Examining How Classroom Communities Developed Practice-Based Epistemologies for Science Through Analysis of Longitudinal Video Data

Christina Krist
University of Illinois at Urbana-Champaign

Recent reforms in science education emphasize having students develop and refine core disciplinary ideas through participation in science knowledge–building practices. Supporting students’ meaningful participation in these practices is challenging, in part because our understanding of how this kind of participation develops is underexplored. This paper characterizes the epistemic dimension of classroom scientific practice—the underlying assumptions and criteria guiding students’ knowledge-building work—in moment-to-moment interactions, and how their use of these criteria shifted over time. These patterns of change illustrate how, by 8th grade, students were more consistently using sophisticated disciplinary forms of epistemic criteria than they were in 6th grade. The shifts toward disciplinary sophistication were cumulative over time and across content areas, suggesting that they reflected a shift in something other than content knowledge gains. Yet these cumulative shifts did not occur in a clean, predictable progression toward sophistication. This study documents the details of these shifts over time and across content areas. By providing an empirical account of the evolution of students’ knowledge-building work in practice over time, this study argues for the centrality of epistemic learning goals in science education and proposes implications for how to measure and support students’ meaningful participation in scientific practices in contextually valid ways.

Educational Impact and Implications Statement

Being a scientist involves not only learning science ideas, but knowing how to use the practices and tools of science to build knowledge. Although we know how students’ science content ideas develop, we know less about how they learn to build science knowledge themselves over time. This study documents how middle school students’ participation in science practices evolved over the course of three years, focusing on how they developed mechanistic accounts and justified claims. It shows that students engaged deeply in this knowledge-building work in the context of a curriculum that supported their engagement in authentic science tasks. In addition, their engagement gradually became more aligned to the work that scientists do, albeit in complex and unexpected ways. This evolution occurred across science content area units and across teachers, suggesting that learning how to develop mechanistic explanations and justify ideas can carry over and continue to build from unit to unit and from classroom to classroom; and it emphasizes the importance of practice-based learning goals, in addition to content-focused learning goals, in science education.

Keywords: epistemic cognition, epistemology, science practices

Supplemental materials: http://dx.doi.org/10.1037/edu0000417.supp

Recent reforms in science education, including the Next Generation Science Standards, emphasize participation in scientific practices as the means by which students develop and use scientific ideas (National Research Council, 2012; NGSS Lead States, 2013). These reforms conceptualize science as knowledge-building practice: an ever-evolving collection of ways of interacting with the world in order to explain natural phenomena (Dunbar, 1999; Nersessian, 1992; Rouse, 2002; Salmon, 1984; Thagard, 1993). Following from theories conceptualizing learning as participation in social practice (Lave & Wenger, 1991; Rogoff, 2003; Sfard, 1998), these reforms advocate a shift in how school science is done: from students learning about to figuring out core science

This research was supported by National Science Foundation Grant DRL-1020316 to Northwestern University and by the NAEd/Spencer Dissertation Fellowship Program. Thanks to Brian Reiser for his invaluable guidance and support in directing this dissertation project; to Kelsey Edwards for her contributions in conducting and managing data processing and coding; and to Andrew Elby, Leema Berland, Joshua Rosenberg, Elizabeth Dyer, and Enrique Suárez for their helpful feedback on earlier drafts of this article.

Correspondence concerning this article should be addressed to Christina Krist, Department of Curriculum and Instruction, University of Illinois at Urbana-Champaign, Champaign, IL 61820. E-mail: ckrist@illinois.edu
ideas by iteratively engaging in investigations, modeling, and explanation-building through argumentative discourse (Schwarz, Passmore, & Reiser, 2017).

Although the idea that students should be involved in authentic scientific inquiry is not new (e.g., Mann, 1912), implementations of inquiry in science classrooms have often been essentialized to the point of becoming rote. For example, the practice of designing and conducting investigations has often been taught as the five-step scientific method: a set of five decontextualized steps that are often applied without regard for the nature of the question being answered or the nature of the knowledge being produced. Though this method is familiar to most of us from our own schooling, such decontextualized and routinized sets of steps resemble little of the actual intellectual activity of science (Lehrer & Schauble, 2006; Windschitl, Thompson, & Braaten, 2008). The practice turn in science education instead recognizes that scientific work involves an ensemble of interrelated social, conceptual, and material practices used in concert to build knowledge about the natural world (Duschl, 2008; Ford, 2015; Forman, 2018; Stroupe, 2015). The emphasis on articulating distinct yet interrelated practices in the reforms such as the NGSS is an attempt to remediate the rote ways of engaging students in inquiry that characterize most of K–12 science education (Bamijlover, Boyd, Pasley, & Weiss, 2006; Lemke, 1990). Thus, the practice turn attempts to inextricably link the doing of science with the ideas of science by having students engage in practices in order to construct ideas about the natural world. Accordingly, science learning goals are no longer only about measuring content knowledge; learning is about coming to participate in these practices in increasingly sophisticated ways.

Engaging students in science knowledge-building practices—and knowing how to determine whether and how students are participating in them meaningfully rather than by rote—is challenging, in part because our understanding of how this kind of participation in practices develops is underexplored. As such, we have some important candidate anchors for what sophisticated science knowledge-building looks like and involves (e.g., Berland et al., 2016; Ryu & Sandoval, 2012; Smith, Maclin, Houghton, & Hennessey, 2000) and evidence that students’ participation can move from lower to upper anchors over time (e.g., Schwarz et al., 2009; Smith et al., 2000); but we have a limited understanding about the processes by which students come to build science knowledge in sophisticated ways, or metrics for measuring that learning-in-progress across practices and contexts. Understanding these processes will help us recognize progress and allow us to develop more specified means of supporting students’ participation in knowledge-building practices.

A key learning goal underpinning current reforms is that in addition to doing scientific practices, students should come to understand the hows and whys undergirding those practices: the practical criteria for how to build knowledge (Sandoval, 2005), how and why those criteria are useful for doing so (Manz, 2015a; Russ, 2014), and when and how to apply those criteria productively (Barzilai & Chinn, 2018; Elby & Hammer, 2010). These how and why understandings are proposed as the thing that students carry with them across science practices and contexts (Berland et al., 2016; Hofer, 2006; Schommer & Walker, 1995). For example, the criteria that mechanistic explanations are typically better than nonmechanistic accounts is useful to draw upon when designing an investigation and when revising a model; and it is useful while working to explain both how and why an odor travels across a room and why there was suddenly a fish kill in Lake Michigan. Accordingly, students’ understanding of these hows and whys is one candidate metric for gauging students’ learning-in-process.

The process of coming to use these hows and why understandings meaningfully can be described as a type of epistemic cognition, or a process of “how people acquire, understand, justify, change, and use knowledge” (Greene, Sandoval, & Bråten, 2016, p. 1). This term encompasses a diversity of scholarship in educational psychology and the learning sciences that ranges from a focus on individuals’ thinking and learning (e.g., personal epistemology: Hofer & Pintrich, 1997; epistemic beliefs: Muis, 2007; Schommer, 1994; metacognition: Hofer, 2004; Kuhn, 1999, 2000) to a focus on the joint construction and negotiation of epistemic criteria and practices through social activity (e.g., social cognition: Clément, 2016; practical epistemologies: Wickman, 2004; epistemic practices: Kelly, 2010). I focus in this paper on the latter: the epistemic criteria used by a group that guide their epistemic practices, and how the collective uses of those criteria shift over time.

This study builds upon decades of design-based research in the learning sciences that has explored how to productively engage learners in disciplinary practices in K–12 classroom settings (e.g., Brown & Campione, 1996; Cobb, Stephan, McClain, & Greameijer, 2001; Engle & Conant, 2002; Lehrer & Schauble, 2006). From studies of learning in these designed contexts, we know that shifts in students’ epistemic beliefs can occur through participation in rich disciplinary practices (Kuhn, Shaw, & Felton, 1997; Smith et al., 2000). In addition, careful research has begun to characterize the epistemic practices that classroom communities develop as they participate in science practices over time (e.g., Manz, 2012, 2015b; Pierson, Clark, & Sherard, 2017; Ryu & Sandoval, 2012; Schwarz et al., 2009). This study contributes to this line of research, adding an empirical description of middle school students’ use of two specific epistemic criteria in moment-to-moment classroom interactions, and how those criteria were taken up and reconfigured by the classroom community over the course of three content-area units across three years.

Moreover, this study demonstrates that it is possible to investigate epistemic practices longitudinally in order to observe incremental patterns of change drawn from close qualitative detail of students’ activity. These patterns of change illustrate how, by 8th grade, students were more consistently using sophisticated disciplinary forms of epistemic criteria. The shifts toward disciplinary sophistication were cumulative over time and across content areas, suggesting that they reflected a shift in something other than content knowledge gains. Yet these cumulative shifts did not occur in a clean, predictable progression toward sophistication. This study contributes to our growing understanding of how classroom communities develop epistemic understandings that guide their practice and has implications for how to support students’ meaningful engagement in scientific practices.

Theoretical Framing

One key decision when defining and measuring epistemic cognition in science is about the substance of that epistemic cognition. What are the epistemic ideas that matter for doing science? A
second key question concerns how those epistemic ideas are developed and expressed. What counts as epistemic change, and how can we see it?

The epistemology of a discipline can be characterized by how that field answers the questions summarized by Duschl and Osborne (2002) and Sandoval (2005): What exactly do we know? How do we know what we know? And why do we believe it? Generally speaking, scholars agree that having some understanding of the scientific answers to these questions should be a central part of science education. A commonly cited problem plaguing science education is the “mile wide, inch deep” curriculum in which students learn science as a set of disconnected facts (National Research Council, 2012; Schmidt, Raisen, Britton, Bianchi, & Wolve, 1997). This is a problem in that it neither prepares students to be scientifically literate citizens (e.g., Feinstein, 2011) nor prepares them to work in STEM fields (e.g., Cooper et al., 2015). In response, there is general consensus that students should gain a sense for how science knowledge comes to be—answers to the epistemic questions identified above—rather than only learning the facts and conclusions of that knowledge.

