Idiosyncratic cases and hopes for general validity: What education research might learn from ecology

David Hammer[a], Julia Gouvea[a], Jessica Watkins[b]

[a]Tufts University; [b]Vanderbilt University (as of August, 2018)

To appear, *Infancia y Aprendizaje*

Abstract

We reflect on our ongoing struggles with rigor and validity in a project based on case study analyses: How can particular instances of learning, with all their idiosyncratic details and dynamics, contribute findings of lasting value to education? For this essay, we look to the field of ecology, which has faced similar challenges, for insights into the goals and methods of research.

Keywords: Case studies; methodology; science education

Running head: Idiosyncratic cases and general validity

Posing the Question

Consider the following episode of children's science, third graders in a United States public school working to understand what affects the motion of a toy car.[1]

The unit had started several weeks earlier with a "launching question" (Hammer, Goldberg, & Fargason, 2012): What are ways to get a toy car moving? The children quickly thought of many possibilities, and over several days with their teacher's guidance they settled into experimenting with ramps and rubber bands. Some of their ideas about ramps centered on rolling, and one student, Isaac, got to thinking about how rolling works.

---

[1] The class was part of a research project in scientific inquiry, focused on "responsive teaching" (Goldberg et al., 2008-2011). Video and description of *Isaac's Wheels* are available at studentsdoingscience.tufts.edu. See also Hammer and Radoff (2014).

The episode, which we call *Isaac's Wheels,* began when he spoke up in class to say that "it matters" if the toy car has wheels. The teacher, Sharon Fargason, asked him to read something she had seen in his journal: "The car goes faster becase its weels ceep track uv the floor" ["The car goes faster because its wheels keep track of the floor"]. He had also made several sketches, of interlocking gears, a bicycle, and a bicycle wheel with ridges or maybe gear teeth on the outside.

With Ms. Fargason's support, Isaac explained his idea about rolling: Each point on a wheel lines up with a point on the ground, and this lets the wheel move forward without dragging. It was not easy to follow his thinking, and his classmates had questions. Jourdan asked what would make the car stop, and Isaac answered that the wheels would get "tired." Scarlett asked, "How could the wheels get tired?" That got Isaac looking for a more tangible mechanism to explain slowing: There is "scratching," he said, between the rubber wheels and the "little strings" (axles) that hold them.

This is one of many examples we have collected and studied, over years of research on students' inquiries. For physicists, *Isaac's Wheels* is impressive and easy to recognize as nascent science: Physicists see Isaac's theory of rolling as correct, recognize Jourdan's question as central, and share Scarlett's concern about Isaac's initial response. Most are amazed at what third graders are able to do, and it motivates study of how this quality of thinking could arise (Radoff, 2017).

We take as a premise that this and other episodes—more precisely the video, audio, written and material records of them—are valuable for research on learning in science. The question is how: How can an episode like this, so clearly an example of something science educators would like to see, contribute to research on learning—rigorous research that could produce findings of general, lasting value to science education?

Existence proofs and theory building.

Of course this question is not new. The education research community has long debated and divided over the respective merits and validities of qualitative and quantitative research. We do not review that literature here; our scholarship is about learning and instruction in science, not about methodology. However, of course, we cannot avoid methodology as it concerns our work. We take the opportunity of this article to reflect on our wrestling over the quality and rigor of our research on episodes like *Isaac's Wheels*.

In some ways, it is clear what studying an episode can accomplish. For one, it can be an existence proof to challenge widely held assumptions. Here, for example, Jourdan was asking for a mechanism to explain why the car *stops,* against the expectation that novice physics involves the robust misconception that motion must have a cause (McCloskey, 1983). Isaac's first answer that the wheels get tired matched the misconception, but Scarlett immediately challenged it as implausible, and Isaac moved on to identify a mechanism for slowing.

There are many examples of this use of episodes in the literature, showing instances of students' abilities for epistemic reasoning (Metz, 2011), reasoning about emergent phenomena (Levy & Wilensky, 2008), algebraic sense-making (Brizuela & Schliemann, 2004), and many others. We have used case studies in this way too, including to document instances of mechanistic reasoning (Louca, Elby, Hammer, & Kagey, 2004; Russ, Scherr, Hammer, & Mikeska, 2008), problem scoping (Watkins, Spencer, & Hammer, 2014), and modeling (Svoboda & Passmore, 2013).

For existence proofs, the answer to our question seems straightforward, since the claim only concerns the case at hand. It would be difficult to contest the finding that Jourdan was thinking about what causes the toy car to stop or that Scarlett did not accept the answer that wheels get tired, and in this way the case is a rigorous demonstration that learners are not *always* driven by the robust misconception that motion needs a cause, as is argued in some literature (e.g., McCloskey, 1983).

Another use of case studies is to develop novel theoretical constructs and conjectures, such as of *meta-representational competence* (MRC, diSessa et al., 1993) and *productive disciplinary engagement* (PDE, Engle & Conant, 2002). Our recent work includes Watkins et al.'s (2018) use of case studies to offer conjectures about professional development for teachers in engineering education; and Radoff, Jaber, and Hammer's (in press) case study of one student's progress to posit the notion of *meta-affective learning*.

There can certainly be lasting value of new theoretical ideas in education research in the form of *generativity:* novel conceptions of how people think and learn expand our ability to imagine what is possible. MRC and PDE have clearly inspired this sort of thinking in the field.

Generativity is a marker of progress across many disciplines. In the arts, creative leaps inspire painters, writers, and musicians in new directions. In the natural sciences, progress is often punctuated by novel constructions that make is possible to think about the world in new ways. It took creative insight to imagine the space between two magnets as occupied by "lines of force," responsible for the push and pull of objects as a distance. This idea opened up a new way of conceptualizing "empty" space (Nersessian, 1992).

We also have aspirations, however, for findings of general lasting *validity*, in a sense we have experienced in the natural sciences. While we value the creative metaphor of *lines of force*, we are also concerned whether the model will endure, whether it will stand up to empirical testing and theoretical critique. *Lines of force* led a to more general and mathematically rigorous theory of "electromagnetic fields," which has stood extensive testing and critique, the 'test of time:' It continues to play a central role in physics, lasting in itself as well as leading to other still more general constructs of field theory.

Our backgrounds are in physics (David and Jessica) and biology (Julia), and they have shaped how we understand the epistemic aims, values, virtues, and processes of research (Chinn, Buckland, & Samarapungavan, 2011). Physics in particular has emphasized the search for fundamental laws, which has had an influence but not been as central in biology. What, we are wondering, are productive aims and values in education research?

Maybe we need to adjust our aspirations. Analytic philosophers aspire to produce findings of lasting validity, but our colleague Nancy Bauer (2015) argues they should not, that progress in philosophy is like progress in the arts. Painters, writers, and musicians develop new ideas and hope they will influence later work, but do not hold those ideas accountable to evidence: Impressionism was progress, but of a different form than the formulation of field theory.

Refining the question.

Within the natural sciences and engineering there are robust methodologies for producing and assessing valid findings. Perhaps the most widely known is the "randomized controlled trial" (RCT), basically dividing a population at random into different conditions and comparing the results. There are excellent reasons for considering RCTs the "gold standard" in some areas, including in medical research: One group gets the treatment and a control group gets a placebo, and the study looks for a difference in measured outcomes, ideally showing a large effect size and statistical significance. There have been many productive uses for RCTs in education research as well, for the study of hypotheses that afford well-defined independent and dependent variables. Scholars have used RCTs or similar methods to show the efficacy of instructional interventions (Carpenter, Fennema, Peterson, Chiang, & Loef, 1989) to test theoretical constructs such as stereotype threat (Steele & Aronson, 1995), and to study "preparation for future learning" (Schwartz & Marin, 2004).

There are also excellent reasons for questioning the gold standard status of RCTs, including that their use and validity depend on a restrictive set of conditions for the target of study – that there is no important variation within the study population for example (Cartwright, 2007).

For our purposes in this essay, the need for well-defined independent and dependent variables—that is, "input" features of the situation and "output" results—highlights the heart of the problem with prospecting for generally valid relationships in cases such as *Isaac's Wheels*: How do we identify and define those variables? What features of this situation are relevant and clearly applicable in other situations?

