-PD1	Post-PD2	Post-PD3	Post-PD4
Traditional beliefs	.759	.818		.833	
Implementation confidence	.880	.835	.872	.823	.861

Based on the factor analysis, I created a total score for each factor by summing the scores of each item in the factor, a common method for creating factor scores from Likert-type items (Netemeyer et al., 2003). Because the eight belief items each had six response options, the total score had a theoretical range of 8-48, with 8 representing lowest levels of traditional beliefs and 48 representing the highest levels. Similarly, the five confidence items had five response options each, meaning the implementation confidence total score had a theoretical range of 5-25, with 5 representing the lowest possible confidence and 25 the highest. Table II.4 shows the descriptive statistics for the outcome variables at each time point.

Table II.4: Descriptive Statistics for Outcome Variables

Factor	N	Mean	SE	SD	Min-Max	95% Confidence Interval	
						Lower Bound	Upper Bound
Traditional beliefs							
Pre-survey	301	22.75	.40	6.97	8-46	21.96	23.54
PD1	207	17.56	.46	6.61	8-48	16.66	18.47
PD3	172	17.64	.56	7.34	8-48	16.54	18.74
Implementation confidence							
Pre-survey	304	19.50	.21	3.69	5-25	19.08	19.92

PD1	209	20.20	.21	3.01	8-25	19.79	20.61
PD2	197	20.78	.23	3.29	9-25	20.31	21.24
PD3	173	21.25	.22	2.87	12-25	20.82	21.68
PD4	162	21.06	.24	3.11	12-25	20.57	21.54

Note. The pre-survey was given in May 2018, PD1 in June 2018, PD2 in January 2019, PD3 in June 2019, and PD4 in January 2020

Predictor Variables

I used two sets of predictor variables in this study: a set of items asking teachers to rate how valuable they found various PD activities and a set of teacher characteristics items. Below I discuss these items in more depth and provide descriptive statistics for each.

Value of PD Activities Items. Each post-PD survey asked teachers to rate on a 3-point Likert scale the degree to which they found specific activities from the PD valuable (i.e., 1 = not at all valuable, 2 = somewhat valuable, and 3 = very valuable). These included items that were consistent across all PD rounds and activities that were unique to each round. I analyzed the three items that were identical for all four PD sessions because they represented activities that happened in each workshop. These items were: “engaging in the anchoring phenomenon for your specific unit,” “conducting investigations from the students’ perspective for your specific unit,” and “building the storyline for your specific unit.”

For each time point, I began by conducting an exploratory factor analysis and found that the three items each loaded onto one factor. Given that these activities were different, I interpreted this factor to represent how valuable the teacher found the PD in general. I created a composite score for each time point by summing the scores from the three items. Although these items were asked across time, I was unsure if participants’ responses changed over time, so I then constructed an unconditional linear growth model of this value construct using HLM 8 software (Raudenbush et al., 2019). I found that the amount teachers valued the PD activities did not vary significantly over time ($\beta = -0.04$, $p = .108$), and therefore concluded that teachers' value of PD activities could be modeled by a between-person variable. I constructed this variable by averaging the summed PD value scores for each participant across the four time points. Descriptive statistics for this final variable are shown in Table II.5.

Teacher Background Characteristics. I included the following teacher characteristics as potential covariates: participants' gender identity, race, years of teaching experience, percent of career spent teaching middle school science, minutes per week they teach science, past amount of NGSS-aligned PD, and how recently their state had adopted the NGSS or standards based on the NGSS. For gender, the sample was heavily weighted towards women, so I constructed a dummy variable with woman as the reference category and compared that to anyone who did not identify as a woman. Because the sample was so predominantly White, there was not enough variability to make meaningful comparisons between any other racial categories. Therefore, for race I constructed one dummy variable comparing participants who identified as White and those who did not identify as White. The pre-survey asked teachers how many years of teaching experience they had prior to the first PD as well as how many years they had taught middle school science. In order to avoid collinearity between these two variables, I calculated the percent of each participant's career that had been spent teaching middle school. The pre-survey asked teachers both how many days per week they taught science and how many minutes each class period was; I multiplied these two values together to find how many minutes per week they taught science class. Teachers self-reported the number of hours of NGSS-aligned PD they had experienced before the first round of PD. In order to facilitate interpretation of this number, I divided it by eight to represent a standard eight-hour "day" of PD. In order to determine how recently each participant's state had adopted NGSS, I referenced each state's standards adoption timeline and constructed a variable representing how many years before the first round of PD the state had adopted NGSS.

Once I constructed all relevant items, I used a boxplot to find outliers of greater than three standard deviations away from the mean (J. W. Osborne, 2010). I inspected each outlier to determine if it was likely due to a data input error and addressed each through either correcting, truncating, or deleting (Aguinis et al., 2013). For example, one teacher wrote they taught 200-minute classes 5 days a week, which was likely an input error because a second teacher at the same school taught 45-minute classes, so I corrected the entries to match. On the other hand, another teacher wrote they taught 60-minute classes 0 days a week; because no other teacher from the sample taught in that same school, I deleted that entry for that variable. Four participants wrote that they had experienced over 1000 hours of NGSS-aligned PD by 2018, which would represent 125 or more days of PD in the five years between the release of the NGSS in 2013 and the time of data collection. Given the large amount of time this represents, I truncated each of these entries to 75 days, which represents 15 8-hour days of PD per year between 2013 and 2018, a value that was still very high but more likely to be realistic (Aguinis et al., 2013). Table II.5 shows the descriptive statistics of all of the predictor items after data cleaning.

Table II.5: Descriptive Statistics for Predictor Variables

Variables	N	Mean/%	SE	SD	Min-Max
Value Placed on PD Activities (average)	311	8.45	0.05	0.80	5-9
Gender	302				
Woman		77.5			

Not a woman		22.5			
Race	303				
White		82.2			
Not white		17.8			
Years Teaching	309	12.23	0.47	8.26	0-39
Percent of career teaching middle school science	302	77.11	1.76	30.61	0-100
Minutes per Week Teaching Science	307	271.62	3.63	63.54	90-450
Previous Days of NGSS-Aligned PD	300	6.60	0.63	10.99	0-75
Years Since NGSS Adoption	322	2.71	0.10	1.73	0-5

Data Analyses

Because the outcome variables were measured at a different number of time points (three for traditional beliefs, five for implementation confidence), I investigated changes in each construct separately. I began by graphing the average change in each construct over the time studied, which is a helpful preliminary step when working with longitudinal data to understand their overall shape (Singer & Willett, 2003). In order to explain variations in initial status and change over time, I constructed a set of two-level hierarchical growth models, with time points nested within individuals (Hedeker & Gibbons, 2006; Raudenbush & Bryk, 2002). This approach is necessary because hierarchical linear modeling (HLM) can account for self-correlation when multiple data points are collected from the same individuals over time (Raudenbush & Bryk, 2002). In addition, it is robust to individuals who do not have a full set of data points over time, a particularly common problem in longitudinal studies (Hedeker & Gibbons, 2006). This approach can also take into account that individuals had different amounts of experience with the PD, with some entering at later rounds than others. A common concern with modeling growth over time is assuming linear growth when change may in fact be curvilinear. To address this concern, I added higher order variables of time to determine if growth was curvilinear (Singer & Willett, 2003).

For each outcome variable, I began by constructing an unconditional, quadratic growth model with random intercept and slopes. For both models, I found that the variability in the slope for $Time^2$ was not significantly different from zero. Therefore, I fixed that slope and my initial unconditional models took the following form:

Level 1: $Y_{ti} = \pi_{0i} + \pi_{1i}(Time) + \pi_{2i}(Time^2) + e_{ti}$

Level 2: $\pi_{0i} = \beta_{00} + r_{0i}$

$\pi_{1i} = \beta_{10} + r_{1i}$

$\pi_{2i} = \beta_{20}$

In this model, Y_{ti} represents the outcome variable (traditional beliefs or implementation confidence) at time *t* for individual *i*. The level 1 intercept (π_{0i}) represents the predicted value of the outcome for individual *i* at time 0. The level 1 coefficient for time (π_{1i}) represents the predicted change in the outcome variable for a 1-unit change in time. The coefficient for $time^2$ (π_{2i}) represents changes in the rate of change of the outcome variable over time. Because I was interested in how beliefs and confidence change after engaging in rounds of PD and curriculum enactment, I coded the pre-survey at time 0 and each subsequent PD as time 1-4.

My goal was to construct the most parsimonious model that could explain differences in traditional beliefs and implementation confidence. Therefore, I employed a forward-selection model building process in which I added the level-2 predictor variables into the model one at a time and assessed the significance of each predictor. Predictors whose coefficients were significantly different from zero at the level of $\alpha < .05$ were retained and those that were non-significant were removed. I also monitored the deviance, reliability, and intraclass correlation coefficient to track the overall quality of the model over time. When I arrived at my final model, I used the variance of the final and unconditional models to calculate the amount of variance the model explained for each random-effects predictor (Raudenbush & Bryk, 2002). The Appendix at the end of this paper provides additional methodological details about the data cleaning, analysis, and results.

Results

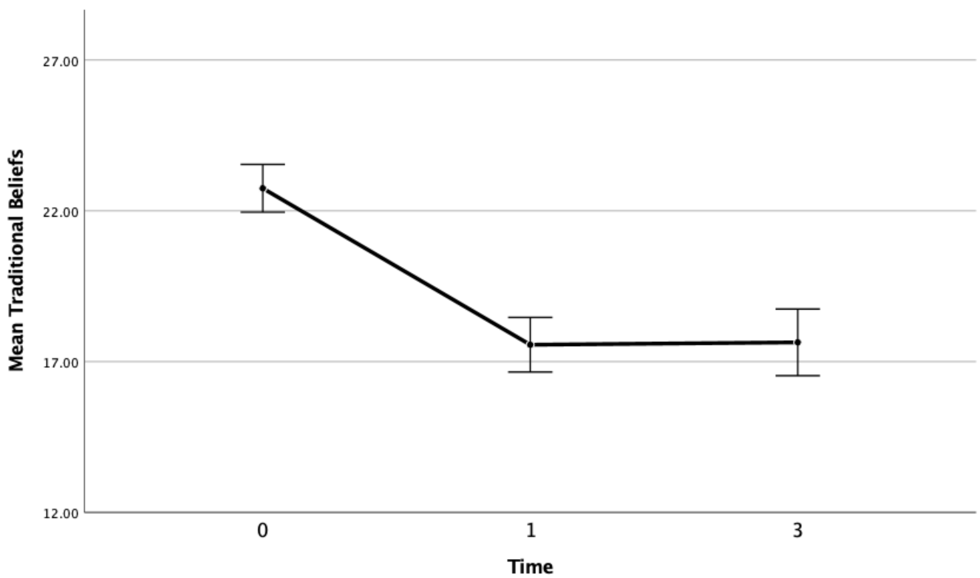
Based on a set of teacher surveys, I tracked how teachers' traditional beliefs and implementation confidence changed over time. For each construct, I began by graphing the mean scores across all of the teachers over time, which showed significant changes that began to level off over time. I then constructed a longitudinal two-level HLM model for both traditional beliefs and confidence to identify teacher characteristics that helped explain differences in these outcomes. I found that teachers with more teaching experience, more past PD, and who found the PD activities more valuable were less likely to have traditional beliefs. Similarly, teachers with more past PD were more likely to have higher implementation confidence and teachers who found the PD more valuable were more likely to increase in confidence more quickly.

Teachers' Traditional Beliefs

Teachers' traditional beliefs showed significant change after the initial round of PD that leveled off over time. Before beginning any PD, the average traditional belief score across all teachers was 22.75 points, which decreased to 17.64 points by the final measured time point after the third round of PD. This corresponds to a shift in average rating from around "slightly disagree" to items like "Teachers should provide students with the outcome of an activity in advance so students know they are on the right track as they do the activity" on the pre-survey towards "moderately disagree" on the final survey (Table II.2 includes all of the belief items).

Figure II.2 is a graph of this change over time, demonstrating that the significant decrease in traditional beliefs occurred between the pre-survey and after the first round of PD with no significant change between rounds 1 and 3. The 95% confidence intervals around the mean, shown in Table II.4, confirm that traditional beliefs decreased significantly between time point 0 and 1 but did not show significant change between time points 1 and 3 (note that the belief items were not asked after PD 2 and 4).

Figure II.2: Mean Traditional Beliefs over Time



Note. Error bars represent 95% confidence intervals

As an overall group, teachers moved away from traditional beliefs after round 1 and stayed at about that same level through subsequent rounds. The problem with looking only at means, however, is it can mask complexity in the sample. In order to see if teachers showed this shift equally and which teacher characteristics predict differences in this shift between teachers, I constructed a two-level longitudinal model (Raudenbush & Bryk, 2002). Table II.6 includes both the unconditional and final model. The within-individual results confirm my conclusions for research question 1a that there was a significant decrease in levels of traditional beliefs over time but that change was not linear; the rate of decrease of traditional beliefs slowed over time. The results at the between-individual level helps to answer research question 1b, providing information about which teacher characteristics predict differences in traditional beliefs. Teachers who had more experience teaching, had experienced more aligned PD, and who placed more value on the PD activities were more likely to disagree with the traditional beliefs statements. None of the variables tested were significant predictors in differences over time.

Table II.6: Traditional Beliefs Models

Effect	Traditional Beliefs	
	Unconditional β (SE β)	Final β (SE β)
	Fixed effects	
Status at presurvey (π_{oi})		
Intercept (β_{00})	22.79^{***} (0.40)	23.24^{***} (0.41)
Years of teaching experience (β_{01})		-0.09^{*} (0.04)
Previous days of PD (β_{02})		-0.09^{**} (0.03)
Value of PD activities (β_{03})		-1.71^{***} (0.45)
Linear rate of change (π_{1i})		
Time (β_{10})	-7.02^{***} (0.58)	-6.94^{***} (0.59)
Quadratic rate of change (π_{2i})		
Time^2 (β_{20})	1.82^{***} (0.18)	1.81^{***} (0.19)
	Random effects	
Variance components		
Level 1 (σ^2)	17.74[a]	17.69[a]
Level 2 intercept ($\tau 00$)	29.31^{***}	23.56^{***}
Level 2 time ($\tau 10$)	0.88^{**}	0.92^{**}
	Model fit statistics	
Deviance	4389.89	4165.60
Intraclass Correlation Coefficient	0.63	0.57

*p<.05; **p<.01; ***p<.001

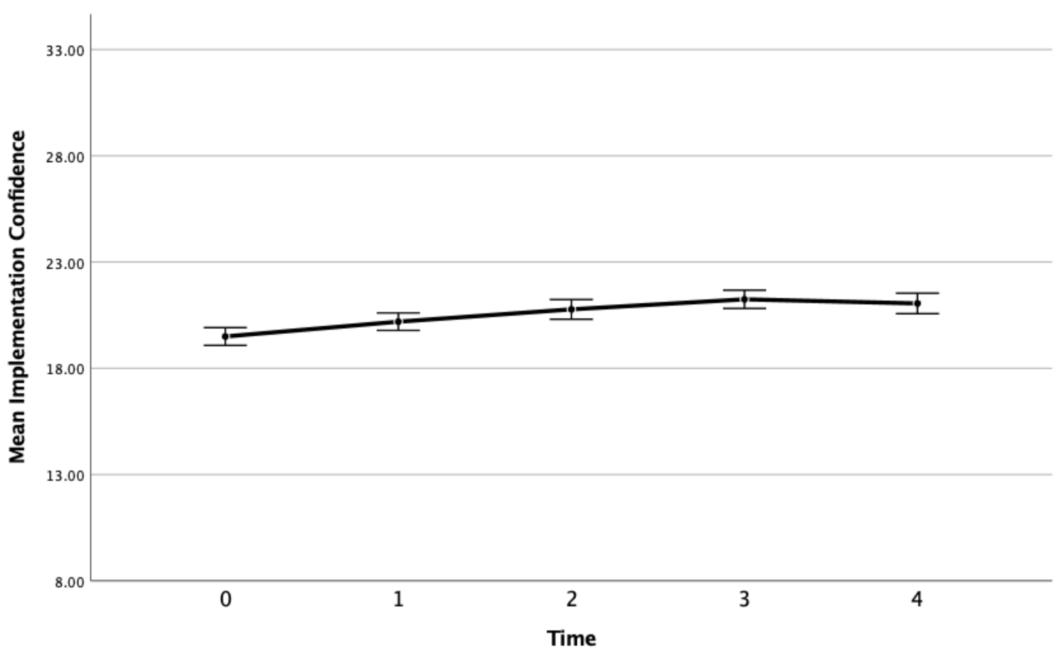
[a]The HLM software does not conduct a hypothesis test for the level 1 variance, so no p-value is reported

The percent of variance the model explained at each level helps to determine how well this model explains differences between teachers. In comparison to the unconditional model, the final model explained 20% of the variance in initial levels of traditional beliefs and 1% or less of the variance in change over time or within people, which is expected given that there were no significant predictors within individuals or for change over time.

Teachers' Implementation Confidence

The pattern for implementation confidence was similar to the one for traditional beliefs: confidence increased over the early rounds of PD and then leveled off over time. Initially, the average score of implementation confidence was 19.50, which corresponded with a rating of "somewhat confident" in response to items like "help students use science practices to figure out pieces of core science ideas" (see all confidence items in Table II.2). By the end of the final PD, the average score was 21.06, which corresponded with a rating between "somewhat" and "very confident." Figure II.3 is a graph of implementation confidence over time, showing a somewhat linear increase between time points 0 and 3 and a slight but not significant decrease between time points 3 and 4. The 95% confidence intervals around each mean (see Table II.4) illustrate that adjacent time points were not always significantly different, but that over time teachers' confidence increased until time point 3 and then plateaued.

Figure II.3: Mean Implementation Confidence over Time



Note. Error bars represent 95% confidence intervals

The mean results for both traditional beliefs and confidence are encouraging. They show that over time teachers did grow in terms of their beliefs around what good science instruction looks like and their self-confidence in implementing OpenSciEd. In addition, the confidence results in particular show the value of engaging in this work iteratively over time; teachers' confidence initially did not change between the pre-survey and the end of PD 1, but it did continue to change for the next year up through round 3. These results support the value of the iterative, long-term approach to science teacher learning taken by the OpenSciEd field test approach.

Similar to the traditional beliefs models, I then constructed a two-level longitudinal model to explain more nuance than can be seen by averages on their own. The within-person level confirmed that there was a significant increase in implementation confidence over time and that the rate of change of confidence decreased over time. The between-person model helped to identify which teacher characteristics predict differences in implementation confidence (research question 2b). Table II.7 shows the unconditional and final models for implementation confidence. The HLM model shows that teachers with more previous related PD were more likely to have higher implementation confidence. In addition, teachers who found the PD activities more valuable were likely to show faster increases in confidence over time.

Table II.7: Implementation Confidence Models

Effect	**Implementation Confidence**	
	Unconditional β (SE β)	Final β (SE β)
	Fixed effects	
Status at presurvey (π_{oi})		
Intercept (β_{00})	19.42^{***} (0.21)	19.47^{***} (0.20)
Previous days of PD (β_{01})		0.05^{***} (0.01)
Linear rate of change (π_{1i})		
Time (β_{10})	0.84^{***} (0.19)	0.71^{***} (0.20)
Value of PD activities (β_{11})		0.33^{***} (0.08)
Quadratic rate of change (π_{2i})		
Time^2 (β_{20})	-0.10^{*} (0.05)	-0.10^{*} (0.05)

Random effects		
Variance components		
Level 1 (σ^2)	5.47[a]	5.27[a]
Level 2 intercept ($\tau00$)	6.68***	5.70***
Level 2 time ($\tau10$)	0.10†	0.10*
Model fit statistics		
Deviance	5208.46	4886.99
Intraclass Correlation Coefficient	0.55	0.52

†p<.10; *p<.05; **p<.01; ***p<.001
[a]The HLM software does not conduct a hypothesis test for the level 1 variance, so no p-value is reported

Once again, the amount of variance explained at each level helps show how well the final model explained variation in implementation confidence. In comparison to the unconditional model, the final model explained 15% of the variance in initial confidence, but less than 1% of the variance in change over time and 4% of the variance within people. This means that there are likely other characteristics that explain differences in implementation confidence.

Summary of Results

Overall, there was a similar pattern of change over time in the survey results related to teachers' confidence about implementing the OpenSciEd curriculum and their traditional beliefs about science teaching. Teachers showed significant gains in confidence and a significant decrease in agreement with traditional science teaching statements after initial rounds of PD. In later rounds, these changes leveled off and teachers' confidence and beliefs did not significantly change. The biggest difference between these two measures was that beliefs showed significant changes after round 1 of PD and then plateaued, whereas confidence did not show significant changes until after round 2. I then constructed multilevel models to identify teacher characteristics that predicted differences between teachers in their changes in beliefs and confidence. Teachers with more experience teaching, more past PD around the NGSS, and who found the PD more valuable were less likely to agree with traditional beliefs about science instruction. Similarly, teachers who had more past PD about the NGSS were more likely to be confident in their OpenSciEd implementation and teachers who found the PD more valuable increased their confidence faster. These models explained 20% or less of the variance in traditional beliefs and implementation confidence, meaning that there are likely other unmeasured variables contributing to these differences between teachers.

Discussion

The study of how the variable of time interacts with PD outcomes is quite popular, but mostly in terms of looking at PD duration, asking if the number of PD contact hours is associated with positive teacher or student learning outcomes (e.g. Blank & de las Alas, 2009; Cohen & Hill, 1998; Crowther & Cannon, 2002; Desimone et al., 2002; Hill & Ball, 2004; Lynch et al., 2019; Yoon et al., 2007). I argue that finding the "correct" number of hours for PD is a pursuit that is both simplistic and likely impossible. It is not simply the existence of many contact hours over a long span of time that supports teacher growth, but rather what actually happens within those hours, why, and how what happens does or does not support teacher learning and growth (Kennedy, 2016; Lynch et al., 2019). Rather than trying to answer the question of how many hours over what duration of time is most "effective," I tracked the same group of teachers over time as they engaged in multiple PD workshops and implementation efforts in their own classroom in order to understand the shape of change in their traditional beliefs and implementation confidence over time. In addition, using an HLM model, I have identified teacher characteristics associated with these changes.

In this section, I highlight three key takeaways from these analyses and discuss their implications for practice: 1) beliefs and confidence showed initial changes that leveled out over time; 2) confidence required more time to change than beliefs; and 3) finding the PD activities valuable was associated with lower traditional beliefs and faster increase in confidence. The section ends with a discussion of limitations of this study and how future research might address these limitations.

Beliefs and Confidence Change and Level Over Time

After initial rounds of PD, teachers expressed fewer traditional beliefs and higher confidence, but those changes leveled off over time. This finding aligns with other longitudinal studies of teacher learning from multiple rounds of PD over time; for example, Heck and colleagues (2006) found that teachers "pedagogical preparedness" and "content preparedness" increased but leveled off over the course of multiple years of curriculum-based PD. Teachers' belief systems are complex and change in beliefs over time is rarely linear, meaning beliefs can become more aligned with reform ideas about instruction, regress, and/or plateau over the course of teachers' careers (Fletcher & Luft, 2011; Pilitsis & Duncan, 2012).

In this group of teachers, on average their beliefs and confidence plateaued in the second year of the study. This pattern of change has implications for future design of long-term curriculum-based PD as it calls into question the value of repeating similar approaches over time. Teachers with different instructional beliefs respond differently to different kinds of teacher learning activities (Pilitsis & Duncan, 2012). Therefore, a similar phenomenon may have occurred in this case in which the PD workshops were helpful for facilitating initial changes in beliefs and confidence but teachers may have needed more differentiated or targeted support as they became more experienced with the curriculum and its implementation in their classroom. More targeted instructional coaching activities such as examining their own students' work or engaging in lesson study might have provided further opportunities to facilitate growth in teachers' beliefs and confidence rather than continuing to use the same workshop model for PD in later rounds (Gibbons & Cobb, 2017).

Confidence Required More Time to Change than Beliefs

While the overall shape of change for both beliefs and confidence was similar, there was a key difference between the two in terms of how many rounds it took to show significant change. Traditional beliefs decreased significantly after one round of PD whereas confidence required two rounds to show significant increases. Multiple studies have shown that PD programs can increase teachers' teaching confidence (Kang et al., 2019; Lo et al., 2014; von Suchodoletz et al., 2018). The results from this study are in line with these previous studies but also demonstrate that confidence may require more time, iteration, or enactment to change than beliefs.

Teacher learning is not linear, but rather happens in a cyclical process of experience, enactment, and reflection (Clarke & Hollingsworth, 2002). These cycles are iterative, and teachers' thinking about their instruction and their instructional practice often change together (Kazemi & Hubbard, 2008). It makes sense, therefore, that confidence did not significantly change until after teachers had had an opportunity to implement an OpenSciEd unit. Teachers can express beliefs that do not align with their instructional practice without even noticing that discrepancy (Fletcher & Luft, 2011; Mills et al., 2019). Confidence, on the other hand, is more context-specific, meaning that teachers develop confidence about particular elements of their practice (Jones & Leagon, 2014). For example, one of the items asked about teachers' confidence "work[ing] with students to motivate the next step in investigating a phenomenon, rather than just telling them what they will do next" (see Table II.2). This is a particular feature of storyline curricula in which classes jointly establish their next steps even though the likely outcomes have been anticipated by the teacher and curricular materials (Reiser et al., 2021). Given its complexity, teachers may have needed opportunities to attempt this kind of joint navigation in their classrooms and then come back together to reflect with colleagues in a second PD to express more confidence in their ability to do this work. Overall, the results from this study show the value of creating PD opportunities that provide teachers with iterative opportunities to learn about, enact, and then reflect on instructional reforms in order to improve teachers' confidence in maintaining those reforms over the long-term.

Finding the PD Valuable was Associated with Lower Traditional Beliefs and Faster Increase in Confidence

Based on the HLM analysis, three teacher characteristics explained differences in teachers' traditional beliefs and implementation confidence. Those characteristics are past related PD experience, years of teaching experience, and finding the PD activities valuable. Teachers with more PD experience, more teaching experience, and who found the PD activities more valuable were more likely to disagree with the traditional belief statements. In terms of confidence, teachers with more past PD experience were more likely to report higher levels of confidence and teachers who found the PD more valuable were more likely to show a faster increase in confidence over time.

Of these variables, teachers' value of the PD is particularly interesting because it is a variable that can be designed for, unlike past PD or teaching experience which are characteristics of teacher participants that cannot be changed. Teacher attitudes about PD can vary widely, and the degree to which teachers find PD valuable impacts their willingness to participate in that PD (Masuda et al., 2013; Torff & Sessions, 2008). Another way to look at this is that teachers need to trust PD designers and facilitators to create learning environments that will support their development.

A key aspect of trust is that it involves being willing to take risks and be vulnerable because of the perceived positive intentions of the party with which one is placing one's trust (Smetana et al., 2016). Well-designed PD involves asking teachers to take intellectual risks in order to push their own thinking forward, so they need to trust the designers and facilitators that those risks will be helpful to their own learning. The data from this study show that teachers who saw that value in the OpenSciEd PD were more likely to quickly increase their confidence, meeting one of the key goals of the PD. Therefore, PD designers should consider how they build that trust with participants so that it is more likely the participants will meaningfully engage with the work of the PD. For example, connecting the work in PD with teachers' organizational contexts and allowing teachers to consider how to implement ideas from PD in their own context could be a powerful strategy in supporting teachers' buy-in to PD (Allen & Heredia, 2021).

In addition to trusting PD designers and facilitators, another aspect influencing teachers' attitudes towards PD is their motivation for attending (Masuda et al., 2013). Some teachers are intrinsically motivated to attend PD in order to develop their own professional practices and skills, whereas others are more extrinsically motivated by licensure or school requirements (Avidov-Ungar, 2016). Extrinsically motivated teachers may be less likely to see PD as valuable, so less likely to engage with the learning activities and therefore less likely to learn from them (Avidov-Ungar, 2016). This similar kind of feedback loop has been seen in other contexts of teacher learning. For example, when provided with educative curriculum materials, teachers who saw them only as a source of activities were less likely to learn from them than teachers who saw the materials as a resource to improve their own practice (Marco-Bujosa et al., 2017). Something similar may have been going on here. Teachers who saw the PD as valuable were more likely to get something out of it. While designers cannot change the reasons teachers attend PD, they can help to support intrinsic motivation for participants by pointing out the rationale behind various PD activities and how they might help teachers' learning. For example, teachers who understand that the goal of rehearsing a discussion is to consider and practice their teacher moves in a low-stakes environment (Lampert et al., 2013) might be more likely to engage meaningfully with that activity and therefore come away with more powerful insights into how they might run a similar discussion in their own classroom.

Limitations and Future Work

In this section I discuss two major limitations of this study and how future work could address those limitations moving forward. First, the design of data collection necessarily conflated teachers' PD and implementation experiences because surveys were only given after each round of PD but teachers taught OpenSciEd units between each round. Second, the HLM models explained relatively little of the variance between teachers in terms of changes in beliefs and confidence. This means the profile of teachers who responded well to the PD is limited and requires more work to better define and understand.

The relationship between teachers' experience in PD and their personal domain of beliefs and ideas is important but far from the only one that influences professional growth (Clarke & Hollingsworth, 2002; Kazemi & Hubbard, 2008). By collecting data on teachers' beliefs and confidence only after each PD, I was not able to see how changes in these constructs were influenced by teachers' experience teaching as compared to their time in PD. Future work could disentangle this by collecting data on teachers' beliefs and confidence during and immediately after implementation as well as after PD to provide a more complex picture of how these change over time.

The HLM models showed that teachers with more teaching and PD experience and who better valued the PD were less likely to hold traditional beliefs and more likely to be confident in their implementation. Unfortunately, these models explained relatively little of the variance in these outcomes, meaning there are certainly other characteristics that can help craft a profile of teachers likely to respond to this curriculum-based PD. Qualitative methods, such as case study research, are particularly well suited for building these profiles as they can help elucidate complexities of teacher learning and thinking that is obscured by large-scale survey analysis (Yin, 2018). Future work, therefore, should include studying teachers as they engage in a similar series of PD workshops and curriculum implementation to try to explain their unique patterns of learning and identify other characteristics that might explain variation between teachers.

Conclusion

We are in an exciting time in science teacher education. The NGSS and similar state standards call for wide scale change in our approach to science education, and we are just recently beginning to see the release of curriculum materials that truly embrace and implement these exciting reforms (Campbell & Lee, 2021). Pairing high-quality curriculum materials with professional development can be a particularly effective strategy to support teacher learning, but we still have much to learn about how teachers respond to these structures over time (Lynch et al., 2019; Short & Hirsh, 2020). The findings from this study provide evidence that curriculum-based professional development can help to positively impact teachers' thinking but also that more differentiated or targeted interventions may be needed as teachers become more experienced with the target curriculum. In addition, they demonstrate that while beliefs changed after one round of PD, confidence required two rounds and an intervening enactment of a new curriculum unit in their classrooms to change their confidence. This difference highlights the importance of iterative rounds of learning, enactment, and reflection in PD design (Kazemi & Hubbard, 2008; McNeill et al., in press). An HLM analysis showed that teachers who have more experience and better value the PD are likely to show quicker changes in confidence and less traditional beliefs, highlighting the importance of ensuring that teachers find PD activities meaningful and valuable for their learning. True instructional change requires a coherent system of teacher support over time (Loucks-Horsley et al., 2010). As we continue to work towards improving science teaching for all students, we should keep in mind how teachers' ideas change over time so that we can design supports that best meet teachers where they are.