However, there is substantial disagreement on the particulars of how and why students should gain a sense of the discipline, on what form that sense should take, and how it can be learned over time. The answers to these questions tend to come from three distinct perspectives, each of which is reflected in research on both epistemic cognition and science education. One set of answers comes from a strand of research focused on the nature of science (NOS), or a description of science as an endeavor that students should come to understand. A second set of answers comes from research on personal epistemology, which focuses on individuals’ beliefs and implicit theories about knowledge and knowing and how those beliefs develop over time within an individual’s development. Finally, a third set of answers comes from research on science as social practice, which emphasizes the contextual and distributed nature of knowledge and knowledge construction. In this section, I describe similarities and differences in each of these theoretical perspectives along three key dimensions: (a) the goal of having students understand the nature of scientific knowledge; (b) what form students’ epistemic conceptions of the nature of the discipline take; and (c) the (implicit) definition of epistemic change.

Understanding the Nature of Science to Learn About Professional Science

Research on the nature of science (NOS) focuses on students’ views about the enterprise of professional science, such as that scientific knowledge is tentative and relies on an empirical basis (Lederman, 2007; Millar & Osborne, 1998; Smith & Scharmann, 1999). The motivation behind having students develop these views is to move away from views of science as factual information and toward viewing it as a human endeavor, often connected to the idea that students need to understand the processes and commitments by which scientific knowledge has come to be in order to participate as informed scientific citizens (Abd-El-Khalic, 2003; Alchlin, 2011; American Association for the Advancement of Science, 1993; Holbrook & Rannikmae, 2007; Lederman, Antink, & Bartos, 2014; National Research Council, 1996).

Despite the insistence on constructivist views of the nature of science (e.g., McComas & Olson, 1998; Tsai, 2002), this perspective tends to hold an empiricist-oriented view of the nature of one’s epistemic conceptions of NOS (Eflin, Glennan, & Reisch, 1999). It assumes discrete universal truths that describe the discipline of science, and that one either knows these truths or does not. Accordingly, this version of NOS is often measured by surveys or interview questions that ask students to produce or choose between options of these discrete statements about the professional scientific enterprise (e.g., Lederman, Abd-El-Khalic, Bell, & Schwartz, 2002; Redish, Saul, & Steinberg, 1998). From these measures, we know that students’ ideas about NOS can and do shift over time, often after direct instruction about NOS (Abd-El-Khalil & Lederman, 2000; Akerson, Nagum-Joshi, Weiland, Pongsanon, & Avsar, 2014; Schwartz, Lederman, & Crawford, 2004).

This view of conceptions of NOS leads to implications for epistemic change that emphasize explicit teaching of NOS ideas. Students are not tabula rasa, but their prior knowledge is not useful for developing such views: because students are not scientists and have not experienced the work of professional scientists, they do not have relevant prior experiences to draw upon to help them develop views that align with the nature of the discipline. Therefore, the only viable process of epistemic change from this perspective is one that provides experiences that call into question students’ nonconstructivist views of NOS and replaces them with constructivist views, akin to cognitive conflict approaches to conceptual change (see Limón, 2001 for a relatively recent review).

A parallel to a cognitive conflict approach is often the implicit line of argument guiding NOS intervention studies examining epistemic change. These studies often portray students and teachers as epistemologically deficient: they lack useful understanding or hold naive views of NOS prior to any instructional intervention (e.g., Akerson & Donnelly, 2010; Bartos & Lederman, 2014; Deng, Chen, Tsai, & Chai, 2011; Hashweh, 1996; Mesci & Schwartz, 2017; Wahbeh & Abd-El-Khalic, 2014). Accordingly, it is argued that NOS needs to be taught through a combination of explicit instruction about NOS (and, in some cases, about conceptual change theories; Tsai, 2006) and extensive reflection on “philosophically valid” inquiry experiences (Abd-El-Khalil, Bell, & Lederman, 1998; Lederman & Lederman, 2014; Schwartz et al., 2004). In other words, instruction should provide students with opportunities to experience the tentativeness and human-centeredness of doing science; reflect on those experiences; and connect...

1 The academic conversation around NOS is certainly multifaceted and includes perspectives and interventions grounded in careful historical and philosophical analyses of science (e.g., Allchin, 2011; Eflin et al., 1999; Feinstein, 2011; Matthews, 2012, 2014). These works often critique the very premise that I use to characterize NOS here: that the goal of understanding the nature of science is to learn about professional science in and of itself, rather than to engage in evaluation and active decision-making around scientific claims. At the same time, the research program founded on NOS that I describe here has continued to be highly influential in science education (Duschl & Grandy, 2013). Even amidst efforts to integrate philosophical and historical critiques into the NOS research agenda (e.g., Erduran & Daghe, 2014), this program stands in strong epistemological opposition to the bodies of work described in the following sections. It is because of the prominence of this perspective that I focus extensively on this particular NOS tradition.
those experiences to an external list of characteristics describing NOS. Studies from these contexts demonstrate that these interventions do impact students’ views of NOS (Bell, Blair, Crawford, & Lederman, 2003; Khishfe & Abd-El-Khalick, 2002; Schwartz et al., 2004). However, because measures of NOS ask about professional science, it is unclear exactly what it is about students’ epistemic understanding that is changing or how it might impact their future inquiry practices or engagements with science.

Using Personal Epistemologies for Doing Science

A second perspective builds on foundational research on personal epistemologies, or individuals’ beliefs and implicit theories about knowledge and knowing and how those beliefs develop over time (Hofer & Pintrich, 1997, 2002). The majority of this work has examined the developmental trajectories of personal epistemologies, often described in terms of a series of stages. People move from an absolutist stage where knowledge is either right or wrong, to a multiplist stage that views knowledge as subjectively constructed and allowing for multiple valid perspectives (Baxter Magolda, 1992; Kuhn, 2001; Perry, 1970). A subset of this work focuses on personal epistemologies in a science context and also describes an evaluativist stage in which knowledge is constructed, but there are established criteria that allow one to make judgments about the relative validity or strength of competing claims (Kuhn, Cheney, & Weinstock, 2001).

Another subset of personal epistemology research conceives of one’s personal epistemology as multidimensional rather than developmentally stage-based (e.g., Hammer & Elby, 2002; Hofer, 2004; Schommer, 1990), though perspectives differ as to how independent the dimensions are from one another. Multidimensional models emphasize the context-dependence of one’s expressed or enacted epistemologies, acknowledging the specifics of the context as influencing the sets of specific epistemological resources or beliefs that are activated or expressed at any given time.

The context-specificity of these models has raised an ongoing debate about the domain specificity of epistemic cognition: are there beliefs about knowledge and knowing that are particular and apply only in the domain of science? Or are personal epistemologies domain-independent, reflecting general views of knowledge and knowing that can be applied in many situations? Elby, Macrander, and Hammer (2016) note that “most current research on personal epistemology in science] reflects the assumption that the epistemological views expressed and enacted in a science context are at least partly views about science in particular as opposed to domain-independent epistemological views that happen to be on display in a science context” (p. 115). In other words, there may often be conflation or a lack of clarity about whether the epistemic beliefs that students express or enact are reflections of scientific epistemologies, or whether they are components of a more general epistemology as it relates to learning in a classroom setting—one’s “personal pedagogy.” Scholars vary as to the extent to which they think this is of concern (Elby, 2009; Lederman, 2007; Sandoval, 2005, 2009). It is likely that both components are at play. For example, students who express epistemological beliefs that reflect Kuhn et al.’s (2001) evaluativist stage may hold domain-general beliefs about knowledge, but leverage criteria for evaluation that are specific to the disciplinary context.

In any case, the goal of having students develop sophisticated epistemologies from multidimensional perspectives of epistemologies is typically because sophisticated epistemologies are useful for both the kind of knowledge-building work that constructivist learning environments aim to support, and the kind of work that scientists do (Elby & Hammer, 2001; Lising & Elby, 2005; Russ, 2014; Scherr & Hammer, 2009). In other words, the goal of having students develop epistemologies for science is to help them become better at constructing knowledge themselves. From this view, epistemologies are fundamentally pragmatic in that they account for what people actually think and how they actually reason in practice. Accordingly, students’ epistemologies are practical, in that they are ideas that tacitly guide action and are therefore visible in students’ activity (Sandoval, 2005); that is, they are both visible in and, account for, the actions and decisions that students make as they are constructing knowledge products. Therefore, an important part of learning science is learning when, how, and why to productively wield the epistemic tools of science—many of which may be at least partially domain-general—for science knowledge building (Berland et al., 2016; Manz, 2015a; Russ, 2014).

Epistemologies for Science as Characteristics of Disciplinary Communities’ Social Practice

Finally, a third perspective maintains the context-dependent (yet often domain-agnostic) nature of epistemologies and the goals of practical utility described in the Using Personal Epistemologies for Doing Science section but theorizes that these “practical epistemologies” exist within the social context rather than within the individual (Kelly, McDonald, & Wickman, 2012; Lundqvist, Almqvist, & Östman, 2009; Östman & Wickman, 2014; Wickman, 2004). Rather than aiming to ascribe particular epistemological beliefs, views, or resources to individuals, this approach investigates how epistemologies are manifested in action and discourse, and are jointly constructed and reconstructed over time through interaction with one another and with features of the context (e.g., physical materials, social norms and expectations, implicit power dynamics). Epistemic change, from this perspective, is reflected in shifts in individuals’ participation (Lave & Wenger, 1991; Wenger, 1998) and in the interactional system (Cole, 1998; Greeno, 2011; Rogoff, 2003) over time.

