

**Reasoning with evidence while modeling: Successes at the middle school level**Lauren A. Barth-Cohen<sup>1,2,\*</sup>, Sarah K. Braden<sup>3</sup>, Tamara G. Young<sup>2</sup>, and Sara Gailey<sup>3</sup><sup>1</sup>*University of Utah, Department of Educational Psychology, Salt Lake City, Utah 84112, USA*<sup>2</sup>*University of Utah, Department of Physics & Astronomy, Salt Lake City, Utah 84112, USA*<sup>3</sup>*Utah State University, School of Teacher Education & Leadership, Logan, Utah 84322, USA*

(Received 21 September 2020; accepted 10 May 2021; published 4 August 2021)

Research in undergraduate physics and in K–12 science education has demonstrated challenges and successes in facilitating student engagement with reasoning practices associated with professional physicists. Here we focus on one important dimension of physics reasoning, using evidence to revise models. While this topic has been explored at the undergraduate level, less is known about younger students’ physics reasoning, especially within the context of modeling and model revision, where measurement tools are less emphasized and where evidence is often observational qualitative data. Here we examine 7th graders’ conversations about pieces of observational scientific evidence in their conceptual models of magnetism. We present a series of cases from a whole class discussion to illustrate students’ productive reasoning using evidence while revising, challenging, and testing their scientific models. These cases illustrate how students reason about evidence in the context of canonical and noncanonical models of magnetism, specifically when different pieces of evidence contradict each other and when running the model is inconclusive. The results capture a variety of complex ways evidence can be productively used in reasoning with models, and, more importantly, show that middle school students’ reasoning shares important similarities with more advanced physics reasoning, which can be leveraged when designing and building future instruction.

DOI: [10.1103/PhysRevPhysEducRes.17.020106](https://doi.org/10.1103/PhysRevPhysEducRes.17.020106)**I. INTRODUCTION**

Physics reasoning processes have become a central area of research in physics education research (PER). Reasoning with a model is one important reasoning process because it offers a means for learning about the physical world and a means of understanding scientists’ models of it. Key to modeling are the practices of testing and revising models in light of empirical evidence. Reasoning with a model has been explored within the context of undergraduate physics [1–3], however, the focus rarely includes how evidence is being used in testing and revising the model. At the K–12 level, there is a larger literature on modeling and the importance of evidence [4], but reasoning processes during modeling and specifically the use of evidence-based reasoning while constructing and evaluating models are relatively unexplored in physics [5].

Understanding the experiences and practices of K–12 students in physics is vital to the success of undergraduate physics programs because undergraduate students’ views

and experiences in science are shaped by early learning experiences [6,7] and those experiences can impact students’ university major choices and career interests [8]. Thus, designing modeling instruction in physics at the undergraduate level that builds on the evidence-based reasoning students exhibit in K–12 requires instructional designers to understand how the reasoning experiences of learners at any age fit into a larger trajectory. This study offers a portrait of 7th graders’ uses of evidence while they discussed models of magnetism that were self-generated. We thereby offer a glimpse into how students at one point in the K–12 pathway use evidence to reason with and revise models. While undergraduate and K–12 learning environments differ in instructional approaches and other factors, including the availability of lab equipment, which likely impact the ways students incorporate scientific evidence into their modeling process, researchers can only begin to explore the impacts of these contextual differences on student reasoning by comparing cases such as the one we offer here with cases described in other contexts.

**A. The importance of evidence within modeling**

The field of K–12 science education has experienced a growing focus on scientific models given the inclusion of the scientific practice “Developing and Using Models” within the Framework for K–12 Science Education [9] and

\*Lauren.BarthCohen@utah.edu

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.*

the Next Generation Science Standards [10]. In this context, models are defined as abstract representations of scientific phenomena that enable the user to construct, represent, and communicate knowledge about a phenomenon. Models can reduce the complexity of the scientific phenomenon while also amplifying other key aspects of the phenomenon. They are by design incomplete as they focus on salient information and leave out extraneous information [11,12]. Constructing model-based explanations has been shown to help learners to build knowledge about science content [13–18]. However, the purpose of one learning how to develop a model is not only to learn about the relevant scientific theories and principles, but also to engage in practices that mirror the cognition of actual scientists [4,19].

At the K–12 level modeling is often conceptualized as occurring in four related steps: (1) making observations and collecting data, (2) constructing the model, (3) revising the model, and (4) evaluating the model [11,20–22]. Here we focus on model revision, which is defined as any point during the larger modeling process where one changes the model, often resulting in an increase in complexity of the model [23,24]. Model revision is important for building knowledge as one comes to recognize the models' limits, strengths, and weaknesses [11] while also building knowledge about how the scientific phenomenon works [25]. Crucially, in order to construct, revise, and/or evaluate a model, students must use evidence to guide their reasoning. During the modeling process, the model comes to represent how the scientific phenomenon works, as additional evidence is generated or collected and considered in relation to the developing model. Eventually the model may account for the underlying mechanism or process for how a cause brings about an effect [26]. Model revision often occurs when learners experience a mismatch between the model and the observed evidence. Revision can be motivated by the existence of anomalous data [21,27], or a desire to incorporate previously unknown information [28].

Within the modeling literature, there is consensus that the model should focus on the underlying mechanism that accounts for the scientific behavior, that is, how the model works, rather than surface level features, such as how the model looks [4,29–31]. As a result, during the modeling process, model revision should ideally result in some growth in the fit between the model and the scientific phenomenon it represents. This growth in fit comes about through rejection of components of a model, and revision or elaboration of a model [20,32]. A strong form of model revision lies with the models explanatory power being enhanced when the model is revised, creating a better fit with the evidence [4,30,31]. For example, this may involve the model better capturing the underlying mechanism or process for how a cause brings about an effect. Though students may also exhibit a weaker form of model revision where they change the surface features of the model unmotivated by the evidence or the underlying

mechanism, for example, changing color without attention to a causal process.

During model revision, evidence can be used in multiple ways to support, challenge, or reject an entire model or part of it. But, how exactly do students generate and use evidence in real classrooms while reasoning with models in physics? Furthermore, what happens when different pieces of evidence contradict each other or when students use evidence to run or test a model? While researchers are keenly aware of the importance of evidence, little is known about how students actually use evidence in conversation in classrooms to support their reasoning with models in physics. Furthermore, learners relative expertise can impact their scientific observations which are the material for “evidence” used in scientific reasoning [33,34]. This paper examines how students deploy evidence in classroom conversation while reasoning with and revising models in physics, specifically involving magnetism.

## B. The importance of evidence within modeling in the PER literature

Despite sharing a recognition of the importance of evidence, the K–12 and PER literature differ in how they address modeling. Early work in PER implemented modeling instruction that emphasized the teaching of modeling skills and validation of models in physics. Results demonstrated improved student learning performance [35,36] and improved student attitudes about learning physics content [37,38]. From a different perspective, a body of research by Etkina and colleagues [39,40] has focused on undergraduate students' reasoning processes in lab classes where students exhibit various dimensions of sense making, including using evidence and constructing models, and results have documented the importance of scaffolds in student reasoning [41]. A different perspective has identified the importance of relational reasoning (linear causal reasoning, and simple causal reasoning) within students' explanatory models of the micro-macro levels [42]. While critical in unpacking how students use models, this work to date has not closely examined how students use evidence while engaging in mechanistic reasoning during modeling.

In addition, measurement is recognized as a key to evidence collection within scientific reasoning [43]. A body of research has built a framework for modeling that addresses how measurement tools are used in advanced physics labs [3,44,45]. This framework emphasizes that students need to construct models of the physical phenomena as well as a model of the measurement tools—where the measurement tools measure not the desired quantity, but a related quantity. The initial work [3] focused on upper division labs where students are often asked to verify given models. Additional research has examined how undergraduates engage in the recursive process of modeling through think-aloud interviews aimed at troubleshooting mock data from a formal experiment [1]. More recently,

Vonk *et al.* [45] applied this framework to lower division labs where modeling the measurement tool is less emphasized and the focus is on students creating models of an unfamiliar physical system or engaging in model breaking to determine if familiar models are valid under nonideal conditions. In many ways the approaches taken in these studies share important similarities with the K–12 modeling literature as students examine a real-world physical system, generate a model, collect data, and then further develop the model based on evidence. However, in these studies data and predictions are often quantitative and numerical, while at the K–12 level, data is often qualitative and predictions, although sometimes relevant, are often less emphasized. Furthermore, at younger grades there may be less need for students to generate a model of the equipment [3], or possibly, because of the everyday material, the notion of a model of the quantitative measurement tools is not highlighted in the instruction. Finally, similar to the K–12 level, in these studies, there is an emphasis on data and evidence, but troubleshooting data from a formal experiment is less common at the K–12 level.

Especially relevant to the current study, Russ and Odden [46] examined how the processes of generating evidence and modeling are intertwined and can mutually reinforce each other. Similar to the K–12 literature, these authors emphasized the importance of mechanistic models that account for how a system runs by describing the underlying processes for how a cause can bring about an effect [26]. Russ and Odden [46] focused on the use of observational and everyday evidence, which is also similar to the kinds of evidence that are common at the K–12 level. In their analysis, Russ and Odden examined both evidence use and modeling, and they identified students' productive strategies for recursively intertwining the two, specifically that of using evidence to build or flesh out a model and using a model to determine what evidence to look for next. Our focus on evidence is similar, but we examine middle school students' reasoning during a whole class discussion as compared to Russ and Odden's research on groups of undergraduates working on tutorial worksheets. Furthermore, we examine student reasoning within the underexamined context of magnetism.