References

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270–301. https://doi.org/10.1177/1094428112470848

Allen, C. D., & Heredia, S. C. (2021). Reframing organizational contexts from barriers to levers for teacher learning in science education reform. *Journal of Science Teacher Education*, *32*(2), 148–166. https://doi.org/10.1080/1046560X.2020.1794292

Allen, C. D., & Penuel, W. R. (2015). Studying teachers' sensemaking to investigate teachers' responses to professional development focused on new standards. *Journal of Teacher Education*, *66*(2), 136–149. https://doi.org/10.1177/0022487114560646

Avidov-Ungar, O. (2016). A model of professional development: Teachers' perceptions of their professional development. *Teachers and Teaching*, *22*(6), 653–669. https://doi.org/10.1080/13540602.2016.1158955

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.

Bang, M., Warren, B., Rosebery, A. S., & Medin, D. (2012). Desettling expectations in science education. *Human Development*, *55*, 302–318. https://doi.org/10.1159/000345322

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. M., Campbell, K. M., & Weiss, A. M. (2013). *Report of the 2012 national survey of science and mathematics education*. Horizon Research, Inc. http://www.horizon-research.com/2012nssme/research-products/reports/technical-report/

Blank, R. K., & de las Alas, N. (2009). *Effects of teacher professional development on gains in student achievement: How meta analysis provides scientific evidence useful to education leaders*. Council of Chief State School Officers. https://eric.ed.gov/?id=ED544700

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3–15. https://doi.org/10.3102/0013189X033008003

Brousseau, B. A., Book, C., & Byers, J. L. (1988). Teacher beliefs and the cultures of teaching. *Journal of Teacher Education*, *39*(6), 33–39. https://doi.org/10.1177/002248718803900607

Campbell, T., & Lee, O. (2021). Instructional materials designed for A Framework for K-12 Science Education and the Next Generation Science Standards: An introduction to the special issue. *Journal of Science Teacher Education*, *32*(7), 727–734. https://doi.org/10.1080/1046560X.2021.1975359

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

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(8), 947–967. https://doi.org/10.1016/S0742-051X(02)00053-7

Cohen, D. K., & Hill, H. C. (1998). *Instructional policy and classroom performance: The mathematics reform in California* (RR-39; CPRE Research Report Series). Consortium for Policy Research in Education. https://eric.ed.gov/?id=ED417942

Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, *44*(4), 613–642. https://doi.org/10.1002/tea.20157

Crawford, B. A., & Capps, D. K. (2018). Teacher cognition of engaging children in scientific practices. In Y. J. Dori, Z. R. Mevarech, & D. R. Baker (Eds.), *Cognition, Metacognition, and Culture in STEM Education* (Vol. 24, pp. 9–32). Springer International Publishing. https://doi.org/10.1007/978-3-319-66659-4_2

Crowther, D. T., & Cannon, J. R. (2002). Professional development models: A comparison of duration and effect. In P. A. Rubba, J. A. Rye, W. J. Di Biase, & B. A. Crawford (Eds.), *Proceedings of the 2002 annual international conference of the Association for the Education of Teachers in Science* (pp. 1397–1411). https://eric.ed.gov/?id=ED465639

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, *34*(3), 3–14. https://doi.org/10.3102/0013189X034003003

Davis, E. A., Palincsar, A. S., Smith, P. S., Arias, A. M., & Kademian, S. M. (2017). Educative curriculum materials: Uptake, impact, and implications for research and design. *Educational Researcher*, *46*(6), 293–304. https://doi.org/10.3102/0013189X17727502

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81–112. https://doi.org/10.3102/01623737024002081

Edelson, D. C., Reiser, B. J., McNeill, K. L., Mohan, A., Novak, M., Mohan, L., Affolter, R., McGill, T. A. W., Buck Bracey, Z. E., Deutch Noll, J., Kowalski, S. M., Novak, D., Lo, A. S., Landel, C., Krumm, A., Penuel, W. R., Van Horne, K., González-Howard, M., & Suárez, E. (2021). Developing research-based instructional materials to support large-scale transformation of science teaching and learning: The approach of the OpenSciEd middle school program. *Journal of Science Teacher Education*, *32*(7), 780–804. https://doi.org/10.1080/1046560X.2021.1877457

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics* (4th edition). Sage.

Fletcher, S. S., & Luft, J. A. (2011). Early career secondary science teachers: A longitudinal study of beliefs in relation to field experiences. *Science Education*, *95*(6), 1124–1146. https://doi.org/10.1002/sce.20450

Furtak, E. M., & Penuel, W. R. (2019). Coming to terms: Addressing the persistence of "hands-on" and other reform terminology in the era of science as practice. *Science Education*, *103*(1), 167–186. https://doi.org/10.1002/sce.21488

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945. https://doi.org/10.3102/00028312038004915

Gibbons, L. K., & Cobb, P. (2017). Focusing on teacher learning opportunities to identify potentially productive coaching activities. *Journal of Teacher Education*, *68*(4), 411–425. https://doi.org/10.1177/0022487117702579

Greenleaf, C. L., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., Schneider, S. A., Madden, S., & Jones, B. (2011). Integrating literacy and science in biology:

Teaching and learning impacts of reading apprenticeship professional development. *American Educational Research Journal*, *48*(3), 647–717. https://doi.org/10.3102/0002831210384839

Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, *26*(1), 5–17. https://doi.org/10.3102/0013189X026001005

Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher*, *15*(5), 5–12. https://doi.org/10.3102/0013189X015005005

Heck, D. J., Rosenberg, S. L., & Crawford, R. A. (2006). *LSC teacher questionnaire study: A longitudinal analysis of data collected between 1997 and 2006*. Horizon Research, Inc. http://www.horizon-research.com/lsc-teacher-questionnaire-study-a-longitudinal-analysis-of-data-collected-between-1997-and-2006

Hedeker, D. R., & Gibbons, R. D. (2006). *Longitudinal data analysis*. Wiley-Interscience.

Hill, H. C., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education*, *35*(5), 330–351. https://doi.org/10.2307/30034819

Jones, M. G., & Leagon, M. (2014). Science teacher attitudes and beliefs: Reforming practice. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of reseaerch on science education* (Vol. 2, pp. 830–847). Routledge.

Kang, E. J. S., McCarthy, M. J., & Donovan, C. (2019). Elementary teachers' enactment of the NGSS science and engineering practices. *Journal of Science Teacher Education*, *30*(7), 788–814. https://doi.org/10.1080/1046560X.2019.1630794

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional development: Attending to the coevolution of teachers' participation across contexts.

Journal of Teacher Education, *59*(5), 428–441. https://doi.org/10.1177/0022487108324330

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*(3), 226–243. https://doi.org/10.1177/0022487112473837

Lo, A., Krist, C., Reiser, B., & Novak, M. (2014, March 31). *Examining shifts in teachers' understanding of NGSS and their impact on planned instruction*. NARST International Conference, Pittsburgh, PA. https://www.researchgate.net/publication/315547774_Examining_Shifts_in_Teachers%27_Understanding_of_NGSS_and_their_Impact_on_Planned_Instruction

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2010). *Designing professional development for teachers of science and mathematics*. Corwin Press. https://doi.org/10.4135/9781452219103

Lowell, B. R., Cherbow, K., & McNeill, K. L. (2021). Redesign or relabel? How a commercial curriculum and its implementation oversimplify key features of the NGSS. *Science Education*, *105*(1), 5–32. https://doi.org/10.1002/sce.21604

Luft, J. A., Firestone, J. B., Wong, S. S., Ortega, I., Adams, K., & Bang, E. (2011). Beginning secondary science teacher induction: A two-year mixed methods study. *Journal of Research in Science Teaching*, *48*(10), 1199–1224. https://doi.org/10.1002/tea.20444

Lynch, K., Hill, H. C., Gonzalez, K. E., & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis*, *41*(3), 260–293. https://doi.org/10.3102/0162373719849044

Maeng, J. L., Whitworth, B. A., Bell, R. L., & Sterling, D. R. (2020). The effect of professional development on elementary science teachers' understanding, confidence, and classroom implementation of reform-based science instruction. *Science Education*, *104*(2), 326–353. https://doi.org/10.1002/sce.21562

Marco-Bujosa, L. M., McNeill, K. L., González-Howard, M., & Loper, S. (2017). An exploration of teacher learning from an educative reform-oriented science curriculum: Case studies of teacher curriculum use. *Journal of Research in Science Teaching*, *54*(2), 141–168. https://doi.org/10.1002/tea.21340

Masuda, A. M., Ebersole, M. M., & Barrett, D. (2013). A qualitative inquiry: Teachers' attitudes and willingness to engage in professional development experiences at different career stages. *Delta Kappa Gamma Bulletin*, *79*(2), 6–14.

McNeill, K. L., Affolter, R., & Reiser, B. J. (in press). Anchoring science professional learning in curriculum materials enactment: Illustrating theories in practice to support teacher learning. In A. Castro Superfine, S. R. Goldman, & M.-L. M. Ko (Eds.), *Teacher learning in changing contexts: Perspectives from the learning sciences*. Routledge.

McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2017). Moving beyond pseudoargumentation: Teachers' enactments of an educative science curriculum focused on argumentation. *Science Education*, *101*(3), 426–457. https://doi.org/10.1002/sce.21274

McNeill, K. L., & Knight, A. M. (2013). Teachers' pedagogical content knowledge of scientific argumentation: The impact of professional development on K-12 teachers. *Science Education*, *97*(6), 936–972. https://doi.org/10.1002/sce.21081

McNeill, K. L., & Reiser, B. J. (2018). Open source for opening minds: New OpenSciEd materials support science standards. *The Learning Professional*, *39*(6), 44–48.

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053–1075. https://doi.org/10.1002/tea.21459

Mills, K., Jass Ketelhut, D., & Gong, X. (2019). Change of teacher beliefs, but not practices, following integration of immersive virtual environment in the classroom. *Journal of Educational Computing Research*, *57*(7), 1786–1811. https://doi.org/10.1177/0735633119854034

National Academies of Sciences, Engineering, and Medicine. (2015). *Science teachers' learning: Enhancing opportunities, creating supportive contexts*. The National Academies Press. https://doi.org/10.17226/21836

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press. https://doi.org/10.17226/13165

National Research Council. (2015). *Guide to Implementing the Next Generation Science Standards*. National Academies Press.

National Science Teaching Association. (n.d.). *About the Next Generation Science Standards*. Retrieved September 16, 2020, from http://ngss.nsta.org/About.aspx

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, *19*(4), 317–328. https://doi.org/10.1080/0022027870190403

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Sage Publications.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. National Academies Press. https://doi.org/10.17226/18290

Odden, T. O. B., & Russ, R. S. (2019). Defining sensemaking: Bringing clarity to a fragmented theoretical construct. *Science Education*, *103*(1), 187–205. https://doi.org/10.1002/sce.21452

OpenSciEd. (2019). *How can a sound make something move?* https://www.openscied.org/instructional-materials/8-2-sound-waves/

OpenSciEd. (2020). *OpenSciEd Teacher Handbook* (Version 3.0). https://www.openscied.org/teacher-handbook/

Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education*, *25*(2), 177–196. https://doi.org/10.1007/s10972-014-9384-1

Osborne, J. W. (2010). Data cleaning basics: Best practices in dealing with extreme scores. *Newborn and Infant Nursing Reviews*, *10*(1), 37–43. https://doi.org/10.1053/j.nainr.2009.12.009

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307–332. https://doi.org/10.3102/00346543062003307

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation.

American Educational Research Journal, *44*(4), 921–958. https://doi.org/10.3102/0002831207308221

Pilitsis, V., & Duncan, R. G. (2012). Changes in belief orientations of preservice teachers and their relation to inquiry activities. *Journal of Science Teacher Education*, *23*(8), 909–936. https://doi.org/10.1007/s10972-012-9303-2

Pruitt, S. L. (2014). The Next Generation Science Standards: The features and challenges. *Journal of Science Teacher Education*, *25*(2), 145–156. https://doi.org/10.1007/s10972-014-9385-0

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed). Sage Publications.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., Jr., & du Toit, M. (2019). *HLM 8: Hierarchical linear and nonlinear modeling* (8.1.4.14) [Computer software]. Scientific Software International.

Reiser, B. J., Michaels, S., Moon, J., Bell, T., Dyer, E., Edwards, K. D., McGill, T. A. W., Novak, M., & Park, A. (2017). Scaling up three-dimensional science learning through teacher-led study groups across a state. *Journal of Teacher Education*, *68*(3), 280–298. https://doi.org/10.1177/0022487117699598

Reiser, B. J., Novak, M., McGill, T. A. W., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal of Science Teacher Education*, *32*(7), 805–829. https://doi.org/10.1080/1046560X.2021.1884784

Ricketts, A. (2014). Preservice elementary teachers' ideas about scientific practices. *Science & Education*, *23*(10), 2119–2135. https://doi.org/10.1007/s11191-014-9709-7

Robertson, A. D., Scherr, R., & Hammer, D. (Eds.). (2015). *Responsive Teaching in Science and Mathematics* (1st ed.). Routledge. https://doi.org/10.4324/9781315689302

Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organization and change*. Jossey-Bass.

Russ, R. S., & Berland, L. K. (2019). Invented science: A framework for discussing a persistent problem of practice. *Journal of the Learning Sciences*, *28*(3), 279–301. https://doi.org/10.1080/10508406.2018.1517354

Sandholtz, J. H., & Ringstaff, C. (2014). Inspiring instructional change in elementary school science: The relationship between enhanced self-efficacy and teacher practices. *Journal of Science Teacher Education*, *25*(6), 729–751. https://doi.org/10.1007/s10972-014-9393-0

Schwarz, C. V., Passmore, C., & Reiser, B. J. (2017). Moving beyond "knowing about" science to making sense of the natural world. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 3–21). NSTA Press.

Short, J., & Hirsh, S. (2020). *The elements: Transforming teaching through curriculum-based professional learning*. Carnegie Corporation of New York. https://www.carnegie.org/publications/elements-transforming-teaching-through-curriculum-based-professional-learning/

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–23.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Smetana, L. K., Wenner, J., Settlage, J., & McCoach, D. B. (2016). Clarifying and capturing "trust" in relation to science education: Dimensions of trustworthiness within schools and associations with equitable student achievement. *Science Education*, *100*(1), 78–95. https://doi.org/10.1002/sce.21195

Smith, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *Journal of the Learning Sciences*, *3*(2), 115–163.

Taie, S., & Goldring, R. (2020). *Characteristics of public and private elementary and secondary school teachers in the United States: Results from the 2017–18 National Teacher and Principal Survey first look* (NCES 2020-142REV). National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020142rev

Taylor, J. A., Getty, S. R., Kowalski, S. M., Wilson, C. D., Carlson, J., & Van Scotter, P. (2015). An efficacy trial of research-based curriculum materials with curriculum-based professional development. *American Educational Research Journal*, *52*(5), 984–1017. https://doi.org/10.3102/0002831215585962

Torff, B., & Sessions, D. (2008). Factors associated with teachers' attitudes about professional development. *Teacher Education Quarterly*, *35*(2), 123–133.

Voet, M., & De Wever, B. (2019). Teachers' adoption of inquiry-based learning activities: The importance of beliefs about education, the self, and the context. *Journal of Teacher Education*, *70*(5), 423–440. https://doi.org/10.1177/0022487117751399

von Suchodoletz, A., Jamil, F. M., Larsen, R. A. A. A., & Hamre, B. K. (2018). Personal and contextual factors associated with growth in preschool teachers' self-efficacy beliefs during a longitudinal professional development study. *Teaching and Teacher Education*, *75*, 278–289. https://doi.org/10.1016/j.tate.2018.07.009

Wilson, S. M. (2013). Professional development for science teachers. *Science*, *340*(6130), 310–313. https://doi.org/10.1126/science.1230725

Yin, R. K. (2018). *Case study research and applications: Design and methods* (Sixth edition). SAGE.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (REL 2007-No. 033; Issues & Answers). US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=70

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

Zivic, A., Smith, J. F., Reiser, B. J., Edwards, K. D., Novak, M., & McGill, T. A. W. (2018). *Negotiating epistemic agency and target learning goals: Supporting coherence from the students' perspective*. International Conference of the Learning Sciences, London, U.K.

Appendix – Further Methods

Here, I provide more details on my methods with respect to the analysis of the outcome variables, the creation of the "Value of PD Activities" variable, and the complete model building process.

Outcome Variables

As described in the main text, I began my analysis of the outcome variables with an exploratory factor analysis at each time point. For each analysis, I used principal axis factoring with varimax rotation and Kaiser normalization. Tables II.8 through II.12 show the results of those analyses at each time point. Note that for the PD2 survey (Table II.10) and the PD4 survey (Table II.12), the belief items were not asked, so only one factor was extracted and therefore no rotation was used.

Table II.8: Factor Analysis Results of Outcome Items at Pre-Survey

	Item	Factor loading	
		1	2
Factor 1: Implementation Confidence (Eigenvalue = 3.435)			
11	Get students to ask questions at the beginning of a unit that guide the lessons that follow?	**.763**	-.033
12	Work with students to motivate the next step in investigating a phenomenon, rather than just telling them what they will do next?	**.794**	-.061
13	Help students use science practices to figure out pieces of core science ideas?	**.803**	.121
14	Push students to go deeper to revise their explanatory models of phenomena?	**.792**	.075
15	Help students put pieces together related to disciplinary core ideas and crosscutting concepts?	**.763**	.155
Factor 2: Traditional Beliefs (Eigenvalue = 3.295)			
1	Teachers should have students do interesting hands-on activities, even if the activities do not relate closely to the concept being studied.	.020	**.351**
2	Hands-on/laboratory activities should be used primarily to reinforce a science idea that the students have already learned.	.001	**.620**
3	At the beginning of instruction on a science idea, students should be provided with definitions for new scientific vocabulary that will be used.	-.014	**.694**

4	Teachers should provide students with the outcome of an activity in advance so students know they are on the right track as they do the activity.	-.029	**.597**
5	Students should do hands-on or laboratory activities, even if they do not have opportunities to discuss them as a class.	-.040	**.298**
6	Teachers should explain an idea to students before having them consider evidence that relates to the idea.	-.091	**.736**
7	Students should know what the results of an experiment are supposed to be before they carry it out.	-.067	**.676**
8	When students do a hands-on activity and the data don't come out right, teachers should tell students what they should have found.	-.114	**.470**
Did not load onto either factor			
9	Students' ideas about a science concept should be deliberately brought to the surface prior to a lesson or unit so that students are aware of their own thinking.	.035	.183
10	It is better for science instruction to focus on ideas in depth, even if that means covering fewer topics	.126	-.085

Table II.9: Factor Analysis Results of Outcome Items at PD 1 Survey

	Item	Factor loading	
		1	2
Factor 1: Traditional Beliefs (Eigenvalue = 4.213)			
1	Teachers should have students do interesting hands-on activities, even if the activities do not relate closely to the concept being studied.	**.347**	.315
2	Hands-on/laboratory activities should be used primarily to reinforce a science idea that the students have already learned.	**.479**	.252
3	At the beginning of instruction on a science idea, students should be provided with definitions for new scientific vocabulary that will be used.	**.719**	.189
4	Teachers should provide students with the outcome of an activity in advance so students know they are on the right track as they do the activity.	**.712**	.406
5	Students should do hands-on or laboratory activities, even if they do not have opportunities to discuss them as a class.	**.426**	.293
6	Teachers should explain an idea to students before having them consider evidence that relates to the idea.	**.753**	.242
7	Students should know what the results of an experiment are supposed to be before they carry it out.	**.652**	.170
8	When students do a hands-on activity and the data don't come out right, teachers should tell students what they should have found.	**.551**	.243
Factor 2: Implementation Confidence (Eigenvalue = 2.848)			
11	Get students to ask questions at the beginning of a unit that guide the lessons that follow?	-.384	**.481**

12	Work with students to motivate the next step in investigating a phenomenon, rather than just telling them what they will do next?	-.434	**.548**
13	Help students use science practices to figure out pieces of core science ideas?	-.416	**.677**
14	Push students to go deeper to revise their explanatory models of phenomena?	-.374	**.574**
15	Help students put pieces together related to disciplinary core ideas and crosscutting concepts?	-.432	**.654**
Did not load onto either factor			
9	Students' ideas about a science concept should be deliberately brought to the surface prior to a lesson or unit so that students are aware of their own thinking.	.115	.213
10	It is better for science instruction to focus on ideas in depth, even if that means covering fewer topics	-.198	.067

Table II.10: Factor Analysis Results of Outcome Items at PD 2 Survey

Item		Factor loading
Factor 1: Implementation Confidence (Eigenvalue = 3.435)		
11	Get students to ask questions at the beginning of a unit that guide the lessons that follow?	**.726**
12	Work with students to motivate the next step in investigating a phenomenon, rather than just telling them what they will do next?	**.793**
13	Help students use science practices to figure out pieces of core science ideas?	**.758**
14	Push students to go deeper to revise their explanatory models of phenomena?	**.826**
15	Help students put pieces together related to disciplinary core ideas and crosscutting concepts?	**.722**

Note. The belief items were not included on this survey. Because only one factor was extracted, this was not rotated.

Table II.11: Factor Analysis Results of Outcome Items at PD 3 Survey

Item		Factor loading	
		1	2
Factor 1: Traditional Beliefs (Eigenvalue = 3.997)			
1	Teachers should have students do interesting hands-on activities, even if the activities do not relate closely to the concept being studied.	**.611**	.075
2	Hands-on/laboratory activities should be used primarily to reinforce a science idea that the students have already learned.	**.621**	-.148
3	At the beginning of instruction on a science idea, students should be provided with definitions for new scientific vocabulary that will be used.	**.748**	-.201

4	Teachers should provide students with the outcome of an activity in advance so students know they are on the right track as they do the activity.	**.693**	-.019
5	Students should do hands-on or laboratory activities, even if they do not have opportunities to discuss them as a class.	**.558**	.052
6	Teachers should explain an idea to students before having them consider evidence that relates to the idea.	**.714**	-.004
7	Students should know what the results of an experiment are supposed to be before they carry it out.	**.611**	-.028
8	When students do a hands-on activity and the data don't come out right, teachers should tell students what they should have found.	**.521**	-.017
Factor 2: Implementation Confidence (Eigenvalue = 2.933)			
11	Get students to ask questions at the beginning of a unit that guide the lessons that follow?	-.018	**.838**
12	Work with students to motivate the next step in investigating a phenomenon, rather than just telling them what they will do next?	-.028	**.767**
13	Help students use science practices to figure out pieces of core science ideas?	.049	**.485**
14	Push students to go deeper to revise their explanatory models of phenomena?	.004	**.480**
15	Help students put pieces together related to disciplinary core ideas and crosscutting concepts?	-.079	**.392**
Did not load onto either factor			
9	Students' ideas about a science concept should be deliberately brought to the surface prior to a lesson or unit so that students are aware of their own thinking.	.135	.051
10	It is better for science instruction to focus on ideas in depth, even if that means covering fewer topics	-.017	.145

Table II.12: Factor Analysis Results of Outcome Items at PD 4 Survey

	Item	Factor loading
Factor 1: Implementation Confidence (Eigenvalue = 3.435)		
11	Get students to ask questions at the beginning of a unit that guide the lessons that follow?	**.645**
12	Work with students to motivate the next step in investigating a phenomenon, rather than just telling them what they will do next?	**.792**
13	Help students use science practices to figure out pieces of core science ideas?	**.712**
14	Push students to go deeper to revise their explanatory models of phenomena?	**.798**
15	Help students put pieces together related to disciplinary core ideas and crosscutting concepts?	**.790**

Note. The belief items were not included on this survey. Because only one factor was extracted, this was not rotated.

The reliability for each of these factors is shown in Table II.3. Based on this factor analysis, as discussed in the main text, I concluded that these were reliable and valid measures of teachers' traditional beliefs and OpenSciEd instructional confidence and I created factor scores for each measure at each time point. I then checked the normality of these items by constructing a histogram of the total scores for each factor at each time point. Figure II.4 shows the histograms for the belief items at all three measured time points. These items appear to be normally distributed at first and then become positively skewed over time, but with a robust distribution throughout the theoretical range of 8-48.

Figure II.4: Histograms of Traditional Beliefs Total Scores

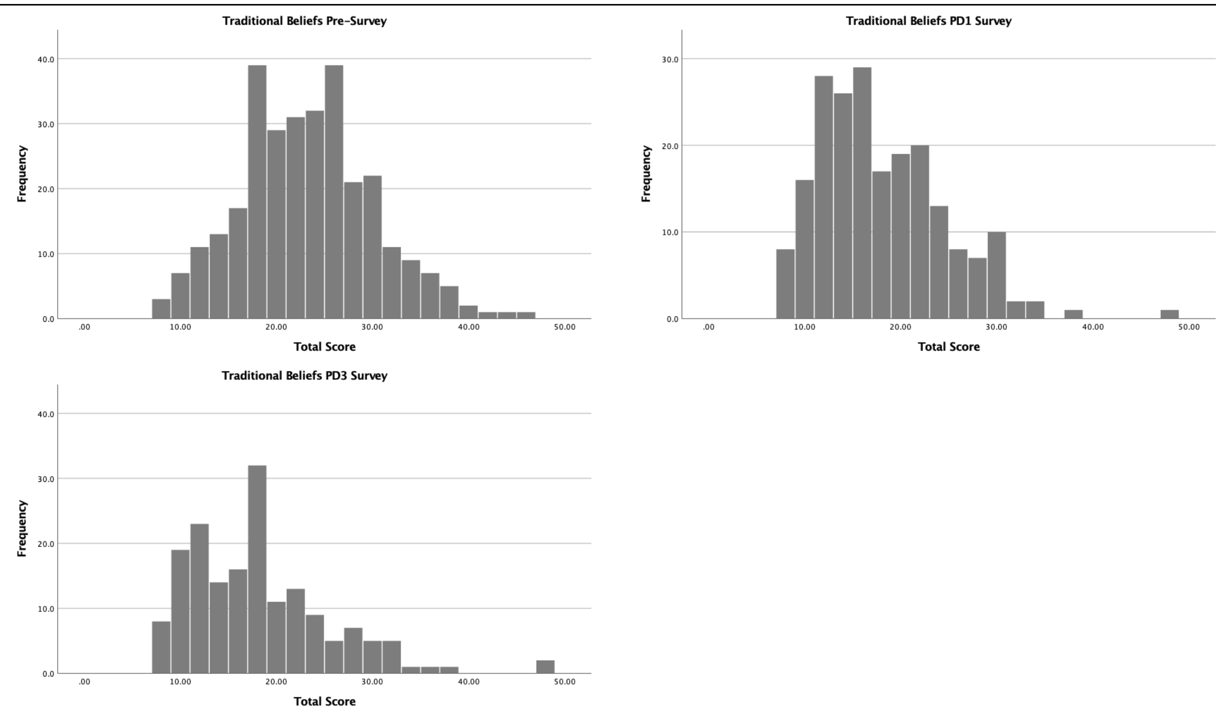
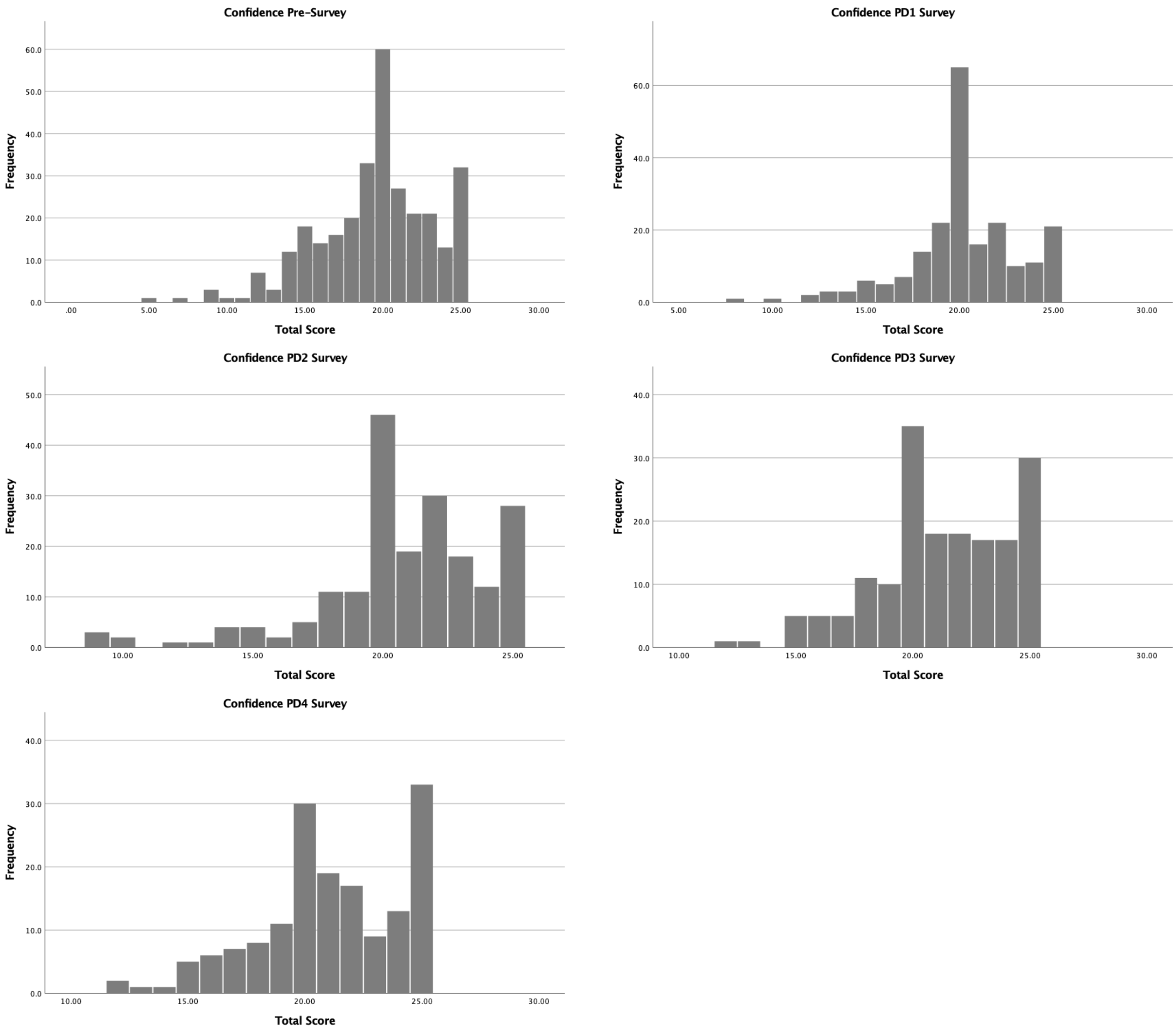


Figure II.5 below shows the five histograms for the implementation confidence factor, demonstrating that they were all somewhat normal but negatively skewed and approaching the upper limit of 25. Based on these initial explorations of the outcome variables, I concluded that they would be appropriate for building growth models.