A situated, practice-based view of epistemology requires accounting for epistemic change at both the individual and group level, and the dynamic interactions between those levels (Elby et al., 2016; Sandoval, 2009, 2014). This theory-building work is still in progress, and has often utilized a science subject-area context. Studies that carefully document students’ science activity in practice, often coupled with analysis of the artifacts they produce, have begun to investigate the productiveness of participation in particular epistemic practices by examining how the nature of students’ practice codevelops with their conceptual knowledge (Manz, 2012; Pierson et al., 2017); how consideration of the purpose and goals of particular contexts influence the nature of students’ participation in science practices (Berland & Crucet, 2016; Chinn, Buckland, & Samarapungavan, 2011; Rosenberg, Hammer, & Phelan, 2006); and what students think about their own scientific practice and products (Chinn & Brewer, 2001; Pluta, Chinn, & Duncan, 2011).
This work has produced some candidate design principles about how to effectively support students’ epistemic cognition. These design conjectures include: framing learning activities such that they activate and build upon students’ epistemological resources for science and for learning (Elby & Hammer, 2010; Hammer & Elby, 2003); providing opportunities to revise knowledge products or introduce complexities from the world in ways that make salient the need for disciplinary epistemic criteria (Manz, 2015b; Pierson et al., 2017; Ryu & Sandoval, 2012); positioning students as the ones who are making key decisions about whether and how these criteria should be developed and used (Ryu & Sandoval, 2012; Stroupe, 2014); and providing repeated opportunities to consistently engage in this kind of system, such that the stabilization, or consistent use, of particular sets of epistemological resources is possible (Elby, 2009; Elby & Hammer, 2010). However, only a handful of studies (Lidar, Lundqvist, & Östman, 2006; Ligozat, Wickman, & Hamza, 2011) examine the interplay between teachers’ epistemological moves and messages in interactions with students and the practical epistemologies of the classroom community (see Russ’ (2018) call for this); and even fewer examine the emergent stabilization of practical epistemologies over extended periods of time (cf. Lehrer, Schauble, & Lucas, 2008; Ryu & Sandoval, 2012). Put simply, we are still very much in the process of building a knowledge base around understanding shifts in students’ epistemic practices, especially over extended time periods.

Perspectives on Epistemology Used in This Paper

This paper focuses on the development of classroom communities’ epistemic practices across three years. It takes a view of the form of epistemologies as situated and evident in students’ interactions and actions. In addition, I focus on epistemologies as characteristics of group interaction: what kind of epistemic ideas does the classroom community take up, reconfigure, and use over time? In adopting this perspective, I do not take the stand that epistemologies are only found in interaction, or that individuals do not hold epistemic ideas or beliefs or engage in individual processes of epistemic cognition. Although this paper does not explore these interactions, examining the relationships between individual beliefs and participation and group-level practices and dynamics is a critical area for future work.

I focus on two specific epistemic ideas that are highlighted as key components of epistemological cognition generally (i.e., that are at least partly domain-general), and emphasize particular versions of these ideas are central to science. Following Berland et al. (2016), I call these two ideas Nature of Account, or the kind of accounts the classroom community is working to provide; and Justification, or how the classroom community justifies the claims they’re working out.

Epistemic practices related to the Nature of Account in science often focus on the construction of mechanistic accounts. Mechanistic accounts move beyond a pure description of the phenomenon by providing a theoretical account of the regularities behind these empirical observations. This theoretical dimension—proposing and reasoning through potential mechanisms—provides specific ideas that can be tested and evaluated (Darden & Craver, 2002; Nersessian, 1992; Russ, Scherr, Hammer, & Mikeska, 2008). Because of its power to produce causal explanations and testable predictions, mechanistic reasoning is a hallmark of modern scientific work (Machamer, Darden, & Craver, 2000; Salmon, 1978). Although these are not the only types of accounts that scientists aim to produce, in school settings, mechanistic accounts are often described in contrast to accounts that reflect “doing school,” or providing sets of disconnected terms or facts (Jimenez-Aleixandre, Rodriguez, & Duschl, 2000). The shifts in practice that we might expect to see that indicate movement toward increasingly constructing mechanistic accounts could be seen in shifts in the causal structure of the accounts that students are working to build (Darden & Craver, 2002; Machamer et al., 2000; Nersessian, 1992; Russ et al., 2008). Shifts we might expect to see that indicate movements away from “doing school” could be reflected in shifts toward increasing work toward causal coherence of those accounts (Bolger, Kobiela, Weinberg, & Lehrer, 2012; Kapon, 2017; Krist, Schwarz, & Reiser, 2019; Sandoval, 2003).

Epistemic beliefs and/or practices related to Justification can be found in nearly every model of epistemic cognition (see Moshman, 2014, Chapter 2 for a recent synthesis). In addition, justification is a central epistemic practice in science, involved in argumentation, explanation-construction, investigation design and enactment, and data analysis. Accordingly, there have been several frameworks developed to characterize justification (see Sampson & Clark, 2008 for an overview of those central to science education).

For the purposes of this paper, I focus on a relatively simplified version of justification: whether and how students provide and interpret sources of evidence in connection with a claim. Because I am interested in whether students come to justify ideas meaningfully, rather than doing so by rote in their classroom, I focus on the comparison of justifications ranging between a “typical school” response that information need not be explicitly justified because the “ready-made science” (Latour, 1987) they experience is trustworthy and must be true; to justifications that treat claims as the result of a series of measured judgments made through a process of considering and weighing alternative ideas and evidence (Greene et al., 2016; Kuhn, Hemberger, & Khait, 2017; Moshman, 2015). It is in these measured judgments that particular epistemic criteria come into play that guide decisions about...
Aims of the Present Study

This paper aims to examine how classroom communities’ epistemic practices develop over time by presenting an empirical study following a cohort of middle school students across three years. It focuses on three middle school classrooms with expert teachers using curriculum materials designed to engage students meaningfully in scientific practices by utilizing the design principles articulated in the Epistemologies for Science as Characteristics of Disciplinary Communities’ Social Practice section (Krajcik, McNeill, & Reiser, 2008). In particular, I trace the classroom communities’ use of two specific epistemic criteria—Nature of Account and Justification—over time to characterize the dynamic shifts and stabilizations of those ideas within group-level discourse.

I examine how these classroom communities develop epistemologies for building scientific knowledge by investigating the question:

How does the classroom community’s use of two epistemic criteria for building scientific ideas (Nature of Account and Justification) shift over the course of three 12-week content-area units across three years?

This study presents a qualitative, longitudinal investigation of classroom communities’ epistemic change, as defined by shifts in a community’s epistemic practices. It captures incremental shifts drawn from close examination of students’ scientific activity. These shifts illustrate how, by 8th grade, students were more consistently using sophisticated disciplinary forms of epistemic criteria. The shifts toward disciplinary sophistication were cumulative over time and across content areas, suggesting that they reflected a shift in something other than content knowledge gains. Yet these cumulative shifts did not occur in a clean, predictable progression toward sophistication. By providing an empirical account of the evolution of students’ knowledge-building work in practice over the course of extended engagement in a rich knowledge-building context, this study contributes to our growing understanding of how classroom communities develop epistemic understandings guiding their practice and has implications for how to support students’ meaningful engagement in scientific practices.

Method

To investigate this question, I conducted a longitudinal instrumental case study (Stake, 1995). I take a situative analytic lens to understanding epistemic cognition within this context (Greeno, 2006, 2011). Rather than focusing on individual cognition and behavior, I aim to examine the larger interactive system of the classroom and consider learning to be an emergent outcome of the interactions between students, teacher, and curriculum materials (along with larger cultural systems of schooling, although I do not consider those analytically within this paper). Accordingly, I take my unit of analysis to be the group; I am looking to see whether and how the classroom community learns. In particular, I examine the epistemic criteria evident in the group’s interaction and how these criteria are taken up and reconfigured over time. The primary data source for the study is a collection of video recordings of selected classroom lessons following the same cohort of students in three content-area units, one in 6th grade (January–April, 2013), one in 7th grade (January–April, 2014), and one in 8th grade (September 2014 – January 2015).

Participants and Context

School. Mountain View School is a K–8 school located in a working- to middle-class suburb of a large Midwestern city. It is a moderately high-achieving school, with 75% of students approaching or exceeding standards on the PARCC tests for math and English/Language Arts and 61% of students meeting or exceeding standards on the new state science assessment. The school is consistently ranked as one of the top 100 elementary and middle schools on sites like niche.com, and district-administered parent surveys report that 90% of parents feel that their students are safe and 89% are very satisfied with the quality of their child’s education. Mountain View also has a relatively low rate of student attrition: in 2012–2013, their 6th grade class had 98 students. In 7th grade, seven new students transferred in. In 8th grade, three transferred in and four transferred out. These low rates of student movement are important for the longitudinal nature of this study.

Students. Demographically, Mountain View is mixed. According to the district web page, the K–8 students enrolled in 2014–2015 were 52% White (including a high proportion of Eastern European immigrants), 31% Asian/Mid-Eastern, 12% Hispanic, 4% multiracial, and 2% African American; 18% of students received free or reduced-price lunches (a commonly used proxy for low socioeconomic status), and 18% were designated English language learners. The majority of students at Mountain View were second-generation immigrants who spoke a language other than English at home. There were 46 languages other than English spoken at home, with Spanish and Urdu the most common.

Informed parent consent and student assent were obtained at the beginning of each school year. Of the 108 total students who were a part of the grade-level cohort observed for this study, 104 consented to be video-recorded as part of this study for at least one year and 94 consented to participate for all three years. There were four classrooms in each grade level, one of which was selected for observation each year. Each classroom observed consisted of between 23 and 28 consenting students.

Curriculum. The science classrooms I observed were all using the IQWST curriculum (Krajcik, Reiser, Sutherland, & Fortus, 2011). This curriculum was designed to engage students meaningfully in scientific practices organized around a driving question investigating sets of real-world phenomena. The key design features of this curriculum that make it a good case site for observing students’ practical epistemologies include an accessible driving question motivating the entire unit (Blumenfeld et al., 1991), structuring of activities that promotes cycles of inquiry motivated by evolving questions (Krajcik et al., 2008), and an emphasis on student-driven model and explanation development and revision (Lehrer et al., 2008).

This comprehensive middle school science curriculum consists of distinct content-area units. For each grade level, there is a physical science unit; a chemistry unit; a life science unit; and an earth science unit. These units have a spiral design: principles from
a 6th grade unit about the particle nature of matter are leveraged in a 7th grade unit on chemical reactions and another on energy transfer and conversion, and so forth.

Although there are four units for each grade level (12 in total), the teachers at Mountain View planned to do three units during each school year during the time of observation. In addition, during the year of data collection in 7th grade, they decided to split the 7th grade biology unit because of time constraints, starting half of it at the end of 7th grade and completing the unit at the beginning of 8th grade. They also chose to enact part of the 7th grade Earth Science unit during 8th grade because of the dearth of Earth Science that this particular cohort of students had received.

The driving questions and content area topics of each of the units enacted at Mountain View are displayed in Table 1.