### C. Reasoning about magnetism

While magnetism is widely covered in K–12 curriculum and standards [10], there is limited research on K–12 students' reasoning about magnetism, with a notable exception being Cheng and Brown [23] who explored how different kinds of scaffolding of scientific models contributed to more sophisticated models of magnetism. Other work has focused on how teachers implement models of magnetism in the classroom [47]. In PER at the undergraduate level, a significant body of research has used the Conceptual Survey of Electricity and Magnetism (CSEM) and related instruments to measure conceptual

understanding [48,49]. A smaller body of literature has examined students' reasoning in the context of magnetism often focusing primarily on student difficulties [50–52] with relatively minimal attempts to capture the potential for success in younger students' reasoning about this topic. Furthermore, the focus on student difficulties has often occurred within the context of traditional instruction, rather than reform-based approaches emphasizing the process of modeling and the role of evidence in that process.

### D. Theoretical framework: Students' knowledge resources related to modeling and evidence

To operationalize our focus on the use of evidence in modeling, we use a conceptual ecology perspective to conceptualize learners' knowledge systems [53,54]. A knowledge system contains a large number of elements (e.g., intuitions, mental models, schemes, etc.) that operate at a variety of grain sizes in different ways across numerous contexts. Within this approach, a pervasive theme is to emphasize learners' capabilities and potential for success, what some have dubbed "the good stuff" [55]. Specifically, with a broad focus on *productive knowledge resources*, we refer to students' ways of thinking about the external world based on existing experiences that can be fruitfully leveraged for future learning [56]. Although prior work has documented learners' productive knowledge resources and intuitions about physics content, along with other work that has focused on epistemological resources [57,58], the key hypothesis undergirding this research is that all students likely enter the modeling task being able to access productive knowledge resources related to reasoning with models and evidence. Although we are not aiming to specifically identify these resources, this approach is applied to conceptualizing student reasoning with models. When developing models about phenomena too small to see, we expect K–12 students to be able to access intuitions and ways of sense making that allow them to generate new ideas and models for those phenomena. For example, in work focused on how 6th grade students drew models of how odor gets from the source to your nose, students' models contained tiny particles too small to see, which they depicted moving and colliding in such a way that the collisions resulted in a change of direction [59]. Similarly, with our specific focus on magnetism, we also expect students to draw on their experiences, both from in and out of school opportunities to generate models of magnetism that involve things too small to see while also considering the organization and dynamics of the unseen level. Thus, a goal of our analysis is to examine students' sense making while modeling and their uses of evidence with an assumption that those ways of thinking can be leveraged in future learning.

### E. Research question

To investigate middle school students' use of evidence when analyzing models in physics, we ask: How do middle

TABLE I. Overview of instructional design steps in the lesson, time allocated to each step, and research data collected. The final discussion, which was a focus of the current analysis, occurred in step 5.

Instructional sequence	Day in lesson	Total time allotted	Data collected
1. Teacher poses a question or presents an anchoring phenomenon (whole class)	1	15 min	Audio and video recordings of groups
2. Students develop initial diagrammatic models and explanations (individual)	1	15 min	Audio and video recordings of groups
3. Students conduct empirical investigations to gather data for model revision at eight stations (small group and individual)	1, 2, 3	3.5 h	Audio and video recordings of groups
4. Students share individual models with their small group and work to develop consensus models (small group)	3, 4, 5	130 min	Audio and video recordings of groups, pictures of consensus models
5. Students review and evaluate peers' models (whole class discussion)	5	60 min	Audio and video recordings of whole class
6. Students apply their models to explaining related phenomena (individual assessment)	5	20 min	Photocopies of individual participants assessments

school students use evidence when navigating the modeling process? Here we present a series of cases of different ways the students use evidence to analyze their models. Across these cases we examine students' uses of evidence while they engage in the process of revising, challenging, and testing a model. The models contain canonical and non-canonical physics elements, and sometimes the use of evidence results in changes to their models, and sometimes it does not. These exemplar cases were chosen to illustrate a variety of complex ways evidence can be productively used in reasoning with models. This research is explored within the context of a 7th grade class engaged in an instructional unit about modeling magnetism. Our hope is that the results of this analysis will contribute to the development of a deeper understanding of what successful reasoning with models and evidence can look like at the middle school level and beyond.

## II. METHODS

### A. Student context and participants

Data come from a single 7th grade science classroom at a Title 1 charter school in the Intermountain West region. The classroom and school have demographics that mirrored the larger district: 46.8% White and 53.2% racial or ethnic minority (40.0% Hispanic, 4.6% Multi-Racial, 4.3% African American, 2% Asian, and 2.3% Pacific Islander). Furthermore, this population is 13% English Language Learners, 10% receive special education services, and 48.9% of families are eligible for free or reduced lunch. The charter school has a science focus and admission is done by lottery.

The 7th grade science teacher, Dr. Streatfield (a pseudonym) holds a Ph.D. in biology and had been teaching for several years prior to participating in this study. From our classroom observations, she had a strong command of classroom management and a good rapport with the students. She generally taught using a reform-based

approach. At the time of the data collection, the state had recently adopted new science standards that mirrored NGSS [10], but they were not yet widely implemented, and Dr. Streatfield expressed interest in learning more about instructional approaches to support students in scientific modeling.

The research came about from initial conversations between the research team (two science education researchers and a graduate student) and the teacher, as the team had a prior relationship with this teacher. A preliminary version of the magnetism curriculum had been designed by the research team and prior to the instruction, Dr. Streatfield and the researchers implemented a variety of curricular modifications. The students had not received prior instruction on magnetism that school year, although magnetism is included in the 5th grade standards in this state.

### B. Building conceptual models of magnetism curriculum

The two-week instructional unit was designed around the states' 7th grade standard for modeling magnetism. The specific instructional sequence was informed by Passmore, Schwarz, and Mankowski [60] and Kenyon, Schwarz, and Hug [61], and contained the following steps and conditions as summarized in Table I.

As the use of evidence is a critical component of developing scientific reasoning in general and in developing models, we maintained students' focus on gathering and discussing evidence in two places in the curriculum. First, in step 3 (Table I), students visited 8 stations exhibiting various phenomena including magnets and were asked to record primarily qualitative data that they would use as evidence to support their models. The stations and the intended student observations are listed in Table II.

After visiting these stations to collect evidence, we asked students to revise their initial individual models in light of the evidence. In addition to this individual thinking, students then worked in small groups to engage in

TABLE II. List of magnetism stations for data collection and the intended student observations at each station.

Station	Description	Intended observation
1	Students test magnetic and nonmagnetic items.	Some objects are attracted to magnets, some are not. Magnetic objects are different than nonmagnetic objects.
2	Students pick up paperclips with a magnet through pieces of cardboard.	The magnetic field extends beyond the magnet and gets weaker with distance. The magnetic field can go through nonmagnetic items.
3	Students stack ring magnets on a pencil.	Magnets have a north pole and a south pole, and can attract or repel other magnets, depending on the orientation.
4	Students pick up paperclips using a magnet and a nail.	The magnetic field extends through magnetic items, magnetic items can become temporary magnets.
5	Students magnetize a nail, then use that to pick up paperclips.	Magnetic items can be temporarily magnetized.
6	Students use a hammer (and safety goggles) to break a magnet.	Breaking a magnet creates 2 new magnets, each with a north and south pole.
7	Students move around iron filings in a jar using magnets.	The iron filings align with the magnetic field. The students can use this station to determine the shape of the magnetic field.
8	Students cause a paperclip to float with a magnet (and string and tape).	The magnet does not have to be touching the paperclip for the force of the magnet to be stronger than the force of gravity. The force of the magnet gets stronger as the paperclip gets closer.

generating evidence to support their consensus models (step 4, Table I). As mentioned previously, students struggled to cite evidence when explicitly asked to do so in their small groups. However, for a variety of reasons, one being the integral role evidence plays in developing and evaluating models, the teacher was able to engage students in explicit discussions about their evidence for their models when student groups presented their models to the whole class in step 5. Thus, the data analyzed for this paper draw largely from the whole class discussion (step 5, Table I) while students were presenting their model to peers. Comments on students' prior conversations are included only as they are relevant for interpreting the ideas discussed in the whole class discussion. For these reasons, the earlier model building process (steps 1–4, Table I), including the social dimensions of the consensus building process and how they impacted students' discussions of evidence in small groups, is reserved for a separate analysis.

In enacting the curriculum, an additional goal was to promote equity for multilingual and multicultural learners, which includes students labeled as “English Language Learners” (ELLs) by the school, and others who may be underrepresented in science, technology, engineering, and mathematics (e.g., women and girls). The curriculum addresses three possible barriers to the equitable participation of traditionally underrepresented students in science conversations: (1) the accessibility of the anchoring phenomenon to students from various cultural backgrounds; (2) lack of familiarity with the grammatical structures and conversational practices required for participation in science conversations [62,63]; and (3) the tendency for peer groups engaged in science conversations to recreate social hierarchies based on race, language background, class, and gender that exist in society which naturalizes the dominance of White male students in science [64–66].

To address these barriers, we selected a phenomenon involving magnetism that would be highly accessible to learners with varying levels of background knowledge—the floating paperclip. We also developed two sets of conversational scaffolds and a rotating leadership position to guide peer group discussions. Classrooms themselves are complex linguistic and cultural systems. By making the “rules of the game” explicit to learners, as opposed to relying on students' implicit conceptions of how to talk to each other, we hoped to provide multilingual students and other traditionally marginalized students with the grammatical material and alternative conversational formats needed to disrupt the power dynamics that enact processes of marginalization. For example, we developed a set of participation cards to scaffold the consensus-building conversation through asking students to (a) identify similarities and differences in the individual models, (b) identify agreements, (c) discuss how to draw the model and drawing it, and (d) supply evidence to support the model. These phases of the conversation were described with instructions for students on color-coded papers which also served to indicate a rotation in who would lead each phase of the conversation—an attempt to take this role out of the control of the students and to disrupt the domination of a single student voice. We also applied other equity-promoting strategies such as supplying enough materials for each person to have access to work with them, and offering all students ample planning time before being required to share their thoughts with their classmates. For more information about the curriculum please see Braden, Barth-Cohen, Gailey, and Young [67].

