

Idiosyncratic cases and hopes for general validity: what education research might learn from ecology / Casos idiosincrásicos y expectativas de validez general: lo que la investigación en educación puede aprender de la ecología

David Hammer, Julia Gouvea & Jessica Watkins

To cite this article: David Hammer, Julia Gouvea & Jessica Watkins (2018) Idiosyncratic cases and hopes for general validity: what education research might learn from ecology / Casos idiosincrásicos y expectativas de validez general: lo que la investigación en educación puede aprender de la ecología, *Infancia y Aprendizaje*, 41:4, 625-673, DOI: [10.1080/02103702.2018.1504887](https://doi.org/10.1080/02103702.2018.1504887)

To link to this article: <https://doi.org/10.1080/02103702.2018.1504887>



Published online: 20 Aug 2018.



Submit your article to this journal [↗](#)



Article views: 1112



View related articles [↗](#)



View Crossmark data [↗](#)



Idiosyncratic cases and hopes for general validity: what education research might learn from ecology / *Casos idiosincrásicos y expectativas de validez general: lo que la investigación en educación puede aprender de la ecología*

David Hammer^a, Julia Gouvea^a, and Jessica Watkins^b

^aTufts University; ^bVanderbilt University

(Received 1 April 2018; accepted 21 June 2018)

Abstract: We reflect on our ongoing struggles with rigour 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

Resumen: Reflexionamos sobre nuestras dificultades con el rigor y la validez en un proyecto basado en análisis de diversos estudios de caso: ¿Cómo pueden aportar a la investigación resultados de validez perdurable las instancias particulares de aprendizaje, con todas sus peculiaridades y sus dinámicas idiosincrásicas? En este artículo, nos enfocamos en el ámbito de la ecología, que se enfrenta a dificultades similares, para aprender sobre sus objetivos y métodos de sus investigaciones.

Palabras clave: estudios de caso; metodología; educación en ciencias

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 car¹.

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 centred on rolling, and one student, Isaac, got to thinking about how rolling works.

English version: pp. 625–644 / *Versión en español*: pp. 645–668

References / *Referencias*: pp. 668–673

Translated from English / *Traducción del inglés*: Mercè Rius.

Authors' Address / *Correspondencia con los autores*: David Hammer, Department of Education, Tufts University, Medford, MA 02155, USA. Email: david.hammer@tufts.edu

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 because 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 rigour 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 modelling (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, Hammer, Sherin, & Kolpakowski 1991) and *productive disciplinary engagement* (PDE, Engle & Conant, 2002). Our recent work includes Watkins, McCormick, 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 it 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 at 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 to a 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 & Martin, 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, rigour 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, 2015–2018). 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 rigour. 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; analysing 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, Thompson, & Braaten, 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 shift 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 do 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 15 candidates to present to the panel, drawing from other projects (especially Goldberg, Hammer, Bendall, & Coffey, 2008–2011) and collecting new data in undergraduate courses at Tufts. All of these 15 involved video as the primary form of data, often supplemented with writings and drawings, such as from Isaac's notebook². 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 analyse. 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 exclude whatever took place 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.

Analysing 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 analyse 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 — 7,500 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., favouring 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 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 & Conant, 2002; Engle, Lam, Meyer, & Nix, 2012), 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, Watkins, & Hammer, 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 cases³.

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, Hammer, 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, Hammer, et al., 2018, p. 577). Every case in our set included evidence of at least some participants positioning themselves 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) analyse 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 et al., 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 analyse 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 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 analyse Scarlett's question focusing on her in that moment, 'Scarlett's question', or we could analyse 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 of 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 analyse).