Figure II.5: Histograms of Implementation Confidence Total Scores



Predictor Variables – Value Placed on PD Activities

As discussed above, in order to create the variable "value placed on PD activities," I looked at the three items that were consistently asked after all 4 PD sessions. An exploratory factor analysis showed those items did all load onto one factor, as shown below in Tables II.13 through II.16.

Table II.13: Factor Analysis Results of Value of PD Activities Items at PD 1 Survey

Item	Factor loading
Factor 1: Value Placed on PD Activities (Eigenvalue = 1.645)	
Engaging in the anchoring phenomenon for your specific unit	**.740**
Conducting investigations from the students' perspective for your specific unit	**.783**
Building the storyline for your specific unit	**.695**

Table II.14: Factor Analysis Results of Value of PD Activities Items at PD 2 Survey

Item	Factor loading
Factor 1: Value Placed on PD Activities (Eigenvalue = 1.903)	
Engaging in the anchoring phenomenon for your specific unit	**.795**
Conducting investigations from the students' perspective for your specific unit	**.812**
Building the storyline for your specific unit	**.782**

Table II.15: Factor Analysis Results of Value of PD Activities Items at PD 3 Survey

Item	Factor loading
Factor 1: Value Placed on PD Activities (Eigenvalue = 1.727)	
Engaging in the anchoring phenomenon for your specific unit	**.804**
Conducting investigations from the students' perspective for your specific unit	**.753**
Building the storyline for your specific unit	**.717**

Table II.16: Factor Analysis Results of Value of PD Activities Items at PD 4 Survey

Item	Factor loading
Factor 1: Value Placed on PD Activities (Eigenvalue = 1.692)	
Engaging in the anchoring phenomenon for your specific unit	**.788**
Conducting investigations from the students' perspective for your specific unit	**.825**
Building the storyline for your specific unit	**.625**

After completing the factor analysis, I did a reliability analysis of each factor at each time point. As shown in Table II.17, these reliabilities are relatively low, but that could be due to the small number of items in each factor and the relative novelty of this construct (Field, 2013). Therefore, I continued to use these factors as a measure of how valuable the teachers found the PD activities. As discussed in the methods section above, the final step for constructing this variable was to determine if it varied over time within individuals, and I found that the variable did not significantly vary over time within people. Therefore, I constructed an average value placed on PD composite score and used it as a level-2, between-people predictor.

Table II.17: Reliability Analysis of Value of PD Activities Items (Cronbach's Alpha)

Factor	Post-PD1	Post-PD2	Post-PD3	Post-PD4
Value Placed on PD Activities	.577	.698	.624	.606

Model Building Process

As discussed in the methods section above, I built each model using a forward-selection process in which predictors that were significant at a level of $\alpha < .05$ were retained. Tables II.6 and II.7 above show the initial and final models. Tables II.18 and II.19 show all of the models I built for traditional beliefs and implementation confidence, respectively.

Table II.18: Traditional Beliefs Models

Fixed Effects β (SE β)	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Presurvey status											
Intercept	22.79^{***} (0.40)	22.79^{***} (0.40)	22.48^{***} (0.49)	22.86^{***} (0.40)	22.88^{***} (0.39)	24.82^{***} (0.80)	24.22^{***} (0.84)	24.93^{***} (1.03)	24.79^{***} (0.87)	24.77^{***} (0.85)	23.24^{***} (0.41)
Gender			0.53 (0.94)								
Nonwhite			1.64 (1.13)								
Years teaching				-0.15^{**} (0.05)	-0.13^{**} (0.04)	-0.14^{**} (0.04)	-0.11^{*} (0.04)	-0.10^{*} (0.04)	-0.09^{*} (0.04)	-0.10^{*} (0.04)	-0.09^{*} (0.04)
% MS Sci					-0.03^{*} (0.01)	-0.03^{*} (0.01)	$-0.02^{\dagger}$ (0.01)	$-0.02^{\dagger}$ (0.01)	$-0.02^{\dagger}$ (0.01)	$-0.02^{\dagger}$ (0.01)	
Mins teach/week						-0.002 (0.006)					
Prev PD days							-0.11^{**} (0.04)	$-0.07^{\dagger}$ (0.04)	-0.09^{**} (0.03)	-0.09^{*} (0.03)	-0.09^{**} (0.03)
Years of NGSS								-0.30 (0.25)			
Value of PD									-1.82^{***} (0.50)	-1.85^{***} (0.45)	-1.72^{***} (0.45)
Time											
Intercept	-6.98^{***} (0.58)	-7.02^{***} (0.58)	-7.61^{***} (0.73)	-7.03^{***} (0.57)	-7.04^{***} (0.65)	-7.20^{***} (0.56)	-7.05^{***} (0.60)	-7.45^{***} (1.30)	-6.71^{***} (0.61)	-6.97^{***} (0.60)	-6.94^{***} (0.59)
Gender			2.05 (1.37)								
Nonwhite			0.34 (1.44)								
Years teaching				0.05 (0.07)							
% MS Sci					0.01 (0.03)						
Mins teach/week						0.01 (0.01)					

Prev PD days							0.10† (0.05)				
Years of NGSS								0.17 (0.36)			
Value of PD									-0.88 (0.82)		
Time^2											
Intercept	1.81*** (0.18)	1.82*** (0.18)	1.99*** (0.23)	1.82*** (0.18)	1.84*** (0.21)	1.87*** (0.18)	1.84*** (0.19)	2.02*** (0.42)	1.69*** (0.19)	1.83*** (0.19)	1.81*** (0.19)
Gender			-0.60 (0.44)								
Nonwhite			-0.11 (0.43)								
Years teaching				-0.01 (0.02)							
% MS Sci					-0.004 (0.008)						
Mins teach/week						-0.003 (0.004)					
Prev PD days							-0.03† (0.02)				
Years of NGSS								-0.07 (0.11)			
Value of PD									0.41 (0.26)		
Random Effects											
Level 1	15.79	17.75	17.77	17.76	17.79	17.56	17.58	17.81	17.99	17.82	17.69
Level 2 presurvey	33.07***	29.31***	29.02***	28.19***	27.71***	27.03***	25.49***	25.26***	22.55***	22.73***	23.56***
Level 2 time	9.95	0.88**	1.09**	0.92**	0.86**	0.84**	0.92**	0.89**	0.74*	0.82*	0.92**
Level 2 time^2	0.52										
Model Fit Statistics											
Deviance	4387.38	4389.89	4240.63	4386.27	4301.67	4271.62	4186.73	4183.30	4093.44	4095.03	4165.60
ICC	0.68	0.63	0.62	0.61	0.61	0.61	0.59	0.59	0.56	0.56	0.57

†p < .10; *p < .05; **p < .01; ***p < .001

Table II.19: Implementation Confidence Models

Fixed Effects β (SE β)	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Presurvey status										
Intercept	19.42*** (0.21)	19.43*** (0.21)	19.32*** (0.24)	19.46*** (0.21)	19.47*** (0.21)	19.43*** (0.21)	19.39*** (0.20)	19.38*** (0.42)	19.32*** (0.24)	19.47*** (0.20)
Gender			0.27 (0.51)							
Nonwhite			0.39 (0.60)							
Years teaching				0.02 (0.03)						
% MS Sci					-0.001 (0.01)					
Mins teach/week						0.0003 (0.003)				
Prev PD days							0.06** (0.02)	0.04*** (0.01)	0.05*** (0.01)	0.05*** (0.01)
Years of NGSS								0.01 (0.12)		
Value of PD									0.49 (0.30)	
Time										
Intercept	0.83*** (0.19)	0.84*** (0.19)	0.73** (0.23)	0.76*** (0.20)	0.91*** (0.19)	0.78*** (0.20)	0.84*** (0.20)	0.76† (0.41)	0.60* (0.23)	0.71*** (0.20)
Gender			0.53 (0.42)							
Nonwhite			-0.04 (0.60)							
Years teaching				0.03 (0.02)						
% MS Sci					-0.01 (0.01)					
Mins teach/week						0.004 (0.003)				

Prev PD days							-0.02 (0.01)			
Years of NGSS								0.03 (0.11)		
Value of PD									0.69* (0.31)	0.33*** (0.08)
Time^2										
Intercept	-0.10* (0.05)	-0.10* (0.05)	-0.06 (0.06)	-0.08 (0.05)	-0.12* (0.05)	-0.08 (0.05)	-0.10* (0.05)	-0.10 (0.10)	-0.05 (0.06)	-0.10* (0.05)
Gender			-0.19† (0.10)							
Nonwhite			0.02 (0.15)							
Years teaching				-0.01 (0.01)						
% MS Sci					0.003 (0.002)					
Mins teach/week						-0.001 (0.001)				
Prev PD days							0.004 (0.003)			
Years of NGSS								0.002 (0.03)		
Value of PD									-0.14† (0.08)	
Random Effects										
Level 1	4.89	5.47	5.50	5.45	5.39	5.48	5.33	5.34	5.24	5.26
Level 2 presurvey	8.58***	6.68***	6.53***	6.62***	6.43***	6.57***	6.29***	6.31***	5.45***	5.70***
Level 2 time	2.71**	0.10†	0.10*	0.10*	0.10†	0.08†	0.10†	0.10†	0.10*	0.10*
Level 2 time^2	0.11									
Model Fit Statistics										
Deviance	5194.79	5208.46	4972.19	5141.24	5050.52	5126.16	4997.37	4990.72	4878.91	4886.99
ICC	0.64	0.55	0.54	0.55	0.54	0.55	0.54	0.54	0.51	0.52

†p < .10; *p < .05; **p < .01; ***p < .001

Section III: Paper #2 - The Student Hat: A New Tool in Practice-Based Professional Development

Recent reforms in science instruction have focused on moving away from teaching science as a set of facts for students to memorize and towards emphasizing science as a process of figuring out the natural world (Schwarz et al., 2017). This movement has gained the strength of policy in most of the United States with the release of the *Next Generation Science Standards* (NGSS) and the *Framework for K-12 Science Education* on which the standards are based (National Research Council, 2012; NGSS Lead States, 2013). This vision of teaching science as a process of figuring out is also focused on providing students with more epistemic agency in class, meaning that their contributions are valued and have meaningful impacts on the sensemaking work done by the class (Miller et al., 2018; Stroupe, 2014). This approach is in contrast to traditional forms of instruction that emphasize students following procedures for activities set by the teacher or curriculum (Stroupe, 2014). The connecting thread behind this vision of science instruction is that science classrooms better represent the work actually done by scientists while also pushing the field of science to be more open to diverse ideas and sensemaking practices (Bang et al., 2017).

This vision is not easy, and teachers find it challenging to shift towards practice that truly supports students' scientific sensemaking (National Academies of Sciences, Engineering, and Medicine, 2015; Pruitt, 2014). This shift is particularly challenging for in-service teachers given that many were not themselves taught science in this way and are coming off an era of accountability focused on supporting students to memorize final-form science knowledge (McNeill & Berland, 2017; National Academies of Sciences, Engineering, and Medicine, 2015). A long-standing strategy that has been used to support teachers in this change is professional development (PD), which has been defined as activities designed to support in-service teachers to learn about and improve their work as teachers (Desimone, 2009). This broad definition can encompass many activities, but the most common generally include a targeted workshop in which teachers leave the classroom to engage in some designed learning activities with facilitators (Wilson, 2013). Research on PD programs has shown that PD can help to increase teachers' knowledge of, beliefs in, and ability to implement new instructional reforms (Luft & Hewson, 2014; Wilson, 2013).

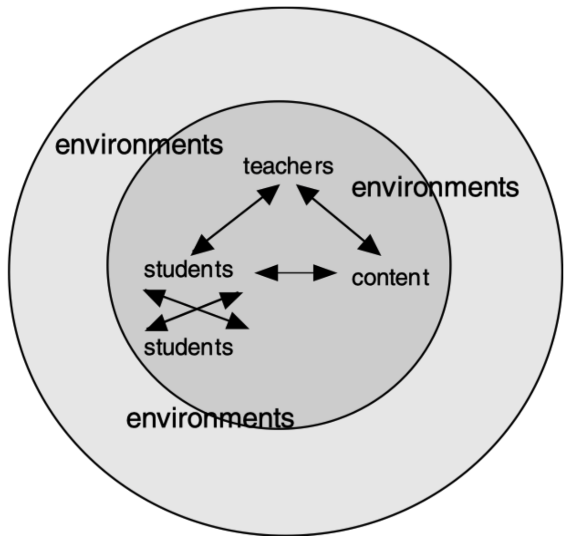
A popular approach to PD is to ground it in teachers' practice, meaning that the PD is situated in the ideas, questions, or problems that teachers face as they engage in the work of teaching (Ball & Cohen, 1999; Borko et al., 2010; National Academies of Sciences, Engineering, and Medicine, 2015). This approach conceptualizes the work of teaching as not a set of knowledge to be gained, but rather as *knowledge in use*, and therefore learning to teach should include applying understandings of strong instruction to teachers' contexts and experiences (Grossman et al., 2018; Lampert, 2010). A hallmark of these practice-based approaches to PD is to include representations or approximations of teacher practice for participants to jointly analyze (Ball & Cohen, 1999). Representations can include classroom video or analysis of student work (e.g. Heller et al., 2012; Roth et al., 2011; Sherin & van Es, 2009); whereas, approximations can include things like rehearsing a particular discussion or activity or engaging in learning activities as they will be presented to students (e.g. Borko et al., 2017; McNeill et al., in press). These approaches align well with practice-based preservice teacher education movements, which have received considerable attention in recent literature (Grossman, 2018; Stroupe et al., 2020).

The body of literature on practice-based teacher education (preservice and in-service) is particularly robust in math (Grossman, 2018). In science, much of the theoretical grounds and methodological approaches towards practice-based teacher learning is borrowed from this work in math (Stroupe et al., 2020). This is a helpful start, but science has particular disciplinary demands that are distinct from math (National Research Council, 2012). In addition, our conceptualization and interpretation of the vision put forth by the NGSS has changed since its publication (Sadler & Brown, 2018). And even as we come to a more robust and complex agreement on that vision, there still is a dearth of instructional materials that can consistently help make that vision a reality (Campbell & Lee, 2021; Sadler & Brown, 2018). As we continue to understand what strong NGSS instruction can and should look like, therefore, we also need to refine our vision for aligned teacher learning (McNeill et al., in press; National Academies of Sciences, Engineering, and Medicine, 2019). This includes considering what practice-based PD can and should look like with science teachers. Based on the evolving needs of science teacher education for the NGSS, in this paper I analyze some of the more common practice-based PD activities focusing on what they are helping teachers to learn about the NGSS classroom. In so doing, I highlight the contributions of one less common activity: thinking like a student (also called “student hat") and argue it should become a more frequently used tool from the practice-based PD toolbox by illustrating its use, along with other practice-based PD activities, in one workshop.

The Instructional Triangle as a Lens for Analyzing PD Activities

Practice-based professional development activities could theoretically address a number of different learning goals, and designers need to consider which ones they want to target when deciding how and when to engage participants in particular activities. Ultimately, the goal of all teacher PD should be to support student learning. In order to determine what goals we have for teacher PD, then, it is helpful to consider what takes place in classrooms to support student learning. For that, I turn to the instructional triangle, which was first proposed by Cohen, Raudenbush, and Ball (2003) to describe the work of teaching. The triangle, shown in Figure III.1, conceives of learning as taking place in the interaction between three key players: the disciplinary content, the teacher, and the students, all of which exist in and interact with the greater learning environment. The key idea here is that teaching is not a one-way street in which teachers bestow knowledge on students, but rather a complex set of interactions (Ball, 2018).

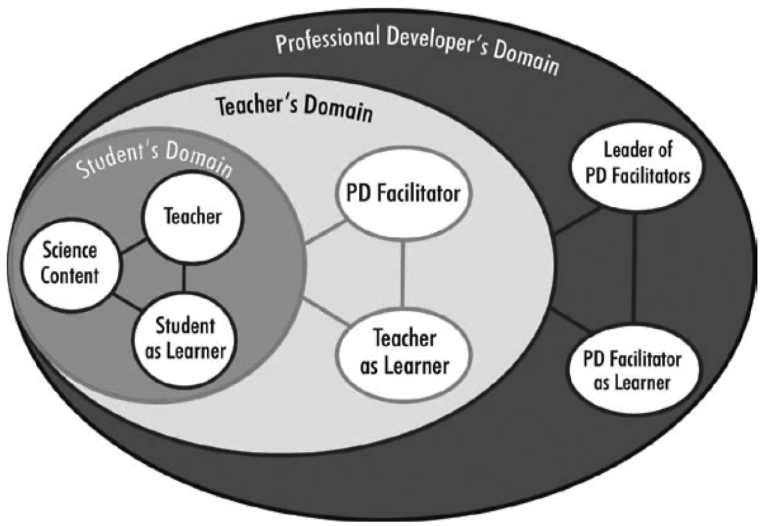
Figure III.1: The Instructional Triangle



Note. From Cohen et al., 2003, Figure 1

Multiple authors have expanded on and pushed back on the triangle, including Ball herself. She used part of her 2018 AERA presidential address to recenter students at the top of the triangle, add arrows to make the interaction between the triangle and the environment clearer, and make the barriers between the classroom and the environment porous to show the interconnected nature of the two (Ball, 2018). Another common revision is to use it to explain teacher learning by expanding it to multiple triangles nested within each other (Lauffer & Lauffer, 2009; Luft & Hewson, 2014; Nipper & Sztajn, 2008; Tekkumru-Kisa & Stein, 2017a). This kind of expansion is based on the idea that teaching and teacher learning happens within a larger system and that in order to see change in instructional practices at the classroom level, there must be coherence in messaging and practice across levels of the system (C. D. Allen & Penuel, 2015; Lauffer, 2010; Marrongelle et al., 2013). This highlights that the design of professional development should align with the instructional shifts advocated for in science classrooms. Figure III.2 shows a representative example of this kind of expansion, from Lauffer and Lauffer (2009).

Figure III.2: The Instructional Triangle as Nested Domains



Note. From Lauffer & Lauffer, 2009, Figure 4.4

In Figure III.2, PD takes place in the middle "teacher's domain" where the teacher is engaging in PD as a learner and interacts with the PD facilitator and the "student's domain." The idea is that what teachers should be learning about during PD is the K-12 learning environment. The K-12 learning environment is represented by the innermost triangle, called the "student's domain." This triangle is most analogous to the original instructional triangle (see Figure III.1) and demonstrates that in the classroom the three major players are the teacher, the students, and the science content. Therefore, over the course of a practice-based PD workshop, teachers should engage in a range of activities that help them attend to *all three* components of the K-12 triangle:

the science content, the teacher and their role in the classroom, and the students and their thinking and experiences in an NGSS class.

Current work in science education reform, as exemplified by the NGSS, provides some vision as to what the role of the science content, the student and the teacher should be in the science classroom. Unlike past models in which the content is viewed as a set of facts to be memorized, content in current science classrooms should include the core ideas of science, the practices that scientists engage in to construct those core ideas, and the crosscutting concepts that inform or are informed by that work (National Research Council, 2012). In this way, science content is not just *what* we know but also *a way of knowing* about the natural world (Driver et al., 2000). In order to do that, the role of the students should shift to be active collaborators in knowledge building, bringing their own backgrounds, experiences, and ideas to inform decisions the class takes in terms of conducting investigations, analyzing evidence, and making claims (Bang et al., 2012; Miller et al., 2018; Stroupe, 2014). Others have called this a focus on students' "epistemic agency," meaning the degree to which they have meaningful control over the idea work done in the classroom (Miller et al., 2018). Finally, the teacher's role in this work is to establish environments in which students can engage meaningfully with natural phenomena, do investigative tasks together, and come to joint agreement about what they understand and how; this is in stark contrast to past views of the teacher as the ultimate source of knowledge to share with students (Lowell et al., 2021; Robertson et al., 2015; C. V. Schwarz et al., 2017). Given the complexity of these views of science, student and teacher, it is helpful for us to look again at the way current PD activities support teachers in developing their understanding of each and how they are related.

Defining and Analyzing Common Practice-Based Professional Development Activities

The practice-based PD and teacher education literature is based in the idea that teachers should engage with representations and approximations of practice during teacher learning in order to prepare them to effectively do that practice in their classroom (Ball & Cohen, 1999; Grossman, 2018). In order to make sense of how PD might help teachers to prepare for NGSS-aligned science instruction, I begin by defining a set of common activities used in practice-based professional development. For each of these activities, I give a brief outline of the rules of engagement for that activity based on how it has been used in the literature and then outline a representative example of how this activity has been used. Finally, I analyze that example to illustrate what this activity might foreground and background for teachers. I discuss five different activities: watching classroom video for student thinking or teacher moves, engaging in rehearsals of target teacher moves, analyzing student work, doing science activities as an "adult learner," and doing science activities with a "student hat." Table III.1 summarizes each of these activities.

Table III.1: Some Common Practice-Based Professional Development Activities

Activity	Structure	What it foregrounds	What it backgrounds
Watching classroom video	Groups of teachers watch video of students engaging in NGSS-designed lessons while attending to some particular component of the learning environment, such as the students' responses or the teacher's decisions. (Roth et al., 2011; Sherin & van Es, 2009; Taylor et al., 2017; Tekkumru-Kisa & Stein, 2017b; Tekkumru-Kisa et al., 2018)	Particular student reactions to instruction and/or teacher moves that align (or do not align) with target instructional approach.	The resources, ideas, and experiences students bring to instruction and how they are (or are not) valued. Complex ideas about science content.

Activity	Structure	What it foregrounds	What it backgrounds
Rehearsing teaching	One teacher practices an instructional activity while others act as students. The rehearsing teacher or facilitator can pause to discuss or change the teacher's decisions mid-rehearsal. (Kazemi et al., 2016; Kelley-Petersen et al., 2018; Kloser & Windschitl, 2020; Lampert et al., 2013)	The relationship between teaching moves and student responses or reactions. How teachers can support student meaningful engagement with class activities.	The ideas, interests, and experiences students bring to the content. Deeper conceptual understanding of the science ideas from an expert lens.
Analyzing student work	Teachers look at student work from an NGSS classroom to determine what ideas students brought to the work, where they were confused, and possible instructional next steps. (Heller et al., 2012; Loucks-Horsley et al., 2010; McNeill & Knight, 2013)	The ideas students share during and at the end of instruction and how they align with canonical forms of those ideas.	The moves teachers did (or did not do) to support students to get to these ideas. How students responded to instruction in the moment.
Doing science as an adult content learner	Teachers engage in science activities in order to improve their understanding of the science content. These activities may be the same ones their students will do and/or address more complex ideas. (Fulp et al., 2009; Heller et al., 2012; Roth et al., 2019)	The teacher's relationship to canonical representations of core ideas, science practices, and/or crosscutting concepts.	The ideas, interests, and experiences students bring to the science. Teacher moves that might support student engagement with science.
Doing science in student hat	Teachers engage in science activities they will do with their students while considering ideas and experiences their students might use to make sense of the lessons. Teachers share those ideas using language the students might use during the activity. (McNeill et al., in press; Reiser, Michaels, et al., 2017)	How students might respond to, think about, or be (un)interested in the content. How students' interests and ideas are (not) used as resources in instruction.	Content ideas as understood by experts, including accepted terminology. Specific teacher moves that support student engagement with science.

Watching Classroom Video

Almost since the invention of video technologies, they have been recognized as helpful for situating teacher learning in the work of teaching, and much has been written about the use of video in teacher education (Brophy, 2004). The rules of engagement for this kind of PD activity are fairly straightforward and consistent throughout the literature. Somebody (either the PD facilitator or a participant) brings a selection of classroom video they have chosen to highlight a particular problem of practice or successful implementation of a target reform, participants hear a brief introduction to the context, and watch the video one or more times to create a common experience for analysis and discussion. What distinguishes the various use of classroom video is what exactly the teachers are asked to attend to during and after watching (Abell & Cennamo, 2004; Sherin, 2004; Sherin & van Es, 2009). While watching classroom video, teachers are frequently asked to attend to *student thinking and response* and/or *teacher moves and choices* during instruction. In fact, many programs report asking teachers to use both of these lenses in succession. For example, Borko and colleagues (2008) reported on a PD program in which teachers first planned a mathematical task, then taught and recorded the lesson and submitted recordings to the PD facilitators. The facilitators then chose some clips for the group to analyze; they spent one session attending to the teacher's role during the task and the next looking at student thinking. Tekkumru-Kisa and Stein (2017b) went the other direction, starting with asking teachers to attend to the cognitive depth of students' talk during a science task and then later watching contrasting videos of how teachers facilitated that task. The important idea throughout here is that classroom video can be a valuable representation of classroom practice, but it can also be complicated and difficult to interpret unless the video is carefully selected and teachers' attention drawn to particular components of instruction that the facilitator wishes to highlight (Chen et al., 2020; Sherin & van Es, 2005).

A well-established example of using video to support teacher PD in science is the Science Teachers Learning from Lesson Analysis (STeLLA) project (Roth et al., 2011, 2017, 2019;

Taylor et al., 2017). The STeLLA PD is a yearlong program for elementary science teachers focused on increasing their ability to recognize and support students' scientific thinking while designing science lessons that logically build science ideas to support students (Roth et al., 2011, 2017). The program is based around a framework of 17 strategies organized based on their ability to consider student thinking and science content development (Roth et al., 2011, 2017). These strategies include asking questions to reveal student thinking or explicitly showing how science ideas are linked together. The focus, therefore, is on helping teachers to understand how science ideas canonically progress from simple to more complex (Jin et al., 2019) and how teachers can help students to share their thinking as they travel along that progression. Teachers begin their work with STeLLA in a summer institute that includes "content deepening" activities to improve their science content knowledge and classroom video analysis sessions, designed to highlight some of the STeLLA strategies (Roth et al., 2017). Before watching the video, teachers are prompted with one or two specific STeLLA strategies to attend to during the video so they can consider how the teacher used those strategies. Teachers are then asked analysis questions to push on what they noticed about the target strategy and how it may have supported student thinking (Roth et al., 2017). Over the course of the year, the program shifts from providing videos to asking participants to bring their own videos and engage in the same analysis. A number of cluster-randomized studies have been done on the program, showing that teachers who engage in the program demonstrate higher increases in content knowledge, pedagogical content knowledge, and effective practices than teachers who participate in a control content-only PD (Roth et al., 2011, 2019). These results have also been extended to students, showing that students of teachers in the STeLLA perform better on science content assessments than students of control teachers (Taylor et al., 2017).

The STeLLA program is a strong example of thoughtful use of classroom video because it shows that by carefully selecting classroom video and providing effective prompts, PD providers can help highlight particular aspects of an NGSS classroom that may be hard to replicate in the PD space. Asking participants to consider how teachers implement specific moves such as asking for students' ideas can help them to see both how and why that teacher move might be helpful in supporting students' epistemic agency (Miller et al., 2018). On the other hand, participants might attend specifically to students, foregrounding for them what students say or do during instruction. This attention to students can help participants see what it looks like when students engage in the complex sensemaking around natural phenomena the NGSS aims for, and, more broadly, demonstrate that all students *can* engage in complex sensemaking, which not all teachers believe (Lebak, 2015). At the same time, however, classroom video might not help teachers see some of the more expansive approaches to sensemaking that students bring to scientific phenomena, particularly students who have traditionally been marginalized by science (Bang et al., 2017). Video helps teachers focus on what students or teachers say, but might make it harder to consider why they said that, how they might be feeling in that moment, the previous instructional and classroom context, or the relationship students have with each other and the science ideas under discussion.

Rehearsing Teaching

The rehearsal, a signature pedagogical tool of core practices in teacher education movement, is usually a 10 to 20-minute activity in which one teacher implements some component of a lesson for other teachers who are role playing as students (Kazemi et al., 2016; Kelley-Petersen et al., 2018; Lampert et al., 2013). On its face, the rehearsal appears very similar to microteaching, which gained popularity in the late 1960s as way for teachers to practice specific teaching moves such as wait time under the guidance of teacher educators (D. W. Allen & Eve, 1968; MacLeod, 1995). The proponents of rehearsals argue there are two key differences between the two activities. First, during rehearsal either the rehearsing teacher or the teacher educator can pause to discuss or even change an instructional decision, which allows the entire group to consider how the moves the teacher makes impact student thinking (Davis et al., 2017; Kloser & Windschitl, 2020; Lampert et al., 2013). Second, some argue that rehearsals focus more on using students' ideas as resources to support discussion and making in-the-moment instructional decisions based on those ideas rather than mastering context-independent teaching moves (Forzani, 2014).

A representative example of rehearsal in science comes from Kloser and Windschitl (2020), who detailed their use of rehearsal in their secondary science methods classes. In rehearsals focusing on supporting preservice teachers to facilitate discussion, they provide teachers with data relating to a key scientific idea and ask them to rehearse a discussion making sense of that data. In one such rehearsal, Kloser provided students with a map of the world with different types of fossils labeled across continents. These data were designed to serve as evidence of tectonic plates as a force shaping Earth's landforms. During the rehearsal, a pre-service teacher acting as a student proposed that there could be multiple similar fossils in different continents because dinosaurs swam from one continent to another. The pre-service teacher acting as the teacher responded with "not quite what I was looking for." Kloser then paused and the group had a discussion about how that framing might close down student thinking and participation. In that conversation, Kloser invited multiple novice teachers in to consider how they might structure the discussion to help students' ideas be heard without getting lost with suggestions that did not move the discussion towards its ultimate goal. By pausing and redirecting these kinds of questions about practice to the entire group, the rehearsal allowed the class to come to a joint understanding of the theories of learning at play rather than simply having the teacher educator provide the "correct" approach to this particular student response.

As shown in this brief example, the focus of rehearsal is on supporting teachers to consider their instructional moves and decisions and how those decisions influence the work of the students. As a result, these activities foreground the interactions between teacher and student while backgrounding the particular science content ideas under discussion. This can be seen in the way that Kloser and Windschitl (2020) provided data for discussion: the idea that tectonic plates cause differences in landforms was used as a context for conversation, but the goal was not to understand *that idea* or even how students engaged with it. Rather, it was focused on teacher moves that might open up space for students to make sense of data more broadly. Rehearsal, therefore, can help teachers to consider the impact their moves might have on student participation and therefore thinking, which is important and valuable work. But they also might background the teacher's and students' relationship with content under discussion. In addition, the focus on teacher moves may limit the amount teachers consider students' affective responses to instruction, particularly those traditionally underserved by science (Jaber & Hammer, 2016). Unless thoughtfully facilitated, this could end up focusing teachers' thinking on moves that only facilitate sensemaking as envisioned by canonical, Western science (Bang et al., 2017).