In Isaac's third grade class, possible common features to study might be the launching question, "what are ways to make the car start moving?" or the use of journals. Indeed, these were shared features across third-grade classes in the project.

But much more about this episode was idiosyncratic. Isaac's analogy to gears happened only in this class, and it was key. If we wanted to include it as part of more general study, we would need to treat it as an instance of a more general category, possibly of analogy (Blanchette & Dunbar, 2000) or of mechanistic reasoning (Russ et al., 2008). That abstraction would involve looking across cases to discover and refine categories, and it generally involves nuanced subjective judgment.

There were also the particular ways Ms. Fargason recognized, valued, and engaged with students' ideas, including Isaac's journal entry and explanation during class, as well as Jourdan's, Scarlett's, and other students' questions. They too were central to the dynamics of this episode, but we could not make them, precisely, the independent or dependent variables of more general study. Again, we would need to look across cases, including of other teachers, to abstract more general categories (Levin, Hammer, & Coffey, 2009; Lineback, 2015; O'Connor & Michaels, 1993; Pierson, 2008).

Part of the problem is that there are so many particular features to the episode we could see as important. Another part is that the features interact with each other in reflexive, complex dynamics (Yackel & Cobb, 1996): That Scarlett felt entitled to ask Isaac "how do wheels get tired" reflects her framing of what they were all doing together and her possible role in that, forming and formed by emergent classroom norms. Those norms reflect how Ms. Fargason had been attending and responding to students' ideas; that the teacher noticed and endorsed Scarlett's question reflected and contributed to those norms as well.

In an earlier essay, Hammer and Sikorski (2015) considered the implications of complexity for research on learning progressions, comparing classroom reflexive relationships to non-linear interactions in physical systems to argue for the possibility of chaotic dynamics. Chaos theory is well-developed in many physical systems, rigor borne out by empirical study: When the dynamics are chaotic, tiny differences between systems can amplify to produce radically different outcomes. If classroom learning involves chaotic dynamics, then it requires approaches other than aggregating data across many instances, such as in RCTs, which are designed to average away small variations.

Here, then, is a refined version of our question for this essay: *How can episodes of students' doing science, with all their particularities and complexities, contribute to research on learning, rigorous research with general lasting validity?*

Aims and plans for this essay.

This question arises for us in the context of our ongoing research to study "the dynamics of learners' engagement and persistence in science" (Hammer, Watkins, & Gouvea, 2015-18). We opened this essay with an instance of students' doing science, as a case in point to ground and motivate our question. In the following section, we describe our current project as a case in point of research methodology. We discuss how we have approached this project and what we see as outstanding problems of method and rigor. We then look to research in ecology, which has tackled similar problems of complexity in efforts to understand biological systems. Finally, we reflect back on what our review of research in ecology might suggest for our research on learning.

The Dynamics of Learners' Engagement in Science

A question for research and the challenges of its pursuit.

Our ongoing project is to study the dynamics of learners' engagement and persistence in doing science. In this we hope to help address a long-standing goal that students learn to engage in constructing understanding themselves, rather than learning science as a static body of knowledge. Despite the work of educators and researchers to achieve this goal, it is still relatively rare to see students' doing science in classrooms (Duschl & Osborne, 2002; Lemke, 1990; Thompson et al., 2016). There has been research to understand barriers or challenges to students' scientific engagement (Chinn &

Malhotra, 2002; Kuhn & Pease, 2008; Tang, Coffey, Elby, & Levin, 2010); in this project we sought to examine moments of its happening, to understand how these moments came to be. Our aim is ambitious, to abstract general patterns while acknowledging particularities and complexities of the dynamics in diverse cases.

Here we describe the methodological challenges and our decisions in three phases of our work: Selecting episodes to study; analyzing the dynamics within each; and comparing across the case studies for patterns and themes.

Selecting episodes to study.

The first challenge we face is in identifying instances to study, which, of course, must reflect a sense of what science is. That has defied clear definition, whether by philosophers (Giere, 1988; Kitcher, 1993), educators (Grandy & Duschl, 2007; Windschitl et al., 2008), or scientists (Mayr, 1998; Wilson, 1999). When the question arises among physicists, the typical response is surrender: "Physics is what physicists do," or "physics is the study of physical phenomena." Part of the challenge is that what science is and how it takes place shifts over time: Quantum mechanics and chaos theory, for two examples, have challenged foundational assumptions about how science proceeds and what it can accomplish. For us, the challenge is also that we are looking for *the beginnings of* science in students' thinking.

Rather than specify a set of criteria, we rely on intersubjective agreement. We first look for episodes in which students seem to be engaged in trying to figure something out about natural phenomena, episodes that seem like they might be science. Then we present these as candidates to a panel of collaborating scientists and science education researchers. We select episodes to study for which there is an easy consensus that the data does show students' doing science, recognizing that in this approach it is likely we exclude some data it would be valuable to study.

In our first pass, we found fifteen candidates to present to the panel, drawing from other projects (especially Goldberg et al., 2008-2011), and collecting new data in undergraduate courses at Tufts. All of these fifteen involved video, as the primary form of data, often supplemented with writings and drawings, such as from Isaac's notebook.[2] Eight of the candidates passed the threshold of clear consensus, split between elementary school levels and college, and mostly physical science. (In subsequent work, we have focused more on finding episodes from life sciences and from middle and high school levels.)

The next challenge is setting bounds on the data to analyze. This is a significant decision: If we set the bounds narrowly around a particular event, we exclude as data for analysis what took place in the class in the days and weeks before, which could certainly have had an influence. Of course, our data excludes whatever took place

[2] Video provides rich data, but it is not strictly necessary. The requirement is that the data be rich enough to provide the panel a sense of what the students were doing as science.

outside of class—students' conversations or experimentation. Recognizing that our data cannot be "complete," we make pragmatic choices. In *Isaac's Wheels*, it seemed straightforward: The episode started with Isaac's entry into the discussion and ended with the class moving on to other work. In other cases, it is difficult to decide what to include and what not, especially for cases from activities that span multiple days.

Analyzing the dynamics of an episode

We designed our approach to analysis with care to consider particularities and complexities, to look at each episode with fresh eyes and without predetermined claims. As much as we are able, we first analyze these cases independently of each other, disallowing explicit comparison between cases and trying to focus on what emerges in the particular data. As well, we distribute the lead role of analysis: The initial eight analyses were divided among four members of the project team. The substance of the analysis spans theoretical perspectives and grain size, including individuals' thinking (diSessa, 1993; Tannen, 1993) and social interactions and activities (Jordan & Henderson, 1995; McDermott, Gospodinoff, & Aron, 1978).

The research leader of each analysis first drafts an account of what took place, citing evidence to support claims. For example, in *Isaac's Wheels*, research leader Jennifer Radoff described that Scarlett was holding Isaac accountable to the emergent norm of mechanistic explanations, and that her question prompted Isaac to look for a mechanistic explanation. Presenting the analysis to the group—we call it "workshopping"—gives others the opportunity to challenge these claims, here to ask if there is sufficient evidence Scarlett was seeking mechanism *per se*, when she contested the idea that wheels can be "tired," or to agree the evidence is strong that, in response, Isaac came up with a mechanism for slowing.

The resulting analyses are rich accounts of the episode with many fine-grained claims like these about its dynamics, supported by evidence in the data. They run several thousand words — 7500 in the case of *Isaac's Wheels*. Often they note multiple possible interpretations that the research team cannot settle based on the evidence available.

The final step for each case is a synthesis of "what contributed" to the students' engagement and persistence in science. We read through the case and identify the main features of the dynamics. In *Isaac's Wheels*, they include the teacher's orchestration of the activity and her attention to student ideas and questions, the students' rich conceptual resources for thinking about motion (e.g., they know the car will come to a stop; they know about dragging), their roles as intellectual agents and the kinds of knowledge they consider (e.g., favoring mechanistic explanations, in contrast to their work earlier in the unit).

Comparing across cases

Once we have a sufficient collection of analyses, we shift to looking across them for commonalities, patterns that abstract beyond the particular dynamics of individual cases. We first draw on the summaries of "what contributed" to brainstorm possibilities, and we proceed from there to sort and select the ideas that seem both amenable to study and important as potential contributions to the literature. The work then becomes to define those ideas more precisely, to identify relevant literature, and to give examples and descriptions of what counts as evidence within the data.