**Teachers.** The three middle school science teachers at Mountain View each taught one grade level, and they each considered science to be their primary subject area (though they were all credentialed in multiple subject areas). These teachers all had long-standing relationships with the university research team with which Christina Krist was affiliated. Because of their involvement in developing and pilot-testing the curriculum materials, they were very familiar with them. At the time of data collection, all three teachers had worked with the IQWST materials for 11 years, and they continued to develop them together: at their request, their administration gave them a shared preparation period every day and allowed them to meet together once every two weeks during school improvement days for an extended time. They used these planning times to revise the materials. As such, these teachers worked as a cohesive instructional team with deep familiarity with the substance and pedagogical goals of the curriculum materials and how to enact them effectively. Additional information specific to each participating teacher can be found in the online supplemental materials.

### Data Collection

All data collection completed as part of this study was done under IRB approval. I observed and video-recorded classroom lessons from the nine content-area units enacted from 6th grade through 8th grade, beginning in 2012. Within each unit, I selected specific lessons for observation in which students explicitly worked to build one of the core principles of the unit. Though all of the activities in the unit were designed to help students work toward these ideas, there were various points at which lessons involved synthesizing activities to pull ideas together (e.g., lessons in which students were drawing and presenting models, revising explanations, or pulling together interpretations from multiple days’ worth of evidence to explain a phenomenon). I posited that these lessons were ones that provided a high potential for eliciting evidence of the epistemic ideas guiding students’ knowledge-building work. Lessons often stretched over multiple days of instruction. I consistently observed the same class period (e.g., Period 4) for the duration of each school year. This resulted in between 23 and 28 of the 104 total consenting students recorded each year. Because students were randomly sorted between class periods each year (Mountain View did not have any kind of tracking structure during the years of data collection), some students were recorded for multiple years.

The recordings captured 61 unique consenting students across the three years, or about 59% of the total student cohort. Of those 61 students, 43 were recorded during one grade only, 16 were recorded during two grades (e.g., 6th and 7th grade), and two students were in the class period selected for recording all three years.

Video recordings were primarily collected using a stationary camera in the back of the classroom. When students were working in small groups, I focused the camera on one particular small group and placed an external audio recorder in that group. I selected the small group based on talkativeness of the group, consent to be recorded, and any comments from the teacher on students who I should avoid record for various other reasons. I continued recording the same group throughout the duration of a recorded lesson for continuity (e.g., all of Lesson 5). However, to avoid fatiguing groups and to obtain a distributed sample, I switched groups when recording a new lesson. In doing so, I attempted to record evenly across groups that had all consenting students. Overall, 42 of the 61 unique consenting students (69%) were part of a focal small group at least once in the video corpus.

### Data Selection and Analysis of Epistemic Considerations

I chose one unit from each grade level for detailed examination. I selected the first complete unit in 6th grade (Chemistry), the last complete unit in 8th grade (Earth Science), and the middle unit

<table>
<thead>
<tr>
<th>Unit</th>
<th>Driving question</th>
<th>Content-area topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sixth grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Physical science</td>
<td>Can I believe my eyes?</td>
<td>Light</td>
</tr>
<tr>
<td>6 Chemistry*</td>
<td>How can I smell things from a distance?</td>
<td>Particle nature of matter</td>
</tr>
<tr>
<td>6 Biology</td>
<td>Where have all the creatures gone?</td>
<td>Ecosystems</td>
</tr>
<tr>
<td>Seventh grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Chemistry</td>
<td>How can I make new stuff from old stuff?</td>
<td>Chemical reactions</td>
</tr>
<tr>
<td>7 Physical science*</td>
<td>Why do some things stop while others keep going?</td>
<td>Force and motion</td>
</tr>
<tr>
<td>7 Biology [Lessons 1–5]</td>
<td>What is going on inside me?</td>
<td>Body and cell systems</td>
</tr>
<tr>
<td>Eighth grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Biology [Lessons 5–8]</td>
<td>What is going on inside me?</td>
<td>Body and cell systems</td>
</tr>
<tr>
<td>8 Earth science*</td>
<td>How is the Earth changing?</td>
<td>Plate tectonics</td>
</tr>
<tr>
<td>7 Earth science [Lessons 1–6]</td>
<td>What makes the weather change?</td>
<td>Weather and seasons</td>
</tr>
</tbody>
</table>

*Note.* Focal units are marked with an asterisk.
completed in 7th grade (Physics). These units are marked with an asterisk and shaded in Table 1. In total, I analyzed 7 days of instruction in 6th grade, 14 days of instruction in 7th grade, and 16 days of instruction in 8th grade, totaling 24 hr and 40 min of video.

**Selection of science knowledge building episodes.** I then further reduced the video corpus by selecting episodes of science knowledge building activity during each recorded class period. A collection of discourse indicators was used to identify potential episodes (see Table 2). These indicators functioned as “sensitizing concepts” (Blumer, 1954; van den Hoomaard, 1997). That is, we used these indicators as a loose indication of the kind of conceptual work—students’ meaningful science knowledge building—that we were aiming to identify and characterize as the goal of this research. We developed these discourse indicators based on Berland and colleagues’ (2016) framework for epistemologies in practice, as this framework describes the range of possible answers to the epistemic considerations (e.g., a range of ways to justify ideas) rather than only describing the disciplinary version of it.

We then defined episodes of science knowledge building with a two-step process: one team member applied indicators and identified episodes, and a second team member verified the episodes. We identified an episode for more in-depth analysis by selecting chunks of time that included a cluster of at least two indicators within a 1- to 3-min span of time (see Figure 1). We then determined the precise starting and ending points of the episode based on alignment with activity structures and made sure that the episode captured the introduction of the activity and the complete arc of the topic of discussion. Team members were trained on the indicators until they reached 100% agreement with experienced indicators until they reached 100% agreement with experienced

After identifying episodes, the research team transcribed each episode by turn of talk (each new speaker received a new line in the transcript) and coded each turn for the epistemic considerations (ECs) evident in that turn. A high proportion of students contributed to the talk in the episodes: 83%, 100%, and 92% of the students in the 6th, 7th, and 8th grade classrooms (respectively) contributed at least one full sentence (i.e., more than a one-word or short phrase response to a teacher question or prompt) during at least one episode, suggesting that treating student talk collectively was relatively representative of the classroom community as opposed to only reflecting the contributions of a handful of students. We used turns of talk as evidence of students’ use of ECs because their talk-in-action demonstrates the actual choices that students made in performing the work of knowledge building. These choices may not necessarily have been deliberate ones, but their talk-in-action reveals an implicit stance about their answers to the ECs.

To determine interrater reliability for the turn-by-turn coding schemes, two researchers double-coded episodes in each grade-level unit until the total amount of time double-coded in each unit reflected at least 20% of the length of all the episodes in the unit and “substantial agreement” was reached (defined by Landis and Koch (1977) as $\kappa = 0.61 – 0.8$; we required $\kappa \geq 0.7$). Because each turn of talk could receive multiple possible codes (i.e., the coding categories were not mutually exclusive) and there were a large number of possible codes per turn, we chose a method of calculating agreement that produced the most conservative measure with these data. We created two additional codes for interrater reliability called *Number and Agreement*. *Number* indicated the number of codes that each coder assigned to an utterance. *Agree-

<table>
<thead>
<tr>
<th>Epistemic consideration</th>
<th>Description of range of implicit answers (Berland et al., 2016)</th>
<th>High-level indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of account: <em>What kind of answer are we working to build?</em></td>
<td>Descriptive accounts, Causal accounts, Mechanistic accounts</td>
<td>Mark any time students or teacher are considering what or what kind of information to include in a model or explanation. This could include:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Describing what happened in a phenomenon</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Identifying key components of a phenomenon</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Making an analogy to another phenomenon</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Making a comparison to another experience</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Identifying a relationship between components</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Offering a how/why explanation or rationale</td>
</tr>
<tr>
<td></td>
<td>Justification: <em>How do we justify the ideas we’re working to build?</em></td>
<td>No justification, Justified by identifying source information, Justified by interpreting and connecting across source information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A student insisting on including or removing information because the teacher said it was right or wrong</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A student or teacher making a comparison to a previous experience in order to support a knowledge claim or decide what/how to include a piece of information in their model or explanation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A student or teacher referring to empirical evidence or an activity done in class as they are considering what/how to include a piece of information in their model or explanation</td>
</tr>
</tbody>
</table>
The extent to which these codes were the same. If both coders provided the same string of codes (i.e., they selected precisely the same combination of possible codes to characterize a turn), that turn received a 1 for Agreement. If the coders did not agree on the precise set of codes, the coder who coded more conservatively (i.e., gave fewer codes or a less sophisticated code) received a 0. Cohen’s kappa, weighted by Number, was then calculated in terms of these additional codes. Each new coder was trained on the coding scheme until they reached “almost perfect agreement” (κ = 0.81–1.0) with an experienced coder on at least 10% of the total dataset.

**Coding for nature of account.** To characterize Nature of Account, or the classroom community’s answer to “What kind of answer are we working to provide?”, we were interested in whether the classroom community was working to provide mechanistic explanatory accounts (Russ et al., 2008) versus other types of accounts (e.g., descriptive, anecdotal, teleological, etc.). We coded each turn of talk within each episode for the elements of mechanistic accounts that it contained. These elements were adapted from Russ et al.’s (2008) framework for mechanistic reasoning as described by Krist et al. (2019). We also identified elements of other types of accounts. These elements are described in Table 3.

Then, because we were looking at the coconstruction of explanatory accounts through group discourse, we aggregated turn-by-turn element codes by episode to get a sense of the kind of account the classroom community was working to provide. We categorized each episode based on whether the majority of student talk turns contained explanatory or other elements. We focused only on student turns to ensure that it was students’ work that was reflected in the episode characterization rather than the teachers’. If the majority of turns of talk contained explanatory elements, we characterized the episode as building an Explanatory account and noted the most sophisticated element that occurred at least three times as an indicator of the completeness and coherence of the account (Identifying < Unpacking < Linking). If the majority of turns of talk contained explanatory elements, we characterized the episode as building an Other type of account. Because these episode-level characterizations were determined algorithmically from the turn-by-turn codes, there was no formal metric of reliability calculated.

**Coding for justification.** To characterize Justification, or the classroom community’s answer to “How do we justify the claims we’re working out?,” we first identified and numbered in chronological order the claims made within each episode. Each turn of talk was coded for whether it contained a justification of any kind, and if it did, which claim the justification supported. We indicated the number of claims that were justified versus unjustified within each episode.