### C. Data collection

Audio and video recordings of all aspects of instruction were collected along with researchers' written field notes.

We also collected students' written work, including their final magnetism models and an individual final assessment. This analysis draws from the audio and video recordings of the whole class discussion in step 5, Table I (approximately 60 min), relevant researchers field notes, and the small groups' final hand drawn models. Furthermore, Dr. Braden was a participant observer in the classroom during the entire unit.

#### D. Data analysis

This study takes a multiple case study approach because the goal is to seek an in-depth understanding of the phenomenon [68], specifically how the students use evidence to analyze their models. The analysis began with reviewing notes and video to identify instances of conversations where students explicitly mentioned evidence in relationship to their developing models. We noticed when reading field notes that throughout the instruction students often struggled to cite evidence to support their models, even when explicitly asked to "list evidence." Furthermore, in the small group discussions students often glossed over disagreements in their evidence and models, all of which suggested that those portions of the data would not be as fruitful for answering the research question. However, we also noticed from that initial review that students used evidence to support their reasoning during the final classroom discussion when they were required to display their model and justify their reasoning to their peers and the teacher.

Based on that observation, we transcribed the video from the final classroom discussion and identified instances in which the students explicitly used evidence to challenge or revise their own or their peers models. From the initial transcript we identified a series of preliminary cases in which evidence was used in different ways to analyze the models. During subsequent review of focal cases, the transcripts were updated following conventions modified from conversation analysis [69] which allow for the inclusion of paralinguistic and nonlinguistic features such as pauses, laughter, and co-occurring gestures that add interpretive cues to the linguistic data (see Table III for transcript conventions).

To facilitate a more complete and coherent understanding of each case, we wrote narrative summaries [70] in which we described the model, the evidence being used, the questions being asked, and the outcome of the discussion. In these summaries we paid careful attention to both the canonical and noncanonical dimensions of the models and focused on the students' perspective of the physics. We also made explicit note of the teachers' questions and comments because they often had a strong influence on the discussion. These narrative summaries allowed us to determine when we needed to consider additional data; for example, reviewing student discussions that occurred prior to the whole class presentations and including nonverbal interactions in the analysis.

TABLE III. Transcript conventions.

Symbol	Meaning
.	End of intonation unit; falling intonation
,	End of intonation unit; fall-rise intonation
?	End of intonation unit; rising intonation
!	Raised pitch and volume throughout the intonation unit
∞	Lower volume
:	Length
=	Latching; no pause between intonation units
–	Self-interruption; break in the word, sound abruptly cut off
@	Laughter; each token marks one pulse
[]	Overlapping speech
()	Uncertain transcription
/	Alternate hearings of uncertain transcription
#	Unintelligible; each token marks one syllable
<>	Transcriber comment, nonvocal noise, gesture, or gaze

From those narrative summaries, we enacted interpretations directly from observation [71] as guided by our research question about how the students use evidence to analyze their models. In doing this, we identified places in our data where similarities and differences emerged in how the students use evidence to analyze their models. This process was enacted iteratively with multiple rounds of discussion among the research team to establish internal consistency. We delved within each case to identify how the evidence was used and when looking across multiple cases, we used a contrasting analysis approach to identify differences [72]. Given a common goal in a multiple case study approach of generalizing theoretical propositions [73], here we are able to document the span of different ways individuals perspectives on evidence are productively brought together when reasoning with models.

### III. ANALYSIS: THREE CASES OF STUDENTS USING EVIDENCE IN MODELING

#### A. Prelude to analysis: Challenges in reasoning about magnetism

During the class discussion the students wrestled with a series of important issues pertaining to magnetism. Below, as a prelude to the analysis we briefly summarize each issue as considered by the students because these challenges serve as the context in which students grappled with evidence:

- (i) **Shape of the magnetic field:** Across the multiple groups models, there are different perspectives on the shape of the magnetic field. In several instances, the field is implicitly or explicitly presented as a circle (or oval) encompassing both the magnet and the paperclip and is often presented as decreasing as one moves away from the magnet (case 2,–Sec. III D, and case 3–Sec. III E). However, interestingly, in one instance the students drew the magnetic field as shaped like a



Jimena's comment here ended the group's presentation and the teacher cued the class to transition to a question and answer period. A few minutes into this phase of the presentation Dr. Streatfield tells students to indicate by show of fingers if they agree or disagree with any part of the model. Shay begins whispering and gesturing to her partners to call on her. Shay, begins, "So I kind of disagree with a little part. We did the magnetic field kinda like flower petals" she continues off camera describing and drawing on the board that she thinks the shape of the magnetic field should be different. The shape of the magnetic field had been discussed previously by this group. First, during earlier conversations within the group, the other students commented on the shape of Jimena's magnetic field, but none of the other students expressed

concerns about the four-petaled flower shape. Also, during a prior presentation, in which different students had drawn the magnetic field in concentric circles (see Fig. 2 in case 2–Sec. III D), Lance had pointed out that when you see pictures of the Earth's magnetic field, the lines come out of the north pole and go in at the south pole. It is unclear how these other discussions may have impacted Jimena's description of how she drew the magnetic field. Also, note that her drawn magnetic field is consistent with a quadrupole, however, the students were using dipole magnets in this class. Thus, if a student had examined the more sophisticated quadrupole arrangement of two bar magnets, the four-petaled flower shaped magnetic field would have been accurate.

After Shay points out her concern about the inefficacy of the four-petaled flower shape, she returns to the group who are still standing in front of their poster and mouths a barely audible "sorry!" to Jimena who appears to mouth back "its okay" before turning quickly to face the class again. Shay whispered something additional to Jimena, who nodded while maintaining her gaze towards her classmates. At this point, it is unclear if there was agreement about Shay's alternative drawing for the magnetic field.

A few minutes later, Dr. Streatfield asks the students if they have anything to add to the presented model. Logan, one of the students observing the presentation provides additional evidence to support Shay's point about the magnetic field.

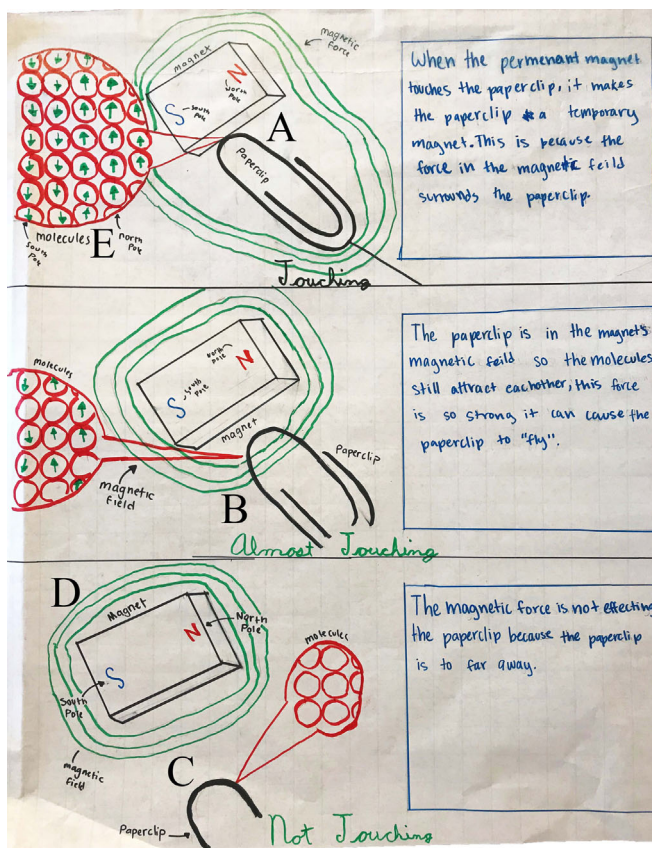


FIG. 2. This model shows three positions of the magnet and paperclip, touching (top, point A), nearby within the magnetic field, but not touching (middle, point B), and outside of the magnetic field (bottom, point C). In each image the magnetic field is drawn with concentric circles surrounding the magnet (e.g., point D) and the small entities inside the magnet are referred to as "molecules" as represented by small circles, some of which have arrows (point E). The north pole is represented by "up" arrows and the south pole by "down" arrows. In the instance when the paperclip is touching, all molecules are magnetic. When the paperclip is nearby but not touching, some of the molecules are magnetic, and when far away the molecules are not magnetic.

453 **Logan:** *(off camera)* Um, so, I would like to add to  
 454 what Shay said, ah, to what Shay said about the  
 455 magnetic field is ah, that shape instead of the  
 456 flower petals. Um because when we were at station  
 457 6,<sup>1</sup> um, you actually, we my group saw that if you  
 458 have the magnet and you put the poles right next  
 459 to the, um, a-a-all the iron fillings in the tube.  
 460 Um, the iron fillings, ah, form this, ah, it  
 461 was—they're like spikes going up, like, um, going up  
 462 as if the lines were coming out of the pole and  
 463 then it was, it was happening on both poles and  
 464 then on the side, instead of forming spikes, the  
 465 um, the fillings kind of, um, smoothed out and  
 466 formed just a little hill.  
 472 **Dr. Streatfield:** That's great evidence. That's  
 473 really great. Awesome.