Analyzing Student Work

Analyzing student work during practice-based PD is the process of systematically looking at a set of artifacts produced by students in order to determine what ideas those students seem to understand, where they are confused, and potential instructional next steps (Loucks-Horsley et al., 2010; Reiser, 2013). A variety of things can count as student work, including written or drawn responses to a task, video or audio recordings the students have produced explaining something, or journals or notes students have taken. What makes student work helpful to analyze, however, is that the prompt is rich enough and the responses long enough to show variety in student thinking (Loucks-Horsley et al., 2010). In other words, it is not particularly helpful to analyze students' multiple-choice tests or a task in which all students provided the same response as those do not provide enough complexity and variation to make claims about what students did and did not understand. This activity also tends to include clear instructions for analysis, usually with a written protocol that directs teachers to consider what ideas they are seeing in the work, how they agree or disagree with what others are seeing, and potential instructional next steps (e.g. Heller et al., 2012; McNeill & Knight, 2013). The student work can be pre-selected by PD facilitators to highlight particular student ideas or brought in by participants. Each of those approaches has its own advantages and disadvantages: facilitator provided work can be easier to supply and help facilitators to ensure particular ideas are surfaced in discussion. Participant-provided work, on the other hand, can be more meaningful for teachers as it is more directly connected to their practice, but at the possible expense of teachers feeling vulnerable sharing work from their classroom or their existing relationships with their students interfering with their interpretation of the work (Loucks-Horsley et al., 2010).

A characteristic example of using analysis of student work comes from Heller and colleagues (2012), who designed a PD workshop designed to improve elementary science teachers' ability to support their students' conceptual understanding of science. Their study involved three different versions of the workshop, all of which included a "science content" component and one of which included a "looking at student work" component. In the student work component, teachers brought in formative assessments they had designed and given to their students. Teachers were given a bank of assessment items they could choose from which included multiple choice questions and accompanying explanations; for example, one item included an image of a battery, light bulb, and wire connected incorrectly and asked students to identify if the light bulb would light up and to explain their thinking. Groups of teachers used a written protocol to analyze the quality of the formative assessment task itself, the ideas that students did or did not show on their responses, implications of those responses for teaching and learning, and the process of engaging in this student work analysis. At the end of the year of PD, the authors found that both teachers' and students' content knowledge and written explanations improved, which they partially attributed to teachers' practice writing and interpreting formative assessments (Heller et al., 2012).

By focusing so tightly on the ideas students express in some written or recorded artifact, the activity of analyzing student work foregrounds for teachers the ideas that students have about a particular topic and how those ideas do (or do not) align with canonical understandings of that topic. This can help teachers to know where their students might not be approaching their goals for understanding and design instruction to help them achieve those goals (Loucks-Horsley et al., 2010). It can also help foreground for teachers the *process* of thinking about student work in a way that highlights student ideas rather than simply looking for a grade or number of points, which is particularly important if teachers are going to be truly responsive to what students are thinking (Robertson et al., 2015). However, the nature of a static artifact can background the processes that took place to get to that artifact, both on the part of the teacher and the student. It might be harder for teachers to consider the moves teachers made during instruction, how students might have responded to them, and how students' own ideas or interests are (or are not) reflected in the work.

Doing Science as an Adult Content Learner

In order to support teachers to develop their understanding of science, many PD programs ask teachers to do science investigations, read texts, and/or engage in sensemaking discussions around content, essentially taking on the role of "science learner" (Fulp et al., 2009; Gibbons & Cobb, 2017; Roth et al., 2019; Taylor et al., 2017). This can mean that teachers do the actual investigations, readings, or activities intended for their students and/or that they do more advanced work in order to learn science beyond the scope that they would expect their students to know. The key unifying feature of this work is that teachers are using their own backgrounds and experiences to make sense of science ideas for themselves. Fulp and colleagues (2009) include an example of how a facilitator might explain this to participants:

> You are going to be wearing two different hats today, teacher hat and adult learner, not a 5th or 6th grade learner…. We are going to go through some power point slides with some content…It is always good when the teacher knows more about what they are teaching than the students…This is not something you would share with you students…Once again this is knowledge for you, not your students. (p. 30)

These instructions make it clear that even when teachers are doing activities intended for students, they should be acting as adults to learn the science content, and they may also be asked to work with content beyond the scope of what is expected for their students.

Studies that have focused on developing teachers' knowledge of science content have found that it does impact teachers' and students' ability to show strong understanding of science (Capps et al., 2012; Yoon et al., 2007). As a representative example, take the P-SELL curriculum and PD, which was designed to support elementary teachers to teach science to all their students in ways that particularly support their emergent multilingual students (Lee et al., 2016). The multi-year project included workshops in which teachers engaged in investigations from the curriculum as well as more general inquiry activities not found in the curriculum (Buxton et al., 2008). The goal of asking teachers to do activities from the curriculum was to develop their understanding of the science topics they would be teaching, such as the particulate nature of matter or the role of weathering and erosion in shaping the Earth's surface. The general inquiry activities, on the other hand, were aimed at improving teachers' understanding of science as a process of figuring out the natural world, which at the time was typically referred to as "inquiry" (Crawford, 2014). Analysis of those teachers' content knowledge through a content assessment found that teachers who engaged with the PD showed significantly higher growth of content knowledge compared to a control group of teachers who did not participate in the PD; in addition, students of teachers with higher content knowledge scored higher on state science tests (Diamond et al., 2014).

By focusing on developing teachers' science knowledge, these PD activities foreground for teachers both the core ideas and the structure of science as a process of figuring out the natural world (National Research Council, 2012). This can be particularly valuable because teachers asked to teach content they do not understand are more likely to rely on simple memorization tasks rather than truly engaging their students in scientific activity as envisioned by the NGSS (Childs & McNicholl, 2007; Napier et al., 2020). A focus only on teacher's understanding of science, however, can background both what specific teacher moves or decisions might help students to come to understand these same ideas and what ideas, interests, or sensemaking practices students might bring to the content. Fulp and colleagues (2009) pointed out the former concern when discussing "adult learner hat," noting that teachers require time after engaging in science investigations as learners to consider the teaching logistics and decisions of implementing those same lessons with their students. The latter concern is particularly important in considering how to open up science spaces for students whose sensemaking practices have traditionally been ignored or pushed to the side by Western science (Bang et al., 2012). By focusing on building teachers' canonical science knowledge, engaging in the discipline might not emphasize the sociocultural assumptions built into the way we structure and define science, and therefore how it continues to exclude many who do not share those same cultural backgrounds (Nasir et al., 2014; Warren et al., 2020).

Doing Science in Student Hat

When teachers do science activities in the student hat, they are engaging in learning activities they will do with their students while attending to the ideas, experiences, and interests that their students might bring to a learning activity and the cognitive and affective responses they might have to that activity. In order to do this, teachers are asked to use language students might use, saying, for example, "I wonder why hail is formed in some places but not others" rather than "My students won't know the factors that cause hail to be formed" (Lowell & McNeill, 2020; McNeill et al., in press). The goal is not to build teachers' own content knowledge nor role play particularly difficult student behaviors to consider how a teacher might address them. Rather, the focus here is on pushing teachers to think about and then say out loud how students might respond to a particular learning activity in order to push teachers to better empathize with their students' relationship with the content and instruction. For this reason, this type of activity has also been called "think like a student" rather than "act like a student" because the focus is on considering how a student *thinks* rather than how they *behave* (Next Generation Science Storylines & Next Generation Science Exemplar System, 2018).

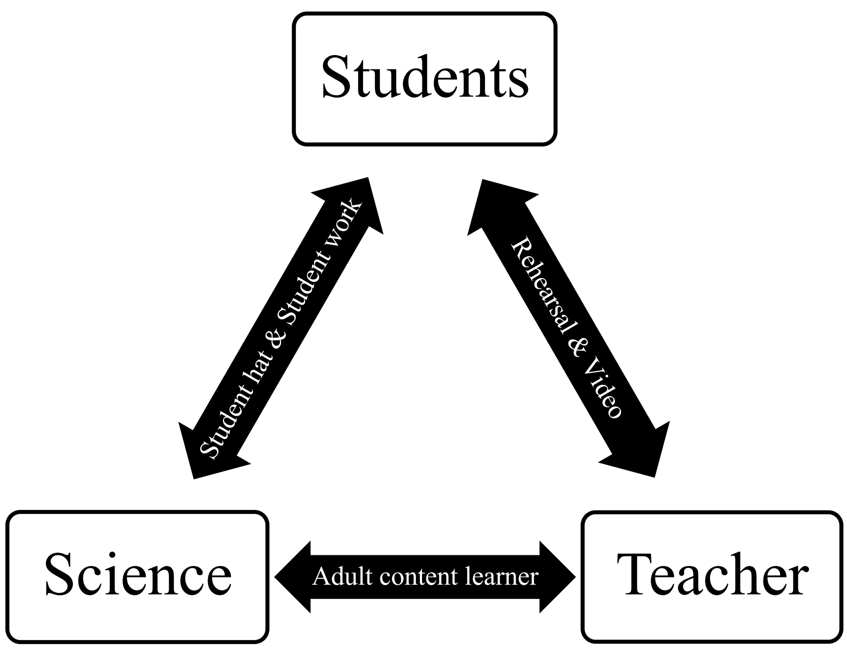
A representative example of this kind of activity in the literature comes from Rosebery and colleagues (2016), who designed and facilitated a PD for early career science teachers focused on supporting teachers to use "expansive pedagogical practices that encourage, make visible, and intentionally build on students' ideas, experiences, questions, and perspectives on scientific phenomena" (Rosebery et al., 2016, p. 1572). They asked participants to engage in investigations around the growth and development of Wisconsin Fast Plants using ideas their students might have rather than their own towards the goal of helping "participants to experience what it was like to problematize, complicate, question, and trouble taken-for-granted meanings in the ways that students often do" (Rosebery et al., 2016, p. 1577). During the PD they had participants track the growth of the Fast Plants in terms of their length, width, and number of leaves in order to construct multi-dimensional representations of the plant's life cycle. This work was substantially more complex than the standard two-dimensional presentation of plant growth as plants getting taller and resulted in teachers pushing on their understanding of how to represent growth and change of living organisms. After engaging in the PD, they found that participants were more likely to see students' sensemaking as generative even if it did not match canonical approaches to talking about science. For example, when one student used gestures and referenced a mural on the wall to explain the life cycle of a pumpkin, teachers saw that as enriching the students' verbal explanations and distributing thinking across the student, class, and local representations (the mural). Rosebery and colleagues concluded that the use of the student hat may have helped teachers to better consider students' potential explanations and expanded for teachers what they saw as students' "legitimate" scientific thinking.

Asking teachers to focus on students' ideas and share them in students' language can help to foreground the cognitive and affective experiences of students in NGSS classrooms. In the study by Rosebery and colleagues (2016), asking teachers to consider the different ways students might interact with the science of plant development helped them to better see that these ideas might develop differently in their students based on each student's personal and cultural resources. Similarly, the experience of engaging in deep sensemaking about a natural phenomenon is different than traditional science and might make students feel uneasy, particularly when encountering moments of uncertainty (Han & Gutierez, 2021; Jaber & Hammer, 2016; Manz, 2018). Foregrounding this experience for teachers could help them to consider what it feels like to engage in complex scientific sensemaking in ways that are more visceral than simply discussing those feelings. At the same time, however, the tight focus on the student's experience might background other teacher learning goals about the content itself or the teacher moves taken to facilitate this thinking. In Rosebery and colleague's program, for example, the focus was so tightly on noticing students' thinking that every teacher came away still wanting to consider what moves they could make to open up space for and value that thinking.

Summary

Going through each of five common practice-based PD activities, I have used the instructional triangle (Cohen et al., 2003) to highlight which components of the classroom learning environment each one foregrounds or background. Figure III.3 shows these activities placed onto the triangle to represent these ideas graphically. Being an adult content learner pushes on teachers' relationship with the science content, helping them to better understand core ideas or the process of figuring out. Watching classroom video and rehearsal both highlight the relationship between the student and the teacher in the classroom, highlighting how student teacher moves might shape student responses or vice-versa, although the key difference between these two is that rehearsal takes place in the moment while video cannot be improvised. The student hat and analyzing student work both help teachers to consider the relationship between students and the science they are learning. Like rehearsal, student hat is improvised in the moment while analyzing student work is more static and post-hoc. This difference is that student hat allows PD facilitators to support teachers in considering the affective experience of engaging in scientific sensemaking (Jaber & Hammer, 2016) in ways that few other PD activities can. This affective experience is a particularly important but overlooked component of supporting teacher learning around what it means to implement the ambitious vision of the NGSS (Jaber, 2021; Jaber et al., 2018). Therefore, given the relative novelty of this student hat approach to teacher learning, it merits taking a closer theoretical look at what specific goals for NGSS teacher learning it might support and how in order to inform future empirical studies of the activity.

Figure III.3: Professional Development Activities Mapped onto the Instructional Triangle



Student Hat: What Does it Do Differently from Other Activities?

On its face, the student hat might look similar to a combination of the other activities. Like being a content learner, it involves teachers doing science investigations and engaging in discussion about those investigations; like rehearsal, it asks participants to consider how students might respond to something while responding like students. Nevertheless, the student hat is fundamentally different from both of these approaches because of the focus of attention. Unlike content learner activities, the goal is not on learning science ideas or explaining them in the most clear and precise way possible, but rather on considering how students novice to the discipline might use their experiences, ideas, and questions to make sense of natural phenomena. Unlike rehearsal, student hat directs teachers' attention to the cognitive and affective responses to being a student in an NGSS classroom rather than teacher moves.

One might reasonably ask if we want to understand how students respond to the NGSS classroom, why not watch video or analyze work of students who have been in such a classroom? These approaches might be preferrable because they involve watching actual students rather than adults pretending to be students. But because both of these are post-hoc activities, taking place after the learning has happened, they do not help teachers consider how students put ideas together piece by piece to come to a deep understanding, or what it feels like to wrestle with uncertainty until finding the joy of success (Jaber & Hammer, 2016). In discussing their version of student hat, Kloser and Windschitl (2020) said that it can help teachers "to feel a growing sense of competence or frustration with a science activity, and perhaps to feel what it is like to have their ideas treated as resources for everyone's reasoning" (p. 70). This is particularly important because it aligns with two major goals of NGSS instruction: supporting students' epistemic agency though responsive teaching and crafting learning experiences that are coherent from the student perspective (Lowell et al., 2021).

Since the release of the NGSS, the science education community has come to agreement that strong science instruction requires that students have some sort of influence over the form or structure of idea work done in class (Campbell & Lee, 2021). In other words, the hope is that, like scientists, students can make scaffolded but meaningful decisions about the investigations they do into the natural world (Miller et al., 2018; Stroupe, 2014). In order to facilitate this epistemic agency, teachers must elicit and listen to the ideas that students bring to the classroom and ensure those ideas are honored and used by the class (Robertson et al., 2015). That requires that teachers consider the multiple ways of knowing that students bring to learning so that they are not only looking out for ideas and thoughts that match teachers' canonical view of science (Warren et al., 2020). This is important because instructional practice that attends to and values the heterogeneity in students' experiences, ideas, and sensemaking practices can begin to counteract traditional disciplinary messages that have excluded minoritized students from science (Bang et al., 2012; Calabrese Barton & Tan, 2019; Rosebery et al., 2010; Warren et al., 2020). Unfortunately, during PD teachers are prone to consider planning and curriculum implementation over student thinking and experiences, even when the PD emphasizes analysis of student thinking and learning (Collins et al., 2019). Because of its affective power, student hat could be an effective tool to focus teachers on the diverse array of student sensemaking in their classroom that will help them support their students' epistemic agency. This idea is based on the theory that if teachers themselves experience idea building as complex, non-linear, and non-standard, they may be more likely to accept unexpected ideas and ways of expressing them from their students (Rosebery et al., 2010, 2016). Nevertheless, this leaves open a number of empirical questions about the influence of student hat on teacher learning and its mechanism for doing so. Does this experience actually help teachers to see multiple ways of knowing in PD, and what about later in their classroom practice? The field would be well served by further investigation of these questions.

A second major feature of strong NGSS curriculum and instruction is that it is coherent *from the student perspective* (Reiser, Novak, et al., 2017; Reiser et al., 2021). Traditional models of coherence have focused on how science content builds logically from simple to more complex ideas, but those models have centered those progressions from the perspective of a content expert who already understands all of the target ideas (Lowell et al., 2021; Reiser, Novak, et al., 2017). Students are not such experts, but teachers often are, and given how long it has been since they were novice science learners, may be likely to forget what exactly novices do and do not know (Nathan & Petrosino, 2003). By asking teachers to limit their contributions in student hat to ideas and information that students might reasonably have at the moment, we can help to foreground the ways that strong learning experiences meaningfully build on students' ideas (National Academies of Sciences, Engineering, and Medicine, 2019; Reiser et al., 2021; Zivic et al., 2018). When helping teachers to understand what it means for curriculum to be coherent from the student perspective and why that it is important, the student hat is uniquely positioned to support teacher learning because of the way it helps teachers see and experience that build first-hand in real time. Nevertheless, this area has limited research and there are legitimate questions about the extent to which student hat can support this idea as compared to more efficient means of watching students build their ideas like analyzing classroom video. This calls for more empirical work investigating the differential benefits of each of these PD approaches on teachers understanding of and ability to support coherence from the student perspective.

Of course, there are possible pitfalls in asking teachers to speak and engage like how their students might. One particularly important one is the potential for teachers to focus on deficit-oriented or stereotyped views of their students' interests or ideas, particularly their students of color. Teachers can bring in deficit-oriented views of their students into PD, which may be because of a history of policy and practice framed around students of color and emerging multilingual learners "lacking" prerequisite academic skills (Battey & Franke, 2015; González-Howard & Suárez, 2021; Lee, 2021). If not thoughtfully facilitated, student hat activities might devolve into representing or reinforcing these negative views of students. In addition, by foregrounding the students' experience with science content, we might background other important components of strong science instruction such as the particular teacher moves that might support this work or the complex science ideas that students are being asked to learn. Taken together, these concerns point to our need to better understand the process of facilitation and how student hat activities in particular can be facilitated well to avoid these pitfalls (or the impact when it is facilitated poorly). What specific moves do facilitators of student hat do to direct teacher thinking towards students without feeling fake, forced, or deficit-oriented? How are those facilitation moves different or similar to other moves used in rehearsal or analysis of classroom videos? These questions remain open for further study.

An Example of Student Hat in Action: The OpenSciEd Curriculum Launch

My approach so far in discussing the student hat has been mostly theoretical, outlining what it may and may not be able to do in terms of supporting teacher learning for the NGSS. In order to make these distinctions clear, it is helpful to consider an illustrative example of the student hat in action during PD and how it can combine effectively with other strategies to address the target outcomes of the PD and the learning needs of the participants (Loucks-Horsley et al., 2010; Reiser, 2013). Therefore, in this section I outline one example of a PD workshop that has used the student hat. I use this example to show how student hat can be integrated with other practice-based PD activities and what makes the student hat different from other similar activities such as engaging as a content learner.

This example comes from the OpenSciEd middle school professional development program, for which I was a member of the PD design team. OpenSciEd is a non-profit organization consisting of curriculum and PD designers from multiple academic and nonprofit institutions. The organization releases open-source curricular materials and associated PD sessions using one particular approach to instructional design called storylines (Edelson et al., 2021). Storyline units are designed around an anchoring phenomenon that students are introduced to at the beginning and motivates the investigations throughout the unit (Reiser et al., 2021; Windschitl et al., 2018). They are called "storylines" because they are built to be coherent from the student perspective, building on students' ideas so that students can construct a story about the anchoring phenomenon that makes sense to them (Reiser et al., 2021). A four-day "curriculum launch" PD workshop introduces teachers to the OpenSciEd curricular approach (OpenSciEd, 2021). The PD agenda is outlined in Table III.2 and includes an introduction to the idea of an anchoring phenomenon, coherence from the student perspective, and two features of instruction teachers often find difficult: facilitating class discussions and assessment. I have italicized in Table III.2 any practice-based PD activity discussed in this paper.

Table III.2: Summary of OpenSciEd Curriculum Launch PD Activities

	Day 1	Day 2	Day 3	Day 4
Whole-group introduction (across grade levels)	Introduction to OpenSciEd *Watch video & analyze student work* from anchoring phenomenon to a common unit (3 hours)	Discuss and *watch classroom video* about supporting coherence from the student perspective (1 hour)	Discuss strategies for supporting productive classroom discussion *Watch video* of classroom discussion (1 hour)	Discuss assessments in OpenSciEd *Analyze student work* from a common unit (1 hour)
Unit-specific (single grade level)	Experience anchoring phenomenon of target unit in *student hat* (3 hours)	Construct summary of unit storyline *Do lessons in student hat* (5 hours)	Plan and *rehearse a classroom discussion* *Do lessons in student hat* (5 hours)	*Analyze student work* from an assessment *Do lessons in student hat* (5 hours)
Whole-group closing (across grade levels)	Reflect on anchoring phenomenon across units (1 hour)	Reflect on supporting coherence from the student perspective (1 hour)	Reflect on discussion norms and facilitating class discussions (1 hour)	Reflect on learning from workshop about OpenSciEd teaching (1 hour)

Given the complexity and novelty of the OpenSciEd approach, the PD workshop is designed to combine a range of practice-based activities in order to meet different goals in support of teacher learning (McNeill et al., in press). In order to show teachers that students can engage in complex sensemaking and the teacher moves that can support that work, the workshop begins with watching classroom video at the beginning of days 1, 2, and 3. Then, to give teachers an opportunity to feel confident for themselves that they can run a class discussion, the session incorporates a rehearsal during day 3. Finally, it asks teachers to analyze student work multiple times in day 4 to help teachers see the kinds of understandings students can communicate when engaged in ongoing sensemaking around an anchoring phenomenon. These are all important goals that hit on all three components of the instructional triangle (students' thinking, teacher moves, science content), but only include part of student thinking because they do not help teachers experience what it looks and feels like to incrementally build science ideas in ways that are coherent to students. For that, the PD includes student hat on each day, which serves to both give teachers a detailed introduction to important lessons and help them understand student coherence in a way that is hard to do without experiencing it.

To illustrate what exactly this student hat might look like in an OpenSciEd PD workshop, I include below a brief transcript from one. This exchange happened during a session focused on teachers preparing to teach a unit on weather systems. In student hat, the participants watched three videos of hail events, drew initial models trying to explain what causes hail, and then participated in a large group conversation to share their ideas and questions. Table III.3 shows what took place.

Table III.3: Example of Student Hat Conversation

	Speaker	**Transcript**
1	Facilitator:	Yeah. Just a few of us if they can share some of their ideas.
2	Mr. Truett:	What I did was, I was wondering, well the question, if the places we saw where the hail occurred might have a darker surface than areas around it. So it would be converting more solar energy to heat and causing a convection cell. So you have a funnel of air rising and when a cloud, a rain cloud comes over that convection cell, if it were gonna drop rain drops they might get blown back up where it's colder and freeze and get coated. This could happen over and over and over so your larger hail stones would have gotten in that updraft from the convection cell and get pushed up into the cold, coated by more water, freezing again and freezing again. And the way we could figure out if that's true is to take one of the hail stones, one of the large hail stones, and literally cut it in half to see if it has rings like a tree.
3	Facilitator:	You used lots of big words

4	Multiple:	Yeah
5	Ms. Vernon	Are you in student hat or teacher hat?
6	Facilitator:	No, we're supposed to be-so let's be in student hat. But, let's ask Mr. Truett some questions about some of those ideas though, right? So, Ms. Howard, do you have any questions?
		[Approximately 1 minute cut discussing Mr. Truett's response]
19	Facilitator:	So we have some ideas that we might need to get some evidence for. We're not quite sure but let's try to see what other folks have. So do you want to start us off, Ms. Townsend?
20	Ms. Townsend:	Yeah, can I just show? I just did it in a cartoon cycle-ish. I'm a really bad drawer but I do land okay. I have a sunny day over the mountains, you see I have the mountains on my mind because I'm going skiing. But it's a sunny day and then I have clouds moving in. White, puffy clouds, because that's what I learned about. And then they got darker, so I colored them in a little bit and had some rain drops. And then I got even darker and I scrubbed down on my pencil and I have bigger dots, meaning that it's the hail, and then I have it come back to being a sunny day again. So my questions were like: Why is it going in that direction? Why are the clouds so dark? Is it, is it, are they full? They look heavy. They look really heavy when they're dark. So are they full of something? I don't know. They look so dark and my question is, what's in the cloud? I don't know what's in the cloud. What makes my white puffy clouds black?
21	Facilitator:	Do you have initial ideas of what was happening in those clouds? I saw you draw some of these pictures of what was happening.
22	Ms. Townsend:	Um, I just drew that they got darker, that they went towards a grey when it was raining and when the hail came down, they got really, really dark. And I'm not sure but all of a sudden it stops, so instead of drawing another picture on the bottom, I made it a cycle. And made it back to the sunny day, so I just ask questions, intensity, was there a force? Was it some sort of pressure thing going on? Wind? Was the wind pushing the storm this way? You know? Because that's what it looked like. That's the kind of storm I usually hide from. I don't look out the window when this kind of storm comes by.
23	Facilitator:	Does anyone, in their models, have anything that talked about what was happening in those clouds? Any models that pointed in that area?
24	Ms. Vernon:	I guess not in the clouds, but I thought something was happening when two clouds meet. Like the lighter clouds and the darker clouds they were meeting.

In this brief snippet of conversation, we see two very different approaches to student hat, which helps to highlight how student hat is different from being a content learner and better aligned with approaches to teaching that value students' own experiences, ideas, and questions. In turn 2, Mr. Truett began by sharing ideas about how hail forms as a teacher or an adult content learner. He used terms such as "convection cell," solar energy," and "updraft" to succinctly explain the process of hail formation. His description matches our canonical understanding of how hail forms, and if the focus of this activity were for teachers to learn the science, this might have been a helpful contribution. But he is not effectively considering ideas and language his novice students are likely to use immediately after being introduced to hail. His engagement in the PD is different from the instructional model in which students begin by making observations and asking questions of a natural phenomenon to drive investigations that make sense to him. Instead, Mr. Truett is focused on science as "final form" knowledge to be shared rather than built through a process of collaborative sensemaking (McNeill & Berland, 2017). The facilitator and another participant signal that this response is different from what is expected in "student hat" in terms of the "big words" that are being used.

Ms. Townsend and Ms. Vernon, on the other hand, provide responses that are aligned with the goals of the student hat activity in that they connect to experiences, ideas and words that middle school students might use if engaged in this experience. Ms. Townsend referenced back to her experiences skiing to help her think about sudden weather events and the difference between clouds that form hail and those that do not. Building off of that, Ms. Vernon proposes that there is something about clouds meeting that is involved. Clearly, these explanations are less complete than Mr. Truett's, but they are more in line with what students might think when doing this work for the first time, and by speaking in a student's potential language, Ms. Townsend and

Ms. Vernon are able to help the group see how students might connect past experiences, initial ideas about a phenomenon, and questions to motivate future investigations. By focusing on this incremental building of ideas rather than canonical explanations of science concepts, student hat can help teachers attend to the importance of hearing and responding to students' thoughts and using them to support the work of the class (Rosebery et al., 2016; Zivic et al., 2018).

Overall, this Curriculum Launch PD session (Table III.2) shows how the student hat can be used in concert with other activities to push on teachers' understanding of sensemaking that honors students' ideas, works with them, and is coherent from their perspective. At the same time, however, the session can also address other goals common among other practice-based PD programs, including considering and practicing specific teacher moves and analyzing student thinking to understand what they might know. The example of Mr. Truett vs. Ms. Townsend also helps to clarify what it sounds like to participate in student hat as compared to engaging in the same experience as an adult learner. Taking on the perspective of a student learner rather than a content expert can help teachers experience instruction more in line with recent shifts in science education in which students' ideas, experiences, and questions drive instruction. This shift is fundamentally different from front loading academic language and canonical scientific explanations of complex natural phenomena. Therefore, by including a range of activities that help teachers to consider the interaction between teachers and students, teachers and science, and students and science, the OpenSciEd curriculum launch supports teachers to implement instruction better aligned with the vision of the NGSS.

Conclusion

Overall, the work in practice-based professional development and teacher education is based in the idea of knowledge in use, meaning that teaching is the work of applying knowledge about students, content, and teaching moves in use to support effective interactions between the three (Ball & Cohen, 1999; Grossman et al., 2018; Lampert, 2010). Therefore, this work can help teachers to consider what it might look like to have a classroom in which students engage in sensemaking around natural phenomena and how teachers might support that work (Stroupe et al., 2020; Windschitl et al., 2018). However, one thing the field continues to struggle with is ensuring that our curriculum and instruction are truly responsive to student ideas, experiences, and interests and allowing those ideas to guide the work of the science classroom (Miller et al., 2018; Robertson et al., 2015). In particular, we have a number of activities that help teachers consider their relationship with content and their relationship with students, but relatively few that push on teachers' understanding of students' relationship with content, and those that do are mostly post-hot analyses of student work (see Figure III.3). This leaves open a need to consider students affective responses to science instruction, which is particularly important for supporting students who have traditionally felt marginalized or ignored by science (Bang et al., 2017; Jaber & Hammer, 2016). Furthermore, helping teachers to actually *feel* the discomfort of not knowing or the joy of seing their idea valued by the group might help them to change their own beliefs about strong science instruction, given that beliefs are tied to teachers' affective responses (Crawford, 2007; Nespor, 1987).

The NGSS, and the broader move towards improving students' scientific sensemaking that it is part of, continues to push teachers and teacher educators to make substantial changes in the forms and grammar of science schooling (Campbell & Lee, 2021; Pruitt, 2014). As teacher educators, we must use every tool in our toolbox to support teachers in making this shift. Student hat offers one potential tool to address this need because it uniquely allows teachers to consider the affective experiences of engaging scientific sensemaking. Given the novelty and rarity of this type of activity, it deserves future empirical work investigating how exactly the student hat might support teacher learning, what some of its downsides or pitfalls are, and how PD facilitators can thoughtfully support teachers like Mr. Truett in engaging in the activity to push on their own thinking. As we come to a better understanding of this particular activity, we can continue to work to improve science teaching and learning for teachers and K-12 students nationwide.

References

Abell, S. K., & Cennamo, K. S. (2004). Videocases in elementary science teacher education. In J. Brophy (Ed.), *Using video in teacher education* (Vol. 10, pp. 103–129). Elsevier.

Allen, C. D., & Penuel, W. R. (2015). Studying teachers' sensemaking to investigate teachers' responses to professional development focused on new standards. *Journal of Teacher Education*, *66*(2), 136–149. https://doi.org/10.1177/0022487114560646

Allen, D. W., & Eve, A. W. (1968). Microteaching. *Theory Into Practice*, *7*(5), 181–185.

Ball, D. L. (2018, April 13–17). *Just dreams and imperatives: The power of teaching in the struggle for public education* [Presidential Address]. AERA Annual Meeting, New York, NY. https://www.youtube.com/watch?v=JGzQ7O_SIYY

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (Vol. 1, pp. 3–22). Jossey-Bass.

Bang, M., Brown, B., Calabrese Barton, A., Rosebery, A., & Warren, B. (2017). Towards more equitable learning in science: Expanding relationships among students, teachers, and science practices. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 33–58). NSTA Press.