To this point, we have gone through this part of the process only once, with our initial set of eight analyses. We had a variety of early ideas about patterns. Almost all cases involved affective displays, e.g., of excitement or frustration; in several cases it was not clear what question the students were trying to answer; in several that seemed to be a principal target for them, to articulate a question; some cases involved students clearly identifying models (such as Isaac's sketches of gears). All cases involved some participant expressing confusion or uncertainty, like Ms. Fargason and several students with respect to Isaac's idea.

Some of these patterns clearly supported ideas already in the literature. Almost all of the cases, for example, involved the teacher's recognizing and engaging with a student's idea or question they had not anticipated (Robertson, Scherr, & Hammer, 2016). Ms. Fargason did that throughout the episode, attending to Isaac's idea and then later to Jourdan's and Scarlett's questions.

Other patterns seemed like they might lead in new directions, and we worked to develop them into more precise constructs. The ambiguity around students' questions and question-seeking connected with two ideas in the literature. One is *problematizing* (Engle et al., 2012; Engle & Conant, 2002), as an aspect of disciplinary engagement in science. The other is a more general notion of *positioning* (Harré & van Langenhove, 1999), "a dynamic alternative to the more static concept of role" (Harré & van Langenhove, 1999, p. 215). Someone may be positioned as "powerful or powerless, confident or apologetic, dominant or submissive, definitive or tentative, authorized or unauthorized, and so on" (p. 217).

We defined *problematizing* as the pursuit of a question, more precisely as the activity of "noticing a gap of understanding, identifying and articulating its precise nature, and motivating a community of its existence and significance" (Phillips et al., 2017). Evidence of problematizing included students' expressing unease that something is missing or amiss, articulating and re-articulating different versions of a question, and arguing that there is a gap or inconsistency they do not understand. We were surprised at how often it was a prominent feature of their inquiries, including six of the initial eight cases.[3]

And we identified *positioning as not-understanding* as social positions in participants' discourse, how they "reveal their uncertainty or confusion about a phenomenon or idea" (Watkins et al., 2018, p. 576). This too afforded new, close analyses of the data, specifically for evidence of students', for example, "displaying puzzlement, asking questions, or pointing out their lack of knowledge"(Watkins et al., 2018, p. 577). Every case in our set included evidence of at least some participant's positioning themselves

[3] *Isaac's Wheels* was one of the two exceptions; we did not see problematizing as a prominent feature. Still, it was evident, most directly in the work by Jourdan, Scarlett and Jamir, supported by Ms. Fargason, of identifying and working to convince Isaac that there was a gap in his reasoning for why the car slows down. The episode also suggests Isaac had done some problematizing for himself earlier, seeing the need to explain how wheels let the car move without dragging.

as not-understanding, in contrast, perhaps, to cases of argumentation among students with conflicting views.

Nagging doubts.

That is how we are working, our current answer to the question, *How can episodes of students' doing science, with all their particularities and complexities, contribute to research on learning, rigorous research with general lasting validity?* It is an answer we can manage at the moment, in order to get on with doing the work, but it is certainly not settled.

In this section we reflect on three interwoven aspects of the project that we continue to see as difficult, in how we (1) delimit the data to study, (2) analyze complex dynamics, and (3) assess the more general validity of our results.

Delimiting the data

As we noted, it seemed straightforward to begin our analysis of *Isaac's Wheels* when the teacher first calls on Isaac, and to end it when the class moved on to another topic. But there were still trade-offs to this bounding that impacted our findings. We chose only to examine the data within those 20 minutes, but those 20 minutes were nested within the day, and within the larger unit on motion that had begun several weeks earlier. It was nested in other ways as well, as a class in an elementary school in San Diego, in a particular community; it was also a class that was part of a research project (Goldberg et al., 2008-2011) on "responsive teaching" (Robertson, Scherr, & Hammer, 2016). No doubt including more data could have given us different insight or perspectives into the dynamics of those 20 minutes.

We made choices as well regarding which details of the data to consider. For a simple example, the video would have allowed us to examine the children's physical placement in the room, who was sitting next to whom, their postures, and so on. Every teacher knows the seating arrangements can matter for classroom dynamics, but they were not part of our analysis. Video is exceedingly rich, and it would not be possible to analyze every aspect of the data, even bounding the episode to 20 minutes.

In both of these ways, looking outward from the episode and inward, it is both consequential and unclear how to delimit the data we study. At the moment, we are doing the best we can, case-by-case, and for each the question comes to whether the data we select supports productive analysis for our research.

Identifying phenomena and entities

However we delimit the data in our episodes, they all have myriad aspects that interact in complex dynamics. This makes for closely related challenges of drawing inferences about the case under study as well as of making conjectures that might be useful elsewhere.

Workshopping the analysis of *Isaac's Wheels,* for example, we considered the inference that Scarlett's asking "how could the wheels get tired" reflected her sense of an emergent epistemological norm, that in this activity explanations should involve tangible mechanisms. Perhaps, though, she was objecting to the idea that a non-living object can be "tired." Although it was not within the bounds of the data we selected for this case, we were aware the class had recently debated whether the toy car had "free will." That discussion involved attention to the difference between living and non-living things; cars do not have "will." Thus we could analyze Scarlett's question focusing on her in that moment, "Scarlett's question," or we could analyze it as part of dynamics that started before this episode and involved the class as a whole.

Here as across the cases, we encounter issues of appropriate scale, of time and of relevant system, how to parse and circumscribe phenomena and units of analysis. That is, we need to choose the appropriate granularity of time and of the entity or system involved. We could focus on Scarlett as an individual, and the particular moment of her posing her question; we could focus on the system of Jourdan, Ms. Fargason, Isaac, and Scarlett over the few minutes they consider the question what causes the car to slow; we could focus on the class as a whole, over patterns of talk and attention over multiple days (if we set broader bounds on the data to analyze).

Assessing the validity of findings

As we described, we worked at first to keep the cases separate, trying to generate and support ideas about what took place based on the data for that episode. The validity of a claim about the dynamics of a particular case, however, does not necessarily recommend it as a valuable finding; perhaps it is idiosyncratic to the moment. It was evidently consequential in this instance that Isaac thought about gears in trying to make sense of rolling; it is not clear from this one case that would be something to declare of generalizable value for third-grade students.

Our look across cases for patterns involves considering a level of abstraction above the particularities. To pursue those ideas further involves developing and refining the ideas, identifying of evidence, iterating analyses and framework development as discussed in accounts of qualitative research (Chi, 1997; Glaser & Strauss, 1967). We also assess the ideas through readings and connections in the literature, considering how they speak to what is known or believed more broadly. Thus we could support our claims of the patterns across cases, of problematizing and of positioning as not-understanding. Each, we could argue, contributes to the current literature.

At the same time, we would not propose these ideas as having general validity, as "laws of scientific thinking." To be sure, it is easy to think of instances of students' doing science that do not show either *problematizing* or *positioning as not-understanding,* such as in episodes of argumentation between students who are confident in their respective positions (e.g., Engle & Conant, 2002; Hammer, 1996).

Comparing to natural science

Writing for the *American Journal of Physics* in 1996, David took up the question of how ideas in education research might compare to ideas in physics:

We tend to assume that, to be useful, research should provide unambiguous, demonstrably valid results, reliable principles and methods, and in terms sufficiently precise so as not to be sensitive to interpretation and judgment. These assumptions might be appropriate if education were already an applied science—we expect physicists to provide engineers with precise, reliable results on which to base their designs—but it is not. We describe students as having *"misconceptions," but we cannot present that account with anything like the precision and confidence with which, for example, we can describe the properties of hydrogen.*

It is possible that these assumptions will never be appropriate for education, that fundamental, epistemological differences between the study of human cognition and the study of the physical world will preclude a physics-like formalism in education. It is also possible that education, like physics, will eventually achieve a productive formalism. (Hammer, 1996, p. 1316)

That was comparing education research to physics; we draw on what we know. Reflecting on the work and challenges of research in our current project, Julia saw more connections to biology. We suspect that may be the more helpful comparison.

Of course, there is an extensive history of connection to biology in cognitive psychology. Piaget recognized intelligence as a biological phenomenon (Piaget, 1950/2005). His model of development, including its concepts of assimilation and accommodation, were based on his own prior study of adaptation and self-regulation. Here we pursue a different connection, in particular to ecology. Ultimately, we expect these considerations will come together with Piaget's, especially in light of more recent progress in dynamic systems models of development (Thelen & Smith, 1994), building on earlier ideas of self-regulation. For now, we focus on ecology.