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 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. 1,316)

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 (Southerland, Johnston, & Sowell, 2006; Strike & Posner, 1992) and of human development (Bronfenbrenner, 1977; O’Connor & McCartney, 2007). We have been more directly influenced by Thelen 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 rigour 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, Dayton, Lonsdale, & Spiess, 1986) and Kelp Forests off the coast of central California (Dayton, Currie, Gerrodette, Keller, Rosenthal, & Tresca, 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 of 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 defences. 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, Inouye, Johnson, Underwood, & Vellend, 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., 2011, 2003; Hughes et al., 2008; Sih, Bell, & Johnson, 2004; Violle et al., 2012).

In the study of animal behaviour, for instance, classical studies reported on the average behaviour 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, Bolnick, Sih, Dirienzo, & Pinter-Wollman, 2016). They first measured the behaviours 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 behaviour 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 behaviour would have prevented ecologists from understanding why some communities persist and others collapse.

A second example concerns mathematical approaches to modelling population sizes, which in some cases can fluctuate dramatically from one year to the next. Theoretical ecologists wondered if these fluctuations could be modelled using chaos theory (Hastings, Hom, Ellner, Turchin, & Godfray, 1993; May, 1974, 1976), which, as we discussed above, can explain with mathematical rigour 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, 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 (Benincà, Ballantine, Ellner, & Huisman, 2015; Bjornstad, 2015; Constantino, Cushing, Dennis, & Desharnais, 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 favour 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 analysing that data? And what should we hope our findings accomplish, beyond particular episodes?

Delimiting the data

We described our approaches and difficulties in bounding the data to analyse 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 make 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 analysing 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 analysing 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 modelling more salient. It has renewed our interest in data-driven interpretation of the relevant scales of dynamics (Conlin, Gupta, & Hammer, 2010; Conlin & Hammer, 2016). 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 analysing 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 & Esmonde, 2005; Saxe et al., 2009).

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 an 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 rigour, 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 scepticism 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 et al., 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 ecosystems, 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 re-examining 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.

Notes

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).
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.
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.

Casos idiosincrásicos y expectativas de validez general: lo que la investigación en educación puede aprender de la ecología

Formulación de la pregunta

Imaginemos la siguiente escena de una clase de ciencia con niños de tercer grado en una escuela pública estadounidense que están tratando de comprender qué ocurre con el movimiento de un coche de juguete¹.

La unidad de estudio había comenzado unas semanas antes con la siguiente ‘pregunta introductoria’ (Hammer, Goldberg, & Fargason, 2012): ¿De qué maneras podemos hacer que el coche de juguete se mueva? Los niños pensaron rápidamente en muchas posibilidades y durante varios días, con la ayuda de la maestra, experimentaron con rampas y gomas elásticas. Algunas de sus ideas con rampas se centraban en que el coche rodase. Uno de los alumnos, Isaac, se puso a pensar en el mecanismo de rodamiento.

Este episodio, al que nos referiremos como *Las ruedas de Isaac*, comienza cuando el niño comenta en la clase que ‘es importante’ que el coche tenga ruedas. La maestra, Sharon Fargason, le pidió que leyese una anotación que había visto en su diario de anotaciones: ‘El coche va más rápido porque las ruedas van siguiendo el suelo’. Isaac también había dibujado diversos bocetos de engranajes, una bicicleta y una rueda de bicicleta con puntas o tal vez con un dentado en el perímetro exterior.

Con la ayuda de la maestra, Isaac explicó su idea sobre el rodamiento: cada punto de una rueda coincide con un punto del suelo y esto hace que la rueda pueda avanzar sin arrastrarse. No resultaba fácil seguir su línea de pensamiento, y sus compañeros hicieron muchas preguntas. Jourdan preguntó qué haría que el coche parase, a lo que Isaac respondió que las ruedas ‘se cansarían’. Scarlett preguntó entonces: ‘¿Cómo pueden cansarse las ruedas?’ Esta pregunta hizo que Isaac buscara un mecanismo más tangible para explicar la desaceleración del coche: ‘Las ruedas de goma *rozan* contra las *cuerdas* (ejes) que las sujetan’.