Bang, M., Warren, B., Rosebery, A. S., & Medin, D. (2012). Desettling expectations in science education. *Human Development*, *55*, 302–318. https://doi.org/10.1159/000345322

Battey, D., & Franke, M. (2015). Integrating professional development on mathematics and equity: Countering deficit views of students of color. *Education and Urban Society*, *47*(4), 433–462. https://doi.org/10.1177/0013124513497788

Borko, H., Carlson, J., Mangram, C., Anderson, R., Fong, A., Million, S., Mozenter, S., & Villa, A. M. (2017). The role of video-based discussion in model for preparing professional development leaders. *International Journal of STEM Education*, *4*(1), 29. https://doi.org/10.1186/s40594-017-0090-3

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education*, *24*(2), 417–436. https://doi.org/10.1016/j.tate.2006.11.012

Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 7, pp. 548–556). Elsevier.

Brophy, J. (Ed.). (2004). *Using video in teacher education* (1st ed). Elsevier/JAI.

Buxton, C., Lee, O., & Santau, A. (2008). Promoting science among English Language Learners: Professional development for today's culturally and linguistically diverse classrooms. *Journal of Science Teacher Education*, *19*(5), 495–511. https://doi.org/10.1007/s10972-008-9103-x

Calabrese Barton, A., & Tan, E. (2019). Designing for rightful presence in STEM: The role of making present practices. *Journal of the Learning Sciences*, *28*(4–5), 616–658. https://doi.org/10.1080/10508406.2019.1591411

Campbell, T., & Lee, O. (2021). Instructional materials designed for A Framework for K-12 Science Education and the Next Generation Science Standards: An introduction to the special issue. *Journal of Science Teacher Education*, *32*(7), 727–734. https://doi.org/10.1080/1046560X.2021.1975359

Capps, D. K., Crawford, B. A., & Constas, M. A. (2012). A review of empirical literature on inquiry professional development: Alignment with best practices and a critique of the findings. *Journal of Science Teacher Education*, *23*(3), 291–318. https://doi.org/10.1007/s10972-012-9275-2

Chen, G., Chan, C. K. K., Chan, K. K. H., Clarke, S. N., & Resnick, L. B. (2020). Efficacy of video-based teacher professional development for increasing classroom discourse and student learning. *Journal of the Learning Sciences*, *29*(4–5), 642–680. https://doi.org/10.1080/10508406.2020.1783269

Childs, A., & McNicholl, J. (2007). Science teachers teaching outside of subject specialism: Challenges, strategies adopted and implications for initial teacher education. *Teacher Development*, *11*(1), 1–20. https://doi.org/10.1080/13664530701194538

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*(2), 119–142. https://doi.org/10.3102/01623737025002119

Collins, L., Cavagnetto, A., Ferry, N., Adesope, O., Baldwin, K., Morrison, J., & Premo, J. (2019). May I have your attention: An analysis of teacher responses during a multi-year professional learning program. *Journal of Science Teacher Education*, *30*(6), 549–566. https://doi.org/10.1080/1046560X.2019.1589846

Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, *44*(4), 613–642. https://doi.org/10.1002/tea.20157

Crawford, B. A. (2014). From inquiry to scientific practices in the science classroom. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of Research on Science Education* (pp. 515–541). Routledge. https://doi.org/10.4324/9780203097267

Davis, E. A., Kloser, M., Wells, A., Windschitl, M., Carlson, J., & Marino, J.-C. (2017). Teaching the practice of leading sense-making discussions in science: Science teacher educators using rehearsals. *Journal of Science Teacher Education*, *28*(3), 275–293. https://doi.org/10.1080/1046560X.2017.1302729

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Diamond, B. S., Maerten-Rivera, J., Rohrer, R. E., & Lee, O. (2014). Effectiveness of a curricular and professional development intervention at improving elementary teachers' science content knowledge and student achievement outcomes: Year 1 results. *Journal of Research in Science Teaching*, *51*(5), 635–658. https://doi.org/10.1002/tea.21148

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287–312. https://doi.org/10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A

Edelson, D. C., Reiser, B. J., McNeill, K. L., Mohan, A., Novak, M., Mohan, L., Affolter, R., McGill, T. A. W., Buck Bracey, Z. E., Deutch Noll, J., Kowalski, S. M., Novak, D., Lo, A. S., Landel, C., Krumm, A., Penuel, W. R., Van Horne, K., González-Howard, M., & Suárez, E. (2021). Developing research-based instructional materials to support large-scale transformation of science teaching and learning: The approach of the OpenSciEd middle school program. *Journal of Science Teacher Education*, *32*(7), 780–804. https://doi.org/10.1080/1046560X.2021.1877457

Forzani, F. M. (2014). Understanding "core practices" and "practice-based" teacher education: Learning from the past. *Journal of Teacher Education*, *65*(4), 357–368. https://doi.org/10.1177/0022487114533800

Fulp, S. L., Warren, C. L., & Banilower, E. R. (2009). *Science: It's Elementary: Year three evaluation report*. Horizon Research, Inc. https://eric.ed.gov/?id=ED518456

Gibbons, L. K., & Cobb, P. (2017). Focusing on teacher learning opportunities to identify potentially productive coaching activities. *Journal of Teacher Education*, *68*(4), 411–425. https://doi.org/10.1177/0022487117702579

González-Howard, M., & Suárez, E. (2021). Retiring the term English language learners: Moving toward linguistic justice through asset-oriented framing. *Journal of Research in Science Teaching*, *58*(5), 749–752. https://doi.org/10.1002/tea.21684

Grossman, P. (Ed.). (2018). *Teaching core practices in teacher education*. Harvard Education Press.

Grossman, P., Kavanagh, S. S., & Pupik Dean, C. G. (2018). The turn towards practice-based teacher education: Introduction to the work of the Core Practice Consortium. In P. Grossman (Ed.), *Teaching core practices in teacher education* (pp. 1–14). Harvard Education Press.

Han, M., & Gutierez, S. B. (2021). Passive elementary student's constructed epistemic emotions and patterns of participation during small group scientific modeling. *Science Education*, *105*(5), 908–937. https://doi.org/10.1002/sce.21665

Heller, J. I., Daehler, K. R., Wong, N., Shinohara, M., & Miratrix, L. W. (2012). Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. *Journal of Research in Science Teaching*, *49*(3), 333–362. https://doi.org/10.1002/tea.21004

Jaber, L. Z. (2021). “He got a glimpse of the joys of understanding” – The role of epistemic empathy in teacher learning. *Journal of the Learning Sciences*, *30*(3), 433–465. https://doi.org/10.1080/10508406.2021.1936534

Jaber, L. Z., & Hammer, D. (2016). Learning to feel like a scientist. *Science Education*, *100*(2), 189–220. https://doi.org/10.1002/sce.21202

Jaber, L. Z., Southerland, S., & Dake, F. (2018). Cultivating epistemic empathy in preservice teacher education. *Teaching and Teacher Education*, *72*, 13–23. https://doi.org/10.1016/j.tate.2018.02.009

Jin, H., Mikeska, J. N., Hokayem, H., & Mavronikolas, E. (2019). Toward coherence in curriculum, instruction, and assessment: A review of learning progression literature. *Science Education*, *103*(5), 1206–1234. https://doi.org/10.1002/sce.21525

Kazemi, E., Ghousseini, H., Cunard, A., & Turrou, A. C. (2016). Getting inside rehearsals: Insights from teacher educators to support work on complex practice. *Journal of Teacher Education*, *67*(1), 18–31. https://doi.org/10.1177/0022487115615191

Kelley-Petersen, M., Davis, E. A., Ghousseini, H., Kloser, M., & Monte-Sano, C. (2018). Rehearsals as examples of approximation. In P. Grossman (Ed.), *Teaching core practices in teacher education* (pp. 85–105). Harvard Education Press.

Kloser, M., & Windschitl, M. (2020). Comparing pedagogies to support core practices in two secondary methods courses. In D. Stroupe, K. Hammerness, & S. McDonald (Eds.), *Preparing science teachers through practice-based teacher education* (pp. 63–80). Harvard Education Press.

Lampert, M. (2010). Learning teaching in, from, and for practice: What do we mean? *Journal of Teacher Education*, *61*(1–2), 21–34. https://doi.org/10.1177/0022487109347321

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*(3), 226–243. https://doi.org/10.1177/0022487112473837

Lauffer, H. B. (2010). SIMPL: A framework for designing and facilitating professional development to change classroom practice. *Science Teacher Education*, *59*, 13–28.

Lauffer, H. B., & Lauffer, D. (2009). Building professional development cadres. In S. Mundry & K. E. Stiles (Eds.), *Professional learning communities for science teaching: Lessons from research and practice* (pp. 55–72). NSTA Press.

Lebak, K. (2015). Unpacking the complex relationship between beliefs, practice, and change related to inquiry-based instruction of one science teacher. *Journal of Science Teacher Education*, *26*(8), 695–713. https://doi.org/10.1007/s10972-015-9445-0

Lee, O. (2021). Asset-oriented framing of science and language learning with multilingual learners. *Journal of Research in Science Teaching*, *58*(7), 1073–1079. https://doi.org/10.1002/tea.21694

Lee, O., O'Connor, C., & Haas, A. (2016). Promoting Science among English Language Learners (P-SELL) model: Curricular and professional development intervention in elementary science instruction with a focus on English Language Learners. In C. Buxton

& M. Allexsaht-Snider (Eds.), *Supporting K-12 English Language Learners in Science* (pp. 15–30). Routledge. https://doi.org/10.4324/9781315659930

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2010). *Designing professional development for teachers of science and mathematics*. Corwin Press. https://doi.org/10.4135/9781452219103

Lowell, B. R., Cherbow, K., & McNeill, K. L. (2021). Redesign or relabel? How a commercial curriculum and its implementation oversimplify key features of the NGSS. *Science Education*, *105*(1), 5–32. https://doi.org/10.1002/sce.21604

Lowell, B. R., & McNeill, K. L. (2020). Using the student hat to push on multiple goals in teacher professional learning. In *Proceedings of the 14th International Conference of the Learning Sciences*. International Society of the Learning Sciences (ISLS). https://repository.isls.org//handle/1/6522

Luft, J. A., & Hewson, P. W. (2014). Research on teacher professional development programs in science. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education: Vol. II* (pp. 889–909). Routledge.

MacLeod, G. (1995). Microteaching in teacher education. In L. W. Anderson (Ed.), *International encyclopedia of teaching and teacher education* (2nd ed., pp. 573–577). Elsevier.

Manz, E. (2018). Designing for and analyzing productive uncertainty in science investigations. In J. Kay & R. Luckin (Eds.), *Rethinking learning in the digital age: Making the learning sciences count, Proceedings of the 13th International Conference of the Learning Sciences* (Vol. 1, pp. 288–295). International Society of the Learning Sciences. https://repository.isls.org//handle/1/686

Marrongelle, K., Sztajn, P., & Smith, M. (2013). Scaling up professional development in an era of common state standards. *Journal of Teacher Education*, *64*(3), 202–211. https://doi.org/10.1177/0022487112473838

McNeill, K. L., Affolter, R., & Reiser, B. J. (in press). Anchoring science professional learning in curriculum materials enactment: Illustrating theories in practice to support teacher learning. In A. Castro Superfine, S. R. Goldman, & M.-L. M. Ko (Eds.), *Teacher learning in changing contexts: Perspectives from the learning sciences*. Routledge.

McNeill, K. L., & Berland, L. (2017). What is (or should be) scientific evidence use in k-12 classrooms? *Journal of Research in Science Teaching*, *54*(5), 672–689. https://doi.org/10.1002/tea.21381

McNeill, K. L., & Knight, A. M. (2013). Teachers' pedagogical content knowledge of scientific argumentation: The impact of professional development on K-12 teachers. *Science Education*, *97*(6), 936–972. https://doi.org/10.1002/sce.21081

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053–1075. https://doi.org/10.1002/tea.21459

Napier, J. B., Luft, J. A., & Singh, H. (2020). In the classrooms of newly hired secondary science teachers: The consequences of teaching in-field or out-of-field. *Journal of Science Teacher Education*, *31*(7), 802–820. https://doi.org/10.1080/1046560X.2020.1800195

Nasir, N. S., Rosebery, A. S., Warren, B., & Lee, C. D. (2014). Learning as a cultural process: Achieving equity through diversity. In R. K. Sawyer (Ed.), *Handbook of the Learning Sciences* (2nd ed., pp. 686–706). Cambridge University Press.

Nathan, M. J., & Petrosino, A. (2003). Expert blind spot among preservice teachers. *American Educational Research Journal*, *40*(4), 905–928. https://doi.org/10.3102/00028312040004905

National Academies of Sciences, Engineering, and Medicine. (2015). *Science teachers' learning: Enhancing opportunities, creating supportive contexts*. The National Academies Press. https://doi.org/10.17226/21836

National Academies of Sciences, Engineering, and Medicine. (2019). *Science and Engineering for Grades 6-12: Investigation and Design at the Center* (B. Moulding, N. Songer, & K. Brenner, Eds.). National Academies Press. https://doi.org/10.17226/25216

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press. https://doi.org/10.17226/13165

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, *19*(4), 317–328. https://doi.org/10.1080/0022027870190403

Next Generation Science Storylines & Next Generation Science Exemplar System. (2018). *Thinking like a kid*. https://www.openscied.org/wp-content/uploads/2021/05/Thinking_like_a_kid_handout_v5.pdf

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. National Academies Press. https://doi.org/10.17226/18290

Nipper, K., & Sztajn, P. (2008). Expanding the instructional triangle: Conceptualizing mathematics teacher development. *Journal of Mathematics Teacher Education*, *11*(4), 333–341. https://doi.org/10.1007/s10857-008-9082-z

OpenSciEd. (2021). *Curriculum launch professional learning*. OpenSciEd. https://www.openscied.org/professional-learning-materials/curriculum-launch/

Pruitt, S. L. (2014). The Next Generation Science Standards: The features and challenges. *Journal of Science Teacher Education*, *25*(2), 145–156. https://doi.org/10.1007/s10972-014-9385-0

Reiser, B. J. (2013, September). *What professional development strategies are needed for successful implementation of the Next Generation Science Standards?* Invitational Research Symposium on Science Assessments. https://www.ets.org/Media/Research/pdf/reiser.pdf

Reiser, B. J., Michaels, S., Moon, J., Bell, T., Dyer, E., Edwards, K. D., McGill, T. A. W., Novak, M., & Park, A. (2017). Scaling up three-dimensional science learning through teacher-led study groups across a state. *Journal of Teacher Education*, *68*(3), 280–298. https://doi.org/10.1177/0022487117699598

Reiser, B. J., Novak, M., & McGill, T. A. W. (2017). *Coherence from the students' perspective: Why the vision of the framework for K-12 science requires more than simply "combining" three dimensions of science learning*. The National Academies of Sciences, Engineering and Medicine, Board on Science Education. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_180270.pdf

Reiser, B. J., Novak, M., McGill, T. A. W., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal of Science Teacher Education*, *32*(7), 805–829. https://doi.org/10.1080/1046560X.2021.1884784

Robertson, A. D., Scherr, R., & Hammer, D. (Eds.). (2015). *Responsive Teaching in Science and Mathematics* (1st ed.). Routledge. https://doi.org/10.4324/9781315689302

Rosebery, A. S., Ogonowski, M., Dischino, M., & Warren, B. (2010). "The coat traps all your body heat": Heterogeneity as fundamental to learning. *Journal of the Learning Sciences*, *19*(3), 322–357. https://doi.org/10.1080/10508406.2010.491752

Rosebery, A. S., Warren, B., & Tucker-Raymond, E. (2016). Developing interpretive power in science teaching. *Journal of Research in Science Teaching*, *53*(10), 1571–1600. https://doi.org/10.1002/tea.21267

Roth, K. J., Bintz, J., Wickler, N. I. Z., Hvidsten, C., Taylor, J., Beardsley, P. M., Caine, A., & Wilson, C. D. (2017). Design principles for effective video-based professional development. *International Journal of STEM Education*, *4*, 31. https://doi.org/10.1186/s40594-017-0091-2

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*, *48*(2), 117–148. https://doi.org/10.1002/tea.20408

Roth, K. J., Wilson, C. D., Taylor, J. A., Stuhlsatz, M. A. M., & Hvidsten, C. (2019). Comparing the Effects of analysis-of-practice and content-based professional development on teacher and student outcomes in science. *American Educational Research Journal*, *56*(4), 1217–1253. https://doi.org/10.3102/0002831218814759

Sadler, T. D., & Brown, D. E. (2018). Introduction to the special issue: A critical examination of the Next Generation Science Standards. *Journal of Research in Science Teaching*, *55*(7), 903–906. https://doi.org/10.1002/tea.21508

Schwarz, C. V., Passmore, C., & Reiser, B. J. (Eds.). (2017). *Helping students make sense of the world using next generation science and engineering practices*. NSTA Press.

Sherin, M. G. (2004). New perspectives on the role of video in teacher education. In J. Brophy (Ed.), *Using video in teacher education* (Vol. 10, pp. 1–27). Elsevier.

Sherin, M. G., & van Es, E. A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education*, *13*(3), 475–491.

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, *60*(1), 20–37. https://doi.org/10.1177/0022487108328155

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. https://doi.org/10.1002/sce.21112

Stroupe, D., Hammerness, K., & McDonald, S. (Eds.). (2020). *Preparing science teachers through practice-based teacher education*. Harvard Education Press.

Taylor, J. A., Roth, K., Wilson, C. D., Stuhlsatz, M. A. M., & Tipton, E. (2017). The Effect of an analysis-of-practice, videocase-based, teacher professional development program on elementary students' science achievement. *Journal of Research on Educational Effectiveness*, *10*(2), 241–271. https://doi.org/10.1080/19345747.2016.1147628

Tekkumru-Kisa, M., & Stein, M. K. (2017a). Designing, facilitating, and scaling-up video-based professional development: Supporting complex forms of teaching in science and mathematics. *International Journal of STEM Education*, *4*, 27. https://doi.org/10.1186/s40594-017-0087-y

Tekkumru-Kisa, M., & Stein, M. K. (2017b). A framework for planning and facilitating video-based professional development. *International Journal of STEM Education*, *4*, 28. https://doi.org/10.1186/s40594-017-0086-z

Tekkumru-Kisa, M., Stein, M. K., & Coker, R. (2018). Teachers' learning to facilitate high-level student thinking: Impact of a video-based professional development. *Journal of Research in Science Teaching*, *55*(4), 479–502. https://doi.org/10.1002/tea.21427

Warren, B., Vossoughi, S., Rosebery, A. S., Bang, M., & Taylor, E. V. (2020). Multiple ways of knowing: Re-imagining disciplinary learning. In N. S. Nasir, C. D. Lee, R. D. Pea, & M. McKinney de Royston (Eds.), *Handbook of the cultural foundations of learning* (1st ed., pp. 277–293). Routledge.

Wilson, S. M. (2013). Professional development for science teachers. *Science*, *340*(6130), 310–313. https://doi.org/10.1126/science.1230725

Windschitl, M., Thompson, J., & Braaten, M. (2018). *Ambitious science teaching*. Harvard Education Press.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (REL 2007-No. 033; Issues & Answers). US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=70

Zivic, A., Smith, J. F., Reiser, B. J., Edwards, K. D., Novak, M., & McGill, T. A. W. (2018). *Negotiating epistemic agency and target learning goals: Supporting coherence from the students' perspective*. International Conference of the Learning Sciences, London, U.K.

<u>Section IV:</u> <u>Paper #3 - The Student Hat in Professional Development: Building Epistemic Empathy to Support Teacher Learning</u>

Recent approaches to the design of learning environments have emphasized the importance of ensuring students' epistemic agency, meaning the degree of control students have over the knowledge building work of the class (Ko & Krist, 2019; Miller et al., 2018; Stroupe, 2014). Knowledge building involves students engaging in a joint construction of understanding of a particular topic, phenomenon, or question (Scardamalia, 2002; Scardamalia & Bereiter, 2006). This is complex and difficult work, and students that have control over their approach to the work are able to more successfully create shared understandings, which leads to more effective long-term learning than traditional approaches (Chuy et al., 2010; Damşa et al., 2010).

As we continue to design for epistemic agency in the classroom, we need to understand how we can support teachers in supporting that agency for students. Traditionally, learning sciences research has focused on student learning rather than teacher learning (Fishman et al., 2014). This is a problem because teachers have shouldered the primary burden of understanding and implementing new approaches to instruction in the current era of standards-based reform (Cuban, 1990; Marrongelle et al., 2013). Furthermore, teachers can implement instructional reforms in ways that undercut or contradict the intention of those reforms (Berland et al., 2016; Cohen, 1990; McNeill et al., 2017). Therefore, it is important that learning scientists and instructional designers understand how teachers learn to implement their designs and what can support that learning.

One of the most common and popular methods for supporting teacher understanding of instructional reforms is professional development (Kennedy, 2016). Recently, Short and Hirsh (2020) argued that it is important to pair professional development (PD) with high quality instructional materials to help teachers understand and implement the key features of those materials; this type of PD is called curriculum-based PD. Curriculum-based PD has the potential to support teachers' own understanding of vital instructional reforms and to increase the likelihood that they will actually implement those reforms, but only if well-designed and targeted towards the teaching practices of interest (Grigg et al., 2013). In other words, in order to support teachers to use particular instructional practices, PD should actually engage teachers in the target practices (Marrongelle et al., 2013). In order to understand how PD might do that, we need to look at the actual activities teachers are asked to do during PD and how those are structured, meaning what participants are asked to do, for what reason, and with what resources. Therefore, this paper looks at one teacher PD session for middle-school science teachers to come to understand what and how it helped teachers to learn about curriculum materials designed to support students' epistemic agency.

Background

This study focuses on how middle school science teachers responded to a curriculum-based PD session designed to support them in implementing new curricular materials designed to support students' epistemic agency. Therefore, I begin by discussing epistemic agency and how recent work in science education has addressed that idea. The particular curricular materials these teachers were learning to implement are called "storyline curricula," so I then describe in more detail what storyline curricula are and how they help address this move towards supporting students' epistemic agency. Next, I turn to PD and outline how curriculum-based PD can help teachers' learning.

Shifting to Support Epistemic Agency

Epistemic agency is a quality of the learning environment defined as "students being positioned with, perceiving, and acting on, opportunities to shape the knowledge building work in their classroom community" (Miller et al., 2018, p. 1058). The idea that students might shape the work being done in the class is in stark contrast with traditional approaches to education in which the teacher or textbook is the primary agent determining which ideas the class will address, when, and how (Hutchison & Hammer, 2010). It is difficult to design learning environments that truly support students' epistemic agency, however. Even when environments are designed to support epistemic agency, they can result in "pseudoagency" in which students are given opportunities to shape knowledge building work, but only the ideas that have already been planned by the teacher or curriculum are taken up (Cherbow & McNeill, 2022; Miller et al., 2018). Often, this pseudoagency happens in response to external accountability requirements determining what ideas students must construct in limited time (Miller et al., 2018).

The literature is still unclear on how to address this dilemma of the balance between true agency and external accountability. One approach that has shown some promise, however, is the idea of being responsive to students' ideas, questions, and experiences and using those to guide classroom work (Robertson et al., 2015; Rosebery et al., 2010). By treating students' ideas as resources to build upon rather than misconceptions to be "fixed," teachers can support students' conceptual understandings and willingness to participate in class, which is a key component of epistemic agency (Furberg & Silseth, 2021). Responsiveness to student ideas as a way to support epistemic agency has taken various forms across the disciplines. For example, in mathematics education, a robust body of literature on noticing students' mathematical thinking emphasizes the importance of understanding students' approaches to a mathematical situation and building from their thinking to develop stronger mathematical skills (Colestock & Sherin, 2015; Sherin et al., 2011; Walkoe, 2015). In history, recent work has shown that when students are positioned as knowledgeable authorities making sense of a compelling problem, they are more likely to engage in productive historical discourse (Freedman, 2020).

In science, efforts to support epistemic agency have focused on engaging students in the process of figuring things out about the natural world (Crawford, 2014; Schwarz et al., 2017). This vision of science education asks teachers and curriculum to use students' own ideas about the natural world as resources to support their engagement in the practices of science as they deepen their explanations of natural phenomena, rather than simply memorize facts presented to them (Bang et al., 2012, 2017; Rosebery et al., 2010). Achieving this vision means changing classroom relationships among student thinking, classroom investigations, and teacher explanation. In traditional models of instruction, a teacher might explain an idea to students and then ask students to carry out an investigation as a way to demonstrate that idea and support engagement through "hands-on" learning (Furtak & Penuel, 2019). Current approaches, in contrast, advocate that students engage with a natural phenomenon that anchors a unit of instruction by inspiring students' questions and motivating a series of investigations, which can then be explained or modeled based on evidence collected over a series of lessons (Windschitl et al., 2018).

In order to do the work of supporting students' epistemic agency, therefore, teachers need to shift the ways they support students' engagement in both investigations of natural phenomena and classroom discussions (Lowell et al., 2022; Manz et al., 2020). These changes are challenging, however, and teachers attempting to implement them may continue to teach in ways that unintentionally reinforce traditional models of science learning that focus on memorizing facts (Berland et al., 2016; McNeill et al., 2017). One way to help teachers avoid this problem is to provide curricular materials that are explicitly designed to support students in learning science by figuring out natural phenomena; unfortunately, there are few of these curricula currently available (Achieve, 2018; Lowell et al., 2021). One relatively recent curricular innovation, designed to support science teachers in making the shifts described above, is storyline curricula (Reiser et al., 2021).

Storyline Curricula

Storyline curricula are materials designed to support students in engaging in science as a process of figuring out rather than simply learning about science ideas (Reiser et al., 2021). Each unit is designed around an anchoring phenomenon (Windschitl et al., 2018), which begins with students observing a phenomenon and then constructing a driving question board to collect their questions about the phenomenon. The driving question board helps to motivate the need for various investigations throughout the unit (Reiser, Brody, et al., 2017).

A key feature of storyline curricula is that they are designed to be coherent from the student perspective (Reiser et al., 2021; Zivic et al., 2018). This means that investigations are designed in a way that is likely to make sense to students as they are developing their understanding of particular science ideas rather than organized in a way that makes sense to somebody who already has a strong command of the discipline (Reiser, Novak, et al., 2017). For example, the structure of cell membranes is often taught as part of a unit on cell structure, but for students, learning about the cell membrane as a generic "part" of the cell provides no motivation to understand why the specialized structures of the membrane are important. (National Academies of Sciences, Engineering, and Medicine, 2019). If students first establish that cells must take in food and eliminate waste, this could lead to them wondering how cells do that without letting anything enter or leave, which would then motivate the need to understand cell membrane structure. The idea here is that the learning activities are designed to address students' questions as they are likely to come up, rather than having students learn ideas simply because the teacher tells them those ideas are important (Reiser, Novak, et al., 2017). Therefore, as students engage in investigations and make sense of them, they jointly construct a "story" about the natural phenomenon anchoring the unit.

This approach to curricular design has been shown to support epistemic agency in that it elicits students' questions and then addresses many of those questions over the course of the unit (Zivic et al., 2018). Although storyline curricula are science specific, they are similar to work in other disciplines. For example, a history curriculum focused on establishing compelling problems for students to discuss has been shown to better support students' productive disciplinary engagement than one that simply recounted historical facts and had students analyze them (Freedman, 2020).

In order to effectively implement storyline curricula, teachers need to have a strong understanding of the science ideas involved and how students might think about them so that teachers can make in-the-moment decisions that support rather than undercut the goal of coherence from the student perspective (Cherbow, 2021; McNeill et al., 2017). This kind of understanding is challenging to develop, however, particularly given that storyline curricula are new and quite different from traditional science curricula (McNeill & Reiser, 2018). One way to support teachers in this work, therefore, is to pair the curriculum materials with professional development focused on effectively implementing those materials (Lynch et al., 2019; Reiser, Michaels, et al., 2017; Short & Hirsh, 2020).

Curriculum-Based Professional Development

Traditional models of PD have included short workshops on general, content-neutral instructional strategies or concerns, such as classroom management, but have shown limited success in actually changing teachers' practice (Kennedy, 2016; Stein et al., 1999). In order to address this problem, a plethora of research has been done on features of PD that make it effective, with one of the most popular being a framework of five key features proposed by Garet and colleagues (2001): content focus, active learning, coherence, duration, and collective participation.

One approach to designing content-focused PD has been to pair the PD with the specific curricular materials teachers are preparing to implement, such that teachers learn both the instructional approach of the target curricular materials and how to implement them during the PD sessions (McNeill & Reiser, 2018; Short & Hirsh, 2020). A strength of this approach is that it is explicitly linked to classroom practice, which can help teachers connect broader instructional reforms with the specific moves and choices they make in the classroom (Ball & Cohen, 1999; Short & Hirsh, 2020). This type of PD, called curriculum-based PD, has been shown to have positive impacts on both teachers' learning and the learning of students whose teachers have engaged in curriculum-based PD (Lynch et al., 2019; Taylor et al., 2015). While attending to these structural features of PD are helpful in design, they do not address what *actual activities* take place during the PD session and how those activities support teacher learning. In order to better understand this level of design, therefore, we need to take a closer look at specific activity structures, how they are used in PD, and what work those structures do to support teacher learning.

A number of activities focused on situating teacher learning in practice are often incorporated in curriculum-based professional development, including watching classroom video, rehearsing teaching, analyzing student work, or doing activities from the discipline such as science investigations or math problems (Gibbons & Cobb, 2017). These activities are valuable because they provide representations and approximations of teachers' practice, helping them to see how some of the ideas or approaches in PD might play out in actual classrooms and practice that work in a low-stakes environment (Ball & Cohen, 1999; Grossman, 2018). One such activity that has received relatively little attention in the literature is the "student hat." During student hat activities, teachers take the role of students experiencing the lessons they will be teaching and are asked to use ideas and experiences their students might have to make sense of the lessons and to share those ideas using language the students might use (McNeill et al., in press; Next Generation Science Storylines & Next Generation Science Exemplar System, 2018). The goal of this kind of activity is to emphasize how ideas build over time from the students' novice perspective rather than from a disciplinary expert's perspective (Edelson et al., 2021; Reiser et al., 2021).

The Student Hat in Teacher Professional Development

The student hat activity structure was conceptualized by Reiser and colleagues (2017) in the Next Generation Science Exemplar project as a way to support teachers to consider how storyline units can support idea building in a way that is coherent from the student perspective. Student hat is different from other similar activities like doing disciplinary activities (e.g. Borko et al., 2005) or rehearsal (e.g. Lampert et al., 2013) because it asks teachers to attend to the way students' thinking might build from their own experiences rather than on developing their teacher content knowledge (in the case of disciplinary activities) or on teacher moves (in the case of rehearsal). Unlike other common PD activities, it is also designed to support teachers to experience the cognitive and affective responses students might have to storyline curricula that are designed to support students' epistemic agency.

Other examples of student hat exist, but there has been relatively little research into this PD structure. For example, Rosebery and colleagues (2016) reported on a PD they developed for early-career science teachers to support the teachers' ability to recognize students' diverse sensemaking practices as generative for science learning and to consider pedagogical practices that encourage and use those practices and ideas in classroom work. A key component of this PD was asking teachers to engage in learning about plant life cycles in order "to experience what it was like to problematize, complicate, question, and trouble taken-for-granted meanings in the ways that students often do" (Rosebery et al., 2016, p. 1577). Rosebery and colleagues found that after the PD, teachers were better attuned to the intellectual power in students thinking, even if it did not reflect traditional scientific thinking and had expanded their conception of what it meant to be "smart" in a science classroom. This work, therefore, supports the idea that student hat can help teachers to consider how students' ideas and experiences are (or are not) valued by instructional practice and what can be done to better support the epistemic agency of all students, especially those who have been traditionally marginalized by science.