We then characterized each justification for the type of source that it drew upon (e.g., authoritative source, personal experience, or collected data) and the complexity of the interpretation of that source information in relation to the claim. The codes for types of sources were generated inductively to capture those that students actually used, and they are not hierarchically ranked or aligned, which allowed us to view their use as interactive and context-dependent (Chinn et al., 2011). The codes for the interpretation of sources are derived from a synthesis of frameworks for analyzing scientific argumentation and argumentative discourse (Felton & Kuhn, 2001; Kelly & Takao, 2002; Kuhn & Udell, 2003; Lawson, 2003; Sandoval & Millwood, 2005). These frameworks were selected and synthesized because they attempt to capture the epistemic dimensions and goals for justifying ideas rather than the structural elements or the content correctness of an argument. These codes are presented in Table 4.

**Identifying longitudinal patterns and shifts.** Once the utterances and episodes had been coded, I summarized the codes by class period. I then visualized the codes by degree of sophistication (i.e., arranged from nonmechanistic elements to preexplanatory elements to explanatory elements for Nature of Account, and from working with source info to simple interpretations to complex interpretations for Justification) and arranged the class-period visualizations chronologically. I looked for patterns within these visualizations such as the appearance of a new code, the disappearance of a code that had been frequent, and marked increases or decreases in particular codes. For the most part, I did not conduct quantitative analysis on these codes. Instead, in presenting my analyses of these findings, I attempt to treat these aggregations of codes as “tallies of claims about the data” rather than data themselves (Hammer & Berland, 2014, p. 43) and present these “tallies of claims” for readers to interpret for themselves as well. I present qualitative excerpts that reflect the identified shifts in order to allow for qualitative interpretations of validity to the patterns (i.e., does a shift in X code “feel” like a different kind of thing in the discourse?). However, there is one exception: I conducted chi-squared analyses of the relationship between source type and type of interpretation, since the nonhierarchical nature of the source
Table 3
Codes for Characterizing Nature of Account in Classroom Discourse

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-explanatory elements (structural precursors to mechanistic reasoning)</td>
<td>Describing a phenomenon: Describing what would be visible or observable to everyone with a naked eye about a phenomenon they are working with. Articulating a sequence: Describing a sequence of events or steps in a process</td>
<td>“Something about ice. I’ve noticed that like, it’s kinda of weird but, you know how like lakes, like only freezes on top. Why doesn’t the whole water freeze? Cause shouldn’t that?”</td>
<td>Describing the target phenomenon</td>
</tr>
<tr>
<td>Explanatory elements (markers of mechanistic reasoning)</td>
<td>Markers of causal structure: Identifying a factor: Identifying an entity, a property, an entity-behavior relationship, or the name of a process, typically at the scalar level below that of the target phenomenon (see Krist, Schwarz, &amp; Reiser, 2019 for elaboration on what “counts” as a factor for phenomena at various scalar levels)</td>
<td>“Well, if you’ve ever, like I’ve actually done like where you make glass. Right. So they heat it up. They heat it up, use a lot of heat. And then it’s a liquid and then it quickly freezes to get to a solid. It quickly does that you’re right. It gets over to the solid.”</td>
<td>Identifying setup conditions and/or chaining: backward and forward</td>
</tr>
<tr>
<td></td>
<td>Unpack factor: Describing the behavior of entity or agent; the consequences of the property; interactions between relationships; and/or chaining of processes</td>
<td>“Ok. You know how you said um the particles make up the air?”</td>
<td>Identifying entities, identifying properties of entities, identifying organization of entities</td>
</tr>
<tr>
<td></td>
<td>Marker of causal coherence: Linking: Explicitly connecting underlying behaviors or interactions to how they play out to cause the aggregate phenomenon</td>
<td>“I just don’t get, cause you said this last time, that um, not last time but just earlier today. You said that um, that the particles are together. I just didn’t get that cause, I don’t think the particles would like stick together and go around in the same way.”</td>
<td>Identifying activities and/or chaining: backward and forward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Excellent. So if it’s warmer to have more energy, it goes faster, and we can smell it.”</td>
<td>Chaining: Backward and forward</td>
</tr>
<tr>
<td>Elements of nonmechanistic accounts</td>
<td>Providing known-answer information</td>
<td>“Um it says the fact that all matter can exist as a solid, liquid, and a gas.”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Giving an illustrative example</td>
<td>“Excellent. That is crazy to think about. So all matter can exist as a solid and a liquid and a gas. So I’m thinking of like, what would be hard to think about, like maybe rocks.”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recalling details from a class activity or personal experience</td>
<td>“We stuck the, um, the pH paper into the liquid and it was an acid.”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Sometimes like, I don’t know that my mom is cooking is brownies, but then I’m just downstairs playing games . . . [inaudible] . . . and um I just come upstairs and I only smell like, a sweet scent, and then I look in the oven sometimes and just see brownies.”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Yeah”; “Josh, call on someone”</td>
<td></td>
</tr>
<tr>
<td>No evidence of account type</td>
<td>Utterance does not provide evidence of what kind of answer or knowledge product the classroom community is working to provide.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Cohen’s kappa = .87. Italics text indicate the names of codes for elements of mechanistic accounts.
Table 4

<table>
<thead>
<tr>
<th>Code</th>
<th>Source of justification</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authoritative source</td>
<td>Source of justification in utterance is the teacher, textbook, a movie, etc.</td>
<td>“I’ve seen in videos, it’s kind of like a circle and tiny line and then a circle and then that’s how they are connected.”</td>
</tr>
<tr>
<td>General everyday experience</td>
<td>Source of justification in utterance is general knowledge that comes from living in the world</td>
<td>“I’m pretty sure lava is a rock [inaudible] melted which is a melted rock. So it just, it gets melted.”</td>
</tr>
<tr>
<td>Personal experience</td>
<td>Source of justification in utterance is a specific personal experience</td>
<td>“I was going to say that—I’d say that when it gets cold it gets really small, but one day I was eating dinner and like, it was after my brother’s birthday party so I was like, playing with a balloon while I was eating dinner—”</td>
</tr>
<tr>
<td>Hypothetical imagined data or scenario</td>
<td>Source of justification in utterance is an imagined “what if” situation or an imagined modification on an experiment</td>
<td>“Uh, I think it’s a solid because what if it never changes, it’s still a solid it doesn’t change form. You have the ability to change it but you really can’t change it like um, you can’t turn it into a liquid really. Because if it was a liquid you could actually kind of go through it easily. And if it were a gas you could easily go through it and you couldn’t feel it or anything. But you can so I think it’s a solid.”</td>
</tr>
<tr>
<td>Collected data</td>
<td>Source of justification in utterance is data that has been collected by students (or by other scientists, which they are using as empirical data)</td>
<td>“Alright, for what I drew first, was like, all the molecules were stuck together and I kind of made, like, water, cause, cause like you know, water, there was water at the bottom of the balloon. So whatchu drew?”</td>
</tr>
<tr>
<td>Agreed-upon idea</td>
<td>Source of justification in utterance is an idea that they had agreed upon as a class previously; e.g., “We said that ...”; “We decided that ...”</td>
<td>“No remember when we learned that it was actually magma underneath the rocks and then the magma goes through and it layers over.”</td>
</tr>
</tbody>
</table>

Note. Cohen’s kappa = .77. Bold text indicates language in the examples providing evidence for the code.
codes made visualizations that supported pattern identification difficult.

Results

The following sections present shifts over time in classroom communities’ use of Nature and Justification as evidenced in their in-the-moment work of building science knowledge. I look at shifts both within and between content-area units at each grade level. By 8th grade, students were more consistently using sophisticated disciplinary forms of epistemic criteria. In addition, the sophisticated work they were doing in 8th grade required and built upon the shifts that occurred during the 6th and 7th grade units. This suggests that, in the aggregate, classroom communities were incrementally developing disciplinarily sophisticated epistemic practices over time. However, the shifts within each unit did not necessarily appear “sophisticated” on their own. In the following sections, I discuss key shifts toward disciplinary sophistication in the aggregate, across all three years. Table 5 summarizes these shifts. I then discuss the shifts within each grade-level unit and show both how these “micro” shifts were not necessarily recognizable as shifts toward sophistication in isolation; but also how these shifts were leveraged in later epistemic work.

Nature of Account: What Kind of Account Are We Working to Build?

The Nature of Account that the classroom community was working to build was characterized first in terms of whether the accounts were mechanistic, a form of knowledge that is characteristic of and central to the discipline of science (Chinn et al., 2011; Hofer & Pintrich, 1997; Schommer, 1990). Here I present shifts toward mechanistic accounts in terms of both the causal structure of the accounts that students were working to build (Darden & Craver, 2002; Machamer et al., 2000; Nersessian, 1992; Russ et al., 2008) and the causal coherence of those accounts, or how well they “hang together” to explain the phenomenon in question (Bolger et al., 2012; Kapon, 2017; Krist et al., 2019; Sandoval, 2003).

Moving from descriptive or teleological accounts to mechanistic accounts. These classroom communities were primarily constructing explanatory accounts, rather than descriptive or teleological accounts, even very early on. With only two exceptions, greater than 50% of the time spent engaging in epistemically rich discussions in each class session was spent constructing explanatory accounts. This aligns with research on mechanistic reasoning that demonstrates that students have the resources to be able to do this, even without formal training (Kapon, 2017; Russ et al., 2008). In other words, the shift from descriptive or teleological accounts to mechanistic ones was not a shift that these classroom communities needed to make, as they were already predominantly working to develop mechanistic accounts.

Moving from simplistic to multidimensional, coherent causal accounts. Although the classroom community was working to build mechanistic accounts, from 6th grade to 8th grade they moved from primarily constructing simple mechanistic accounts to primarily constructing accounts with complex causal chains (i.e., a more complex causal structure) that were internally coherent. Figure 2 presents more detailed characterizations of the Explanatory episodes detailing the complexity of the explanatory accounts they were working to provide. In particular, it highlights whether the structure of the account constructed during the episode utilized one or more than one mechanistic factor; and whether the account confirmed coherence by linking the underlying mechanistic factors and behaviors to the aggregate phenomenon in question.