There is also a history of reference to ecology, in education research. A number of authors have argued for ecological models of conceptual understanding (Strike & Posner, 1992; Southerland, Johnston, & Sowell, 2006) and of human development (Bronfenbrenner, 1977; O'Connor & McCartney, 2007). We have been more directly influenced by Thelen's and Smith's (1994) "dynamics systems approach," which they describe through discussions of both physical and biological systems. The latter may be more apt, as we discuss below, but regardless there is clear overlap with prior characterizations of ecological models. Indeed, the epistemic and methodological challenges we describe above reflect our ecological theoretical perspectives on the dynamics of classroom activity and student learning.

What we have not seen in prior calls for ecological models in education research is direct consideration of research in ecology. How do ecologists manage particularities and complexities in their work; how do they conceptualize and produce rigorous research with lasting validity? We turn to that now, and we will then return to reflect on how ecology research can inform our work.

Research in Ecology

Ecologists study complex systems comprised of populations of evolving species interacting in changing environments. Ecology is a relatively young field, compared to physics, certainly, but also to other sub-disciplines of biology such as genetics and evolution. As the field has grown there has been formal discussion in the literature aimed at understanding and developing a productive epistemology of ecology. The challenges of its study are in many ways similar to the challenges we experience in education, including in particular of complexity across multiple scales of phenomena.

Within ecology it is important to distinguish notions of *complicated* and *intricate*, both of which apply to biological systems, from *complexity.* One could liken an ecosystem to a *complicated*, *intricate* machine, like a clock, which could be studied by examining the structure and properties of its parts and how the parts can be combined. Indeed, in biology textbooks, many biological systems are described in terms of parts and steps: mitosis, food webs, or the Krebs cycle.

But the analogy to a clock obscures key aspects of ecosystem *complexity*. As ecologist Charles Elton famously described, in order for a clock to resemble an ecosystem, the gears would have to be allowed to develop and evolve over time. The components would not just be linked; the linkages, through processes analogous with predation and competition, would change the components. And finally, a part of the system "retains the right to arise and migrate and settle down in another clock" (Elton, quoted in Hardy, 1968, p. 5). The systems that ecologists study are much more complex than clocks because the components and interactions among them shift and evolve over time and move across heterogeneous landscapes. As a result, ecosystems are highly variable across time and space. Any specific ecosystem is a product of local geography and climate as well as the particular histories of the organisms that happened to arrive and evolve in that place.

With all this complexity and variability, ecologists have struggled to generate lasting, rigorous knowledge. Looking to physics, some scholars have sought value in enduring principles that might qualify as "ecological laws" (Lange, 2005; Lawton, 1999; Murray, 1992) or simple, elegant mathematical models (May, 1974, 1976). Others have argued that general ecological theory has no practical utility, and that the field should embrace its status as "a science of case studies" (Hansson, 2003; Schrader-Freschette & McCoy 1993). Thus ecologists have been grappling with similar questions to ours, and we may learn from what they have done to conceptualize and re-conceptualize their discipline.

For the remainder of this section, we digress from our own research to review several interrelated themes of progress in ecology: The aims are becoming less about general laws and more about mid-level constructs of possible mechanisms; the methods are attending more to dynamics at multiple scales and shifting to embrace variation and complexity. We then return in the next section to consider how these trends in ecology might help us with our research in education.

Letting go of laws.

The pursuit of "laws of ecology," was motivated by comparisons with physical laws: general, abstract, principles applicable across systems. In this pursuit, ecologists were advised to not be "distracted by the biological equivalents of friction" (Murray, 1992, p. 596). In physics, the move to ignore friction made it possible to formulate Newton's Laws, and part of their success was their ability later to account for friction as a kind of force. To bring rigor to ecology, some argued the need for similar moves to ignore variability and inconsequential fluctuation, hopefully making it possible to formulate basic laws.

Some ecologists have worked in this way, and they have considered candidate laws. One example is the law of exponential growth: "A population will grow (or decline) exponentially as long as the environment experienced by all individuals in the population remains constant" (Turchin, 2001, p. 18). Like Newton's Laws, this law can be written out in an equation; it emerges from the simple physical rules of organismal interactions, simple rules of birth and death, and it is an idealization—it ignores aspects of real systems such as resource and space-limitations. As with Newton's Laws, real systems deviate from the idealization; in real populations, resources and space run out quickly, so no natural population grows exponentially for very long. Unlike Newton's Laws, however, the modifications needed to make the basic equation fit any actual population outside of the laboratory make it practically unusable. The simple law, while general, has limited explanatory or predictive power for real systems.

The fate has been similar for other possible laws of ecology, and over time the hunt for them has waned. Many ecologists have acknowledged that even if laws of ecology could be articulated, they were unlikely to be useful (e.g., Hansson, 2003; Lange, 2005; O'Hara, 2005; Simberloff, 2004).

Others have been reluctant to let go of laws. Lawton (1999) famously called research in community ecology "a mess" because of the number of interacting factors that impact the dynamics. "The contingency becomes overwhelmingly complicated at intermediate scales, characteristic of community ecology, where there are a large number of case histories, and very little other than weak, fuzzy generalization" (Lawton, 1999, p. 177). For Lawton (1999), the implication was that ecology should move on to research at levels that would allow the discovery of "general patterns, laws, and rules" (p. 177).

Balancing global and local knowledge: Mid-level theory.

Simberloff (2004) disagreed with Lawton in two ways. First, he maintained that understanding ecology at the scale of communities has great value for society, such as for the management of fisheries or forests. He agreed that identifying the factors influencing the dynamics in a single community would not allow one to predict which of those factors would be important in a new community. However, Simberloff's second disagreement was that it is a mistake to presume that general laws are the only form of useful knowledge. Rather, ecology has been advancing significantly through the development of local causal mechanisms and approaches to testing for their occurrence in systems.

Knowledge of possible mechanisms (as opposed to necessary ones) represents a mid-level of theory building—not universal law, but still of general value. A classic example is Dayton's (1984) review of research on species diversity in marine ecosystems. It showed how and when possible mechanisms can apply, mechanisms such as competition, predation, environmental disturbance, and niche creation.

For example, in the "rocky intertidal" – communities of species that live in the rocks and pools in the tidal zone – competition for space dominates the dynamics of species diversity. Many organisms such as barnacles and mussels encrust the rock spaces, excluding one another. Competition among these multiple species prevents a single species from dominating. Working in concert with competition is disturbance (from waves or predators) that exposes new clean rock space, allowing for species replacement that keeps the community perpetually diverse.

In the arctic sea, where all species are relatively rare, the overall rate of interaction among species is also rare, and therefore neither competitive exclusion nor frequent disturbance is essential to maintaining diversity. Instead, in these communities, dominance by a single species, like sea urchins, may be kept under control by a very rare, but very efficient specialist predator, such as a starfish. The absence of this single predator would drastically reduce the overall community diversity. In other communities, for example the sea floor, positive feedback processes can increase organismal diversity: Worms drill into coral rock, for example, creating new habitats that other organisms can occupy.

Dayton and other ecologists built this body of knowledge through intensive study of specific systems: sponges in the McMurdo Sound, Antarctica (Dayton, 1972), coral on the Jasper Sea Mount in the Pacific (Genin et al., 1986), and Kelp Forests off the coast of central California (Dayton et al., 1984). The results are rich descriptions of possible mechanisms and when and why they might apply, and, as Simberloff (2004), they have proven useful for ecologists' understanding local ecosystems.

Attending to dynamics across scales.

Shifting the focus to possible mechanisms and when they might apply motivates attention to the scope and scale of effects.

For instance, a major contribution in ecology has been specifying the scale-dependence in the relationship between species diversity and ecosystem productivity—the amount of biomass produced per area (Purvis & Hector, 2000). At a global scale, productivity is positively related to diversity: A densely packed rainforest produces much more biomass per area than a temperate forest. Yet, at the scale of a single community, diversity can decrease productivity (Fraser et al., 2015; Gaston, 2000). A larger number of species creates more competition, both for space and through the production of inhibitory chemicals and other defenses. Plants must divert some of their resources to this competition rather than investing it in growth, leading to a decrease in productivity overall. The pattern reverses once again at the level of a single species. If individuals are too similar yields can decrease when the population is attacked by herbivores or disease. At this level, genetic diversity will once again tend to increase productivity overall (Hughes et al., 2008).