Logan adds to Shay's prior comment by mentioning new evidence. Specifically, he describes a station where he observed the behavior of the iron filings near the poles of the magnet as compared to near the edge of the magnet (lines 456–466). During this explanation he describes the filings as "spikes" (line 461) that were "coming out of the

<sup>1</sup>Logan's evidence about the iron filings likely came from station 7.



pole” (line 462) and that this was happening at “both poles” (line 463) and results in “a little hill” (line 466). This description represents more of a canonical view of a magnetic field rather than the four-petaled flower shape as illustrated in Fig. 1. Thus, Logan offered additional evidence to support Shay’s proposed change to the way the magnetic field is represented in the model.

All of this suggests a conceptual revision in the model from a magnetic field drawn in the shape of a four-petaled flower to something more resembling the dipole magnetic canonical shape. Interestingly, during this discussion, nobody challenges the suggested change to the shape of the magnetic field. The silence from the other group members and the absence of additional revisions to the shape of the magnetic field may reflect implicit agreement about this change in the model. While the understanding and views of the silent partners cannot be definitely inferred from the conversation, in the assessment following the presentations, the students in this group draw more canonical magnetic fields, which suggests that possibly the silent group members incorporated the ideas from this discussion into their understanding. During this discussion, observational evidence related to the behavior and orientation of iron filings was critical for the revision to the model. Although the model contained several inconsistencies and noncanonical ways of representing information, we see in this example that the students were able to revise one important part of the model using evidence to support their reasoning, and that this process pushed towards a more canonical representation of magnetism.

#### D. Case 2: Using evidence to challenge a model

In the prior case we saw instances of using evidence to support and revise a model with relatively few complications; in contrast, in this case evidence is used to more explicitly challenge a model. This challenge occurs at the very beginning of the whole class discussion when there was a disagreement about temporary magnetism that was not resolved. During the discussion a question arose about whether or not the model should show that the paperclip has a north and south pole. One student (Lance) argued that the paperclip should be drawn with north and south poles, citing evidence from the station where a nail and a permanent magnet are used to pick up paperclips (see Table II, station 4). Another student (Emma) argued that the paperclip should not be drawn like this because temporary magnets don’t have north and south poles. Emma argued with evidence from the station where students created a temporary magnet (see Table II, station 5), that the temporarily magnetized paperclip or nail will stick to another magnet no matter which way it is oriented. Thus, the different pieces of evidence were used to challenge the model and the discussion ended with no

apparent consensus about the existence and representation of poles on temporary magnets.

In this case the students have a disagreement about whether or not the model should show that the paperclip has a north and south pole as indicated by the arrows in the enlargement of bubbles for the views of the magnet touching the paperclip and the floating paperclip (see point E in Fig. 2). Although there are several noncanonical dimensions of this model, this discussion and the disagreement center on temporary magnetism based on different evidence from the stations.

During this group’s presentation, Raven and Emma explain that when situated within a magnetic field, the paperclip becomes a temporary magnet, which has a north and south pole, illustrated by the up and down arrows in the molecules (Fig. 2). The students make the following claims without citing evidence:

- 
- 25 **Emma:** Okay, so basically the first one is showing  
 26 what happens when the magnet and paper clip are  
 27 touching. So, when the permanent magnet touches  
 28 the paper clip, it ah makes the paperclip turn into  
 29 a temporary magnet. This is because the force in  
 30 the magnet—the magnetic field surrounds the  
 31 paperclip and ah attracts the molecules *(reading*  
 32 *from poster; then turns to Raven)*.  
 33 **Raven:** So, when the paperclip is touching the magnet  
 34 the um molecules show the, the arrows  
 35 going up *(pointing)* is the north pole and the  
 36 arrows going down is the south pole.
- 

Continuing the presentation, Lance then provides evidence from station 4 (see Table II) where the students create a temporary magnet with a nail, to support Emma’s prior claim (lines 27–29) and extend the idea to the situation where the magnet and paperclip are not touching.

- 
- 45 **Lance:** Okay, so ah, we got, we believe that the  
 46 touching that the, paperclip essentially becomes  
 47 like a magnet and so in, number 4 it was *(Dr.*  
 48 *Streatfield interjects “Station 4”)*, station 4 ah, we  
 49 had to the nail and when you touched it to a  
 50 magnet, it caused things, like a paperclip, to be  
 51 able to be picked up without it touching the  
 52 magnet. So, that shows that the molecules inside  
 53 the nail have become like a magnet, so it can pick  
 54 up other, um, metal objects that. Certain metal  
 55 objects. Um, and, yeah, that’s ### *(softly*  
 56 *mumbling and trailing off)*.
- 

Later during the question and answer time, Lance was asked to clarify his thinking and he claims that the paperclip becomes a temporary magnet in both the touching and nearly touching situations.

100 **Wesley:** Alright, so, ah the very first thing you said  
 101 that when the magnet was touching the paperclip.  
 102 You said that when its touching it becomes a  
 103 temporary magnet. So does that mean, when you  
 104 touch it, ah it, it also becomes a magnet and  
 105 whatever like, whatever is ferrous that touches the  
 106 paperclip it'll also act like a magnet?  
 107 **Lance:** Yeah, um, that's essentially what's  
 108 happening. Cause like I was explaining this, um,  
 109 becomes a magnet so it can pick up different  
 110 objects, and it c-it also works when it's not  
 111 completely touching it's almost touching. We tried  
 112 and, had it (dangling up at this guy) and touched  
 113 another paperclip to it, and it's a temporary  
 114 magnet because if these are apart are for a while  
 115 *(spreads hands indicating distance)*, this isn't able  
 116 to, pick up anything that's ferrous, but when its  
 117 touching its becomes a magnet until they are  
 118 separated.

At this point evidence related to touching a ferrous object to a floating paperclip has been offered in lines 107–118 to support the claims about temporary magnetism and the north and south poles drawn on the zoom in bubble for the floating paperclip in this group's model. Later, Damian disagrees with Lance's explanation for how the molecules are drawn, questioning the results of the nail station.

170 **Damian:** *(off camera)* Um, I kinda question, your,  
 171 um, discussion about how like the nail is evidence,  
 172 how that, how you, ah, kind of like, attached the  
 173 magnet to the nail and makes it magnetic.  
 174 **Dan:** Do you want to show (him)? *(looking at Lance)*  
 175 **Dr. Streatfield:** *(off camera)* So, you disagree that  
 176 that happens? Or you disagree about their  
 177 explanation? For why it happens.  
 178 **Damian:** *(off camera)* Ah, about their explanation  
 179 for why it happens.  
 180 **Dr. Streatfield:** *(off camera)* Okay, do you have  
 181 something to add to it? Like how you would  
 182 explain it?  
 183 **Damian:** *(off camera)* Um, I don't know, I could  
 184 almost say, like, the magnetic um *(Emma raises*  
 185 *hand)* pull, will ah, almost goes through the nail,  
 186 through the object attached to the nail.

Here Damian argues that the reason the paperclips stick to the nail is not that the nail has become magnetic, but rather that the magnetic field of the magnet goes through the nail (lines 184–185). This reasoning is consistent with the observations students made at station 2 (see Table II) in which they pick up paperclips that are separated from the magnet with a piece of cardboard. In this case, the cardboard does not become a temporary magnet. So, Damian is using evidence consistent with what he observed at station 2 to question, is the paperclip actually a temporary magnet? Or is the magnetic field of the magnet just traveling through the

paperclip? However, it is unclear whether or not Damian explicitly connected his hypothesis to his observations from station 2 as this is not stated in the conversation.

At this point, Emma offers evidence to support Damian's perspective. She uses evidence that could have come from any number of the stations that involved magnets and nails, but this time states that because temporary magnets do not repel other magnets, they do not have north and south poles.

188 **Emma:** Um, I kinda disagree with our #, because  
 189 um *(laughing)* ah um you can't make another  
 190 object at the north or south pole, like, we tried it  
 191 a few times, and like, if you make a temporary  
 192 magnet, then it won't have a north or south pole  
 193 because, like they don't, like not attract  
 194 something else.  
 195 **Dr. Streatfield:** *(off camera)* Oh, so you don't feel  
 196 that repelling.  
 197 **Emma:** Yeah, I think that was wrong.  
 198 **Dr. Streatfield:** *(off camera)* Okay, interesting.

Emma's evidence does not resolve the disagreement among students about temporary magnetism. At this point, Lance has argued that the paperclip becomes a temporary magnet because it will pick up ferrous objects (line 116–117). Damian questioned this argument, offering an alternative explanation that the magnetic "pull," or field, goes through the object (paperclip) when attached to the nail (lines 184–186). While Emma has argued that a temporary magnet will not create a north and south pole in the nail because "we tried it a few times" and the temporary magnet "does not attract something else" (lines 188–194). Meaning that temporary magnets do not repel other magnets as clarified by Dr. Streatfield (lines 194–195).

In this discussion there is an explicit disagreement about temporary magnetism and it is not resolved, despite all of the students supplying evidence to justify their reasoning. The challenge is that the two perspectives on temporary magnetism are built from different evidence and the students and teacher do not work to resolve the differences. Some of the students presenting the model argued that the paperclip is a temporary magnet with a north and south pole, while other students, including those in the presenting group challenged this aspect of the model with evidence that temporary magnets do not have north and south poles and do not repel other things. Part of the complexity in this case is that the students do not negotiate meaning to resolve how the conceptual model could be altered to resolve this conflict between two evidence-based chains of reasoning. That is, there are two different pieces of evidence and the talk does not highlight how the pieces of evidence could be related to each other. There is no questioning of the evidence and no discussion of the quality or criteria by which to judge the evidence. By the end, this disagreement about temporary magnetism was not resolved and there was no apparent consensus.

### E. Case 3: Using evidence to test a model

In the final, and more complicated case, students are testing a model by generating new evidence to resolve questions about the model. Building on the prior conversation from case 2 (Sec. III D), here the students continue to discuss temporary magnetism, anchored by student generated questions about whether a paperclip has a north and south pole. Recall in case 2, the focus was on using evidence to explicitly challenge a model. Here the group attempts to use evidence to test a model, which involves asking questions about the model and considering different scenarios and evidence that might provide insights.