Causal structure: One factor to multiple factors. The most dramatic shift over time is one from constructing accounts that draw on one mechanistic factor, to accounts that draw on multiple mechanistic factors (Figure 2a).

To illustrate what this distinction looked like in practice, consider an episode from Lesson 11 of the 6th grade unit. Students focused on one key explanatory factor—temperature—and “played out” how that key factor mattered for their particle model of matter. After spending almost an entire class period tinkering with a computer model that visualized particle motion at different temperatures, Niral asked, “Isn’t one of the reasons why squirrels look for their nuts before winter is because when the ground freezes it’s harder to smell nuts?” Katerina agreed with him, unpacking the implications of that factor (temperature) on particle motion: “Because when the nuts freeze, the molecules, you can’t smell them as easily [. . .] because when it’s warmer the molecules speed up and then it has more energy and can get to [the squirrel] faster, but in winter it can’t.” In this episode, Niral and Katerina provided an explanatory account that focuses on one key explanatory factor: particle temperature.

In contrast, in the following excerpt from Lesson 1 of the 8th grade unit, Achmed and Lindy identified some possible factors that contributed to where earthquakes were found: a specific geographic location (California) and the texture of the land (bumpiness), respectively. Theo then proposed an alternative factor—plates and plate interactions—and unpacked the consequences of that factor. In doing so, he coordinated together the factors that Achmed and Lindy had proposed:

Mr. M: Let’s share out some predictions. Achmed will you share yours first. What did you predict for where do you think they’re found?

Achmed: Um, I said they’re usually found in like, California.

Mr. M: California. Lindy what did you predict?

Lindy: Oh um, I do not know if this is for earthquakes or tornadoes, but my sister and my mom said [we] do not really get big earthquakes because our land is flat and other lands are bumpy.

Mr. M: Ok so you are predicting flat land is—you do not get earthquakes where flat land is?

Lindy: Yeah.

4 Note that what counted as a mechanistic factor was determined relative to the phenomenon of interest (specifically, an element that was at least one scalar level below that of the target phenomenon). In addition, theoretical entities were also counted as factors. As a result, depending on the context of the phenomenon, factors were sometimes microscopic, sometimes macroscopic, and sometimes theoretical. See Krist, Schwarz, and Reiser (2019) for further elaboration of how factors were identified in this way.
Mr. M: Oh that’s cool, but are you thinking you would get it where it’s bumpy?

Lindy: Yeah.

Theo: We would still get earthquakes but—we still would get earthquakes. Just not as bad.

Mr. M: So [inaudible]

Theo: Like in California. California’s on top of um, two plates that have collided.

Mr. M: Yeah.

Theo: So now they’re like moving but they cannot move because they’re rubbing against each other. So it’s like shaking everything. And that’s like how I think it gets bumpy there, because the plates are like, overlapping. And I think Illinois, most of it, is under a solid plate. Like it’s [inaudible]. But in California, two of the plates have collided. [...] So they’re trying to move but they cannot so they make an earthquake.

Here, Theo built an account that coordinated together the multiple factors they had identified (a location in California, bumpy ground, and the Earth’s plates) to build a coherent mechanistic account. It is significant that he does so just moments after Achmed and Lindy shared their ideas: he was hearing their ideas about California and bumpy ground for the first time as they shared them with the class. He brought in his idea about plates by responding to Lindy’s idea and drawing in Achmed’s, weaving these pieces together in the moment into an explanation that accounted for both of the factors they had identified.

Theo’s content knowledge supported him in being able to do this (though the notion that something called “plates” existed

<table>
<thead>
<tr>
<th>Measure</th>
<th>Sixth chemistry</th>
<th>Seventh physics</th>
<th>Eighth earth science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of account</td>
<td>Unpacking and linking one factor</td>
<td>Unpacking multiple factors and articulating connections</td>
<td>Unpacking and coordinating multiple factors</td>
</tr>
<tr>
<td>Justification</td>
<td>Interpret source for your audience</td>
<td>Select careful observations and/or empirical details to support claim</td>
<td>Complex argumentation and source differentiation</td>
</tr>
</tbody>
</table>

Table 5
Summary of Shifts in Classroom Communities’ Use of Epistemic Considerations Over Three Years

![Graph](a)

![Graph](b)

Figure 2. Student discourse characterization in explanatory episodes over time. Key shifts include (a) moving from working with predominantly one factor to predominantly multiple factors and (b) transitioning between unpacking factors to linking factors. Lessons in 6th grade each lasted only one day. Gaps in Lesson Day sequences are attributable to either days not being recorded because they were noncontent days (e.g., school assemblies, substitute teachers) or because there were no analytic episodes identified, usually because of classroom activities consisting largely of independent student work (e.g., quizzes). See the online article for the color version of this figure.
remained contentious and not a given in this classroom until after Lesson 2. But Theo could have simply stated his content knowledge: earthquakes happen where plates meet. He did not need to coordinate his idea with Achmed’s and Lindy’s to make a contribution. In doing so, he indicated that his underlying assumption about the kind of account they were working to provide was one that accounted for multiple identified factors.

**Causal coherence: Unpacking only to Unpacking + Linking.**

A second key shift undergirding the classroom communities’ move from simplistic to multidimensional, coherent causal accounts are the transitions between unpacking factors to linking factors (Figure 2b). This difference seems small: it is between only identifying and “playing out” how a factor or factors behave (e.g., unpacking that molecules are moving faster at higher temps), and unpacking plus iteratively connecting that unpacking back to the aggregate phenomenon to “check” for causal coherence (e.g., do faster molecules account for an odor traveling farther more quickly?). Though subtle, this distinction is epistemically important because it reflects an explicit “check” for causal coherence of the account.

Both of the examples already presented involved “checking” for causal coherence of either one factor (in the squirrels example from 6th grade) or multiple factors (in the plate tectonics example from 8th grade). In fact, most of the episodes in 6th and 8th grade involved linking rather than Unpacking Only. However, in 7th grade, most of the episodes involved only Unpacking, without the “checking” for causal coherence as part of the conversation.

For example, in this excerpt from Lesson 5 in 7th grade, Ms. K’s class was trying to explain why a ball does not bounce as high as it was dropped from, and why it eventually stops. Several students had proposed possible ideas; to wrap up the conversation, Ms. K called on Paul as the “final thought”:

**Paul:** I think it has something to do with like friction, in the air when it loses some of its energy.

**Ms. K:** To what?

**Paul:** Just like, cause the air is kind of like pushing back on it when it’s trying to fall. So I think the gravity that’s like, I think it just loses energy when it’s falling.

**Ms. K:** But I like what you said about air. We haven’t thought about that yet.

**Paul:** The air’s just kind of there. It’s pushing against it.

**Ms. K:** But is air matter?

**Paul:** Yeah.

**Ms. K:** Are air molecules moving?

**Paul:** Yeah.

**Ms. K:** Ok so kind of like connected into what Theo said [about some of the ball’s energy getting trapped in the ground], could the ball be sharing a little bit of its energy with the air molecules that it bumps into? I mean, that’s connected to what you’re saying, right?

**Paul:** Yeah.

Here we see Paul identifying a possible explanatory factor—the air—and unpacking how it might be involved—by pushing against the ball. This episode was not coded as Linking because neither Paul nor Ms. K explicitly connected the “sharing” of energy to the air to the slowing down and stopping of the ball (e.g., they did not play out the outcome of sharing and sharing and sharing energy, gradually slowing down until it runs out and stops).

This seems like a minor difference; clearly this was the unstated implication they were trying to explain. But instead of fully playing out the unpacking by linking it back to the aggregate phenomenon, like Katerina did with the squirrels, we see hints of the kind of coordinating that Theo did in 8th grade being set up. Rather than picking one factor (temperature) and “running” it through their model to completion, they are identifying a few possibilities, testing them out part way (in terms of the mechanism) and seeing how they jive with other ideas (at the underlying scalar level). In other words, the Unpacking accounts in the 7th grade unit suggest that the classroom community has begun to develop value for an account that not only can be “played out” logically, but also pulls together multiple observations and factors.

**Looking at individual units: How did these shifts build up over time?** I now turn to look at these two patterns together (Figure 2a and 2b) for each specific grade-level unit in order to see how these aggregate-level shifts toward disciplinary sophistication by 8th grade were built up over time.

**6th grade: Unpacking and linking one factor, then introducing more factors.** First, the beginning part of the 6th grade unit consisted of the classroom community constructing accounts that primarily consisted of one explanatory factor (Figure 2a), and that Unpacked and Linked that factor in a coherent account (Figure 2b; recall that “Linking” requires that “Unpacking” has been done). Then, in Lesson 6, all episodes contained mention of multiple explanatory factors (Figure 2a), though only about half of those episodes Unpacked or Linked those factors (Figure 2b), meaning that sometimes these explanatory factors were named but not made sense of in any significant way. The episodes in Lesson 12 reflect a mixture of these approaches: about 1/3 of the episodes contained one explanatory factor and 2/3 contained multiple factors; and some episodes Unpacked those factors and some Linked. A small number (represented by the difference in total size of the bars for Lesson 12 in Figure 2a and Figure 2b) neither Unpacked nor Linked.

These patterns seem to suggest something like this story: early on, the classroom community was able to construct structurally simple, yet causally coherent, mechanistic accounts. However, by Lesson 6, their stories became too complicated and they needed to include more factors in order to explain their phenomena. Sometimes these multiple factors could be easily integrated into a coherent account; but other times they could not. We do not see much of what happened between Lesson 6 and Lesson 12, but the mixture of types of accounts constructed in Lesson 12 suggests that the classroom community was committed to including multiple factors, and to constructing coherent accounts; but that it was not always possible. They were still working through how to incorporate multiple factors in a coherent, explanatory way.

**7th grade: Identifying multiple factors and considering connections between them.** Then, in 7th grade, we see that the use of multiple factors persisted to a greater extent than in 6th grade (Figure 2a). During each class period, there was a greater propor-
tion of episodes that contained multiple factors than episodes
containing only one factor (with the exception of Lesson 3 Day 2,
where there were never multiple factors, and Lesson 5 Day 4,
where there was an equivalent proportion of episodes). Interest-
ingly, the increased identification of multiple factors corresponded
with less Linking compared to 6th grade (Figure 2b). In other
words, the activities and behaviors of the factors were not neces-
sarily getting connected in conversation back to show how they
contributed to the aggregate phenomenon. This is one pattern that
does not, on its own, reflect a clear shift toward disciplinary
sophistication: while the classroom community was considering
multiple factors (which is necessary for more complex mechanistic
accounts), they were doing less of the work of Unpacking and
Linking those factors, or showing how those factors were behaving
in order to explain the phenomenon.