Scale has also been critical to understanding species' persistence and extinction. A population viewed on a small scale of time and space may go extinct, raising alarm for conservationists. However a single population may be understood as part of a larger network of populations linked by migration—a "metapopulation" (Hanski, 1998). With this view, persistence of the species may still be possible as long as the extinct population is replaced through the birth or growth of a population somewhere else in space. If conservation efforts focus on saving individual populations rather than managing larger metapopulations, resources may be allocated at the wrong scale to little effect.

Embracing the study of variation and complexity.

Because ecological systems are so complex, ecologists have had to make simplifying assumptions in order to study them. These assumptions seemed necessary for generating general knowledge that could rise above the inherent variation. However, as the field has matured, researchers have begun to re-examine the validity of many of these assumptions and to embrace the study of variation and complexity.

One recent shift concerns how to deal with within-species variation. Ecologists know that individual variation is pervasive, and of course that variation is of central concern with respect to natural selection and evolution. Still, with respect to questions of population and community dynamics, for decades it has been acceptable to treat individuals within a population as roughly interchangeable. Recent research has called that assumption into question (Bolnick et al., 2003; Bolnick et al., 2011, Hughes et al., 2008; Sih et al., 2004, Violle et al., 2012).

In the study of animal behavior, for instance, classical studies reported on the average behavior of a species, ignoring any differences at the individual-level. New research that has focused on measuring and studying the impacts of individual variation has found that this variation can matter for populations and communities.

For example, not all individual predators within a single species behave in exactly the same way when confronted with prey. Some are consistently more "bold," attacking without hesitation, and others are more "shy," attacking less frequently and less vigorously (Sih et al., 2004). The specific composition of individual types, not just the number of predators, can impact the matter and energy dynamics of the larger system. Researchers documented this effect experimentally in spider communities (Pruitt et al., 2016). They first measured the behaviors of individual spiders of various species and then combined different individuals to create communities with the same species composition but more or less variation in behavior from individuals. They found that communities in which individual variation was higher gained more mass and were less likely to disband over time. Treating all individuals as having the same average behavior would have prevented ecologists from understanding why some communities persist and others collapse.

A second example concerns mathematical approaches to modeling population sizes, which in some cases can fluctuate dramatically from one year to the next. Theoretical ecologists wondered if these fluctuations could be modeled using chaos theory (Hastings et al., 1993; May, 1974, 1976), which, as we discussed above, can explain with mathematical rigor how tiny differences can amplify to produce radically different outcomes. Ecologists hoped that if they were able to find the signatures of chaos in natural populations that they could use the mathematical tools of chaos theory to understand system dynamics (Hastings et al., 1993; May, 1974, 1976). But those signatures have been difficult to find except in carefully controlled laboratory experiments or very simple natural systems (Beninca et al., 2015; Bjornstad, 2015; Constantino et al., 1995). Even some of the most optimistic ecologists have since expressed that predictable dynamics, even if chaotic, are likely to be rare in most ecological systems (Hastings 2001).

And so the problems of how to address the variability and complexity are ongoing areas of research: it is part of ecological inquiry to be interested in studying these dynamics – it is no longer acceptable to simply average them away.

Back to Education Research

We opened this essay with *Isaac's Wheels* and used it to motivate the question: *How can episodes of students' doing science, with all their particularities and complexities, contribute to research on learning, rigorous research with general lasting validity?* We then turned to the methodology of our current project, *The dynamics of learners' engagement and persistence in science*, which we used to instantiate choices and challenges in education research.

From our brief, virtual visit with ecologists, we can see they have been grappling with and debating about how to handle the idiosyncrasies of phenomena—including whether they should "continue to devote so much time and effort to traditional studies in community ecology" (Lawton, 1999, quoted in Simberloff, 2004, p. 787). The similarities between their challenges and ours are especially striking given the relative advantages of research in ecology compared with education research: Many of the outcomes ecologists are interested to understand—populations, biomass productivity, biodiversity, etc— are much easier to recognize and quantify than, say, conceptual understanding, motivation, or engagement in disciplinary practices.

But these differences in the fields are much of what makes the comparison so useful. For one, given the relative ease of quantification in ecology, and the limited value of chaos theory to their work, we are less inclined to pursue it in ours. More generally, the tangibility of the targets of study in ecology makes it especially helpful as an analogical base for thinking about education. The themes we have reviewed of epistemological and methodological progress in ecology—letting go of universal laws in favor of mid-level possible mechanisms, looking across scales, and embracing variation—all seem relevant to our questions.

In this final section of the essay, we reflect back on our project as a case-in-point of education research, returning to the areas of nagging doubt we discussed above: How do we delimit the data to study? How do we identify phenomena and entities, in analyzing that data? And what should we hope our findings to accomplish, beyond particular episodes?

Delimiting the data.

We described our approaches and difficulties in bounding the data to analyze as well as, within those bounds, choosing what details to consider. There are corresponding issues within ecology, which are related to the scales of phenomena and system.

Ecologists' setting out to study the stability of an ecological community, for example, seems broadly analogous to our working to study the dynamics of students' engagement during a class. These choices of the scale of the system, its elements, and the interval of time have implications for identifying phenomena and entities. The ecologists need to decide over what time scale measurements of stability makes sense; they need to decide whether to include interactions with nearby communities. We may find it helpful to look beyond the 20 minutes we selected for study, to see the role of an insight from a previous discussion or to trace the origin and flow of ideas from other contexts.

Ecologists' choosing what details to consider, such as whether to keep track of individual spiders over time, seems broadly analogous to our choices of what needs attention within an episode. There too, the initial choices have implications: Individual spiders' 'personalities' could only become apparent in studies attending to those details.

As we discussed, we began analyses of case studies with decisions about selecting the data to study, but research in ecology motivates us to reconsider that approach. Perhaps we should plan for iteration in how we delimit the data, in ways analogous to iteration in the refinement and application of coding schemes (e.g., Chi, 1997; Glaser & Strauss, 1967). To some extent, we have been examining and re-examining data within the bounds, as group members notice and argue the relevance of details the lead analyst had not considered in their initial pass.

We have not, however, been adjusting the boundaries around episodes, in iteration with analyses. We suspected, for example, that the earlier discussion of "free will" might have influenced Scarlett's challenging Isaac, "how can wheels get tired;" perhaps it would make sense to expand the boundaries of the data to include that discussion. This, of course, would add significantly to the work of analyzing the case, which at first look seemed simple to contain, but the added effort may lead to more robust findings. In other episodes, there is no immediately clear choice, and it has been difficult to set boundaries; iteration may limit the repercussions of the initial decisions.

Identifying phenomena and entities.

There are two related challenges we face in identifying phenomena and entities, one of parsing complex scaling dynamics when analyzing a particular episode and the other of determining which aspects of the episode may be of general relevance.

The first concerns the scales, in time and system, of what we understand as the phenomena in the episode, and it is tightly entangled with the question of how to delimit the data. In one example above, considering a wider scope of data supported ecologists' developing the construct of a metapopulation (Hanski, 1998). Similarly, our expanding the scope of data could support our considering the possibility that Scarlett's question was part of a larger phenomenon of her resisting anthropomorphic reasoning, continuing from the earlier debate over "free will." And we could consider whether to attribute the concern to her as an individual or to the class, as the students formed a shared sense and norm of mechanistic explanations. Perhaps both scales are relevant, analogous to ecologists' identifying the reflexive relationship between individual spiders' personalities and the resilience of spider communities (Pruitt et al., 2016).

The comparison of our work to ecologists' has made these matters of modeling more salient. It has renewed our interest in data-driven interpretation of the relevant scales of dynamics (Conlin, Gupta, & Hammer, 2010; Conlin & Hammer, 2015). More broadly, it supports calls in the learning sciences for research that looks across multiple scales of time and system (e.g., Jacobson, Kapur, & Reimann, 2016). Cobb and colleagues have developed methodology for analyzing reflexive relationships between individual and social levels in mathematical activity and practices (Cobb, Stephan, McClain, & Gravemeijer, 2001; McClain & Cobb, 2001; Yackel & Cobb, 1996). Saxe and colleagues have proposed an approach to studying the development of mathematical ideas that coordinates moment-to-moment interactions with patterns of change over time for individuals and communities (Saxe et al., 2009; Saxe & Esmonde, 2005).