In her study on an outdoor teaching PD for science teachers, Glackin (2019) engaged participants in "simulated modelling" of outdoor activities so that they could experience what it felt like from the student perspective to learn science outdoors. Based on interviews and surveys with her participants, Glackin found that these activities could increase teachers' empathy for what students experience in the classroom, demonstrating that student hat can help teachers to consider students' affective response to instruction.

These studies provide some initial evidence that student hat can have positive impacts on teacher learning in terms of understanding how students might experience reform-oriented instruction both cognitively and affectively. Nevertheless, the literature is fairly thin on both what specific teacher-learning goals the student hat supports during PD and how it does so. Therefore, this study asked: What does the student hat activity structure help teachers learn during curriculum-based professional development? How does it support that learning?

Theoretical Framework

Existing studies of the student hat in PD have so far taken a mostly high-level approach to investigating its impact on teacher learning, using pre-post assessments and/or overall interviews to determine changes in teacher learning after participating in PD that used student hat activities (Glackin, 2019; Rosebery et al., 2016). These are helpful approaches to understanding *what* teachers might have learned during the PD, but less useful in highlighting *how* they did so. In order to better understand the mechanism of learning, we need to look specifically at the learning as it takes place (Walkoe & Luna, 2020). In order to identify moments that likely contributed to teachers' learning, I adopt the framework of productive struggle, which posits that one of the times learning takes place is during moments that learners struggle with the target learning outcome (Hiebert & Grouws, 2007; Warshauer, 2015).

The idea of productive struggle is particularly popular in mathematics education and refers to the process by which students "expend effort to make sense of mathematics, to figure something out that is not immediately apparent" (Hiebert & Grouws, 2007, p. 387). The key piece here is that students are engaged in some kind of problem solving and struggling to find the solution. What makes struggles productive is that they are challenging but doable and can advance students' thinking about a topic (Kapur & Bielaczyc, 2012; Sengupta-Irving & Agarwal, 2017; Warshauer, 2015). Productive struggle as a key lever for supporting student learning has become so influential in mathematics education that it has made it into recommendation reports from the National Council of Teachers of Mathematics (2014), where the "problems" students are trying to solve are literally math problems. As other disciplines have taken up this framework, however, they have expanded the conception of what can be "problem solving" for students to productively struggle with. For example, in science Chen (2021) conceived of the "problem" as modeling a natural phenomenon. He performed a case study of how one 5th grade class's uncertainty around modeling stimulated a productive struggle that helped them to make sense of and represent how the human respiratory, muscular, and circulatory systems work together to help people breathe.

While popular as a construct for looking at student learning, productive struggle has not been used as much to look at *teacher* learning. In their study of middle grades teachers engaging in curriculum design, Trinter and Hughes (2021) considered curriculum design as a problem-solving process in which teachers were engaging in productive struggle. Through analyzing curriculum planning sessions and interviews, Trinter and Hughes found that their participants' struggle to design interdisciplinary curriculum helped them to rethink the goal of curriculum design, moving away from a vision of teachers as adopters of expert curriculum towards one in which they improvise from curricular materials to meet the needs of their students. In struggling to design curriculum, therefore, they increased their capacity to act as strong designers for their own students. Manz and Suárez (2018) took up a construct similar to productive struggle when they looked at teachers' pedagogical uncertainty during PD sessions. They worked with a group of elementary science teachers as the teachers designed lessons that engaged students with investigating the natural world. In so doing, the teachers came across moments in which neither the teacher or students knew the canonically correct answer to a scientific question, which inspired pedagogical uncertainty in teachers who were used to a vision of teaching as giving the students correct answers. Manz and Suárez showed that by embracing this pedagogical uncertainty, they were able to help the teachers to consider how to support students in making sense of natural phenomena in ways better aligned with how scientists do so. Their idea of "pedagogical uncertainty" connects to productive struggle in that it shows how the moments in which teachers struggled in their practice helped them push their own pedagogical thinking forward.

Based on these ideas of productive struggle as a site for learning, I apply the framework to my study here. In my case, I consider the work of considering and embodying students' ideas as the "problem-solving" that teachers were being asked to do. In looking at moments in which they struggled to effectively consider or communicate those ideas in student language, I hope to investigate what teachers were learning and how.

Methods

This article reports on a case study focused of one two-and-a-half-day curriculum-based PD session for middle school science teachers implementing a new storyline curriculum unit from OpenSciEd. Case study is an appropriate approach to develop deep understanding about a specific, bounded event or system (Stake, 2005). Here, the PD session is regarded as a case of the use of the student hat in teacher PD, and analysis of it can help to shed light on the broader phenomenon of how student hat activities support teacher learning. This is an instrumental case study because I aim to develop a deep understanding of this particular case in order to provide insight into the more general phenomenon of teacher learning (Stake, 2005). I use an embedded single-case design for this case study, which is a study of one case that has multiple subunits embedded within it that can be compared and contrasted in order to develop deep insights into the case as a whole (Yin, 2018). Here, the PD session is the broader case within which were embedded multiple moments of productive struggle, which serve as subunits to be compared.

Case studies take place in a particular context, so I begin by describing the context of the curriculum-based PD, the participants, and their background, followed by the data sources and analysis procedures.

Curriculum-Based Professional Development Context

OpenSciEd is a national non-profit organization designing storyline science curricula to be released as open educational resources. A consortium of developers consisting of groups from BSCS Science Learning, Northwestern University, Boston College, The Dana Center at UT Austin, and University of Colorado Boulder began developing middle school (grades 6-8) curriculum and associated PD in January of 2018. Each OpenSciEd unit undergoes a multistep development process that includes initial writing, field testing in 10 states, external review, and revision before release to the public. Each unit also is accompanied by a 2-4 day PD session that is field tested, revised, and then released (Edelson et al., 2021). As a member of the Boston College team, I participated in the design, facilitation, and revision of the field test PD for all of the middle school units.

During the field test, units and their accompanying PD were released every six months beginning in June of 2018 until January of 2021. Field test teachers agreed to attend each PD session, implement the accompanying unit, and provide feedback to the design team. This study focuses on one field test PD session that took place during summer 2019. The session, which was 2.5 days long, was the third OpenSciED PD and focused on helping teachers to support students' epistemic agency through discussion in their OpenSciEd classroom. The session included teachers from all three grade levels and therefore had some activities that took place with all of the participants and some that took place in grade-specific groups as teachers learned more about the OpenSciEd unit they would implement that fall. This PD session, like all of the OpenSciED

PD sessions, was designed based on a framework that emphasized facilitating teacher learning by offering the student perspective, providing images of classroom instruction and contrasting curricular cases, and encouraging cycles of teacher enactment and reflection (Edelson et al., 2021; McNeill et al., in press). Consistent with these principles, the PD session asked teachers to reflect on their current practice, watch classroom video, and consider written scenarios of instruction. In their unit-specific groups, the teachers constructed a summary of the lessons in their unit, engaged in a few key lessons in student hat, and planned for and rehearsed a sensemaking discussion. Table IV.1 outlines the activities that took place during each day of the PD.

Table IV.1: Summary of OpenSciEd Round 3 PD Activities

	Day 1 (half day)	**Day 2**	**Day 3**
Whole-group (across grade levels)	Analyze video & written scenarios for students' diverse ideas and resources (1 hour)	Examine resources and watch classroom video considering how to support equitable classroom culture. (2 hours)	Read discussion planning tool, watch video of a teacher planning & implementing a discussion. (2 hours)
Unit-specific (single grade level)	Experience anchoring phenomenon of target unit as a student (2 hours)	Construct summary of unit storyline, experience key lessons as a student (5 hours)	Plan and rehearse a classroom discussion, experience key lessons as a student (5 hours)

Because all of the student hat moments occurred during the unit specific time, this study focuses on those portions of the PD. The three units addressed in this PD session were titled[1]: *Why do we sometimes see different things when looking at the same object? [Light & matter]* (OpenSciEd, 2021), *Why do things sometimes get damaged when they hit each other? [Contact*

[1] Each OpenSciEd unit has two titles: a long title in the form of a question and a nickname title. Here, I include the nickname title in brackets after each long title. For the rest of the paper, I refer to units by their nickname.

forces] (OpenSciEd, 2020b), and *How can we make something new that was not there before? [Chemical reactions & matter]* (OpenSciEd, 2020a). The sixth-grade teachers engaged with the light & matter unit, which was based around figuring out how two people looking at different sides of a one-way mirror see different things. The seventh-grade teachers worked with the contact forces unit, which was based around explaining what happens when you drop your cell phone and it gets damaged. Finally, the eighth-grade teachers engaged with the chemical reactions & matter unit, which had them figuring out how a bath bomb can make bubbles, color, and smell when added to water.

Participants

The PD session included 32 participants; 29 were middle school science teachers and the other three were a district-level science director, school-level instructional coach, and staff member from a local science museum. All of the participants taught in the same state in either the largest city in the state or a nearby smaller city. Among the participants, 21 (66%) identified as women and 11 (34%) identified as men, which is roughly similar to the overall population of middle school science teachers in the United States, which is 71% women and 28% men (Banilower et al., 2018). This PD session was the third OpenSciEd field test PD session, and this particular state had chosen to increase the number of teachers participating in the field test after the first year. Therefore, 13 of the teachers had never taught an OpenSciEd unit before and 16 had. All of the new teachers had attended a four-day introductory PD session earlier in the same summer, so all 32 participants had had some previous experience working in the student hat. At the end of the PD session, 12 teachers agreed to be interviewed about their experience participating in the student hat activities. The background of each participant is given in Appendix A.

Data Sources

The data from this study came from two main sources: video recordings of the PD sessions and interviews with the interview participants. All of the PD session across two and a half days was recorded. During the whole group sessions, multiple cameras were set up to capture both the entire room and a few small group discussions. During the unit specific sessions, each grade level had one camera that focused on the entire group. The recordings included five hours of whole group time and 12 hours of unit specific time for each of the three units, totaling 36 hours of unit specific video. As discussed above, because the whole group time did not involve any student hat activities, this study used only the 36 hours of unit specific video.

Between one and two months after the PD, each of the 12 volunteers participated in a semi-structured interview in which the questions and topics were pre-determined but the interviewer asked follow up questions. This interview approach allows interviewers to probe into participants' understanding of a particular idea or phenomenon (Rossman & Rallis, 2017), making it suitable for this study of the teachers' experience in the student hat. The interviews, which lasted 30-60 minutes, covered a range of topics, including general feedback about the PD sessions, the classroom video used, the student hat activity structure, and the focus on the ideas students bring to the classroom. Table IV.2 includes the relevant questions asked about the student hat (see Appendix B for the complete interview protocol).

Table IV.2: Questions from Interview Protocol

This PD asked you to think like a student or be in "student hat" for a significant portion of the training. How does this compare to other non-OpenSciEd PD experiences you have had? *Follow up if they have done it before* - What types of things did you do in the student hat in previous PD experiences? How was that similar or different to what we did in the OpenSciEd PD? *Follow-up follow-up if they don't mention discussions* - In OpenSciEd we do sensemaking discussions in the student hat as well as investigations. In what ways is that similar to or different from your other PD experiences using student hat?

What were some benefits of thinking in the student hat?
What were some challenges of thinking in the student hat?

Data Analyses

Data analysis began with the transcription of each interview. I built descriptive and in vivo codes (Miles et al., 2020) from what the participants said about the potential learning outcomes of using the student hat. This took place in multiple rounds: a preliminary round to identify potential codes, a second round to apply those codes to the interview data, and a third round to determine the interrater reliability of those codes and modify them as necessary. After the first round, I identified three main codes from the interview data. After coding the entire corpus of interview data, I recruited a second rater to review that codebook and apply it independently to three of the 12 interviews, representing 25% of the data. We coded each turn, defined as one section of uninterrupted speech from the interviewee. We applied each code a maximum of one time per turn if that idea appeared and permitted turns to have multiple codes. We then met to share our codes, calculate interrater reliability by percent agreement of codes, and modified the codebook based on our understandings. We went through this process twice until we achieved an interrater reliability of 90%. At this point, I re-coded the remaining interviews based on the revised codebook, which is shown in Table IV.3.

Table IV.3: Final Codebook for Interview Analysis

Code	Definition	Example
Empathy for students' experiences	Teachers discussed empathizing with how students might experience an OpenSciEd lesson	"Now I got to step back and pretend I'm that kid who I have the most trouble connecting with in a class, and see what that student is going to be reacting to. And just that's an empathy building thing." (Ms. Newcomb)
The instructional approach	Teachers discussed how the student hat helped them learn about the OpenSciEd instructional approach	"I think you can see where they're going to get stuck and where the frustration is going to build in at the beginning. And then you're hopefully also going to be able to see times where they, or you as a student, or they as a student, can string a few consecutive ideas together. And so that is helpful because that's the kind of bridges that you want to build for your own students. And if you've experienced it yourself, I think there's a higher chance that you will see it happening out in the class when it's happening." (Ms. Jennings)
Challenges	Teachers discussed how or why the student hat was challenging or difficult	"Sometimes I think that some people go down too far. They go down to like, " I don't know, is that like a bond?" We go down to eighth grade, they have some knowledge of some things and then there are some people that just can't get down to the student level. Their thinking just can't get them down there." (Ms. Townsend)

The interview data were used to develop and support initial claims about how the student hat impacted teacher learning. The video data of teachers engaging in the student hat were then used to triangulate those claims across multiple units of analysis and forms of data to improve their explanatory power (Yin, 2018). I began the video analysis process by watching all 36 hours of unit-specific PD video. During the first pass, I identified times when the teachers were engaging in student hat activities and their speaking was intelligible on the video. Because there was only one microphone in each room, the teachers were rarely audible when speaking in small groups. Based on this first pass, I reduced the total amount of video to eight hours (three hours from light & matter, and 2.5 each from contact forces and chemical reactions & matter). In order to facilitate my second round of video analysis, I transcribed what participants said in the remaining video. I then re-watched the video while following along with the transcript looking for moments in which teachers struggled to think like a student. As I watched, I clustered (Miles et al., 2020) these moments based on the kinds of things teachers were struggling with and the ideas I saw teachers developing during or immediately after that struggle. These clusters allowed me to construct themes describing the ways in which student hat was helping to support teacher learning (Creswell & Poth, 2018). Finally, I compared the themes I had constructed from the video data with the initial claims from the interview data to find areas of agreement or disagreement in order to create finalized themes describing how the student hat activity supported teacher learning (Miles et al., 2020).

Results

My analysis led to four themes. The first theme, based only on interview data, was that teachers found the student hat *difficult.* The next three themes were based on video analysis triangulated with interview data that the student hat helped teachers to learn about. Those themes were that student hat helped teachers to learn *science content ideas*, it helped teachers to learn about *their students,* and it helped teachers to learn about the OpenSciEd *instructional approach*.

Difficulty: "We Know Too Much and Yet Not Enough"

All twelve of the teachers interviewed noted that they found student hat activities difficult or challenging. In order to better understand what they found challenging, I took the interview segments coded as "challenges" and sorted them into subcodes based on what specifically teachers found difficult about the student hat, as shown in Table IV.4.

Table IV.4: Challenges Subcodes

Difficulty	Interview N (total N = 12)	Example
Teachers struggled to separate the teacher's knowledge and what students might know	8 = 66.67%	"If I have background knowledge, right or wrong, in the content area, as we almost always do, then it's really hard to push that aside and say, 'Okay, I don't know that the answer to this is this,' so how am I going to go through the activity with fidelity knowing what I already know in the back of my head?" (Mr. Bernard)
Teachers felt like the activity was fake or inauthentic to how things might happen in the classroom	6 = 50%	"I also think it's sometimes just hard to suspend your disbelief for certain periods of time. So certain periods of time you're just—the adults in the room will look at you and be like, 'What are we doing?'" (Mr. Emerson)
The student hat limited teachers' ability to think about effective teacher moves to facilitate the lesson	3 = 25%	"So you can feel that it's different, but you're not necessarily going to know what the moves are that made it happen, or understand the value of those moves, so you've got to be able to switch back and forth." (Mr. Kimball)
The student hat takes much longer than other PD activities	2 = 16.67%	"It takes a lot more time because you're doing the experience, the learning itself and that takes time." (Mr. Richter)
Participants focused on acting like an unruly student rather than considering what a student might think	1 = 8.33%	"That ended up being a lot of times also teachers behaving, not thinking, you're the eighth graders, so now all of a sudden they're fooling around, this, that, the other thing." (Mr. Morse)
The student hat was so new and different from other PD activities teachers had done	1 = 8.33%	"To actually play the role of a student hat was totally different than any of our training that I've been to for science." (Ms. Mendell)

The most common reason the teachers found the student hat difficult was because they struggled to separate their own knowledge and experiences from the knowledge and background that students might bring into the classroom. The example in Table IV.4 exemplifies this concern, as Mr. Bernard pointed out that he found it hard to engage in activities from the curriculum without bringing his own content knowledge to the activity. Ms. Mendell also expressed that she found it difficult to set aside her own understanding of science content, but in her case the focus was on letting go of the technical scientific terminology she had been trained to use. She said:

> *One of the challenges for me for thinking in student hat was trying not to be so right and just being myself and trying to say what I felt a student would say and not have to be so technical about the choices of the words that I use.*

This comment from Ms. Mendell suggests that in other PD situations, she had felt it important to use technical vocabulary to demonstrate strong content knowledge, which might have limited her ability to consider how to support her students' learning given that students are developing their own understanding of both concepts and terms.

A common pattern across the teachers' discussion of this difficulty was balancing what they knew about the target science ideas with what they expected students to know about those ideas. Many teachers discovered they did not know as much about their students' knowledge as they initially thought. In the below excerpt, Ms. Jackson summarized that tension:

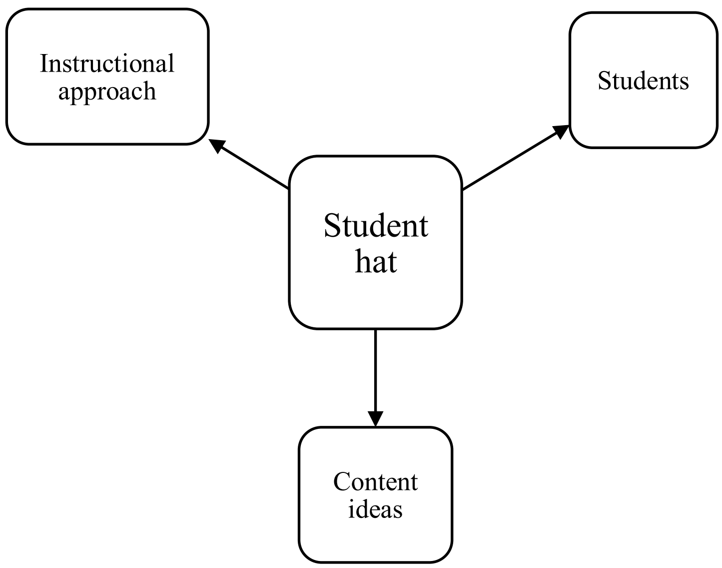
> *I've been teaching older grades, so I'm not as familiar with what students should know before this. So there were some times where teachers are using, seems to me like they're using their teacher brain to bring in the teacher knowledge and I'm like wait, hold on, we're doing a student discussion. Would an 11-year-old really know this idea that you're bringing up, or be comfortable with the vocabulary? I think it's really difficult as teachers to model a student discussion, because we know too much, and yet not enough.*

The idea of knowing "too much, and yet not enough" describes how teachers often brought significant content knowledge towards the teacher learning environments but realized that they needed to consider more the ideas that their students might bring to similar learning experiences.

Supporting Teacher Learning: Content, Students, Instructional Approach

Taken together, the teacher interviews show that teachers struggled to think like a student during student hat activities, which they often referred to as "breaking out of student hat" or "using my teacher hat." Nevertheless, teachers also found the student hat helpful for their learning. Based on the idea that teachers might be learning in those moments of struggle (Trinter & Hughes, 2021; Warshauer, 2015), I looked at moments during the PD session when teachers clearly struggled to think like a student, resulting in three themes around how the student hat activities supported teacher learning. I found that the student hat supported teachers to learn science content ideas, about their students, and about the OpenSciEd instructional approach. Figure IV.1 summarizes these three outcomes.

Figure IV.1: Student Hat Teacher Learning Outcomes



The student hat helped some teachers to develop their own content understanding. In other moments, the student hat was supporting teachers to learn more about their students, specifically building empathy with how students might think, speak, or feel while experiencing OpenSciEd. The student hat also helped teachers to learn about the OpenSciEd instructional approach, meaning both the goals for student learning and teacher student interactions built into the curricular materials and methods and logistics for implementing those materials.

Using Student Hat to Learn Content Ideas

Past work has shown that asking teachers to "engage in the discipline" can help them to better understand the disciplinary ideas they will be teaching to students (Gibbons & Cobb, 2017). In line with this idea, I found that one teacher, Mr. Richter, said that the student hat helped him to learn content ideas:

> *It actually stops and makes you think for yourself as a scientist. Like how does this actually work. It makes you be able to explain that and in other situations you never go through that process and so it's like, yeah, yeah, yeah, we understand that. It's cool. We understand. We're on the same page. But then you get to the deeper level of like, okay actually how do you explain that. How can you tell someone else about it or how can we figure this out together actually makes all the difference. And also it levels the playing field a lot as a new teacher to be on the same page as everyone else that's experiencing something new as well instead of everyone else being like, okay yeah we already figured that out as a veteran teachers.*

Here, Mr. Richter pointed out that often teachers are likely to claim that they understand content they are expected to teach, but that being forced to engage in the learning experiences and consider what students are being asked to do could actually force them to develop a deeper understanding of that content. While Mr. Richter was the only teacher who explicitly discussed this learning takeaway of the student hat, there were multiple moments in which teachers broke out of student hat and the subsequent discussion made clear that they did so because they were wrestling with the content ideas themselves.

One example of this took place in the chemical reactions group during the second day of the PD. The chemical reactions unit is focused on explaining how a bath bomb works, so an early lesson involves observing the ingredients in bath bombs to try to determine what properties those ingredients have and then to begin to sort those properties based on how helpful they are in identifying substances.

Canonically, middle school science focuses on organizing properties as chemical or physical. Chemical properties are characteristics of substances associated with changes of those substances such as flammability and reactivity. Note that chemical properties of a substance can only be observed while that substance is undergoing a chemical change; simply observing a substance on its own cannot reveal its chemical properties. Physical properties, on the other hand, are characteristics that are not associated with changes such as color or mass. Because they are not associated with a chemical change, they can be determined by observing a substance on its own. Physical properties can be further subdivided into extensive and intensive properties. Extensive properties depend on the amount of substance present, for example length or volume. Intensive properties are the same no matter how much of the substance is present, for example density or boiling point.

The goal of the lesson was to focus on this latter difference, between extensive and intensive, by building on students' experiences with the properties rather than simply introducing vocabulary. Because the investigation involved looking at materials on their own, the only possible properties they could have observed were physical. Chemical properties would not be introduced until later lessons. During the discussion, however, teachers demonstrated that they were confused about the relationship between chemical, physical, extensive, and intensive properties, which led them to break out of student hat.

Table IV.5 shows this moment of struggle that led Ms. Jasper to stop thinking like a student. In turns 1 through 12, teachers were discussing whether size and texture are properties that are characteristic of a substance no matter the amount of the substance. In turn 13, the facilitator summarized the group's thinking, but Ms. Jasper interrupted in turn 16 to ask a question using her teacher knowledge. She highlighted her concern with their working definition of properties because it did not align with her understanding of properties being characterized as physical or chemical, even though all of the properties under discussion so far were physical.

Table IV.5: Excerpt from Chemical Reactions Discussion Part I

	Speaker	Transcript
1	Ms. Newcomb:	So, is size a property or not?
2	Ms. Vernon:	No.
3	Facilitator:	I'm hearing Ms. Vernon say it's not because it's still water, no matter how much it has.
4	Ms. Newcomb:	Okay, well then how about we talk about the fact that corn starch and baking soda, the pieces look really different than Epsom salt, salt, table salt or sugar?
5	Ms. Kent:	Texture.
6	Facilitator:	Texture is our word for, uh a property of, uh, a word that we could use to describe what those little particles sort of feel like? Because you were saying different shapes of those little grains and what they feel like, because the corn starch feels different than the sugar, feels really grainy.
7	Ms. Vernon:	Yeah but you could take a hammer to the sugar and turn it into really fine powder-
8	Facilitator:	-and now the texture’s changed. So maybe it's not as useful. It is a physical-
9	Ms. Vernon:	Oh, powdered sugar. But that's a mixture, isn't it? Is it?
10	Multiple:	No.
11	Mr. Kimball:	My eighth-grade brain thinks it's just messed up sugar.
12	Ms. Townsend:	Badly processed sugar.
13	Facilitator:	So eighth graders, I'm hearing that maybe texture might be a useful property. But it might not be, just kind of like color, it's not as helpful at distinguishing between different substances. And it sounds like we're going to need some more properties to be able to start distinguish between these because if we're going to try to figure out what the heck that gas is, we're going to need to have some clues to help us figure which thing it could possibly be. Right. We're just starting to build this list of properties and we're going to see if we can add to it. Can somebody write down a question on a post-it note is the amount of stuff, is the size or the volume of the thing of property? Somebody write that down on a post-it note so that we can stick it on our DQB [driving question board].
14	Ms. Vernon:	Is anybody doing that? I'll do it if no one else is.
15	Facilitator:	Okay, thank you. All right, eighth graders in our last five minutes before we break, what do you think?

16	Ms. Jasper:	Can I ask a teacher question? When we define property as characteristic of a substance that doesn't change. I think that's really confusing between physical and chemical properties. And do you say that now? Because we just talked about all these properties that can change. Do you know what I mean? So like are we saying they're not properties?
17	Facilitator:	It doesn't change, even if you change the amount of substance you have. So water is water is water no matter how much you have.

Immediately after the facilitator's sentence in turn 17, the participants began many simultaneous small group conversations in which they were discussing their understanding of what a property is. After giving them some time to speak in small groups, the facilitator called them back and they continued to discuss this content as teachers. Near the end of the discussion, the teachers were able to move away from chemical vs. physical and towards the idea of extensive vs. intensive as shown in the exchange in Table IV.6, which began with the facilitator re-orienting the group towards the concept of intensive vs. extensive. In turn 48, Mr. Truett introduced the term "intensive property" and tried to connect it to the physical vs. chemical organization scheme by saying that an intensive physical property has the "weight of a chemical property." In turns 52 and 54, the facilitator seemed to pick up on the confusion in the room by pointing out that the goal was not to look at physical and chemical properties at all in this discussion.

Table IV.6: Excerpt from Chemical Reactions Discussion Part II

	Speaker	**Transcript**
41	Facilitator:	I actually think that you need to talk about, doesn't change no matter how much you have when you change. When you change the quantity, the volume of it, it's still that substance.
42	Ms. Kent:	And that's something they do cover in elementary school, if you have a little droplet of water versus a bucket of water, it's still water.
43	Facilitator:	It's still water.
44	Ms. Kent:	That's a concept that they come across.
45	Facilitator:	Yeah, that we're helping to develop. Mr. Truett?

46	Mr. Truett:	We're teacher hats now?
47	Facilitator:	Yeah, we're teacher.
48	Mr. Truett:	Okay. Melting point, boiling point, they are physical properties because you don't change the chemical nature of the substance, but they are called intensive physical properties. So they are characteristic of that substance and not change regardless of quantity.
49	Facilitator:	[To Mr. Morse] That's not different than what you're saying, I don't think is it?
50	Mr. Morse:	No.
51	Mr. Truett:	So an intensive physical property has the weight of a chemical property because-
52	Facilitator:	Oh, we're not distinguishing between chemical and physical properties with the students. We're just trying to say that these are characteristic of the substance and it helps us identify-
53	Mr. Truett:	I don't know where you distinguish between chemical and physical change in this unit or your units. I know where it happens in [my district].
54	Facilitator:	Well, we'd start to distinguish between it in the second lesson set. Physical and chemical change, but not physical and chemical properties per se.
55	Mr. Truett:	Okay, yeah.

Overall, this moment of struggle shows that teachers came into the discussion with one organizational scheme for properties – chemical or physical – but seemed to confuse that with the different organization scheme being addressed by the lesson: intensive or extensive. The student hat activity structure led to teachers struggling with these content ideas, helping them to better consider what a “property” really is and how they can be organized.

In his comments, Mr. Richter noted that student hat “levels the playing field” between teachers because they are all asked to contribute ideas their students might have. That leveling is clear in this example, where multiple teachers were trying to explain what does and does not count as a property. Teacher PD can include relational power dynamics in which the teacher perceived as having the most content knowledge has the most status (Finkelstein et al., 2019), but in this case teachers clearly felt safe in sharing their confusion about the definition of the term "property." By allowing this, the student hat supported these teachers in developing more complex understanding of science ideas.

Using Student Hat to Learn About Their Students

All 12 of the teachers who were interviewed discussed how the student hat structure helped them to learn about their students, specifically by developing their empathy for how students might experience an OpenSciEd lesson. The participants discussed three distinct forms of empathy when talking about the student hat: empathizing with the way students might *speak, think,* and *feel* when engaging in OpenSciEd lessons.

Of the 12 interviewees, four of them discussed that the student hat helped them to think about and empathize with how a student might speak during class. An example of what this looked like came from Ms. Vernon, who said, "[I was] trying to make sure I said something while I had the student hat, how would my student really say this? [….] Yeah, so that was to really think of language that they might use." This reflection shows the focus Ms. Vernon put on what exactly students might say in terms of the words they would choose when contributing in class.

Seven of the interviewees discussed how the student hat helped them to empathize with how a student might think during an OpenSciEd lesson. For example, when asked about student hat, Mr. Morse said the following:

> *I think [the PD] really focuses on what's important about being in student hat is what's the response? What's the thought process of the students going to be during this? […] So you could say, well my student would say, "I don't know." But really when they say, I don't know, they were processing something in order to come to the realization they didn't know.*

This reflection emphasized the idea that even though students might say "I don't know" in class, they are likely doing some important cognitive work that they are hesitant to share. Therefore, engaging in the student hat during PD helped Mr. Morse to consider *why* a student might claim not to know something, such as not feeling confident in their current thinking or believing that only completely correct answers will be accepted by the teacher.

The final form of empathy teachers discussed was empathizing with how students might feel while engaging in OpenSciEd lessons. In total, four of the 12 teachers mentioned how the student hat helped them to empathize with how students might feel. For example, Ms. Jackson stated “it's good to have the same feeling that the students have because it helps me empathize as a teacher with my student experience.” Mr. Emerson built on this idea by discussing how the student hat helped him to feel what students might feel and therefore he could better consider how to support his students to persist through feelings of anxiety or stress that might arise when participating in the OpenSciEd curriculum:

> *Sitting in the student role is the most powerful and most fun part of it because you're—the way it was set up, it felt like I was actually sitting in that role for a moment and you felt anxiety and stress in a way that I think kids probably do. […] I'm always trying to prepare kids for or help them understand that they're going to feel a certain way through understanding or lack of understanding of something, and how to then respond to that.*

The OpenSciEd approach to learning science is different from the way most students were likely used to learning in science class, and that unfamiliarity could potentially cause stress or anxiety (Han & Gutierez, 2021). Being put in the place of a learner himself helped Mr. Emerson to realize the power of this affective response to instruction and pinpoint when and how his students might experience those feelings.