8th grade: Unpacking and linking multiple factors in complex
accounts. By the 8th grade unit, every episode contained multi-
ple explanatory factors (Figure 2a) and most of those episodes
Unpacked and Linked those factors into coherent mechanistic
accounts (Figure 2b). Though we do not see what occurred be-
tween the end of the unit in the middle of 7th grade and the start
of a unit in the middle of 8th grade, these patterns suggest that (a)
the work of learning to construct structurally complex accounts
that coherently leverage multiple explanatory factors takes time,
and (b) that part of that learning process may include segments of
time in which students explore how multiple factors might be
interacting with each other, without necessarily worrying about
constructing full, coherent explanatory accounts during that explo-
ration.

Summary. From 6th to 8th grade, the classroom community
gradually increased the complexity of the causal structure of their
accounts by incorporating multiple factors. However, the consider-
ation and incorporation of new factors often required pulling
back on the causal coherence of the accounts: the implications of
each factor was not always fully explored or linked back to the
aggregate phenomenon. Though such pulling back does not look as
sophisticated on the surface, when sandwiched between the ac-
t-account construction patterns in 6th and 8th grade, the utility of
exploring multiple factors (even without producing fully coherent
accounts) becomes more compelling as contributing to progress
toward disciplinary sophistication.

Justification: How do We Justify the Claims We’re
Working Out?
The justifications that the classroom community developed were
characterized in contrast to a lack of justification, or the “typical
school” response that information need not be explicitly justified
because the “ready-made science” (Latour, 1987) that students
experience is trustworthy and must be true. When Justifications
were provided for claims, I characterized how the claim was
justified in terms of the types of sources that students provided
(which were not ranked according to relative sophistication); and
the nature of the justification provided, or how the source was
interpreted as evidence in relation to the claim (Buehl, 2008;
Kendeou et al., 2016; Manz, 2015a). Here I present shifts toward
complex justifications of claims and shifts over time in the rela-
tionships between source use and complexity of justification.

Moving from unjustified statements of fact to justified
claims. One possibility for a shift toward disciplinary sophisti-
cation from 6th to 8th grade is that the classroom community
begins to justify a higher proportion of claims. That is, there are
fewer knowledge claims over time that are treated as givens or
positivistic truths that do not require justifying. Figure 3 shows the
proportion of justified claims in each class period. Notably, there
is neither a dramatic increase nor decrease in the proportion of
claims justified in each class period. On average, 71%, 66%, and
73% of claims were justified in the 6th, 7th, and 8th grade units,
respectively. In addition, there were not obvious increases or
decreases in the proportion of justified claims over time. This
suggests that shifting from unjustified claims to justifying those
claims was not the primary epistemic shift in these classroom
communities.

Moving from all-sources-are-good to differentiated source
use. A second possible shift toward disciplinary sophistication is
in terms of the justifications themselves—in particular, the sources
they use and the complexity of the justification interpreting those
sources in relation to the claim. There were significant differences

Figure 3. Proportion of justified claims during each class period. See the online article for the color version
of this figure.
by grade level in terms of the types of sources that students drew upon in order to justify their claims, \( \chi^2(10, N = 700) = 85.67, p < .001 \). Table 6 displays the frequencies of students’ use of each type of source at each grade level, shaded by relative frequency. It is important to note that students drew on all source types at some point during each grade level unit (with the exception of agreed-upon ideas in 6th grade). In 6th grade, students drew broadly on all source types, with slightly higher frequencies of sources from personal experience (27.3%), hypothetical imagined data or scenarios (22.2%), and collected data (25.3%). In 7th and 8th grade, their source choices were more focused around hypothetical imagined data or scenarios (25.6% and 24.2%, respectively) and collected data (40.2% and 46.3%).

Moving from simple to complex interpretations. There were also shifts in terms of the nature of the justifications themselves. In 6th grade, most of students’ justifications were working with the source information (65.7%), either identifying the source itself or identifying the relevant information from that source (e.g., the specific line from a reading that supports a claim, or the specific pattern observed in an experiment that is relevant for the claim at hand) and some simple interpretations (30.3%; see Table 6). This pattern persisted in 7th grade, with slightly more complex interpretations than in 6th grade (14.5%, vs. 4.0%). By 8th grade, there were even more complex interpretations (21.7%), especially using hypothetical imagined data or scenarios and collected data.

Relationship between source type and nature of justification. In addition to shifts in frequencies of source types and the nature of justifications, by 8th grade there was also a significant relationship between these two variables. That is, in 8th grade, certain types of sources tended to be used more or less frequently with certain types of justifications. This was not the case in 6th or 7th grade (Table 7; note that the 6th and 7th grade values are nearly statistically significant, and it is possible that the statistical significance in 8th grade is attributable to the fact that the 8th grade data contain nearly five times more counts than either the 6th or 7th grade data).

To look more specifically and which types of sources tended to be coupled with which justifications, I examined the standard residuals from the chi-squared test for sources and justifications in all grades (Agresti, 2003), displayed in Table 8. The residuals with asterisks are those that are different than expected. In particular, authoritative sources were less likely to be used as part of complex interpretations and more likely to simply be information that students identify, and hypothetical imagined data or scenarios were less likely to be sources that students simply named or identified and were more likely to be part of complex interpretations.

Looking at individual units: How did these shifts build up over time? I now turn to look more closely at these patterns descriptively within each specific grade-level unit in order to see how these aggregate-level shifts toward disciplinary sophistication by 8th grade were built up over time.

6th grade: Using many source types and gradually building simple interpretations. In the 6th grade unit, students tended to justify by identifying relevant information and by providing simple interpretations of that source information throughout the unit. They used slightly higher proportions of simple interpretations as the unit progressed (see Figure 4). For example, in response to the idea that not everything needed to be cold to freeze (a claim they were working to support or refute in Lesson 3), Stephanie said, “When like- I kinda did this once, I stuck my finger in some melted wax.” She identified the relevant source—a personal experience—and identified the relevant piece of that personal experience: that she stuck her finger in some melted wax. She described in detail what happened, continuing to identify relevant pieces of information, such as that after about 10 s it “started to dry up and turn hard.” Ms. L then provided a simple interpretation of her experience: “Wow, so it went from a liquid to a solid pretty quickly didn’t it?” Notably, the way the teacher provided the simple interpretation for Stephanie’s claim in this example was a characteristic pattern in the first three lessons of the unit: students identified the relevant source information, and the teacher interpreted that information in light of the claim.

<table>
<thead>
<tr>
<th>Source</th>
<th>Authoritative source</th>
<th>General everyday experience</th>
<th>Personal experience</th>
<th>Hypothetical imagined data or scenario</th>
<th>Collected data</th>
<th>Agreed-upon idea</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sixth grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work with source info</td>
<td>7</td>
<td>10</td>
<td>20</td>
<td>12</td>
<td>16</td>
<td>0</td>
<td>65</td>
</tr>
<tr>
<td>Simple interpretation</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>Complex interpretation</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Sixth totals</td>
<td>12</td>
<td>13</td>
<td>27</td>
<td>22</td>
<td>25</td>
<td>0</td>
<td>99</td>
</tr>
<tr>
<td>Seventh grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work with source info</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td>19</td>
<td>29</td>
<td>13</td>
<td>75</td>
</tr>
<tr>
<td>Simple interpretation</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>Complex interpretation</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Seventh totals</td>
<td>10</td>
<td>8</td>
<td>4</td>
<td>30</td>
<td>47</td>
<td>18</td>
<td>117</td>
</tr>
<tr>
<td>Eighth grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work with source info</td>
<td>28</td>
<td>27</td>
<td>20</td>
<td>53</td>
<td>145</td>
<td>22</td>
<td>295</td>
</tr>
<tr>
<td>Simple interpretation</td>
<td>4</td>
<td>17</td>
<td>3</td>
<td>18</td>
<td>39</td>
<td>3</td>
<td>84</td>
</tr>
<tr>
<td>Complex interpretation</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>10</td>
<td>40</td>
<td>8</td>
<td>105</td>
</tr>
<tr>
<td>Eighth totals</td>
<td>34</td>
<td>51</td>
<td>25</td>
<td>117</td>
<td>224</td>
<td>33</td>
<td>484</td>
</tr>
<tr>
<td>Total</td>
<td>56</td>
<td>72</td>
<td>56</td>
<td>169</td>
<td>296</td>
<td>51</td>
<td>700</td>
</tr>
</tbody>
</table>

Note. Shading represents relative frequency within each grade level.
As the unit progressed, students gradually took up the interpretation work themselves. For example, during Lesson 12, Dylan also provided a personal experience as a justification to argue against the idea that when a balloon gets colder, it will shrink. He provided both the relevant information (underlined) and the simple interpretation (bolded) of his source:

“I was going to say that - I’d say that when it gets cold it gets really small, but one day I was eating dinner and like, it was after my brother’s birthday party so I was like, playing with a balloon while I was eating dinner, but then it was like really hot, and I, like, play with it, and as I’m playing with it it gets smaller and smaller and it starts to get wrinkly. So maybe it gets small when it’s hot too.”

This type of shift from identifying source information (and the teacher often then providing the interpretation) to providing a simple interpretation of the source itself was similar when students used authoritative sources, general everyday experiences, hypothetical imagined data or scenarios, and collected data. Students’ gradual uptake of these types of interpretations suggests that students were becoming more explicit about how relevant information supported their argument, rather than letting the data or information speak for itself.

7th grade: Carefully selecting empirical details and eventually incorporating them in complex interpretations. The use of collected data as a source was more common in 7th grade than in 6th. Coupled with justifications that primarily identified the relevant source information, one could assume that this classroom community was performing a “rote school” version of justification: “How do you know the water evaporated?” “Because the water in the cup decreased from 5cm to 3cm.” Looking more closely, however, their interactions around claims and justifications did not take on the flavor. Instead, the teacher (Ms. K) frequently emphasized their simple and complex justifications.