Philip and colleagues have studied intersections between individual and moment-to-moment dynamics of learning with systemic structures of power and oppression (Philip, 2011; Philip, Gupta, Elby, & Turpen, 2018). That work inspires us to study how interactions between broader systemic dynamics of power around race, class, and gender might affect local dynamics of students' *problematizing* and *positioning*.

The second facet to the difficulty of identifying phenomena and entities concerns relevance beyond particular episodes, closely entangled with the question of how to assess the value and validity of findings. The epistemic aim of general validity could inform how we choose to focus our analyses, to consider whether it is plausible a phenomenon we see in one case might appear in others. If we expect that deciding validity will ultimately involve a gold-standard test, such as a RCT, we should pick targets of study that will afford such testing.

We designed our approach, by contrast, specifically to keep us focused on the data at hand, case-by-case, to hold off comparing across cases until we had a collection of them completed. It is an attempt at qualitative rigor, to resist seeing the data in ways we already anticipate. Comparing our work with ecologists' encourages us to continue in this way, to study particular episodes for what they show, deferring concern for generalizability.

Ecologists have recognized the importance of idiosyncrasies, such as the arrival of a single rare predator, a chance distance migration, or an aberrantly aggressive individual in a community. Part of their scholarly productivity has involved abstracting from the particularities: the importance of variation among individuals in one community inspires considering the possibilities elsewhere. In our project, the phenomena of *problematizing* and *positioning as not-understanding* appeared most clearly in individual cases, then emerged as patterns to study when we compared across cases.

Assessing the validity of findings.

Ecologists have come to a general consensus that there are not general laws, specifically in the "intermediate scales" of community ecology, where the "contingency becomes overwhelmingly complicated" (Lawton, 1999, p. 177). Rather than abandon research at these scales, ecologists have mostly chosen with Simberloff (2004) to let go of laws as the epistemic aims of research. They are, rather, seeking and producing seminal findings in the form of mid-level constructs of *possible mechanisms* that occur in some systems. These are not findings of reliable laws, but they may have lasting validity as contingent truths that can support understanding particular systems. Knowledge of possible dynamics supports researchers or managers (such as of forests or fisheries) in constructing local, particular models.

That is how we understand the possibly general validity of our findings, as possible aspects of students' doing science. We do not expect *problematizing* or *positioning as not-understanding* to be evident in all episodes, but we are confident they occur in some. We hope that, as findings, they support researchers or instructors in making sense of and supporting what they see in their students.

There is room in this account for forms of predictive power. In the traditional sense, researchers might design studies explicitly to support problematizing and positioning as not-understanding, and test the predictions that these conditions lead to more frequent moments of students' doing science.

But reflecting on the shift in ecology to finding possible mechanisms and embracing variation and complexity has us thinking of predictive power on a local scale: Might we evaluate findings as useful if they help researchers or educators better understand their students, with the particular features in their situations? Might we ask how findings are useful in constructing local, particular models that generate predictive ideas for how to support particular students' learning?

At the same time, visiting ecology has raised both our skepticism over fixed criteria for assessing validity and our patience for assessing general applicability. There are claims we can make with great confidence about single instances, and of course we should consider the possible value of those claims for the literature as we consider dissemination, but we may not know until much later whether they will have that value. How we assess validity needs to vary with the nature of the claims. As Cartwright (2007) argued in challenging the status of RCTs, "there is no gold-standard" means of assessing validity.

Still Seeking

Of course we have not arrived at a clear, final answer to our question. Visiting ecology has shown us again how science itself evolves and adapts, in its aims, values, virtues, and processes (Chinn, Buckland, & Samarapungavan, 2011). We noted that as part of the challenge in our project, of having a clear definition of science to guide our identifying its happening in students. The same must be true of research on learning.

Our tour of ecology has us further support the value of case studies of learning, including to look for patterns but even to recognize the possible use of a single instance. We are experimenting with iteration as an approach to delimiting data and identifying phenomena and entities. It has helped us recognize other ways findings can have lasting validity. For instance, just as there is value for conservationists in Dayton's work to identify multiple possible mechanisms for maintaining species diversity in marine eco-systems, identifying multiple possible classroom dynamics that promote or inhibit students' doing science could support educators' pedagogical sense-making and decisions. Our review of ecology has also encouraged us to design follow-up research to examine when and under what conditions these possible dynamics may occur.

Mostly, though, it encourages us to build further on this conversation across disciplines. We plan to recruit ecologists to join us in a cross-disciplinary study group, to read and talk about methodology. That will include our reexamining accounts in education, such as by Cobb and Saxe and their colleagues, cited above, design-based research (Barab & Squire, 2004; Brown, 1992), and a range of work by scholars grappling with similar matters of scale and situativity (Bang, 2015; Maxwell, 2004; Sandoval, 2013; Tabak, 2004). Their work, we know, already offers much of what we found visiting ecology; we look forward to reading it with new appreciation.

Acknowledgements

We are grateful to *Infancia y Aprendizaje* and Bárbara Brizuela for inviting this piece and for helpful feedback. And we are extremely grateful to the Gordon and Betty Moore Foundation and Janet Coffey for funding the project (under grant numbers GBMF3475 and 3475.01). Thanks also to Dan Levin for comments on a draft, and, absolutely, to all of the students and colleagues who have worked with us in the project.

References

Bang, M. (2015). Culture, Learning, and Development and the Natural World: The Influences of Situative Perspectives. *Educational Psychologist, 50*(3), 220-233. doi:10.1080/00461520.2015.1075402

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1-14.

Bauer, N. (2015). *How to do things with pornography*. Cambridge, Massachusetts: Harvard University Press.

Benincà, E., Ballantine B., Ellner, S. P., Huisman, J. (2015). Species fluctuations sustained by a cyclic succession at the edge of chaos. *Proceedings of the National Academy of Sciences USA 112*:6389–6394.

Bjørnstad, O. N. (2015). Nonlinearity and chaos in ecological dynamics revisited. *Proceedings of the National Academy of Sciences*, *112*(20), 6252–6253. doi:10.1073/pnas.1507708112

Blanchette, I., & Dunbar, K. (2000). How analogies are generated : The roles of structural and supeIficial similarity. *Memory & Cognition*, *28*(I), 108–124.

Bolnick, D. I., Svanbäck, R., Fordyce, J. A., Yang, L. H., Davis, J. M., Hulsey, C. D., & Forister, M. L. (2003). The Ecology of Individuals: Incidence and Implications of Individual Specialization. *The American Naturalist*, *161*(1), 1–28. doi:10.1086/343878

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, Volker H.W., Schreiber, Sebastian J., Urban, Mark C., & Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology and Evolution, 26*(4), 183–192.

Brizuela, B., & Schliemann, A. (2004). Ten-Year-Old Students Solving Linear Equations. *For the Learning of Mathematics*, *24*(2), 33–40.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*(7), 513-531.

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*(2), 141-178.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, *26*(April 1988), 499–531. doi:10.3102/00028312026004499

Cartwright, N. (2007). Are RCTs the gold standard? *BioSocieties, 2*(1), 11-20.

Chi, M. T. H. (1997). Quantifying Qualitative Analyses of Verbal Data: A Practical Guide. *The Journal of the Learning Sciences, 6*(3), 271-315.

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the Dimensions of Epistemic Cognition: Arguments From Philosophy and Psychology. *Educational Psychologist*, *46*(3), 141–167. doi:10.1080/00461520.2011.587722

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, *86*(2), 175–218. doi:10.1002/sce.10001

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in Classroom Mathematical Practices. *Journal of the Learning Sciences*, *10*(1/2), 113–163.

Conlin, L. D., Gupta, A., & Hammer, D. (2010b). Where to find the mind: Identifying the scale of cognitive dynamics. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the disciplines: Proceedings of the 9th International Conference of 13 the Learning Sciences* (Vol. 1, pp. 277-284). Chicago, IL: International Society of the Learning Sciences

Conlin, L. D., & Hammer, D. (2016). From the individual to the ensemble and back again. In A. A. DiSessa, M. Levin, & N. J. S. Brown (Eds.), *Knowledge and interaction : a synthetic agenda for the learning sciences* (pp. 311-325). New York ; London: Routledge.