This conversation is launched by a different group presenting a model that shows that the magnetic field of the whole magnet is a result of the magnetic fields of the smaller things inside the magnets. In the student drawings, the entities inside the magnet are represented by two bars, a red one and a blue one (see Fig. 3). Similarly, the magnets are also represented by the same red and blue bar such that the same phenomenon is occurring at the macro and micro levels; the students have correctly identified similar behaviors across the macro and micro levels, which is a known reasoning difficulty [74]. In this model, the two sides of the magnet are identified as having a north and south pole, but this alignment is separate from the colored bars and there are no similar labels on the paperclip. During the discussion below, the students discuss the north and south poles and

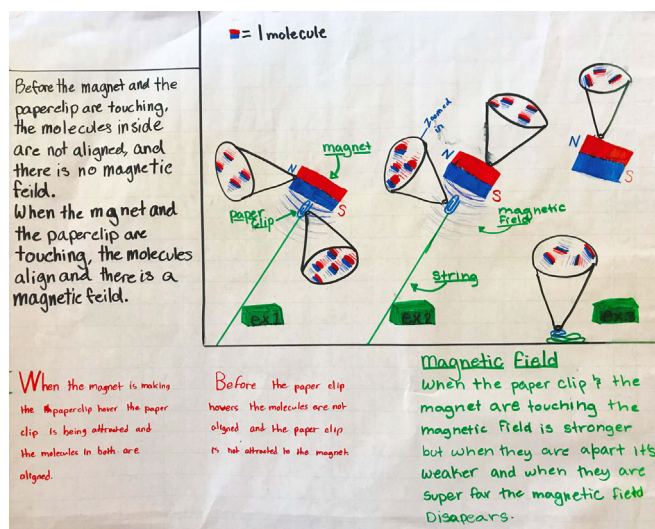


FIG. 3. This model shows three positions, with the paperclip touching the magnet (left), nearby and within the magnetic field (middle), and far away (right). In all instances the magnetic field of the whole magnet is a result of the magnetic fields of the smaller things inside the magnets, in both instances, represented by two bars, a red one and a blue one at the macro and micro levels. The lines indicating the magnetic field are shown getting lighter and lighter in density as you move away from the magnet, demonstrating a decrease in the magnetic field with distance from the magnet.

various pieces of evidence along with other reformulations of the key question. During this process they are testing the given model and considering multiple possibilities, yet, by the end, the main question about temporary magnets having poles is still unresolved.

The discussion begins with Lance asking the key question about whether the paperclip has a north and south pole.

545 **Lance:** *(off camera)* Okay, so, I have a question  
546 about your molecules on the paperclips. So, you  
547 can see that you have a blue and a red side on the  
548 paperclips and I'm assuming, unless I'm blind and  
549 can't see very well. But I was figured that that  
550 means north pole and south pole. So, does the  
551 paperclip, this is rhetorical question, I guess. Does  
552 the paper clip have poles by itself? Or does it  
553 [gain poles when it?  
554 **Diana:** [Well, not necessarily

In asking this question about the paperclips poles, Lance has referred to the two sides of the magnet in the drawing, which he assumes signifies the north and south poles of the magnet. Related to this observation, he then asks if the paperclip has a north and south pole, either "by itself" (line 552), which potentially implies permanent poles, or if it will gain poles (line 553) which potentially implies something that would happen when it is temporarily magnetized. By labeling this question as "rhetorical" (line 551) there are some unknowns. Lance could be assuming the answer is obvious. Or, Lance could be assuming that this question has already been fully answered for others based on the previous conversation that took place during his presentation, thus deflecting potential embarrassment at revealing that for him the question remains unanswered. While Diana begins to answer Lance (line 554) she does not provide a substantive answer to his question. Two other students then present arguments in support of two different options—the paperclip changing poles and the paperclip not having poles:

560 **(Carter):** *(off camera)* Um, I would say that, it has  
561 a north and south pole, but only when another  
562 magnet is on it and it can change which side is  
563 north and which side is south.  
564 **(Lance):** *(off camera)* Okay. I, I don't it's, I don't  
565 think paperclips have north and south poles  
566 because when we go like this, when its connected,  
567 um, like we tried this, we have one paperclip on  
568 each side. Um, and they don't retract, they don't  
569 push each other away. Like they would if it were  
570 two magnets that don't go together. So, I don't  
571 think they have north poles and south poles.  
572 **(Carter):** *(off camera)* Well, because of magnets,  
573 because, I think that, well since we think that the  
574 paperclips can change which side is north and  
575 which side is south. The second one that is, will  
576 be put on, will just switch to the whichever is  
577 attractive.

In lines 560–563, Carter, one of the presenting students, suggests that the paperclip has a north and south pole and when placed near a magnet the poles can change. Following that, Lance, who initially posed the question, disagrees and offers some new evidence in support of a new idea that the paperclip does not have poles. In lines 566–568 Lance cites a scenario that involves “one paperclip on each side” suggesting a magnet with two paperclips, one on each side of it. In this scenario Lance explained that the paperclips do not “retract” or “push away” (lines 568–569). Lance then compares that first scenario to a different scenario, in which two magnets repel each other (lines 569–570). Although there is some ambiguity in this line because the exact arrangements of the paperclips and magnets are unclear, the key point is that Lance claims that paperclips do not repel the same way magnets do. Following that, in lines 572–577 Carter correctly explains that because of the presence of a magnet, the paperclip can change which side is north and which is south and that it will switch in such a way as to be attracted to the magnet.

Following that, Emma immediately questions Carter’s statement with, “why would the paperclip be able to change which way is north and which way is south?” (line 579). Then Logan addresses how a paperclip changes poles:

---

585        **Logan:** Okay um, so the reason is, is because the  
 586        um, the magnet can basically just switch around  
 587        the orientations of the molecules inside the  
 588        paperclip. So um, so if there, the magnet is facing  
 589        one direction, and its next to the paperclip, the  
 590        orientation will be one way. And if it’s facing  
 591        *(pause)* and if it’s facing the other direction, the  
 592        orientation will be a different way. And there  
 593        really is no like boundary of how many times or  
 594        what direction the orientation can be. And so,  
 595        because it’s a temporary magnet, it can be  
 596        attracted to both sides.

---

Here Logan addresses how the paperclip’s poles can change orientations. He explains that the magnet switches the orientation of the molecules inside the paperclip (lines 586–587). As the magnet is facing different directions, the paperclip’s orientation will change (lines 588–592). Following that explanation, over several turns of talk, Emma continues to disagree, arguing that the paperclip cannot switch its poles. Then Eliot joins the conversation (line 611) while maintaining the focus on flipping poles and adding a new explanation for how this may occur, with some accompanying evidence (lines 613–617):

---

611        **Eliot:** *(off camera)* Ah two things. So, like if you  
 612        flipped the orientation of the paperclip wouldn’t  
 613        that mean that the paperclip physically like morph  
 614        in your hands? Cause if its made up so like it has,  
 615        like the metal fillings, it would flip around the

---

*(Table continued)*

*(Continued)*

---

616        poles, you would physically see the little filings flip  
 617        around.  
 618        **Dr. Streatfield:** How many of you saw that? In the  
 619        station with the little filings, if you flipped the  
 620        magnet, did you see the little filings flip over?  
 621        *(some students raise hands)* Okay.  
 622        **Eliot:** *(off camera)* So, would it be that, if you had  
 623        the paperclip, (and you had it hanging) and  
 624        flipped it. Would the paperclip like completely flip  
 625        itself around? Or?  
 626        **Dr. Streatfield:** Thats a really good question.

---

Eliot clarifies the focus on the original question of the paperclip flipping poles and explains that the paperclip would “physically morph in your hand” (lines 613–614). The meaning of this phrase is unclear, with the possibility that the student is referring to an internal reorganization of the paperclip or possibly a more macro level change with the paperclip rotating around. Then in lines 614–617 Eliot continues with new evidence from the station where students had previously explored the behavior and arrangements of iron filings around a magnet (see Table II, station 7). Perhaps the student is hypothesizing that at the microscopic level the inside of the paperclip behaves similarly to the iron filings which flip around. Although some ideas are left ambiguous, as will be shown below, these ideas become foundational for a later discussion. Following Eliot’s initial comment (lines 611–617), Dr. Streatfield and the responding students validate Eliot’s ideas by asking if others saw the iron filings flip over (lines 618–620). Eliot then asks a new question about whether the paperclip would completely flip itself around in certain conditions (lines 622–625). This is a different question than what was being discussed previously about whether the paperclip has a north and south pole. Potentially this reformulation of the question, with its emphasis on the physical movement of the paperclip is easier to answer with observable evidence than the prior question about poles. Dr. Streatfield positively evaluates Eliot’s question (line 626) and goes on to ask the students how they might gather evidence to answer Eliot’s question resulting in students proposing two possible experiments.