Because building empathy is such an internal process, it can be hard to see it taking place when examining video, but there were examples of teachers expressing this kind of empathy for how students might feel during or just after moments in which they struggled to stay in the student hat. For example, during the third day of the PD the contact forces group engaged in a student hat activity as they used spring scales, carts, washers, and the slow-motion video function on their smartphones to investigate the relative forces on objects of different masses or materials as those objects collided. Given that they were trying to measure force at the instant of two objects colliding, the participants found this data challenging to collect and interpret. As a result, they broke out of student hat to discuss their confusions and frustrations. During that time, Mr. Clarke proposed that he pre-mark the lines on the spring scale with white-out to make them easier to see. In order to justify that proposal, he said:

> *I guess my thought is that I don't think I'm affecting their design of the experiment, I'm just enhancing their ability to get the data. Because we were trying to read that, and you know we're trying to move forward, trying to get somewhere so we don't walk out feeling "yeah, we got to the start of the thing and didn't get anything out of it." And that's frustrating, and that's frustrating for kids. And I'm just thinking if we had better indicators, then that would make a difference. We could have moved faster on collecting data.*

When he said "that's frustrating," Mr. Clarke pointed emphatically at the spring scale on the table and added a staccato tone to his voice, suggesting that he also felt that frustration when engaging in the investigation, and, as Mr. Emerson had discussed, was considering how he might help to mitigate those feelings.

Using Student Hat to Learn About the Instructional Approach

Ten of the twelve interviewed teachers discussed how the student hat helped them to learn about the OpenSciEd instructional approach. These comments fell into two broad categories: how the student hat helped teachers to better understand the *broader goals* *and philosophy* behind the curricular materials and how the student hat helped them to understand *effective implementation logistics and moves*.

The OpenSciEd curriculum is designed with an instructional model that focuses on students and the teacher working together in community to explore an anchoring phenomenon, ask questions about it to motivate future investigations, and then use the results of those questions to both answer their questions and motivate future work (Edelson et al., 2021; Reiser et al., 2021). This approach is fundamentally different from traditional approaches to instruction in which the teacher tells the students what to learn and when (Hutchison & Hammer, 2010). Four teachers discussed how the student hat helped them to better understand one or more of these goals of the OpenSciEd instructional model. For example, Ms. Mendell said the following:

> *I had to think like my students, and a lot of times I don't think like my students. I want my students to think like me. So, trying to wear the student hat and allowing the students to be a student opposed to allowing the students to think like Ms. [Mendell] was interesting because I never really thought about the student hat. I always thought about the teacher hat and what they needed to know.*

Here we see that student hat helped Ms. Mendell to re-orient her understanding of the goal of science education, moving towards celebrating her students' ways of thinking and making sense of science phenomena rather than a more traditional approach to identifying and correcting their misconceptions in order to get them to know "what they needed to know."

In addition to understanding the goal of the OpenSciEd instructional approach, nine teachers discussed how the student hat helped them consider effective implementation logistics or moves, which included things like the best way to introduce an idea or discussion or how to set up a particular investigation. For example, Ms. Moffitt said that student hat helped her, "think about when I was going to teach it, I could think about what might be frustrating to a student or what might be a pitfall that they would fall into or myth and how to address that." On a more logistical level, Mr. Richter noted that the student hat helped him to "think about how the actual activities work[…]or setting up instruments to work."

In the PD video, there were many moments in which teachers were struggling to think like a student but seemed to be using that struggle to learn about the OpenSciEd instructional approach. Within these moments, there was a distinct pattern between the times in which teachers were pushing on their understanding of the goals of the instructional model as compared to considering implementation moves or logistics. When considering instructional logistics, the teachers tended to intentionally break away from student hat whereas when learning about the broader goals, the breakout from student hat appeared more unintentional.

Intentional Breaks from Student Hat. When teachers appeared to intentionally break out of student hat, they often used similar linguistic and paralinguistic signals, which included lowering their voice, changing their tone to make it more serious, and/or using the phrase "teacher hat question" to preface their comment. These moves support the idea that these teachers knew they were not following the norms of the student hat activity when making these comments but chose to do so anyway.

For example, in the afternoon of the second day of the PD, the group preparing to teach the light & matter unit had engaged in an investigation in which they observed what happened to light from a flashlight that was directed at a mirror, a piece of plastic, and a one-way mirror. After completing that investigation, small groups drew models of how they thought light behaved and then the entire group engaged in a discussion sharing out those models in order to come to consensus on how they might represent things like the direction and intensity of light. During that discussion, Mr. Donahue intentionally left student hat as shown in Table IV.7.

Table IV.7: First Excerpt from Light Discussion

	Speaker	Transcript
1	Facilitator:	What are some similarities that we see between all three ways? So we can all kinda show them really quick. So, [Mr. Richter], if you could hold yours up while we're talking here. So, are there similar ways in which we tried to represent light and how it moves?
2	Mr. Bernard:	Everybody used lines.
3	Facilitator:	Everybody used lines.
4	Mr. Bernard:	And arrows.
5	Facilitator:	And arrows. Where the lines like curvy? Where they, like how--?
6	Ms. Murray:	They were all straight
7	Facilitator:	They were all straight lines.
8	Mr. Bernard:	As straight as we could do without a ruler.
9	Facilitator:	So do we all agree that that’s a good way of representing light?
10	Multiple:	Mmm-hmmm [affirmative]
11	Mr. Donahue:	[softly] At this point, has it been represented with a laser yet?
12	Facilitator:	[softly, shakes head] No.
13	Ms. Murray:	[softly] In this lesson? No.
14	Mr. Donahue:	[softly] Gotcha [nods head].
15	Facilitator:	[full volume] So, in terms of, so why does it make sense for us to use straight lines for us to represent light? Why did you all do that?

In this excerpt, the facilitator, Mr. Bernard, and Ms. Murray were all speaking in the student hat when Mr. Donahue interrupted at turn 11 to ask about the ordering of student experiences in the curriculum. The way that Mr. Donahue changed to a softer tone and the fact that the facilitator and Ms. Murray both mirrored that tone when answering implied that all three participants knew they were breaking out of student hat but did it anyway. Immediately afterwards, the facilitator returned to a full volume to continue to ask questions in the student hat. This brief moment demonstrated how Mr. Donahue was trying to consider if he might reference a laser when discussing this idea with his students, or, as Ms. Moffitt said in her interview, “how to address” a potential place of student confusion.

This example illustrates a teacher considering issues of planning and implementation rather than student thinking, which is a common issue in PD (Collins et al., 2019). For teachers like Mr. Donahue who are concerned with how they will implement a particular discussion or activity, they may break out of student hat to ask these kinds of logistical questions. These are certainly reasonable questions and important for teachers who will be implementing these lessons to consider, but could be better addressed during different activities focused on teacher facilitation rather than during student hat activities, which are designed to support teachers in understanding how students experience storyline curricula.

Unintentional Breaks from Student Hat. Unlike the intentional breaks, the times when teachers unintentionally broke from student hat were rarely accompanied by a change in tone or volume or a verbal marker. In fact, often a teacher would unintentionally say something their students would never say, immediately realize they had done so, and then the teacher and others would laugh in response to that break out. These events most often occurred when teachers were drawing on their own experiences to make sense of a scientific phenomenon, which is another key feature of the OpenSciEd instructional model.

In the same light discussion discussed above, the same teacher, Mr. Donahue, also unintentionally broke out of student hat when explaining how his group had shown light of different intensity on their model. That exchange in shown in Table IV.8.

Table IV.8: Second Excerpt from Light Discussion

	Speaker	Transcript
1	Facilitator:	Do the different colors matter then?
2	Mr. Donahue:	Yeah. This would be like, this is a--
3	Ms. Murray:	Still kinda bright
4	Mr. Donahue:	-- This would be like a, a Bud, and this would be like a Bud Light, and this would be like a Bud Dry—some strange thing that you wouldn't really make. [Facilitator makes facial expression indicating confusion] [louder] Sorry we're in 6th grade! [many participants laugh] This is like a Coke, this is like a Diet Coke, and this is--[laughter].
5	Many:	Coke Zero! [laughter]
6	Facilitator:	I'm horrified. Horrified. I need to call your guidance counselor. [laughter]
7	Ms. Murray:	You probably need to talk to the nurse, and the principal, and-- [laughter]
8	Ms. Goh:	Our partying is not the principal's business [laughter].
9	Facilitator:	Ok, so back to our normal 6th grade--

In turn 4, Mr. Donahue was trying to connect different kinds of light on his model with another example of different kinds of the same thing, he reached for types of Budweiser—context a sixth grader would most likely not have. He did not change his tone or volume at all until noticing the facilitator making a facial expression indicating surprise, then he spoke notably louder when he recognized that he had slipped out of student hat. This was followed by sustained laughter by both him and the rest of the group, implying that the group found the use of types of beer or the unintentional break from student has as funny.

Although he was referencing experiences that his sixth-grade students likely would not have, the way that Mr. Donahue called on personal experiences outside of the domain of light science to make sense of the model reflects an approach to sensemaking that is common in student thinking (Rosebery et al., 2010). In order to support students' epistemic agency, the OpenSciEd instructional model is intentionally designed to invite students to use their own experiences and backgrounds to make sense of natural phenomena (Reiser et al., 2021). Therefore, this unintentional break from student hat seemed to actually help teachers experience this key design feature of the curriculum in that they were recruiting their own experiences and ways of thinking in order to make sense of a natural phenomenon.

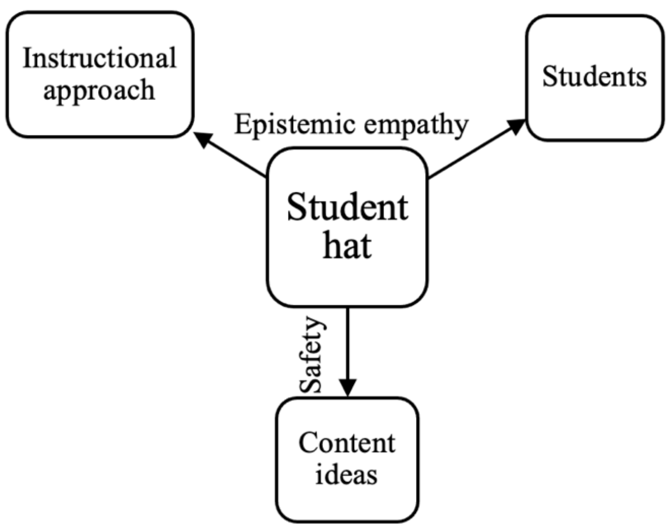
Discussion

This study suggests that the student hat can be used as a pedagogical tool to push on multiple objectives, including developing teachers' understanding of content, their students, and target instructional approaches. By attending to what teachers said both during and after student hat activities, we can begin to consider *how* the student hat helped to facilitate these three teacher learning goals. In the first part of this section, I do that, proposing that student hat provides teachers a sense of *safety* that facilitates content learning and supports teachers to develop their *epistemic empathy* (Jaber, 2021; Jaber et al., 2018) for students, which helps them learn about their students and the OpenSciEd instructional approach. In addition to considering how the student hat supports teacher learning, this study also has implications for effectively designing for the use of the student hat activity structure in PD. In the second part of this section, I revisit the difference between the moments when teachers intentionally and unintentionally stopped using the student to highlight how future PD designers might build trust with their participants to limit the intentional moments while supporting metacognition to demonstrate the value of the unintentional ones.

Safety and Epistemic Empathy to Support Teacher Learning

There is a wealth of literature on the relationship between various PD features and teacher or student learning outcomes (e.g. Darling-Hammond et al., 2017; Desimone et al., 2002; Lynch et al., 2019; Penuel et al., 2011; Yoon et al., 2007). Many of the standard approaches to PD research focus on objectives for teacher learning—change in teachers' beliefs, content knowledge, or practice—at the expense of understanding the *mechanism* for teacher learning, meaning how PD features or activities actually achieve those goals (Walkoe & Luna, 2020). My results imply two mechanisms to support teacher learning: providing teachers a *safe environment to be wrong* allowed them to develop their understanding of science content and building their *epistemic empathy for their students* facilitated learning about both their students and the OpenSciEd instructional approach (See Figure IV.2)

Figure IV.2: Student Hat Teacher Learning Outcomes and Mechanisms



Discussions around science content can bring out relational power dynamics among teachers based on who is perceived as a "content expert" and those dynamics can serve to shut down effective teacher learning (Finkelstein et al., 2019). The student hat might help to avoid that concern by, as Mr. Richter put it, "level[ing] the playing field" so that teachers feel more comfortable expressing confusion about science ideas. This was echoed by Ms. Mendell, who said that that student hat helped her “not have to be so technical about the choices of the words that I use.” Even though the primary goal of student hat is not to push on teachers’ content knowledge, this study suggests that it still might do so by mitigating power dynamics in PD around preexisting content understanding to allow all teachers to better learn the content ideas under discussion.

One of the most consistent comments from teachers was the idea of student hat activities building their empathy. Empathy is a complex concept that has been defined differently across various academic fields (Oxley, 2011), but a common definition is the state of “imagining how another is thinking and feeling” (Batson, 2009). To achieve this state, there are two parts of empathy: a cognitive part that involves thinking about how another might feel and an emotional part that involves taking on those feelings (C. A. Warren, 2018). Most of the early work in empathy focused on empathizing with another’s thoughts or feelings in the process of carrying out everyday life activities, but as the study of empathy became more popular in education, a move began to consider how one empathizes with another’s knowledge building experiences (Horsthemke, 2015). Jaber and her colleagues, therefore, define a particular type of empathy, *epistemic empathy,* as “understanding and appreciating someone's cognitive and emotional experience within an epistemic activity, meaning an activity aimed at the construction, communication, and critique of knowledge” (Jaber et al., 2018, p. 14). Although the teachers in this study did not use the term “epistemic empathy,” they clearly built an appreciation for their students’ cognitive and emotional experiences within OpenSciEd, which helped them to better understand both their students and the OpenSciEd approach.

Some preliminary investigations have suggested that epistemic empathy can be built by providing teachers opportunities to witness student sensemaking, express their own concerns about new approaches to instruction, and to critically reflect on their own assumptions about strong learning (Jaber, 2021). Given the prevalence of empathy across my interview data and the focus on knowledge-building exercises during the PD, I propose that the student hat might also be supporting teachers' epistemic empathy as a mechanism for learning about their students and the instructional approach. For example, when Mr. Emerson discussed feeling the anxiety related to uncertainty around the content that he expected his students to feel, he was building his empathy with them while making sense of a key aspect of scientific sensemaking. Similarly, when Mr. Donahue used the varieties of Budweiser beer to discuss types of light, he was empathizing with the *kind* of idea work his students might engage in, even if not the exact *ideas* they may bring to the work.

As part of building teachers' epistemic empathy, student hat can also help to open up teachers' understanding of what "counts" as legitimate scientific sensemaking (Rosebery et al., 2016). This is particularly important to help support equity in the science classroom, as many science classes have historically marginalized students whose ways of thinking or describing the natural world that do not align with Western approaches to science (Bang et al., 2012; B. Warren et al., 2020). By helping teachers to build their epistemic empathy for the heterogeneity in students' thinking, the student hat can begin to shift teachers' understanding of the value of that heterogeneity in supporting all students' learning (Rosebery et al., 2010).

Designing Professional Development to Use the Student Hat

While the student hat can support multiple goals for teacher learning, it is not a panacea and PD designers should be thoughtful about how, when, and why they choose to use this activity structure. In order to help determine when the student hat would be an effective PD tool, it is helpful to look at the different ways that teachers struggled with the structure. In this study, sometimes teachers *intentionally* left the student hat, changing their tone and volume to ask a "teacher hat question;" whereas other times teachers *unintentionally* called on information or experiences they have had as adults that students could never have and only later recognized they had broken out of student hat.

When teachers intentionally left the student hat, they often did so to ask questions about the logistics of implementation of a particular classroom activity, such as when Mr. Donahue asked if students had used lasers to see the path of light or when Ms. Jasper pushed back on how the curriculum defined "properties." This attention to the logistics of instruction is reasonable because teachers make sense of PD through the lens of their own implementation and accountability context (Allen & Penuel, 2015). Unfortunately, these kinds of breakouts pull away from building teachers' epistemic empathy for students, one of the unique affordances of the student hat. In other words, there are other easier and more efficient ways teachers could discuss the logistics of implementation during PD, including watching video of teachers implementing, reading and discussing the curricular materials, or collaboratively planning for implementation in their context (Allen & Heredia, 2021; Gibbons & Cobb, 2017; Loucks-Horsley et al., 2010).

Given the importance of implementation in supporting teachers' learning from PD (Allen & Penuel, 2015), the fact that these intentional breakouts happened with some frequency, therefore, implies that teachers may have been concerned that they would not have another opportunity to discuss these questions of implementation logistics. A key aspect of strong PD facilitation is building a community in which the participants trust each other and the facilitator to support their mutual learning needs (Borko et al., 2014; Roth et al., 2017). These intentional moves away from student hat may have been a sign that participants were still developing that trust. Therefore, future PD designers should create structures that attend to teachers' questions around implementation logistics so that teachers trust that those questions will be answered. For example, they could have a parking lot where teachers post questions, which are addressed at the end of each day. Creating these structures might allow teachers to stay in student hat, build epistemic empathy, and better understand the target instructional approach. This point does require further research, however, and future studies could investigate this hypothesis by tracking how these intentional breakouts from student hat change (or remain consistent) over time as teachers become more familiar and experienced with the student hat structure.

The unintentional breaks from student hat had a significantly different character than the intentional ones because they often involved teachers using their own adult experiences and background to make sense of a scientific phenomenon. Teachers may have considered these moments when discussing how difficult the student hat was because they were frequently more clearly marked during the discussion, often with laughter or other humorous comments, such as the facilitator responding to Mr. Donahue's unintentional reference to beer by jokingly saying he would "call your guidance counselor." Unlike the intentional break outs, however, these moments demonstrate a key aspect of building epistemic empathy and therefore an understanding of the fundamental approach to science learning that drives the OpenSciEd materials. The instructional materials are built to value students' diverse experiences and sensemaking practices as a way to improve equity and epistemic agency (Bang et al., 2017; Rosebery et al., 2010; Zivic et al., 2018). In addition, they aim to support students in motivating future questions and investigations based on the experiences they bring to the classroom and the work collaboratively done there (Reiser, Novak, et al., 2017; Reiser et al., 2021). During these moments, the teachers were truly doing both of those things, which hopefully helped them design for those moments in their own classroom.

The implication for PD design and implementation, therefore, is that not all difficulty with the student hat is created equal. As teachers struggle to remain in student hat because they are using their own experiences to make sense of phenomena, they may be engaging in the important learning work the student hat can best support. That kind of learning can be hard for teachers to notice, however, especially if they perceive themselves as violating the established structures or dynamics of the PD (Finkelstein et al., 2019). Therefore, providing opportunities for teachers to engage in meta-cognitive reflection on their experience in the student hat and how it helped them to learn about their students and the targeted instructional reform may help teachers to better understand the learning value of the student hat. This is an area where future design work could help to elucidate the relationship between teachers' perception of their own takeaways from student hat and observed teacher learning outcomes.

By providing a safe space for teacher productive struggle and building their epistemic empathy for their students, the student hat can support teachers to learn about disciplinary content ideas, their students, and novel instructional approaches. As we continue to support teacher growth, and by extension student learning, this activity, if thoughtfully implemented, can be a powerful feature of strong PD design.

References

Achieve. (2018). *NGSS alignment claims: How publishers talk about the Next Generation Science Standards*. https://www.nextgenscience.org/NGSSclaims

Allen, C. D., & Heredia, S. C. (2021). Reframing organizational contexts from barriers to levers for teacher learning in science education reform. *Journal of Science Teacher Education*, *32*(2), 148–166. https://doi.org/10.1080/1046560X.2020.1794292

Allen, C. D., & Penuel, W. R. (2015). Studying teachers' sensemaking to investigate teachers' responses to professional development focused on new standards. *Journal of Teacher Education*, *66*(2), 136–149. https://doi.org/10.1177/0022487114560646

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (Vol. 1, pp. 3–22). Jossey-Bass.

Bang, M., Brown, B., Calabrese Barton, A., Rosebery, A., & Warren, B. (2017). Towards more equitable learning in science: Expanding relationships among students, teachers, and science practices. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 33–58). NSTA Press.

Bang, M., Warren, B., Rosebery, A. S., & Medin, D. (2012). Desettling expectations in science education. *Human Development*, *55*, 302–318. https://doi.org/10.1159/000345322

Banilower, E. R., Smith, P. S., Malzahn, K. M., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Horizon Research, Inc.

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), *The Social Neuroscience of Empathy* (pp. 3–16). The MIT Press. https://doi.org/10.7551/mitpress/9780262012973.003.0002

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*, *53*(7), 1082–1112. https://doi.org/10.1002/tea.21257

Borko, H., Frykholm, J., Pittman, M., Eiteljorg, E., Nelson, M., Jacobs, J., Koellner-Clark, K., & Schneider, C. (2005). Preparing teachers to foster algebraic thinking. *ZDM*, *37*(1), 43–52. https://doi.org/10.1007/BF02655896

Borko, H., Koellner, K., & Jacobs, J. (2014). Examining novice teacher leaders' facilitation of mathematics professional development. *The Journal of Mathematical Behavior*, *33*, 149–167. https://doi.org/10.1016/j.jmathb.2013.11.003

Chen, Y.-C. (2021). Epistemic uncertainty and the support of productive struggle during scientific modeling for knowledge co-development. *Journal of Research in Science Teaching*. Advance online publication. https://doi.org/10.1002/tea.21732

Cherbow, K. (2021). *Exploring the teacher-storyline relationship: Curricular design and enactment for coherence from the student perspective* (28322545) [Doctoral dissertation, Boston College]. ProQuest Dissertations and Theses Global.

Cherbow, K., & McNeill, K. L. (2022). Planning for student-driven discussions: A revelatory case of curricular sensemaking for epistemic agency. *Journal of the Learning Sciences*. Advance online publication. https://doi.org/10.1080/10508406.2021.2024433

Chuy, M., Scardamalia, M., Bereiter, C., Prinsen, F., Resendes, M., Messina, R., Hunsburger, W., Teplovs, C., Chow, A., & Chow, A. (2010). Understanding the nature of science and scientific progress: A theory-building approach. *Canadian Journal of Learning and Technology / La Revue Canadienne de l'apprentissage et de La Technologie*, *36*(1). https://www.learntechlib.org/p/43129/

Cohen, D. K. (1990). A Revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, *12*(3), 311–329. https://doi.org/10.3102/01623737012003311

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

Collins, L., Cavagnetto, A., Ferry, N., Adesope, O., Baldwin, K., Morrison, J., & Premo, J. (2019). May I have your attention: An analysis of teacher responses during a multi-year professional learning program. *Journal of Science Teacher Education*, *30*(6), 549–566. https://doi.org/10.1080/1046560X.2019.1589846

Crawford, B. A. (2014). From inquiry to scientific practices in the science classroom. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of Research on Science Education* (pp. 515–541). Routledge. https://doi.org/10.4324/9780203097267

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (Fourth edition). SAGE.

Cuban, L. (1990). A fundamental puzzle of school reform. In A. Lieberman (Ed.), *Schools as collaborative cultures: Creating the future now* (pp. 71–77). The Falmer Press.

Damşa, C. I., Kirschner, P. A., Andriessen, J. E. B., Erkens, G., & Sins, P. H. M. (2010). Shared epistemic agency: An empirical etudy of an emergent construct. *Journal of the Learning Sciences*, *19*(2), 143–186. https://doi.org/10.1080/10508401003708381

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81–112. https://doi.org/10.3102/01623737024002081

Edelson, D. C., Reiser, B. J., McNeill, K. L., Mohan, A., Novak, M., Mohan, L., Affolter, R., McGill, T. A. W., Buck Bracey, Z. E., Deutch Noll, J., Kowalski, S. M., Novak, D., Lo, A. S., Landel, C., Krumm, A., Penuel, W. R., Van Horne, K., González-Howard, M., & Suárez, E. (2021). Developing research-based instructional materials to support large-scale transformation of science teaching and learning: The approach of the OpenSciEd middle school program. *Journal of Science Teacher Education*, *32*(7), 780–804. https://doi.org/10.1080/1046560X.2021.1877457

Finkelstein, C., Jaber, L. Z., & Dini, V. (2019). "Do I feel threatened? No… I'm learning!"—Affective and relational dynamics in science professional development. *Science Education*, *103*(2), 338–361. https://doi.org/10.1002/sce.21489

Fishman, B. J., Davis, E. A., & Chan, C. K. K. (2014). A learning sciences perspective on teacher learning research. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 707–725). Cambridge University Press.

Freedman, E. B. (2020). When discussions sputter or take flight: Comparing productive disciplinary engagement in two history classes. *Journal of the Learning Sciences*, *29*(3), 385–429. https://doi.org/10.1080/10508406.2020.1744442

Furberg, A., & Silseth, K. (2021). Invoking student resources in whole-class conversations in science education: A sociocultural perspective. *Journal of the Learning Sciences*, 1–39. https://doi.org/10.1080/10508406.2021.1954521

Furtak, E. M., & Penuel, W. R. (2019). Coming to terms: Addressing the persistence of "hands-on" and other reform terminology in the era of science as practice. *Science Education*, *103*(1), 167–186. https://doi.org/10.1002/sce.21488

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945. https://doi.org/10.3102/00028312038004915

Gibbons, L. K., & Cobb, P. (2017). Focusing on teacher learning opportunities to identify potentially productive coaching activities. *Journal of Teacher Education*, *68*(4), 411–425. https://doi.org/10.1177/0022487117702579

Glackin, M. (2019). 'It's more than a prop': Professional development session strategies as sources of teachers' self-efficacy and motivation to teach outside the classroom. *Professional Development in Education*, *45*(3), 372–389. https://doi.org/10.1080/19415257.2018.1490917

Grigg, J., Kelly, K. A., Gamoran, A., & Borman, G. D. (2013). Effects of two scientific inquiry professional development interventions on teaching practice. *Educational Evaluation and Policy Analysis*, *35*(1), 38–56. https://doi.org/10.3102/0162373712461851

Grossman, P. (Ed.). (2018). *Teaching core practices in teacher education*. Harvard Education Press.

Han, M., & Gutierez, S. B. (2021). Passive elementary student's constructed epistemic emotions and patterns of participation during small group scientific modeling. *Science Education*, *105*(5), 908–937. https://doi.org/10.1002/sce.21665

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age Publishing.

Horsthemke, K. (2015). Epistemic empathy in childrearing and education. *Ethics & Education*, *10*(1), 61–72. https://doi.org/10.1080/17449642.2014.998025

Hutchison, P., & Hammer, D. (2010). Attending to student epistemological framing in a science classroom. *Science Education*, *94*(3), 506–524. https://doi.org/10.1002/sce.20373

Jaber, L. Z. (2021). "He got a glimpse of the joys of understanding" – The role of epistemic empathy in teacher learning. *Journal of the Learning Sciences*, *30*(3), 433–465. https://doi.org/10.1080/10508406.2021.1936534

Jaber, L. Z., Southerland, S., & Dake, F. (2018). Cultivating epistemic empathy in preservice teacher education. *Teaching and Teacher Education*, *72*, 13–23. https://doi.org/10.1016/j.tate.2018.02.009

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences*, *21*(1), 45–83. https://doi.org/10.1080/10508406.2011.591717

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Ko, M.-L. M., & Krist, C. (2019). Opening up curricula to redistribute epistemic agency: A framework for supporting science teaching. *Science Education*, *103*(4), 979–1010. https://doi.org/10.1002/sce.21511

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*(3), 226–243. https://doi.org/10.1177/0022487112473837

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2010). *Designing professional development for teachers of science and mathematics*. Corwin Press. https://doi.org/10.4135/9781452219103

Lowell, B. R., Cherbow, K., & McNeill, K. L. (2021). Redesign or relabel? How a commercial curriculum and its implementation oversimplify key features of the NGSS. *Science Education*, *105*(1), 5–32. https://doi.org/10.1002/sce.21604

Lowell, B. R., Cherbow, K., & McNeill, K. L. (2022). Considering discussion types to support collective sensemaking during a storyline unit. *Journal of Research in Science Teaching*, *59*(2), 195–222. https://doi.org/10.1002/tea.21725

Lynch, K., Hill, H. C., Gonzalez, K. E., & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis*, *41*(3), 260–293. https://doi.org/10.3102/0162373719849044

Manz, E., Lehrer, R., & Schauble, L. (2020). Rethinking the classroom science investigation. *Journal of Research in Science Teaching*, *57*(7), 1148–1174. https://doi.org/10.1002/tea.21625

Manz, E., & Suárez, E. (2018). Supporting teachers to negotiate uncertainty for science, students, and teaching. *Science Education*, *102*(4), 771–795. https://doi.org/10.1002/sce.21343

Marrongelle, K., Sztajn, P., & Smith, M. (2013). Scaling up professional development in an era of common state standards. *Journal of Teacher Education*, *64*(3), 202–211. https://doi.org/10.1177/0022487112473838

McNeill, K. L., Affolter, R., & Reiser, B. J. (in press). Anchoring science professional learning in curriculum materials enactment: Illustrating theories in practice to support teacher learning. In A. Castro Superfine, S. R. Goldman, & M.-L. M. Ko (Eds.), *Teacher learning in changing contexts: Perspectives from the learning sciences*. Routledge.

McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2017). Moving beyond pseudoargumentation: Teachers' enactments of an educative science curriculum focused on argumentation. *Science Education*, *101*(3), 426–457. https://doi.org/10.1002/sce.21274

McNeill, K. L., & Reiser, B. J. (2018). Open source for opening minds: New OpenSciEd materials support science standards. *The Learning Professional*, *39*(6), 44–48.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (Fourth edition). SAGE.

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053–1075. https://doi.org/10.1002/tea.21459

National Academies of Sciences, Engineering, and Medicine. (2019). *Science and Engineering for Grades 6-12: Investigation and Design at the Center* (B. Moulding, N. Songer, & K. Brenner, Eds.). National Academies Press. https://doi.org/10.17226/25216

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all.* NCTM.

Next Generation Science Storylines & Next Generation Science Exemplar System. (2018). *Thinking like a kid*. https://www.openscied.org/wp-content/uploads/2021/05/Thinking_like_a_kid_handout_v5.pdf

OpenSciEd. (2020a). *How can we make something new that was not there before?* https://www.openscied.org/instructional-materials/7-1-chemical-reactions-matter/

OpenSciEd. (2020b). *Why do things sometimes get damaged when they hit each other?* https://www.openscied.org/instructional-materials/8-1-contact-forces/

OpenSciEd. (2021). *Why do we sometimes see different things when looking at the same object?* https://www.openscied.org/instructional-materials/6-1-light-matter/

Oxley, J. (2011). *The moral dimensions of empathy: Limits and applications in ethical theory and practice* (2011 edition). Palgrave Macmillan.