Constantino, R. F., Cushing, J. M., Dennis, B., & Desharnais, R. A. (1995). Experimentally induced transitions in the dynamic behaviour of insect populations. *Nature*, *375*, 227–230.

Dayton, P. K. (1972). Toward an understanding of community resilience and the potential effects of enrichments to the benthos at McMurdo Sound, Antarctica. In *Proceedings of the colloquium on conservation problems in Antarctica* (pp. 81-96). Lawrence, KS: Allen Press.

Dayton, P. (1984). Processes Structuring Some Marine Communities: Are They General? In D. R. Strong, D. S. Simberloff, L. G. Abele, & A. B. Thistle (Eds.), *Ecological*

Communities: Conceptual Issues and the Evidence (pp. 1–613). Princeton, NJ: Princeton University Press.

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

diSessa, A., Hammer, D., & Sherin, B. (1991). Inventing Graphing: Meta-Representational Expertise in Children. *Journal of Mathematical Behavior*, *10*, 117–160.

Duschl, R. A., & Osborne, J. (2002). Supporting and Promoting Argumentation Discourse in Science Education. *Studies in Science Education*, *38*(1), 39–72. doi:10.1080/03057260208560187

Engle, R. A., & Conant, F. R. (2002). Guiding Principles for Fostering Productive Disciplinary Engagement: Explaining an Emergent Argument in a Community of Learners Classroom. *Cognition and Instruction*, *20*(4), 399–483.

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How Does Expansive Framing Promote Transfer? Several Proposed Explanations and a Research Agenda for Investigating Them. *Educational Psychologist, 47*(3), 215-231. doi:10.1080/00461520.2012.695678

Fraser, L. H., Pither, J., Jentsch, A., Sternberg, M., Zobel, M., Askarizadeh, D., … Bennett, J. A. (2015). Worldwide evidence of a unimodal relationship between productivity and plant species richness. *Science*, *349*, 302–306.

Gaston, K. J. (2000). Global patterns in biodiversity. *Nature*, *405*, 220–227.

Genin, A., Dayton, P. K., Lonsdale, P. F., & Spiess, F. N. (1986). Corals on seamount peaks provide evidence of current acceleration over deep-sea topography. *Nature, 322*, 59. doi:10.1038/322059a0

Giere, R. N. (1988). *Explaining Science: A Cognitive Approach*. University of Chicago Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory; strategies for qualitative research*. Chicago, IL: Aldine Pub. Co.

Goldberg, F. M., Hammer, D. M., Bendall, S., & Coffey, J. (2008-2011). *Learning Progression for Scientific Inquiry: A Model Implementation in the Context of Energy*. United States National Science Foundation Grant #DRL 0732233.

Grandy, R. E., & Duschl, R. A. (2007). Reconsidering the Character and Role of Inquiry in School Science: Analysis of a Conference. *Science & Education*, *16*(2), 141–166.

Grimm, V. (1999). Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling*, *115*, 129–148.

Grimm, V., Ayllón, D., & Railsback, S. F. (2017). Next-Generation Individual-Based Models Integrate Biodiversity and Ecosystems: Yes We Can, and Yes We Must. *Ecosystems*, *20*(2), 229–236. doi:10.1007/s10021-016-0071-2

Hammer, D. (1996). More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research. *American Journal of Physics*, *64*(10), 1316-1325.

Hammer, D., Goldberg, F., & Fargason, S. (2012). Responsive teaching and the beginnings of energy in a third grade classroom. *Review of Science, Mathematics and ICT Education*, *6*(1), 51–72.

Hammer, D. and Radoff, J. (2014) *Children doing science: Essential idiosyncrasy and the challenges of assessment.* Commissioned by the Committee on Successful Out-of-School STEM Learning. National Research Council, Board on Science Education. http://sites.nationalacademies.org/DBASSE/BOSE/CurrentProjects/DBASSE_086842

Hammer, D., & Sikorski, T.-R. (2015). Learning Progressions Implications of Complexity for Research on Learning Progressions. *Science Education*, *99*(3), 424–431. doi:10.1002/sce.21165

Hammer, D., Watkins, J. & Gouvea, J. S. (2015-2018) *Investigating factors that contribute to engagement and persistence in science learning*. Grant #GBMF 3475.01 by the Gordon and Betty Moore Foundation.

Hanski, I. (1998). Metapopulation dynamics. *Nature*, *396*(6706), 41–49. doi:10.1038/23876

Hansson, L. (2003). Why ecology fails at application: Should we consider variability more than regularity? *Oikos*, *100*(3), 624–627. doi:10.1034/j.1600-0706.2003.12479.x

Hardy, A. (1968). Foreword: Charles Elton's Influence in Ecology. *Journal of Animal Ecology*, *37*(1), 3-8.

Harré, R., & van Langenhove, L. (1999). *Positioning theory : moral contexts of intentional action*. Oxford; Malden, Mass: Blackwell.

Hastings, A., Hom, C. L., Ellner, S., Turchin, P., & Godfray, H. C. J. (1993). Chaos in Ecology: Is Mother Nature a Strange Attractor? *Annual Review of Ecology and Systematics*, *24*(1993), 1–33. doi:10.1017/CBO9781107415324.004

Hastings A (2001) Transient dynamics and persistence of ecological systems. *Ecology Letters* *4*(3):215–220.

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters, 11*(6), 609–623. doi:10.1111/j.1461-0248.2008.01179.

Jacobson, M. J., Kapur, M., & Reimann, P. (2016). Conceptualizing Debates in Learning and Educational Research: Toward a Complex Systems Conceptual Framework of

Learning. *Educational Psychologist, 51*(2), 210-218. doi:10.1080/00461520.2016.1166963

Jaber, L. Z., Radoff, J., & Hammer, D. (under review). “It’s scary but it’s also exciting”: Evidence of meta-affective learning in science.

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and Practice. *Journal of the Learning Sciences*, *4*(1), 39–103.

Kitcher, P. (1993). *The advancement of science: science without legend, objectivity without illusions*. New York, NY: Oxford University Press.

Kuhn, D., & Pease, M. (2008). What Needs to Develop in the Development of Inquiry Skills? *Cognition and Instruction*, *26*(4), 512–559. doi:10.1080/07370000802391745

Lange, M. (2005). Ecological laws: what would they be and why would they matter? *Oikos, 110*(2), 394–403.

Lawton, J. H. (1999). Are There General Laws in Ecology? *Oikos, 84*(2), 177–192.

Lemke, J. L. (1990). *Talking Science: Language, Learning, and Values*. Westport, CT: Ablex Publishing.

Levin, D. M., Hammer, D., & Coffey, J. E. (2009). Novice Teachers’ Attention to Student Thinking. *Journal of Teacher Education*, *60*(2), 142–154. doi:10.1177/0022487108330245

Levy, S. T., & Wilensky, U. (2008). Inventing a “Mid Level” to Make Ends Meet: Reasoning between the Levels of Complexity. *Cognition and Instruction*, *26*(1), 1–47. doi:10.1080/07370000701798479

Lineback, J. E. (2015). The Redirection: An Indicator of How Teachers Respond to Student Thinking. *Journal of the Learning Sciences*, *24*(3), 419–460. doi:10.1080/10508406.2014.930707

Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological Resources: Applying a New Epistemological Framework to Science Instruction. *Educational Psychologist*, *39*(1), 57–68. doi:10.1207/s15326985ep3901

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3-11.

May, D. B., Hammer, D., & Roy, P. (2006). Children’s Analogical Reasoning in a Third-Grade Science Discussion. *Science Education*, *90*(2), 316–330. doi:10.1002/sce.20116

May, R. M. (1974). Biological populations with nonoverlapping generations: Stable points, stable cycles, and chaos. *Science, 186*(4164), 645–648.

May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature, 261*, 459–467.

Mayr, E. (1998). *What biology is: The science of the living world*. Cambridge, MA: Harvard University Press.

McClain, K., & Cobb, P. (2001). An Analysis of Development of Sociomathematical Norms in One First-Grade Classroom. *Journal for Research in Mathematics Education*, *32*(3), 236. doi:10.2307/749827

McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models*. Hillsdale, NJ: Erlbaum.