---

628        **Dr. Streatfield:** Does anyone have a way we could  
 629        gather some evidence for this? I have an idea but  
 630        I’m actually not entirely sure if my idea is right.  
 631        Anyone have an idea how we could gather evidence  
 632        to figure this out?  
 633        **Andrea:** We:ll, so you’re saying that *(pause)* if we  
 634        (just hold) the paperclip, the paperclip (turns  
 635        north in our hands)? *(turning to look at student*  
 636        *behind her)*  
 637        **Eliot:** *(off camera)* So, if we had a paperclip  
 638        attached to like the north side of it, but we flipped

---

*(Table continued)*

*(Continued)*

639 the papers—and we flip the magnet to the outside.  
 640 Would the paperclip like, change directions? Or  
 641 was it able to ## the magnet.  
 642 **Andrea:** We could use a protractor and and it  
 643 something to see if how the ##### *(loud sound*  
 644 *of chair moving and low volume of speech)*.  
 645 **Dr. Streatfield:** We, I mean we could put a  
 646 paperclip on a magnet and flip it right now. Do  
 647 you want me to grab a magnet and paperclip?  
 648 **Unidentified Student:** We got paperclips and  
 649 magnets here.  
 650 **Dr. Streatfield:** Paperclips and magnets here.  
 651 Okay, so, do we know which is north and south  
 652 pole on these guys?  
 653 **Unidentified Students:** *(talking at same time)*  
 654 **Dr. Streatfield:** Okay so you're saying *(holding*  
 655 *magnet in right hand and paperclip in left hand)*,  
 656 if we have a paperclip, have our magnet, and now  
 657 you want me to flip it to the other side. Like, flip  
 658 this magnet over and see what happens to the  
 659 paperclip? *(pointing and gesturing with paperclip*  
 660 *and magnet)* Is that the idea?  
 661 **Unidentified Student:** Wait what? *(additional*  
 662 *murmurs of confusion)*  
 663 **Dr. Streatfield:** Like, will it not stick? *(holding*  
 664 *magnet and paperclip. Dr. Braden raises hand)* Oh  
 665 yeah you have your hand up *(pointing at*  
 666 *Dr. Braden)*.  
 667 **Dr. Braden:** Well, were you saying, ah, Lance,  
 668 yeah, were you saying that if you take two  
 669 magnets that are repelling each other, and you put  
 670 a paperclip in the middle, the paperclip sticks on  
 671 both sides?  
 672 **Lance:** Yeah, it sticks on both sides.  
 673 **Dr. Braden:** So I think thats=  
 674 **Lance:** =I don't think it has a north or south pole  
 675 because, when you think about it. A paperclip, if  
 676 it had a north or south pole, if you tried to touch  
 677 the north pole of the paperclip to the north pole of  
 678 the magnet, in theory, it would repel. But it  
 679 doesn't. You can do it either way.  
 680 *(mumbling, sidetalk, loud slamming sound)*  
 681 **Dr. Streatfield:** Yeah these ones should be  
 682 repelling each other like I can see them pushing  
 683 they are repelling right now. So yeah these guys  
 684 are repelling and you say that if I stick a mag—or  
 685 a paperclip they really do stick between the two.  
 686 *(teacher demonstrating for a few seconds)*  
 687 Oh, goodness, yes. I love this conversation. I do I  
 688 really do. *(transitions to talking about time)*

In lines 633–636 Andrea appears to attempt to answer Dr. Streatfield's request for a way to gather relevant evidence to support their collective reasoning. However, as she starts talking she appears confused and looks at Eliot for confirmation indicating she may be offering a reformulation of what she understands Eliot's ideas to be. Eliot responds to Andrea by proposing an experiment to

answer his prior question (lines 637–641) about whether the paperclip would flip around. In this experiment, one would flip the magnet while a paperclip is within its field and see if the paperclip changes directions. Andrea immediately adds on (lines 642–644) suggesting that during this experiment one could use a protractor to measure the extent of the paperclip's rotation. In lines 645–663 Dr. Streatfield engages in a back and forth conversation with multiple students to collect the necessary materials and to work out how to perform the experiment, which led to some confusion expressed by students (e.g., lines 661–662).

In response to student confusion, Dr. Braden proposed an alternative experiment which was previously mentioned by Lance. Dr. Braden suggested taking two repelling magnets and placing a paperclip in the middle, to examine if it will be attracted to both sides (lines 668–671). Notice that this suggested experiment is building on a prior mention (lines 566–571) of two magnets repelling. To build on that experimental suggestion, Lance adds that he expects the paperclip would stick to both sides of the two magnets because the paperclip does not have a north or south pole (lines 674–679). At that point Lance provides a thought experiment in that if a paperclip had a north or south pole, and you touched it to the relevant pole of the magnet, then it would repel. But this is not what happens. The paperclip sticks to both sides of the magnet.

Finally, Dr. Streatfield physically enacts the experiment that was proposed by Lance and Dr. Braden (lines 668–671). Lance had stated that two magnets, held together such that identical poles were adjacent, would repel; however, if you inserted a paperclip between the repelling magnets, the paperclip would stick to both magnets (line 672). Dr. Streatfield performs this experiment (Fig. 4) placing a paperclip between two repelling magnets, showing that the paperclip is attracted to both magnets (line 686), even when the magnets are repelling.

In this experiment, a paperclip is placed between two repelling magnets, and sticks to both magnets. However, the fundamental question about whether the paperclip has a north and south pole is left unresolved. The conversation ends as Dr. Streatfield administers the final magnetism assessment which was completed by students individually in silence. It is likely that time pressure forced this conversation to end prematurely and thus contributed to the key question being unresolved.

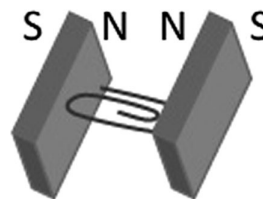


FIG. 4. The experiment performed by Dr. Streatfield.

To summarize, in this complicated case the students initially generate a question about whether the paperclip has a north and south pole, which is based in part on the model showing that a magnet only has a magnetic field when it is touching something else. In response to this question, there was an extended discussion with various arguments proposed pertaining to different arrangements of the magnets and paperclips along with evidence from the observation of metal filings that easily flip around. There was also an attempt at reformulating the key question into something more easily answerable about whether the paperclip can flip around. During this discussion the students proposed and then ran an experiment that involved testing the poles of a paperclip while it is inside of a magnetic field. Although they tested the model, there was no explicit model revision in the whole class conversation. Running the experiment in class may have been completed for demonstration of the possibilities for experimentation, rather than for the model authors to genuinely revise their model or sense making about magnetism.

This case illustrates an instance of the students successfully using evidence to analyze a scientific model, what is often referred to in the literature as testing the model [22]. Furthermore, this instance is unique among the cases presented here because the students use evidence to generate a question and a new observation about the flipping iron filings, and then collect new evidence to support their larger reasoning processes. Additionally, in this instance the teacher played an important role by asking questions throughout the conversation and at the end, while running the experiment, being explicit about what she observed, which is especially important because in a large classroom some might have missed the key observation. This case was more complicated than the prior cases as students were going beyond simply using evidence to support or revise a model and were testing the model with new evidence in an attempt to resolve a question. Similar to case 2 (Sec. III D), this case also contained a focus on the issue of the paperclips temporary magnetism, but here the focus is on using evidence to test a model, while case 2 involved using evidence to challenge a model. Finally, although this case of testing the model didnt result in model revisions, it was still a productive discussion because of the refinement of the foundational question and related evidence and experiments.

#### IV. DISCUSSION

In summarizing the cases, we see that the students used evidence in productive and sophisticated ways to revise, challenge, and test their models. The first case involved straightforward use of evidence to revise a model. In the second case the evidence was used to challenge a model. This case was complicated by multiple pieces of conflicting evidence, and resulted in no resolution. Finally, in the third case the students were testing the model. In doing this, they

generated new evidence in an attempt to resolve a question about the model, but there was no resolution. The second and third cases differ in that in the second case there was an explicit challenge to the model, while in the third case students were testing the model by asking questions and considering different scenarios that might provide new insights. Importantly, these cases show instances of how students use evidence in classroom conversations and what happens when contradicting evidence isn't resolved.

As described previously, we approached this analysis from a resources framework [57] in which we worked from an assumption that students likely enter the modeling task being able to access productive resources related to models and evidence. This perspective allowed us to illustrate students' sense making with models and evidence in the context of magnetism. Prior work on magnetism has often focused on conceptual understanding *vis-à-vis* surveys [48] and students' reasoning difficulties [51]; comparably this approach allowed us to document some of the nuances in how students reasoned about temporary magnetism, magnetic fields, dipole domains, and polarity. We saw an instance of students using observational evidence from magnetic filings to arrive at a magnetic field that is all together reasonable, despite not being canonical. Similarly, students struggled with temporary magnetism and various factors complicated their evidence, resulting in an unresolved disagreement, but yet they wrestled with important issues related to ferromagnetic materials. These nuances in middle school students' reasoning about magnetism are important for conceptualizing the ways in which our entering undergraduate physics students might conceptualize this topic and how we build on their existing understanding in future instruction.

More broadly, the results show that middle school students are able to use scientific evidence in various ways when reasoning with models at a level of nuance and complexity that shares important similarities with more advanced physics reasoning [3,44–46]. Our results directly support those from Russ and Odden [46], who identified undergraduate students' productive strategies for recursively interweaving evidence throughout the modeling process. Specifically, the conversations in this analysis demonstrated use of evidence to build or flesh out a model and use of a model to determine what evidence to look for next. Similarly, these results build on Vonk *et al.* [45] and Dounas-Frazer *et al.* [1], given the common recognition of the importance of generating models and collecting data to further reasoning with models based on evidence. Here we found similar results with middle school students whose models and evidence were all qualitative, which suggests that despite differences in the details of the models and evidence, similar reasoning is seen across levels. Furthermore, our analysis shows how through interaction and conversation, evidence is layered into the talk of learners. In particular, we saw in the data many ways that

students used evidence, not only to revise, or refute a model, but to generate new evidence to resolve questions, to challenge a model, and to test a model. In all of these cases, the use of evidence is layered into the conversation, and sometimes the subject of unresolved disagreements. Documenting that undergraduates and middle school students use evidence in similar ways when engaging in modeling in physics is important for recognizing parallels between younger students' and undergraduates' physics reasoning.