Este es uno de los múltiples ejemplos que hemos acumulado y estudiado durante años de investigación sobre las indagaciones de los niños. Para los físicos, *Las ruedas de Isaac* es un ejemplo impresionante y fácilmente identificable como ciencia incipiente: los físicos consideran la teoría del rodamiento de Isaac como correcta, reconocen la pregunta de Jourdan como esencial y comparten la preocupación de Scarlett sobre la respuesta inicial de Isaac. La mayoría se muestran

sorprendidos ante la capacidad de los alumnos de tercer grado, y les motiva a analizar cómo surge este tipo de pensamiento (Radoff, 2017).

Asumimos la premisa de que este episodio y muchos otros — en particular, sus registros en vídeo, audio y escritos — tienen un gran valor para la investigación y el aprendizaje de las ciencias. La pregunta es ¿Cómo puede contribuir un episodio de este tipo, un ejemplo tan claro de lo que los educadores de ciencias persiguen, a la investigación sobre el aprendizaje, una investigación rigurosa que podría producir resultados de valor general y duradero para la educación en ciencias?

Pruebas de existencia y construcción de teoría

Por supuesto, esta pregunta no es nueva. Durante mucho tiempo, los investigadores de la educación han debatido y se han mostrado divididos sobre los méritos y la validez tanto de la investigación cualitativa como de la cuantitativa. En este artículo no revisamos esa literatura; nuestro estudio es sobre el aprendizaje y la enseñanza de las ciencias, no sobre la metodología. No obstante, está claro que no podemos obviar hablar sobre la metodología que concierne a nuestro trabajo. Aprovechamos la oportunidad de escribir este artículo para reflexionar sobre nuestra lucha por la calidad y el rigor en nuestras investigaciones de episodios como *Las ruedas de Isaac*.

En cierta manera, está claro qué podemos conseguir con el estudio de un episodio concreto. En primer lugar, este puede constituir una prueba de existencia con la que desafiar algunos supuestos ampliamente extendidos. En este caso, por ejemplo, Jourdan preguntó sobre un mecanismo que explicase cómo *se para* el coche, contradiciendo la expectativa de que los estudiantes novatos de física sostienen la idea errónea pero arraigada de que el movimiento debe tener una causa (McCloskey, 1983). La primera respuesta de Isaac de que las ruedas ‘se cansan’ corrobora esa idea errónea, pero Scarlett inmediatamente la cuestiona como poco creíble, por lo que Isaac procede a identificar un mecanismo de desaceleración.

Existen muchos ejemplos del uso de este tipo de episodios en la literatura, en los que se muestra ejemplos de la capacidad de los estudiantes para el razonamiento epistémico (Metz, 2011), el razonamiento sobre fenómenos emergentes (Levy & Wilensky, 2008), la construcción de sentido en álgebra (Brizuela & Schliemann, 2004), y muchas otras capacidades. Nosotros también hemos utilizado los estudios de caso de este modo, incluso para documentar ejemplos de razonamiento sobre mecanismos (Louca, Elby, Hammer, & Kagey, 2004; Russ, Scherr, Hammer, & Mikeska, 2008), el alcance de los problemas (Watkins, Spencer, & Hammer, 2014) y construcción de modelos (Svoboda & Passmore, 2013).

Por lo que respecta a las pruebas de existencia, la respuesta a nuestra pregunta parece clara, puesto que la afirmación solo concierne el caso en cuestión. Sería difícil rebatir que Jourdan está reflexionando qué hace que el coche se pare, o que Scarlett no acepta la respuesta de que las ruedas ‘se cansan’, así que este caso es

una demostración rigurosa de que los estudiantes no *siempre* se dejan llevar por la idea errónea de que el movimiento requiere una causa, como se afirma en la literatura (e.g., McCloskey, 1983).