8th grade: Complex interpretations that leveraged hypothetical imagined data or scenarios, collected data, and agreed-upon ideas to form counterarguments. By 8th grade, students consistently produced complex justifications and the relative proportion of them increased, especially by the end of the unit (see Figure 4): their justifications not only drew together interpretations of mul-

Table 7
Chi-Squared Results and p Values for Relationship Between Nature and Source of Justification at Each Grade Level

<table>
<thead>
<tr>
<th>Source of Justification</th>
<th>Sixth grade</th>
<th>Seventh grade</th>
<th>Eighth grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authoritative source</td>
<td>2.915*</td>
<td>2.931</td>
<td>2.931</td>
</tr>
<tr>
<td>General everyday</td>
<td>1.924</td>
<td>1.833</td>
<td>1.833</td>
</tr>
<tr>
<td>experience</td>
<td>1.924</td>
<td>1.833</td>
<td>1.833</td>
</tr>
<tr>
<td>Personal experience</td>
<td>1.924</td>
<td>1.833</td>
<td>1.833</td>
</tr>
<tr>
<td>Hypothetical imagined</td>
<td>1.924</td>
<td>1.833</td>
<td>1.833</td>
</tr>
<tr>
<td>data or scenario</td>
<td>1.924</td>
<td>1.833</td>
<td>1.833</td>
</tr>
<tr>
<td>Collected data</td>
<td>2.234</td>
<td>2.568</td>
<td>2.568</td>
</tr>
<tr>
<td>Agreed-upon idea</td>
<td>2.234</td>
<td>2.568</td>
<td>2.568</td>
</tr>
</tbody>
</table>

Note. Residuals of Z = about |2| or |3| (Agresti, 2003) are considered to be significantly different than expected and are marked with an asterisk.
Students often used these complex interpretations when forming counterarguments, a part of an argumentation practice that was not done explicitly or in a notable way in 6th or 7th grade. To form these counterarguments, students typically drew on hypothetical imagined data or scenarios, collected data and/or agreed-upon ideas. For example, during a discussion about how the geographic features of Yellowstone were formed, Sara drew together ideas from multiple sources to construct a counterargument against the idea that Yellowstone was formed by a subduction boundary:

I disagree with all of you, I actually think that Yellowstone’s a hotspot, because like Logan said when you were ta—listen! [Laughs] About Iceland, how he said that like, most of those volcanoes were like away—like miles away from a plate boundary, you know, then it had to be a hotspot? Um, but Yellowstone is even further away from the nearest plate boundary, so then, how could it be like, subduction, if there’s not even a plate boundary there?

Here, Sara utilized all three of the Complex Interpretation subcodes (see Table 4). She drew upon multiple sources—an agreed-upon idea about Iceland, collected data mapping volcano locations from the past 10,000 years, an existing authoritative model labeling plate boundaries, and their agreed-upon model of subduction that required a plate boundary. She interpreted these together in a complex chain of logic—in this case, using conditional reasoning—to construct her counterargument that Yellowstone was a hotspot.

This shift toward increasingly complex interpretations of sources reflects a growing attention to persuading one another about a claim. In addition, students’ interpretations of empirical evidence, and agreed-upon ideas that had been built from interpretations empirical evidence in the past, did become powerful sources for persuading others in this classroom community; but also that the use of imagining as a “thought experiment” was also a valued, and valuable, tool for generating and defending claims during the knowledge construction process.

Discussion

Summary of Shifts in Nature of Account and Justification

This study identified and characterized classroom communities’ use of epistemic considerations related to the nature of the accounts they were working to build and how they justified their ideas-in-progress during their knowledge building activities over the course of three years. By 8th grade, students were more consistently using sophisticated disciplinary forms of each epistemic criteria: they were constructing coherent, causal mechanistic accounts and justifying claims with complex interpretations of sources. In addition, the sophisticated work they were doing in 8th grade required and built upon the shifts that occurred during the 6th and 7th grade units. However, the shifts within each unit did not necessarily appear “sophisticated” on their own. The classroom community pulled back from providing fully coherent accounts as they introduced more explanatory factors into those accounts, and they pulled back from offering interpretation of source information as they focused on making careful observations of a few types of sources. These areas that were pulled back were then incorporated, resulting in more sophisticated accounts and justifications than earlier. These were patterns of a classroom community toggling between adding complexity to different aspects of each epistemic consideration (e.g., between Identifying Factors and Linking; or between source type and interpretation of sources) that resulted in cumulatively building up that complexity over time.

How Do These Shifts Contribute to Our Understanding of Epistemic Cognition?

On one hand, these patterns are expected: over time, the classroom community came to build accounts and justify claims in ways that align with normative definitions for what science knowledge building is. On the other hand, there are elements of this study that make the characterization of these shifts a novel contribution.

In particular, the toggling between and cumulative shifts toward sophistication occurred across all three units, in three very different content areas. Although it is likely that the specifics of the different content areas, the natures of the different task structures present in the

![Figure 4](image-url)
observed lessons, and differences in teachers at each grade level contributed to some of the particulars of the observed changes, the overall shifts toward sophistication suggest that these findings reflect one possible epistemic learning trajectory. Most notably, the classroom community did not start back over at “square one” when making observations of a new phenomenon at the beginning of a new unit. For example, despite a lack of prior instruction on types of energy, the classroom community interpreted and triangulated sources of evidence to explain why the radiometer was spinning at the beginning of the 7th grade unit, maintaining the shift to provide simple interpretations of sources that occurred during the 6th grade unit. Similarly, despite a lack of prior instruction on plate tectonics, the classroom community identified, unpacked, and linked multiple factors at the beginning of the 8th grade unit, expanding from the shift to unpack and link careful empirical observations that occurred during the 7th grade unit.

In other words, these epistemic shifts were “sticky”: they were general enough to be useful across content areas, and the classroom community did not have to relearn them when beginning work in a new domain. The nature of these shifts adds support to the argument that criteria for building mechanistic accounts and for providing justifications for claims that interpret various sources of evidence are epistemic considerations that students carry with them across content areas, and that allow them to make progress in constructing scientific knowledge (Berland et al., 2016).

However, I do make this claim with a note of caution. I present this trajectory as one possible trajectory toward epistemic sophistication. The details and nuances of it, such as the specific shift observed in 7th grade was moving from triangulation of sources to simple interpretation of sources, are also inextricably tied to the nature of the enacted unit. Had this unit been a life sciences unit, for example, there may have been other facets of epistemic criteria that were emphasized, resulting in a slightly different trajectory documented in this study. Thus, I do not intend to present this trajectory as the pathway toward epistemic sophistication, with mandates for whether Linking a single factor needs to occur before Unpacking multiple factors can occur, for instance. Instead, it is a pathway; and I use the existence of a pathway to argue that longitudinal characterizations of epistemic learning—as something different than science content learning—are empirically identifiable in practice.

To state it more broadly, this study demonstrates that the theoretical blending of a context-dependent epistemological resources and stability across contexts is a viable model of epistemic learning. That is, each of the particular utterances that I coded were able in practice.

The qualitatively derived results here suggest the need to maintain contextual validity in our methodological approaches (see also Allchin, 2011). Importantly, students’ professed epistemologies or views of the nature of professional science should not be expected to align with students’ enacted experiences of doing science themselves in their own local, classroom-based practices in any predictable way. Whether professed views align with in-the-moment practice is an empirical question to be studied and not a valid a priori assumption. In addition, these methods should not assume to identify and measure a universal set of intermediate stages and their order.

It is likely that future methods and measures will need to elucidate relationships that go beyond linear causal dynamics. The kinds of patterns of epistemological development we might expect to see may resemble complex systems dynamics (Thelen & Smith,
This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of individual users and is not to be disseminated broadly.

1998; Witherington, 2011), such as feedback loops and recursive sequences. In addition, identifying the different levels at which these forms of complexity might matter (e.g., the group and the individual, but also ontological levels between epistemic ideas, such as aims and values versus particular disciplinary criteria; Chinn et al., 2011) is an important area for future work.

Implications for Supporting Students’ Meaningful Participation in Science Practices

Figuring out how to effectively support epistemic learning is a central concern for science educators and education researchers. In this article, I focused on the classroom community, by which I mean the teacher and students. Although this definition emphasizes the interactions between teacher and students, it also ignores interactions between the teacher and the curriculum materials (e.g., Brown, 2011), or between students and the material and natural world (e.g., Lehrer & Schauble, 2006). An analysis of how these types of interactions influenced the observed shifts was beyond the scope of this paper. However, such an analysis is critical for understanding how and why the observed shifts happen, allowing us to better design ways of supporting them.

There are a few contextual elements that motivated the selection of this case for study that we know are essential for supporting shifts in use of epistemic considerations. First, the curriculum materials were intentionally designed to support students’ participation in science practices in meaningful ways. That is, there was already a driving phenomenon that students were working to explain, and the activities were structured such that students were doing the “heavy lifting” of figuring out the core science ideas. The participating teachers were also committed to these goals, and their enactments of the curriculum included modifications that even better reflected them. This created a context in which science practices (and the epistemic understandings that guide these practices) are likely important drivers of the kinds of dyadic shifts in use of epistemic considerations. Forthcoming work will connect these processes in detail. For example, all three teachers in this study explicitly introduced the elements of a mechanistic account or the different levels of quality of justification. Instead, the sophisticated versions that gradually emerged were connected to teachers’ careful framing and responses to students’ ideas in moment-to-moment interactions. Current work is examining the dynamics of these processes in detail. For example, all three teachers in this study implicitly framed activities in ways that (a) created a need to figure something out and (b) created comfort for dealing with acknowledging not knowing something. Forthcoming work will connect these teaching moves to the emergence and refinement of epistemic criteria more specifically.

The practice turn in science education emphasizes students’ meaningful participation in science practices: participating in them to figure something out about the natural world, leveraging the tools and criteria of disciplinary science as they are useful and productive for helping to do so. Supporting students in developing the practical epistemic understandings that guide these practices is essential for participation that allows students to drive the knowledge building. This study begins to contribute to our understanding of how classroom communities develop and use these epistemic ideas in practice. Such an understanding is critical for shaping the design of curriculum materials, classroom cultures, and teacher professional development that supports students’ epistemic learning in science.

References