In all cases, we see that the students attempt to explain the mechanism for the macroscopic polarity of the magnet by illustrating and describing the behavior of the microscopic elements, with noticeable ambiguity (case 1–Sec. III C), as pairs of monopoles (case 2–Sec. III D) and as dipoles (case 3–Sec. III E). The difficulty in determining the dipole nature of the unseen elements is consistent with how the topic is approached in physics curriculums [75]. In both the second and third case the students struggle with the nature of temporary magnetism, specifically, does the paperclip become a temporary magnet in the presence of a magnetic field? As described in Sec. III A, the degree to which a ferromagnetic material becomes temporarily magnetized can vary greatly, which may result in conflicting observations. While the students do not resolve this issue, they use conflicting evidence to support and refute this. More generally, the students' struggles with magnetism may be an instance of the known challenge in reasoning about the relationship between the unseen microscopic and macroscopic levels in complex systems [74].

We see that while in all cases the students do not draw the shape of the magnetic field correctly, there are some canonical aspects to it, and in case 1 (Sec. III E), the students are able to use evidence to revise their model to include a more canonical magnetic field shape. While some research has documented that undergraduates can struggle with the shape of the magnetic field in the context of vector fields [76], our work showed that younger students can also struggle with the shape of magnetic fields due to difficulties interpreting evidence. Yet, we also saw productive sense making in this space.

Future research is warranted to look at student understanding of temporary magnetism and magnetic fields. Our theoretical framework enabled us to examine students' sense making by attuning the focus of the analysis to how students reason with evidence in various modeling processes. However, future work might capture learners' specific knowledge resources related to evidence and models [57] in order to document changes in learners' knowledge systems during the processes involved with reasoning with models. This type of analysis would go a step beyond the current work by examining not only middle school students using evidence to revise a model, but how that process is impacted by and with their knowledge resources, which in turn contributes to longer term learning. This analysis

demonstrates empirically the complexities and productive ways in which evidence was used while students reasoned with models and is therefore foundational for additional theoretical work that may follow.

These results also have implications for instruction that incorporates modeling and reasoning with evidence. We noticed that rich evidence-based discussions occurred less frequently in small group conversations where students struggled to cite evidence, even when explicitly asked to do so. Comparably, the whole class conversation, which included the teacher's thoughtful questions and encouragement of student questions, was central for students' reasoning with evidence. For others aiming to implement similarly structured conversations that foreground evidence, it can be important to scaffold the process of students sharing their own models, listening to peers present their models, and finally questioning peers' evidence [67]. More generally, scaffolds that encourage the sharing of evidence while taking into account the social dimensions of the consensus building process may be key to this success. Additionally, instructors may benefit from supporting students in judging the relevance and quality of the evidence, both at the middle school level and beyond.

Given the nature of the data analyzed and the classroom context, there are some limitations of this analysis. As students worked in small groups there were complicated social dynamics that impacted their use of evidence in conversation and shaped the ideas that became depicted in their final models. These dynamics included issues such as unresolved disagreements about evidence and dimensions of the models, as well as some silence from some students. While instructionally the models were representations from small groups—the models did not necessarily represent actual consensus of previously negotiated ideas. We also observed students struggle with scientific vocabulary and with making the intended scientific observations at each station, these factors likely impacted their discussions and overall reasoning processes. Also, the teachers questions and scaffolds had a notable impact on students' reasoning processes as they pinpointed key pieces of evidence and models and thereby impacted the trajectory of the whole class discussion and students' reasoning. In applying these results to other settings, it is important to remember details of the classroom context, including the teachers role in guiding students' conversations, and the larger purpose of the instructional unit which in this case was implemented through a collaboration between educational researchers and classroom teachers.

Despite these limitations, the results demonstrate that middle school students are capable of sophisticated uses of evidence to support their reasoning when building and evaluating models. Our student models were qualitative and conceptual, and they contain well-reasoned scientific ideas about magnetism and support key physics reasoning skills, including questioning, predicting, and running a model.

Furthermore, we saw that such scientific reasoning can come about even when aspects of the model are noncanonical, which suggests that future work needs to examine students' models holistically and consider multiple dimensions of a model, rather than reducing them to a binary of correct or incorrect. While similar results about what is possible with K–12 student reasoning about models has been documented in the science education literature [18], documenting this reasoning in physics is important for further considerations about how K–12 physics instruction can impact undergraduate physics instruction. Specifically, these results suggest that students can enter their undergraduate physics courses with prior experiences with sophisticated physics reasoning and those prior experiences

could be a potential resource for future learning. Furthermore, these results, in agreement with other literature focused on the K–12 level [5], show that younger students are capable of complex and nuanced physics reasoning. Nonetheless, there continue to be open questions about this understudied population which must be explored in future research.

## ACKNOWLEDGMENTS

We thank the students and teachers who helped contribute to this research by opening up their classroom and embracing this collaboration. We would like to thank Taylor Dexter for assisting with data collection.

- 
- [1] D. R. Dounas-Frazer, K. L. Van De Bogart, M. R. Stetzer, and H. Lewandowski, Investigating the role of model-based reasoning while troubleshooting an electric circuit, *Phys. Rev. Phys. Educ. Res.* **12**, 010137 (2016).
- [2] T. O. B. Odden and R. S. Russ, Sensemaking epistemic game: A model of student sensemaking processes in introductory physics, *Phys. Rev. Phys. Educ. Res.* **14**, 020122 (2018).
- [3] B. M. Zwickl, D. Hu, N. Finkelstein, and H. Lewandowski, Model-based reasoning in the physics laboratory: Framework and initial results, *Phys. Rev. ST Phys. Educ. Res.* **11**, 020113 (2015).
- [4] C. V. Schwarz, B. J. Reiser, E. A. Davis, L. Kenyon, A. Achér, D. Fortus, Y. Shwartz, B. Hug, and J. Krajcik, Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners, *J. Res. Sci. Teach.* **46**, 632 (2009).
- [5] A. L. Stephens and J. J. Clement, Documenting the use of expert scientific reasoning processes by high school physics students, *Phys. Rev. ST Phys. Educ. Res.* **6**, 020122 (2010).
- [6] L. A. Bryan, Nestedness of beliefs: Examining a prospective elementary teacher's belief system about science teaching and learning, *J. Res. Sci. Teach.* **40**, 835 (2003).
- [7] C. J. Eick and C. J. Reed, What makes an inquiry-oriented science teacher? the influence of learning histories on student teacher role identity and practice, *Sci. Educ.* **86**, 401 (2002).
- [8] P. R. Aschbacher, E. Li, and E. J. Roth, Is science me? high school students' identities, participation and aspirations in science, engineering, and medicine, *J. Res. Sci. Teach.* **47**, 564 (2010).
- [9] National Research Council, *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Academies Press, Washington, DC, 2012).
- [10] NGSS Lead States, *Next Generation Science Standards: For States, by States* (National Academies Press, Washington, DC, 2013).
- [11] R. Lehrer and L. Schauble, *Cultivating Model-Based Reasoning in Science Education* (Cambridge University Press, Cambridge, England, 2006).
- [12] R. Duschl, Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals, *Rev. Res. Educ.* **32**, 268 (2008).
- [13] C. T. Forbes, L. Zangori, and C. V. Schwarz, Empirical validation of integrated learning performances for hydrologic phenomena: 3rd-grade students' model-driven explanation construction, *J. Res. Sci. Teach.* **52**, 895 (2015).
- [14] J. L. Jenkins and E. M. Howard, Implementation of modeling instruction in a high school chemistry unit on energy and states of matter, *Sci. Educ. Int.* **30**, 97 (2019).
- [15] R. Lehrer and L. Schauble, Seeding evolutionary thinking by engaging children in modeling its foundations, *Sci. Educ.* **96**, 701 (2012).
- [16] E. Manz, Understanding the codevelopment of modeling practice and ecological knowledge, *Sci. Educ.* **96**, 1071 (2012).
- [17] C. V. Schwarz and B. Y. White, Metamodeling knowledge: Developing students' understanding of scientific modeling, *Cognit. Instr.* **23**, 165 (2005).
- [18] L. Zangori and C. T. Forbes, Development of an empirically based learning performances framework for third-grade students model-based explanations about plant processes, *Sci. Educ.* **100**, 961 (2016).
- [19] M. Won, H. Yoon, and D. F. Treagust, Students learning strategies with multiple representations: Explanations of the human breathing mechanism, *Sci. Educ.* **98**, 840 (2014).
- [20] L. T. Louca and Z. C. Zacharia, Modeling-based learning in science education: cognitive, metacognitive, social, material and epistemological contributions, *Educ. Rev.* **64**, 471 (2012).
- [21] J. Stewart, J. L. Cartier, and C. M. Passmore, Developing understanding through model-based inquiry, *How Students Learn* (National Academies Press, Washington, DC, 2005), pp. 515–565.