Penuel, W. R., Fishman, B. J., Haugan Cheng, B., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, *40*(7), 331–337. https://doi.org/10.3102/0013189X11421826

Reiser, B. J., Brody, L., Novak, M., Tipton, K., & Adams, L. S. (2017). Asking questions. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 87–108). NSTA Press.

Reiser, B. J., Michaels, S., Moon, J., Bell, T., Dyer, E., Edwards, K. D., McGill, T. A. W., Novak, M., & Park, A. (2017). Scaling up three-dimensional science learning through teacher-led study groups across a state. *Journal of Teacher Education*, *68*(3), 280–298. https://doi.org/10.1177/0022487117699598

Reiser, B. J., Novak, M., & McGill, T. A. W. (2017). *Coherence from the students' perspective: Why the vision of the framework for K-12 science requires more than simply "combining" three dimensions of science learning*. The National Academies of Sciences, Engineering and Medicine, Board on Science Education. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_180270.pdf

Reiser, B. J., Novak, M., McGill, T. A. W., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal of Science Teacher Education*, *32*(7), 805–829. https://doi.org/10.1080/1046560X.2021.1884784

Robertson, A. D., Scherr, R., & Hammer, D. (Eds.). (2015). *Responsive Teaching in Science and Mathematics* (1st ed.). Routledge. https://doi.org/10.4324/9781315689302

Rosebery, A. S., Ogonowski, M., Dischino, M., & Warren, B. (2010). "The coat traps all your body heat": Heterogeneity as fundamental to learning. *Journal of the Learning Sciences*, *19*(3), 322–357. https://doi.org/10.1080/10508406.2010491752

Rosebery, A. S., Warren, B., & Tucker-Raymond, E. (2016). Developing interpretive power in science teaching. *Journal of Research in Science Teaching*, *53*(10), 1571–1600. https://doi.org/10.1002/tea.21267

Rossman, G. B., & Rallis, S. F. (2017). *An introduction to qualitative research: Learning in the field* (Fourth). SAGE Publications. https://doi.org/10.4135/9781071802694

Roth, K. J., Bintz, J., Wickler, N. I. Z., Hvidsten, C., Taylor, J., Beardsley, P. M., Caine, A., & Wilson, C. D. (2017). Design principles for effective video-based professional development. *International Journal of STEM Education*, *4*, 31. https://doi.org/10.1186/s40594-017-0091-2

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. *Liberal Education in a Knowledge Society*, *97*, 67–98.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (2nd ed., pp. 397–417). Cambridge University Press.

Schwarz, C. V., Passmore, C., & Reiser, B. J. (2017). Moving beyond "knowing about" science to making sense of the natural world. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 3–21). NSTA Press.

Sengupta-Irving, T., & Agarwal, P. (2017). Conceptualizing perseverance in problem solving as collective enterprise. *Mathematical Thinking and Learning*, *19*(2), 115–138. https://doi.org/10.1080/10986065.2017.1295417

Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (Eds.). (2011). *Mathematics teacher noticing: Seeing through teachers' eyes*. Routledge.

Short, J., & Hirsh, S. (2020). *The elements: Transforming teaching through curriculum-based professional learning*. Carnegie Corporation of New York.

https://www.carnegie.org/publications/elements-transforming-teaching-through-curriculum-based-professional-learning/

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 443–466). SAGE Publications.

Stein, M. K., Smith, M. S., & Silver, E. (1999). The development of professional developers: Learning to assist teachers in new settings in new ways. *Harvard Educational Review*, *69*(3), 237–270. https://doi.org/10.17763/haer.69.3.h2267130727v6878

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. https://doi.org/10.1002/sce.21112

Taylor, J. A., Getty, S. R., Kowalski, S. M., Wilson, C. D., Carlson, J., & Van Scotter, P. (2015). An efficacy trial of research-based curriculum materials with curriculum-based professional development. *American Educational Research Journal*, *52*(5), 984–1017. https://doi.org/10.3102/0002831215585962

Trinter, C. P., & Hughes, H. E. (2021). Teachers as curriculum designers: Inviting teachers into the productive struggle. *RMLE Online*, *44*(3), 1–16. https://doi.org/10.1080/19404476.2021.1878417

Walkoe, J. (2015). Exploring teacher noticing of student algebraic thinking in a video club. *Journal of Mathematics Teacher Education*, *18*(6), 523–550. https://doi.org/10.1007/s10857-014-9289-0

Walkoe, J. D. K., & Luna, M. J. (2020). What we are missing in studies of teacher learning: A call for microgenetic, interactional analyses to examine teacher learning processes.

Journal of the Learning Sciences, *29*(2), 285–307. https://doi.org/10.1080/10508406.2019.1681998

Warren, B., Vossoughi, S., Rosebery, A. S., Bang, M., & Taylor, E. V. (2020). Multiple ways of knowing: Re-imagining disciplinary learning. In N. S. Nasir, C. D. Lee, R. D. Pea, & M. McKinney de Royston (Eds.), *Handbook of the cultural foundations of learning* (1st ed., pp. 277–293). Routledge.

Warren, C. A. (2018). Empathy, teacher dispositions, and preparation for culturally responsive pedagogy. *Journal of Teacher Education*, *69*(2), 169–183. https://doi.org/10.1177/0022487117712487

Warshauer, H. K. (2015). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education*, *18*(4), 375–400. https://doi.org/10.1007/s10857-014-9286-3

Windschitl, M., Thompson, J., & Braaten, M. (2018). *Ambitious science teaching*. Harvard Education Press.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (Sixth edition). SAGE.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (REL 2007-No. 033; Issues & Answers). US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=70

Zivic, A., Smith, J. F., Reiser, B. J., Edwards, K. D., Novak, M., & McGill, T. A. W. (2018). *Negotiating epistemic agency and target learning goals: Supporting coherence from the students' perspective*. International Conference of the Learning Sciences, London, U.K.

Appendix A: Participants' Background

Teacher	Grade Level	Role	District	School (grades)	OpenSciEd Experience	Interview
Mr. Bernard	6th	Teacher	Large City	School K (K-8)	Yes	Yes
Mr. Donahue	6th	Teacher	Small City	School A (5-6)	No	No
Mr. Edwards	6th	Teacher	Large City	School S (K-8)	No	No
Mr. Emerson	6th	Teacher	Large City	School G (K-8)	Yes	Yes
Mr. Ganbe	6th	Teacher	Large City	School U (K-8)	Yes	No
Ms. Goh	6th	Teacher	Large City	School L (K-8)	Yes	No
Ms. Murray	6th	Large City Science Supervisor			n/a	No
Mr. Richter	6th	Teacher	Large City	School M (6-8)	No	Yes
Ms. Anson	7th	Teacher	Large City	School D (6-12)	No	No
Mr. Clarke	7th	Teacher	Small City	School C (7-8)	Yes	No
Ms. Jackson	7th	Teacher	Large City	School U (K-8)	No	Yes
Ms Jennings	7th	Teacher	Large City	School J (7-12)	No	Yes
Ms. Judson	7th	Teacher	Small City	School C (7-8)	Yes	No
Ms. Knott	7th	Teacher	Large City	School S (K-8)	Yes	No
Ms. Larkin	7th	Local Science Museum Education Staff			n/a	No
Ms. Mendell	7th	Teacher	Large City	School M (K-8)	No	Yes
Ms. Moffitt	7th	Teacher	Large City	School Y (K-8)	Yes	Yes

Ms. Reed	7th	Teacher	Large City	School K (K-8)	No	No
Ms. Arnold	8th	Teacher	Small City	School C (7-8)	Yes	No
Ms. Brown	8th	Teacher	Large City	School J (7-12)	No	No
Ms. Ibarra	8th	Science Coach	Large City	School D (6-12)	n/a	No
Ms. Jasper	8th	Teacher	Large City	School R (K-8)	Yes	No
Ms. Jefferson	8th	Teacher	Small City	School C (7-8)	No	No
Ms. Kent	8th	Teacher	Large City	School D (6-12)	No	No
Mr. Kimball	8th	Teacher	Large City	School M (6-8)	Yes	Yes
Mr. Morse	8th	Teacher	Large City	School J (7-12)	Yes	Yes
Ms. Newcomb	8th	Teacher	Large City	School U (K-8)	No	Yes
Ms. Reyes	8th	Teacher	Large City	School H (K-8)	Yes	No
Mr. Rivera	8th	Teacher	Large City	School S (K-8)	No	No
Ms. Townsend	8th	Teacher	Small City	School C (7-8)	Yes	Yes
Mr. Truett	8th	Teacher	Large City	School B (7-12)	Yes	No
Ms. Vernon	8th	Teacher	Large City	School H (K-8)	Yes	Yes

Note. All names are pseudonyms

Appendix B: Full Interview Protocol

General Questions

1. What did you enjoy most about the PD? Why?
2. What aspects of the professional development do you feel had the greatest impact on your own learning? Why?
3. How has your vision of strong middle school science instruction changed (if at all) as a result of participating in OpenSciEd? Why?

Use of Video

4. In the PD we watched a number of videos, what did you think of the use of video in this PD?
5. We used a number of video clips from one teacher's classroom. I am going to ask you about some of those clips and how those clips may have helped your thinking. Think back to the video you watched of Kris' classroom about the features of classroom culture, how did this video clip influence your understanding of OpenSciEd?
6. Think back to the montage video you watched of Kris' classroom about teachers' actions to support equitable classroom culture, how did this video clip influence your understanding of OpenSciEd?
7. Think back to the video clips you watched of a consensus discussion and discussion planning in Kris' classroom. How did these video clips influence your understanding of OpenSciEd?

Student hat

8. This PD asked you to think like a student or be in "student hat" for a significant portion of the training. How does this compare to other non-OpenSciEd PD experiences you have had?
    - *Follow up if they have done it before* - What types of things did you do in the student hat in previous PD experiences? How was that similar or different to what we did in the OpenSciEd PD?
        - *Follow-up follow-up if they don't mention discussions* - In OpenSciEd we do sensemaking discussions in the student hat as well as investigations. In what ways is that similar to or different from your other PD experiences using student hat?
9. What were some benefits of thinking in the student hat?
10. What were some challenges of thinking in the student hat?

Student resources

11. There was a significant amount of time in this PD spent talking about "student resources". What is your understanding of what "student resources" are?
    - *Probe*: if they talk about handouts and the like, remind them that it was Day 1, talking about what students bring into the classroom and how teachers can use that to build
12. In the past, how did you typically respond when students gave an incorrect or "off-base" response in class?
13. Do you have any new ideas about what student participation could look like in your class?
    - *Follow up probes if they do not have any ideas or ask what we mean by "student participation"* - Besides raising hand/talking, what other forms can participation take?

Closing question

13. Do you have anything else you want to say about the OpenSciEd professional development or curriculum that you have not had the opportunity to share?

Section V: Synthesis and Conclusion

Nationally, K-12 science education is in the midst of great change, spurred on by the adoption of the Next Generation Science Standards or standards based off of them in many states (Bybee, 2014). The vision of science teaching and learning this reform advocates for is one in which students engage in the practices of science to figure things out about the natural world (Schwarz et al., 2017). In order to help students to engage in this figuring out work, science instruction must be phenomenon-based, three-dimensional, support student epistemic agency, and coherent from the student perspective (Lowell et al., 2021). These shifts can be challenging for teachers, particularly ones who have taught many years under accountability systems that favored memorization and recitation of lists of facts (National Academies of Sciences, Engineering, and Medicine, 2015). The three papers in this dissertation work together to help address that challenge by adding to our expanding understanding of how we can support teacher learning through curriculum-based professional development. Specifically, this work helps to clarify the particular ways science teachers can be supported to implement the new relationships between themselves, their students, and science content called for by current reforms in science teaching.

By taking a large-scale quantitative approach to teacher learning over two years, the first paper highlighted that curriculum-based professional development can be an effective tool to change teachers' beliefs about what good instruction looks like and their confidence in implementing that kind of instruction. Many studies of professional development are time-limited, looking only before and after a short PD workshop (Borko, 2004). By tracking the same teachers over multiple rounds of professional development and implementation, I was able to highlight some patterns in change over time, showing that teachers' beliefs and confidence changed initially but then leveled off over time, that confidence took longer to show significant changes than beliefs did, and that teachers who valued PD activities showed faster changes than those who did not. These findings will be helpful for the science education community to consider as they design comprehensive systems of professional learning over time to help teachers with these shifts (Loucks-Horsley et al., 2010). For example, designers should consider the value of personalized or differentiated support as teachers become more familiar with the shifts, such as coaching or lesson study (Gibbons & Cobb, 2017). In addition, teacher educators should be sure to include opportunities for enactment and reflection and build trust with teachers that the work is valuable and important, as both of these may support faster increases in teacher confidence over time (Avidov-Ungar, 2016; McNeill et al., in press).

While the first paper helped to clarify *what* teachers learned over the course of the OpenSciEd field test, the second two papers helped to explain *how* they did so. There is already a robust body of literature on the idea that professional development should be embedded in the practice of teaching, providing teachers an understanding of how the ideas from the PD are implemented in the classroom (Ball & Cohen, 1999; Grossman, 2018; Lampert, 2010). This focus has resulted in describing a number of PD activities that can be particularly helpful for teachers considering their practice, including watching classroom video to notice students' thinking and teacher moves (Brophy, 2004; Chen et al., 2020; Sherin et al., 2011; Sherin & van Es, 2005), rehearsing discussions or other classroom activities during PD (Kelley-Petersen et al., 2018; Lampert et al., 2013; Stroupe et al., 2020), or analyzing student work (Loucks-Horsley et al., 2010; McNeill & Knight, 2013). Just like practice-based PD asks us to ensure teacher learning is grounded in practice with students, so too should we also make sure that PD design is grounded in practice with teachers. To that end, my second paper took another look at common practice-based PD activities and considered how they helped science teachers to learn about what takes place in a classroom focused on supporting students to make sense of natural phenomena. Based on that analysis, I proposed that we add another tool to the practice-based PD toolkit: the student hat as a way to support teachers to consider the cognitive and affective experiences of students as they do this work.

While the arguments in paper two were conceptual, the third paper provided empirical backing to the idea that the student hat can help teachers learn about students' experiences during storyline curricular units. While I proposed that student hat might help teachers to consider their students' experiences in a storyline unit, I found that it actually helped teachers to learn about their students' experiences, the science ideas, and the OpenSciEd curricular approach. It helped teachers reach these goals in two ways: first, by creating a safe space for questions and confusions so that teachers could learn about science ideas without feeling judged by the power dynamics frequent in teacher PD (Finkelstein et al., 2019), and second by building their epistemic empathy for students (Jaber et al., 2018). Epistemic empathy helped teachers to consider what it feels like to engage in scientific sensemaking as a student, which can be both uncomfortable and joyful given how different the curricular approach is from what students and teachers are used to (Han & Gutierez, 2021; Jaber & Hammer, 2016). The construct of epistemic empathy is relatively new, and current work suggests that it can be built by giving teachers opportunities to witness and critically reflect on new forms of instruction (Jaber, 2021). Paper three expands on that work, proposing that the student hat is another way to build epistemic empathy and therefore support teachers to consider how to engage their students in true scientific sensemaking in their classroom.

Implications for Research, Practice, and Policy

Findings from this dissertation have implications in terms of how we continue to research in-service science teacher learning, the practice of teacher education, and policies around supporting instructional reform more broadly. In terms of research, this work demonstrates that we need to better understand the process of learning about and adopting new science curricula and how practice-based activities like the student hat can support that work over time. In terms of practice, this work points out some of the gaps in our existing systems of in-service teacher learning and suggests ways we might address those gaps to continue to support teachers as they engage in teaching and instructional reform. Finally, it is important to note that all of this teacher learning work takes place in a policy context that might value or work against moves towards effective science teaching. As teacher educators and researchers, we must keep in mind how we can advocate for policies that help teachers do the best work they can, which will ultimately support students across the K-12 education system.

Implications for Research

The process of teacher learning is complex, multifaceted, and involves more than simply accumulating knowledge of effective teaching practice (Ball & Cohen, 1999; Clarke & Hollingsworth, 2002). It makes sense, therefore, that we need to understand how in-service teachers develop their practice through both formal professional development experiences and in the normal course of teaching. Based on the work I have done here, I see implications for research in both of these areas.

In terms of better understanding the processes of learning during formal teacher professional development, there is much more that needs to be done around the student hat activity, what it provides for teacher learning, and how it can be facilitated. I have proposed that it might help teachers to develop epistemic empathy for their students (Jaber et al., 2018), and in so doing push on questions of how students feel when engaged in complex scientific sensemaking that are often ignored by science teachers and teacher educators (Jaber & Hammer, 2016). Given the novelty of this activity, however, we need to know more about what the student hat can do, how to use it, and when. For example, the process of facilitation of the student hat is an open question. How do facilitators lead student hat activities differently than other kinds of activities and what impact does that have on the way teachers experience them? Are there particular moves or choices that support or inhibit teachers from building epistemic empathy during PD sessions using the student hat? What are the prerequisites for creating a space in which teachers feel comfortable engaging in student hat and see its value for their learning? These are all open questions about the facilitation of professional development that will help us to better understand the actual use of these kinds of activity in formal teacher learning.

In addition to formal teacher learning opportunities, this dissertation demonstrates the need for us to do more research on professional learning systems more broadly and how they can be designed to support teachers engaging in instructional reform aligned with the NGSS. All of the data here was taken during or shortly after formal PD sessions, but we know that teachers' thinking is heavily influenced by their experiences implementing curricular materials (Kazemi & Hubbard, 2008; Webster-Wright, 2009). This means we need to understand how teachers implement new materials like OpenSciEd, what resources they access or do not access, and which ideas introduced in professional development they use to support their own understanding of strong practice. The professional development workshop is an important but ultimately relatively brief component of the broad array of tasks that make up teacher learning (Loucks-Horsley et al., 2010). In order to understand how work during the PD session might impact and be impacted by classroom practice, we must spend time in classrooms themselves. Future research should include following teachers into their classrooms after the PD in order to understand which specific ideas and practices they take up from PD and how, which resources from PD they often turn to for support and why, and the characteristics of organizational contexts that support or impede teachers in making the ambitious instructional reforms the science education community is advocating for.

Implications for Practice

By taking a range of methodological approaches to studying teacher learning, I was able to look at it over multiple time scales, providing implications for practice at both system and workshop levels. The large-scale study of teacher beliefs and confidence over time helped to clarify how the particular OpenSciEd professional learning *system* did and did not support teacher growth, providing us with valuable insights as we continue to plan such systems in the future. Zooming into the *workshop* level, the analysis of the student hat provides some guidance for practice-based PD providers to add another activity tool to their kit that is particularly targeted at helping teachers to understand the experience of engaging in scientific sensemaking.

Teacher learning is best supported when it is part of a thoughtful, planned, and comprehensive system of support that comes in a diverse array of forms, including PD workshops, teacher learning groups, and teacher research activities (Cochran-Smith & Lytle, 2015; Horn & Kane, 2015; Loucks-Horsley et al., 2010; Webster-Wright, 2009). The studies in this dissertation took place in the context of one such system, in which teachers participated in a series of multi-day workshops while implementing new curricular materials over the course of multiple years. This system clearly had benefits. Even the existence of multiple touch points for teacher learning over time is beneficial, and a clear departure from many PD workshops that exist as independent, unrelated, one-offs with little connection to the work teachers are doing in their classrooms (Darling-Hammond et al., 2017). In this particular case, teachers' beliefs and confidence changed over time, demonstrating that teachers did benefit from repeated engagement in OpenSciEd professional development. At the same time, however, these benefits did not change at the same rate, reinforcing the value of including iterative cycles of learning, enactment, and reflection in PD systems in order to facilitate change in teachers' confidence (Clarke & Hollingsworth, 2002; McNeill et al., in press). In addition, I found that teachers who rated the PD activities more valuable showed lower levels of traditional beliefs and faster increases in confidence. This implies that PD providers should establish trust with their participants so that participants understand the reasoning and goals of practice-based activities, as that might help teachers to learn more from those activities over time (Avidov-Ungar, 2016; Masuda et al., 2013).

At the workshop level, this work has pointed out where our current suite of practice-based professional development activities might have a gap: ways to help teachers actively consider what it is like for students to bring their own experiences, ideas, and questions to science class and have those used as resources for collaborative sensemaking rather than misconceptions to be "fixed" (Bang et al., 2017; Campbell et al., 2016; Rosebery et al., 2016). Student hat can potentially fill that gap, supporting teachers to expand their conceptions of what scientific sensemaking might look like from students and increasing their ability to notice and honor that sensemaking in all students, but particularly those who have been historically marginalized by Western science (Bang et al., 2012; Rosebery et al., 2016). Student hat is not a panacea; it cannot serve every possible teacher learning goal in a PD workshop. Nevertheless, student hat can be a valuable way to help teachers to understand what it looks and feels like to do true scientific sensemaking in the K-12 classroom.

Implications for Policy

While this dissertation has focused on teacher learning during and across PD workshops, it is important to note that this learning takes place in a larger schooling and accountability system. As a science education community, we are developing a shared vision of K-12 science instruction that better reflects the work actually done by scientists while also valuing the ideas, experiences, and sensemaking practices that students bring to the classroom with them (Campbell & Lee, 2021; National Research Council, 2012). This is a powerful vision and can be used to help teachers support instruction that is more equitable and meaningful for students (Bang et al., 2017). We need to understand how to best support teachers in achieving that vision, but we also need to recognize that professional learning activities alone will not be sufficient to meaningfully support that change. We know that teachers make sense of learning given the accountability contexts in which they work, and teachers will shift their instruction to match the particular forms of accountability to which they are subject (Allen & Penuel, 2015; Lowenhaupt et al., 2021). As science educators, therefore, we need to advocate for policies and structures at the school, district, and state levels that truly value classrooms in which students are making sense of natural phenomena, and in so doing ensure that the work done in PD is supported and extended throughout teachers' work.

Conclusion

This is an important time in science education in this country. While there is broad agreement in the science education research community on what strong science instruction looks like, we still fail to see that implemented widely in classrooms in the United States (Campbell &

Lee, 2021). Part of this could be due to lack of available curricular materials that actually support teachers in teaching science as a process of figuring out the natural world (Achieve, 2018; Campbell & Lee, 2021; Lowell et al., 2021). Thankfully, recent curricular materials have shown promise in terms of their quality and availability to teachers across the country. While the work of curriculum design is certainly not finished, this availability means we can consider in earnest how to support teachers in implementing these curricula in ways that accurately reflect the vision they were designed for.

Without support, teachers may implement even high-quality curricular materials in traditional ways (Alozie et al., 2010; McNeill et al., 2017). Therefore, these early days of curriculum design and adoption are the perfect time to ensure we apply our understanding of teacher learning to avoid these pitfalls. We can build on the work that has already been done in practice-based teacher professional development (Ball & Cohen, 1999; Lampert, 2010) to consider how we can support teachers to develop their relationship with science as a process of figuring out and with their students in the classroom. But we also have an opportunity to help teachers understand what it *feels like* when students bring their own ideas and experiences into a classroom, those ideas are used as legitimate tools to support science sensemaking, and classrooms build together a mutual understanding of some interesting or unexpected natural phenomenon. If we can do that work in the teacher learning space, we are more likely to see it take place in K-12 classrooms, and we can continue on in the work of supporting all students in seeing themselves as valuable scientific thinkers.

References

Achieve. (2018). *NGSS alignment claims: How publishers talk about the Next Generation Science Standards*. https://www.nextgenscience.org/NGSSclaims

Allen, C. D., & Penuel, W. R. (2015). Studying teachers' sensemaking to investigate teachers' responses to professional development focused on new standards. *Journal of Teacher Education*, *66*(2), 136–149. https://doi.org/10.1177/0022487114560646

Alozie, N. M., Moje, E. B., & Krajcik, J. S. (2010). An analysis of the supports and constraints for scientific discussion in high school project-based science. *Science Education*, *94*(3), 395–427. https://doi.org/10.1002/sce.20365

Avidov-Ungar, O. (2016). A model of professional development: Teachers' perceptions of their professional development. *Teachers and Teaching*, *22*(6), 653–669. https://doi.org/10.1080/13540602.2016.1158955

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (Vol. 1, pp. 3–22). Jossey-Bass.

Bang, M., Brown, B., Calabrese Barton, A., Rosebery, A., & Warren, B. (2017). Towards more equitable learning in science: Expanding relationships among students, teachers, and science practices. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 33–58). NSTA Press.

Bang, M., Warren, B., Rosebery, A. S., & Medin, D. (2012). Desettling expectations in science education. *Human Development*, *55*, 302–318. https://doi.org/10.1159/000345322

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3–15. https://doi.org/10.3102/0013189X033008003

Brophy, J. (Ed.). (2004). *Using video in teacher education* (1st ed). Elsevier/JAI.

Bybee, R. W. (2014). NGSS and the next generation of science teachers. *Journal of Science Teacher Education*, *25*(2), 211–221. https://doi.org/10.1007/s10972-014-9381-4

Campbell, T., & Lee, O. (2021). Instructional materials designed for A Framework for K-12 Science Education and the Next Generation Science Standards: An introduction to the special issue. *Journal of Science Teacher Education*, *32*(7), 727–734. https://doi.org/10.1080/1046560X.2021.1975359

Campbell, T., Schwarz, C., & Windschitl, M. (2016). What we call misconceptions may be necessary stepping-stones toward making sense of the world. *Science Scope*, *39*(7), 19–24.

Chen, G., Chan, C. K. K., Chan, K. K. H., Clarke, S. N., & Resnick, L. B. (2020). Efficacy of video-based teacher professional development for increasing classroom discourse and student learning. *Journal of the Learning Sciences*, *29*(4–5), 642–680. https://doi.org/10.1080/10508406.2020.1783269

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(8), 947–967. https://doi.org/10.1016/S0742-051X(02)00053-7

Cochran-Smith, M., & Lytle, S. L. (2015). *Inquiry as stance: Practitioner research for the next generation*. Teachers College Press.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Finkelstein, C., Jaber, L. Z., & Dini, V. (2019). “Do I feel threatened? No… I’m learning!”—Affective and relational dynamics in science professional development. *Science Education*, *103*(2), 338–361. https://doi.org/10.1002/sce.21489

Gibbons, L. K., & Cobb, P. (2017). Focusing on teacher learning opportunities to identify potentially productive coaching activities. *Journal of Teacher Education*, *68*(4), 411–425. https://doi.org/10.1177/0022487117702579

Grossman, P. (Ed.). (2018). *Teaching core practices in teacher education*. Harvard Education Press.

Han, M., & Gutierez, S. B. (2021). Passive elementary student’s constructed epistemic emotions and patterns of participation during small group scientific modeling. *Science Education*, *105*(5), 908–937. https://doi.org/10.1002/sce.21665

Horn, I. S., & Kane, B. D. (2015). Opportunities for professional learning in mathematics teacher workgroup conversations: Relationships to instructional expertise. *Journal of the Learning Sciences*, *24*(3), 373–418. https://doi.org/10.1080/10508406.2015.1034865

Jaber, L. Z. (2021). “He got a glimpse of the joys of understanding” – The role of epistemic empathy in teacher learning. *Journal of the Learning Sciences*, *30*(3), 433–465. https://doi.org/10.1080/10508406.2021.1936534

Jaber, L. Z., & Hammer, D. (2016). Learning to feel like a scientist. *Science Education*, *100*(2), 189–220. https://doi.org/10.1002/sce.21202

Jaber, L. Z., Southerland, S., & Dake, F. (2018). Cultivating epistemic empathy in preservice teacher education. *Teaching and Teacher Education*, *72*, 13–23. https://doi.org/10.1016/j.tate.2018.02.009

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional development: Attending to the coevolution of teachers' participation across contexts. *Journal of Teacher Education*, *59*(5), 428–441. https://doi.org/10.1177/0022487108324330

Kelley-Petersen, M., Davis, E. A., Ghousseini, H., Kloser, M., & Monte-Sano, C. (2018). Rehearsals as examples of approximation. In P. Grossman (Ed.), *Teaching core practices in teacher education* (pp. 85–105). Harvard Education Press.

Lampert, M. (2010). Learning teaching in, from, and for practice: What do we mean? *Journal of Teacher Education*, *61*(1–2), 21–34. https://doi.org/10.1177/0022487109347321

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*(3), 226–243. https://doi.org/10.1177/0022487112473837

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2010). *Designing professional development for teachers of science and mathematics*. Corwin Press. https://doi.org/10.4135/9781452219103

Lowell, B. R., Cherbow, K., & McNeill, K. L. (2021). Redesign or relabel? How a commercial curriculum and its implementation oversimplify key features of the NGSS. *Science Education*, *105*(1), 5–32. https://doi.org/10.1002/sce.21604

Lowenhaupt, R., McNeill, K. L., Katsh-Singer, R., Lowell, B. R., & Cherbow, K. (2021). *The instructional leader's guide to implementing K-8 science practices*. ASCD.

Masuda, A. M., Ebersole, M. M., & Barrett, D. (2013). A qualitative inquiry: Teachers' attitudes and willingness to engage in professional development experiences at different career stages. *Delta Kappa Gamma Bulletin*, *79*(2), 6–14.

McNeill, K. L., Affolter, R., & Reiser, B. J. (in press). Anchoring science professional learning in curriculum materials enactment: Illustrating theories in practice to support teacher learning. In A. Castro Superfine, S. R. Goldman, & M.-L. M. Ko (Eds.), *Teacher learning in changing contexts: Perspectives from the learning sciences*. Routledge.

McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2017). Moving beyond pseudoargumentation: Teachers' enactments of an educative science curriculum focused on argumentation. *Science Education*, *101*(3), 426–457. https://doi.org/10.1002/sce.21274

McNeill, K. L., & Knight, A. M. (2013). Teachers' pedagogical content knowledge of scientific argumentation: The impact of professional development on K-12 teachers. *Science Education*, *97*(6), 936–972. https://doi.org/10.1002/sce.21081

National Academies of Sciences, Engineering, and Medicine. (2015). *Science teachers' learning: Enhancing opportunities, creating supportive contexts*. The National Academies Press. https://doi.org/10.17226/21836

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press. https://doi.org/10.17226/13165

Rosebery, A. S., Warren, B., & Tucker-Raymond, E. (2016). Developing interpretive power in science teaching. *Journal of Research in Science Teaching*, *53*(10), 1571–1600. https://doi.org/10.1002/tea.21267

Schwarz, C. V., Passmore, C., & Reiser, B. J. (2017). Moving beyond "knowing about" science to making sense of the natural world. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 3–21). NSTA Press.

Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (Eds.). (2011). *Mathematics teacher noticing: Seeing through teachers' eyes*. Routledge.

Sherin, M. G., & van Es, E. A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education*, *13*(3), 475–491.

Stroupe, D., Hammerness, K., & McDonald, S. (Eds.). (2020). *Preparing science teachers through practice-based teacher education*. Harvard Education Press.

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research*, *79*(2), 702–739. https://doi.org/10.3102/0034654308330970