McDermott, R., Gospodinoff, K., & Aron, J. (1978). Criteria for an ethnographically adequate descripton of concerted activities and their contexts. *Semiotica*, *24*(3–4), 245–276.

Metz, K. E. (2011). Disentangling Robust Developmental Constraints From the Instructionally Mutable: Young Children's Epistemic Reasoning About a Study of Their Own Design. *Journal of the Learning Sciences*, *20*(1), 50–110. doi:10.1080/10508406.2011.529325

Murray, B. (1992). Research Methods in Physics and Biology. *Oikos*, *64*(3), 594-596. doi:10.2307/3545180

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. N. Giere (Ed.), *Cognitive Models of Science* (Vol. XV, pp. 3-44). Minneapolis: University of Minneapolis Press.

O'Connor, E., & McCartney, K. (2007). Examining teacher–child relationships and achievement as part of an ecological model of development. *American Educational Research Journal*, *44*(2), 340-369.

O'Connor, M. C., & Michaels, S. (1993). Aligning Academic Task and Participation Status through Revoicing: Analysis of a Classroom Discourse Strategy. *Anthropology and Education Quarterly*, *24*(4), 318–335. doi:10.1525/aeq.1993.24.4.04x0063k

O'Hara, R. (2005). The anarchist's guide to ecological theory. Or, we don't need no stinkin'laws. *Oikos*, *110*(2), 390–393. doi:10.1111/j.0030-1299.2005.13959.x/full

Philip, T. M. (2011). An "Ideology in Pieces" Approach to Studying Change in Teachers' Sensemaking About Race, Racism, and Racial Justice. *Cognition and Instruction*, *29*(3), 297–329. doi:10.1080/07370008.2011.583369

Philip, T. M., Gupta, A., Elby, A., & Turpen, C. (2018). Why Ideology Matters for Learning: A Case of Ideological Convergence in an Engineering Ethics Classroom Discussion on Drone Warfare. *Journal of the Learning Sciences*, *27*(2), 183-223. doi:10.1080/10508406.2017.1381964

Phillips, A., Watkins, J., & Hammer, D. (2017). Problematizing as a scientific endeavor. *Physical Review Physics Education Research*, *13*(20107), 1–13.

Piaget, J. (1950/2005). The psychology of intelligence: Routledge.

Pierson, J. L. (2008). *The Relationship Between Patterns of Classroom Discourse and Mathematics Learning* (Doctoral dissertation). Retrieved from Texas ScholarWorks. (hdl.handle.net/2152/17898)

Pruitt, J. N., Bolnick, D. I., Sih, A., Dirienzo, N., & Pinter-Wollman, N. (2016). Behavioural hypervolumes of spider communities predict community performance and disbandment. *Proceedings of the Royal Society B, 283*, 1–7.

Purvis, A., & Hector, A. (2000). Getting the measure of biodiversity. *Nature*, *405*, 212–219. doi:10.1126/science.306.5700.1309

Radoff, J. (2017). *Dynamics contributing to the emergence and stability of students' scientific engagement over multiple timescales* (Doctoral dissertation). Retrieved from Proquest (1986008525).

Radoff, J., Jaber L. Z., & Hammer, D. (in press). "It's scary but it's also exciting": Evidence of meta-affective learning in science. *Cognition and Instruction*.

Robertson, A. D., Scherr, R. E., & Hammer, D. (2016). *Responsive teaching in science and mathematics*. New York: Routledge, Taylor & Francis Group.

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, *92*(3), 499–525. doi:10.1002/sce.20264

Sandoval, W. (2013). Conjecture Mapping: An Approach to Systematic Educational Design Research. *Journal of the Learning Sciences, 23*(1), 18-36. doi:10.1080/10508406.2013.778204

Saxe, G. B., & Esmonde, I. (2005). Studying Cognition in Flux: A Historical Treatment of Fu in the Shifting Structure of Oksapmin Mathematics. *Mind, Culture, and Activity*, *12*(3), 171–225. doi:10.1207/s15327884mca123&4_2

Saxe, G. B., Gearhart, M., Shaughnessy, M., Earnest, D., Cremer, S., Sitabkhan, Y., … Young, A. (2009). A methodological framework and empirical techniques for studying the travel of ideas in classroom communities. In R. Schwartz, B. Dreyfus, & T. Hershkowitz (Ed.), *Transformation of knowledge in classroom interaction* (pp. 203–222). New York, NY: Routledge

Schrader-Freschette, K. S. & McCoy, E. D. (1993). *Method in Ecology: Strategies for Conservation.* New York, NY: Cambridge University Press.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*(2), 129-184.

Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: An ecological and evolutionary overview. *Trends in Ecology and Evolution, 19*(7), 372–378. doi:10.1016/j.tree.2004.04.009

Simberloff, D. (2004). Community Ecology : Is It Time to Move On ? (An American Society of Naturalists Presidential Address). *The American Naturalist*, *163*(6), 787–799.

Southerland, S. A., Johnston, A., & Sowell, S. (2006). Describing teachers' conceptual ecologies for the nature of science. *Science Education, 90*(5), 874-906.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5). 797-811.

Stillman, R. A., Railsback, S. F., Giske, J., Berger, U., & Grimm, V. (2015). Making predictions in a changing world: The benefits of individual-based ecology. *BioScience*, *65*(2), 140–150. doi:10.1093/biosci/biu192

Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of Science, Cognitive Psychology, and Educational Theory and Practice* (pp. 147-176). Albany: State University of New York Press.

Svoboda, J., & Passmore, C. (2013). The Strategies of Modeling in Biology Education. *Science & Education*, *22*, 119–142. doi:10.1007/s11191-011-9425-5

Tabak, I. (2004). Reconstructing context: Negotiating the tension between exogenous and endogenous educational design. *Educational Psychologist, 39*(4), 225-233.

Tang, X., Coffey, J. E., Elby, A., & Levin, D. M. (2010). The scientific method and scientific inquiry: Tensions in teaching and learning. *Science Education*, *94*(1), 29–47. doi:10.1002/sce.20366

Tannen, D. (1993). What's in a Frame? Surface Evidence for Underlying Expectations. In D. Tannen (Ed.), *Framing in Discourse* (pp. 14–56). New York, NY: Oxford University Press.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, Mass.: MIT Press.

Thompson, J., Hagenah, S., Kang, H., Stroupe, D., Braaten, M., Colley, C., & Windschitl, M. (2016). Rigor and Responsiveness in Classroom Activity. *Teachers College Record*, *118*(5), 1–58.

Turchin, P. (2001). Does Population Ecology Have Laws? *Oikos*, *94*(1), 17–26. doi:10.1034/j.1600-0706.2001.11310.x

Violle, C., Enquist, B. J., Mcgill, B. J., Jiang, L., Albert, C. H., Hulshof, C., … Messier, J. (2012). The return of the variance : intraspecific variability in community ecology. *Trends in Ecology & Evolution*, *27*(4), 244–252. doi:10.1016/j.tree.2011.11.014

Watkins, J., Hammer, D., Radoff, J., Jaber, L. Z., & Phillips, A. M. (2018). Positioning as not-understanding : The value of showing uncertainty for engaging in science. *Journal of Research in Science Teaching*, *55*, 573–599. doi:10.1002/tea.21431

Watkins, J., McCormick, M. E., Milto, E., Portsmore, M., Spencer, K., Wendell, K. B., & Hammer, D. (2018). Data-based conjectures for supporting responsive teaching in engineering design with elementary teachers. *Science Education*, *102*(3), 548-570.

Watkins, J., Spencer, K., & Hammer, D. (2014). Examining Young Students' Problem Scoping in Engineering Design. *Journal of Pre-College Engineering Education Research*, *4*(1), 43–53.

Wilson, E. O. (1999). *Consilience: The unity of knowledge.* New York, New York: Vintage Books.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the Scientific Method : Model-Based Inquiry as a New Paradigm of Preference for School Science Investigations. *Science Education*, *92*(5), 941-967.

Yackel, E., & Cobb, P. (1996). Sociomathematical Norms, Argumentation, and Autonomy in Mathematics. *Journal for Research in Mathematics Education*, *27*(4), 458. doi:10.2307/749877