- [22] M. Windschitl, J. Thompson, and M. Braaten, Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations, *Sci. Educ.* **92**, 941 (2008).
- [23] M. F. Cheng and D. E. Brown, The role of scientific modeling criteria in advancing students' explanatory ideas of magnetism, *J. Res. Sci. Teach.* **52**, 1053 (2015).
- [24] J. P. Gutwill, J. R. Frederiksen, and B. Y. White, Making their own connections: Students' understanding of multiple models in basic electricity, *Cognit. Instr.* **17**, 249 (1999).
- [25] E. Manz, Understanding the codevelopment of modeling practice and ecological knowledge, *Sci. Educ.* **96**, 1071 (2012).
- [26] R. S. Russ, R. E. Scherr, D. Hammer, and J. Mikeska, Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science, *Sci. Educ.* **92**, 499 (2008).
- [27] R. N. Balliet, E. M. Riggs, and A. V. Maltese, Students' problem solving approaches for developing geologic models in the field, *J. Res. Sci. Teach.* **52**, 1109 (2015).
- [28] R. G. Tobin, S. J. Lacy, S. Crissman, and N. Haddad, Model-based reasoning about energy: A fourth-grade case study, *J. Res. Sci. Teach.* **55**, 1134 (2018).
- [29] L. K. Berland, C. V. Schwarz, C. Krist, L. Kenyon, A. S. Lo, and B. J. Reiser, Epistemologies in practice: Making scientific practices meaningful for students, *J. Res. Sci. Teach.* **53**, 1082 (2016).
- [30] D. E. Penner, N. D. Giles, R. Lehrer, and L. Schauble, Building functional models: Designing an elbow, *J. Res. Sci. Teach.* **34**, 125 (1997).
- [31] C. Schwarz, B. J. Reiser, A. Acher, L. Kenyon, and D. Fortus, Models: Challenges in defining a learning progression for scientific modeling, *Learning Progressions in Science* (Brill Sense, Rotterdam, 2012), pp. 101–137.
- [32] J. Stewart, R. Hafner, S. Johnson, and E. Finkel, Science as model building: Computers and high-school genetics, *Educ. Psych.* **27**, 317 (1992).
- [33] L. Barth-Cohen and S. K. Braden, Unpacking the complexity in learning to observe in field geology, *Cogn. Instr.* (to be published).
- [34] C. Eberbach and K. Crowley, From everyday to scientific observation: How children learn to observe the biologists world, *Rev. Educ. Res.* **79**, 39 (2009).
- [35] I. A. Halloun and D. Hestenes, Modeling instruction in mechanics, *Am. J. Phys.* **55**, 455 (1987).
- [36] I. Halloun, Schematic modeling for meaningful learning of physics, *J. Res. Sci. Teach.* **33**, 1019 (1996).
- [37] E. Brewster, L. Kramer, and G. O'Brien, Modeling instruction: Positive attitudinal shifts in introductory physics measured with class, *Phys. Rev. ST Phys. Educ. Res.* **5**, 013102 (2009).
- [38] M. B. Brewer, *Intergroup relations* (Oxford University Press, 2010).
- [39] E. Etkina, A. Warren, and M. Gentile, The role of models in physics instruction, *Phys. Teach.* **44**, 34 (2006).
- [40] E. Etkina, A. Karelina, and M. Ruibal-Villasenor, How long does it take? A study of student acquisition of scientific abilities, *Phys. Rev. ST Phys. Educ. Res.* **4**, 020108 (2008).
- [41] A. Karelina and E. Etkina, Acting like a physicist: Student approach study to experimental design, *Phys. Rev. ST Phys. Educ. Res.* **3**, 020106 (2007).
- [42] A. Leniz, K. Zuza, and J. Guisasaola, Students reasoning when tackling electric field and potential in explanation of dc resistive circuits, *Phys. Rev. Phys. Educ. Res.* **13**, 010128 (2017).
- [43] N. Holmes and D. Bonn, Doing science or doing a lab? Engaging students with scientific reasoning during physics lab experiments, in *Proceedings of the 2013 Physics Education Research Conference* (AIP, New York, 2013).
- [44] J. T. Stanley, W. Su, and H. Lewandowski, Using lab notebooks to examine students engagement in modeling in an upper-division electronics lab course, *Phys. Rev. Phys. Educ. Res.* **13**, 020127 (2017).
- [45] M. Vonk, P. Bohacek, C. Militello, and E. Iverson, Developing model-making and model-breaking skills using direct measurement video-based activities, *Phys. Rev. Phys. Educ. Res.* **13**, 020106 (2017).
- [46] R. S. Russ and T. O. B. Odden, Intertwining evidence-and model-based reasoning in physics sensemaking: An example from electrostatics, *Phys. Rev. Phys. Educ. Res.* **13**, 020105 (2017).
- [47] D. B. Harlow, Structures and improvisation for inquiry-based science instruction: A teacher's adaptation of a model of magnetism activity, *Sci. Educ.* **94**, 142 (2010).
- [48] M. W. McColgan, R. A. Finn, D. L. Broder, and G. E. Hassel, Assessing students conceptual knowledge of electricity and magnetism, *Phys. Rev. Phys. Educ. Res.* **13**, 020121 (2017).
- [49] D. P. Maloney, T. L. OKuma, C. J. Hieggelke, and A. Van Heuvelen, Surveying students conceptual knowledge of electricity and magnetism, *Am. J. Phys.* **69**, S12 (2001).
- [50] E. Bagno and B.-S. Eylon, From problem solving to a knowledge structure: An example from the domain of electromagnetism, *Am. J. Phys.* **65**, 726 (1997).
- [51] K. Jelacic, M. Planinic, and G. Planinsic, Analyzing high school students reasoning about electromagnetic induction, *Phys. Rev. Phys. Educ. Res.* **13**, 010112 (2017).
- [52] J. Park, Modelling analysis of students processes of generating scientific explanatory hypotheses, *Int. J. Sci. Educ.* **28**, 469 (2006).
- [53] D. E. Brown and D. Hammer, Conceptual change in physics, in *International Handbook of Research on Conceptual Change* (Routledge, London, 2009), pp. 155–182.
- [54] A. A. diSessa, Why conceptual ecology is a good idea, in *Reconsidering conceptual change: Issues in theory and practice* (Springer, New York, 2002), pp. 28–60.
- [55] D. Hammer and E. van Zee, The beginnings of scientific reasoning, in *Seeing the Science in Children's Thinking: Case Studies of Student Inquiry in Physical Science* (Heinemann, New York, 2006), pp. 13–37.
- [56] D. Hammer, Student resources for learning introductory physics, *Am. J. Phys.* **68**, S52 (2000).
- [57] D. Hammer, A. Elby, R. E. Scherr, and E. F. Redish, Resources, framing, and transfer, in *Transfer of Learning from a Modern Multidisciplinary Perspective* (Age Publishing, Greenwich, CT, 2005), Vol. 89.
- [58] D. Hammer and A. Elby, On the form of a personal epistemology, in *Personal Epistemology: The Psychology*

- of Beliefs about Knowledge and Knowing* (Routledge, Mahwah NJ, 2002), Vol. 169190.
- [59] J. Krajcik and J. Merritt, Engaging students in scientific practices: What does constructing and revising models look like in the science classroom?, *Sci. Teach.* **79**, 38 (2012), <https://www.proquest.com/docview/1000411241?pq-origsite=gscholar&fromopenview=true>.
- [60] C. Passmore, C. V. Schwarz, and J. Mankowski, Developing and using models, in *Helping Students make Sense of the World using Next Generation Science and Engineering Practices* (NSTA Press, Arlington VA, 2017), p. 109.
- [61] L. Kenyon, C. Schwarz, and B. Hug, The benefits of scientific modeling, *Sci. Child.* **46**, 40 (2008), <https://www.proquest.com/docview/236902580?pq-origsite=gscholar&fromopenview=true>.
- [62] S. Michaels and C. O'Connor, Conceptualizing talk moves as tools: Professional development approaches for academically productive discussion, in *Socializing Intelligence through Talk and Dialogue* (2015), p. 347.
- [63] D. Fisher, C. Rothenberg, and N. Frey, *Content-Area Conversations: How to Plan Discussion-Based Lessons for Diverse Language Learners* (ASCD, Alexandria, VA, 2008).
- [64] K. O'Connor, Communicative practice, cultural production, and situated learning: Constructing and contesting identities of expertise in a heterogeneous learning context, in *Linguistic Anthropology of Education*, Praeger, Westport, CT, edited by S. Wortham and B. Rymes (2003), 61.
- [65] S. K. Braden, Linguistic expertise, mockery, and appropriateness in the construction of identities: A case study from 9th grade physics, *Bilingual Res. J.* **42**, 432 (2019).
- [66] S. K. Braden, Scientists can't really talk to people: Unpacking students metacommentary on the racialized and gendered science nerd trope, *Int. J. Multicult. Educ.* **22**, 87 (2020).
- [67] S. K. Braden, L. Barth-Cohen, S. Gailey, and T. G. Young, Modeling magnetism with the floating paper clip: Supporting diverse learners with visual and linguistic scaffolds, *Sci. Scope* (to be published).
- [68] J. Creswell and C. Poth, *Qualitative Inquiry and Research Design*, 4th ed. (Sage Publishing, Thousand Oaks, CA, 2018).
- [69] J. DuBois and S. Cumming, Discourse transcription, in *Santa Barbara Papers in Linguistics*, edited by S. Schuetze-Coburn and D. Paolino (Department of Linguistics, University of California, Santa Barbara, CA, 1992), Vol. 4.
- [70] D. E. Polkinghorne, Narrative configuration in qualitative analysis, *Int. J. Qual. Stud. Educ.* **8**, 5 (1995).
- [71] R. E. Stake, *The Art of Case Study Research* (Sage Publishing, Thousand Oaks, CA, 1995).
- [72] M. B. Miles and A. M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook* (Sage Publishing, Thousand Oaks, CA, 1994).
- [73] R. K. Yin, *Case Study Research and Applications: Design and Methods* (Sage Publishing, Thousand Oaks, CA, 2017).
- [74] P. Sengupta and U. Wilensky, Learning electricity with Niels: Thinking with electrons and thinking in levels, *Int. J. Comp. Math. Learn.* **14**, 21 (2009).
- [75] F. M. Goldberg, S. Robinson, and V. K. Otero, *Physics for Elementary Teachers* (It's about time, Mount Kisco, New York, 2006).
- [76] L. Bollen, P. van Kampen, C. Baily, M. Kelly, and M. De Cock, Student difficulties regarding symbolic and graphical representations of vector fields, *Phys. Rev. Phys. Educ. Res.* **13**, 020109 (2017).