Otra función de los estudios de caso es desarrollar nuevos constructos teóricos y nuevas hipótesis, como la *competencia meta-representacional* (MRC, diSessa, Hammer, Sherin, & Kolpakowski, 1991) y el *involucramiento disciplinario productivo* (PDE, Engle & Conant, 2002). Nuestro trabajo más reciente incluye el uso que Watkins, McCormick, et al. (2018) hacen de los estudios de caso para formular hipótesis sobre la formación continua de profesores en el campo de la ingeniería; y el estudio de caso de Radoff, Jaber, y Hammer (en prensa) sobre el progreso de un estudiante para plantear la noción de *aprendizaje meta-afectivo*.

Por supuesto, puede existir un valor perdurable de las nuevas ideas teóricas en la investigación sobre la educación en términos de su *generatividad*: las nuevas ideas sobre cómo las personas piensan y aprenden amplían nuestra capacidad de imaginar qué es posible. La MRC y la PDE han inspirado claramente este tipo de pensamiento en este campo.

La generatividad es un marcador de progreso en diversas disciplinas. En las artes, los saltos creativos inspiran a pintores, escritores y músicos a avanzar en nuevas direcciones. En las ciencias naturales, el progreso suele ir acompañado de novedosas construcciones que hacen posible pensar sobre el mundo de nuevas maneras. Hizo falta una mirada creativa para imaginar el espacio entre dos imanes ocupado por ‘líneas de fuerza’, responsables de la atracción y la repulsión de los objetos. Esta idea abrió una nueva senda para la conceptualización del espacio ‘vacío’ (Nersessian, 1992).

No obstante, también aspiramos a obtener resultados de una *validez* general duradera, como los obtenidos en el ámbito de las ciencias naturales. Aunque valoramos la metáfora creativa de las *líneas de fuerza*, también nos preocupa que el modelo perdure, si resistirá las comprobaciones empíricas y la crítica teórica. Las *líneas de fuerza* derivaron en la teoría más general y matemáticamente más rigurosa de los ‘campos electromagnéticos’, que ha resistido verificaciones y críticas exhaustivas, así como la ‘prueba del tiempo’. Esta teoría sigue teniendo un papel central en la física, perdurando por sí misma y facilitando el desarrollo de otros constructos más generales en la teoría de campos.

Nuestras trayectorias provienen de la física (David y Jessica) y de la biología (Julia), ámbitos que dan forma a cómo concebimos los objetivos epistémicos, los valores, las virtudes y los procesos de la investigación (Chinn, Buckland, & Samarapungavan, 2011). La física, particularmente, ha priorizado la búsqueda de leyes fundamentales, que ha tenido influencia pero no se considera central en la biología. ¿Cuáles son, nos preguntamos, los objetivos y valores productivos en la investigación sobre la educación?

Tal vez sea necesario ajustar nuestras aspiraciones. Los filósofos analíticos aspiran a producir resultados de validez duradera, pero nuestra colega Nancy Bauer (2015) afirma que no deberían hacerlo, que el progreso en filosofía es como el progreso en las artes. Pintores, escritores y músicos desarrollan nuevas ideas y confían en que estas ejerzan cierta influencia sobre trabajos posteriores,

pero no someten esas ideas a las exigencias de la evidencia: el impresionismo fue progreso, pero un tipo de progreso distinto de la formulación de la teoría de los campos.

Refinamiento de la pregunta

En las ciencias naturales y en la ingeniería existen metodologías arraigadas para producir y evaluar resultados válidos. Tal vez la más extendida y reconocida sea la del ‘ensayo aleatorio controlado’ (RCT, por su denominación en inglés), que básicamente divide una población de manera aleatoria en dos situaciones o contextos distintos y compara los resultados obtenidos. Existen motivos excelentes para considerar el RCT como el ‘estándar de oro’ en algunos ámbitos, entre los que incluimos la investigación médica: un grupo recibe un tratamiento, mientras que el otro recibe un placebo, y el estudio analiza las diferencias en resultados medidos, que pretende demostrar un efecto idealmente de gran tamaño y significatividad estadística. Este tipo de estudios han tenido usos muy productivos también en el ámbito de la educación para el estudio de hipótesis que permiten definir variables dependientes e independientes muy claras. Los investigadores han recurrido al RCT o a métodos similares para demostrar la eficacia de ciertas intervenciones educativas (Carpenter, Fennema, Peterson, Chiang, & Loef, 1989) para comprobar constructos teóricos tales como la amenaza del estereotipo (Steele & Aronson, 1995) y para estudiar la ‘preparación para el aprendizaje futuro’ (Schwartz & Martin, 2004).

También existen razones excelentes para cuestionar el estatus de ‘estándar de oro’ del método RCT, incluida la alegación de que su uso y validez depende de un conjunto restrictivo de condiciones para el objeto de estudio como, por ejemplo, que no exista una variación importante dentro de la población del estudio, por ejemplo (Cartwright, 2007).

Para los propósitos de este artículo, la necesidad de variables dependientes e independientes bien definidas — es decir, aspectos ‘input’ o de entrada de la situación y resultados ‘output’ o de salida — pone de relieve la clave del problema de la búsqueda de relaciones válidas en general en casos como el de *Las ruedas de Isaac*: ¿Cómo identificamos y definimos estas variables? ¿Qué aspectos de la situación son relevantes y claramente aplicables en otras situaciones?

En la clase de tercer grado de Isaac, algunos aspectos comunes que podríamos estudiar podrían ser la pregunta introductoria, ‘¿Qué maneras hay de hacer que el coche se mueva?’ o el uso de las anotaciones en sus diarios. De hecho, estos eran aspectos comunes en todas las clases de tercer grado del proyecto.

Pero mucho más en este episodio era idiosincrásico. La analogía de Isaac con los engranajes solo ocurrió en esta clase, y fue clave. Si quisiéramos incluirla como parte de un estudio más general, deberíamos tratarla como un ejemplo de una categoría más general, posiblemente de analogía (Blanchette & Dunbar, 2000) o de un razonamiento sobre mecanismos (Russ et al., 2008). Esa abstracción

implicaría revisar otros casos para identificar y refinar categorías, lo que suele requerir un juicio subjetivo matizado.

También hay que tener en cuenta el modo en que la maestra reconoció, valoró y se involucró con las ideas de los estudiantes, incluida la entrada de Isaac en su diario y su explicación en la clase, así como las preguntas de Jourdan, Scarlett y otros estudiantes. Estos también fueron clave en la dinámica de este episodio, pero no podemos convertirlos, exactamente, en variables dependientes ni independientes para ser estudiadas de modo más general. De nuevo, tendríamos que revisar otros casos, incluso con otros profesores, para extraer categorías más generales (Levin, Hammer, & Coffey, 2009; Lineback, 2015; O'Connor & Michaels, 1993; Pierson, 2008).

Parte del problema es que hay tantos aspectos particulares en el episodio que podríamos considerar como importantes. Otra parte es que estos aspectos se relacionan entre ellos en una dinámica reflexiva y compleja (Yackel & Cobb, 1996): el hecho que Scarlett sintiera que podía preguntarle a Isaac '¿Cómo se cansan las ruedas?' refleja su manera de enmarcar lo que estaban haciendo todos juntos y su posible papel en ese conjunto, formado y tomando forma a través de unas normas emergentes de la propia clase. Estas normas reflejan que la maestra Fargason había estado prestando atención y respondiendo a las ideas de los estudiantes; y que la maestra tomó nota y respaldó la pregunta de Scarlett reflejó y contribuyó también a esas normas.

En un artículo anterior, Hammer y Sikorski (2015) consideraron las implicaciones de la complejidad en la investigación en progresiones de aprendizaje (*learning progressions*), comparando las relaciones reflexivas en el aula a las interacciones no lineales en sistemas físicos para argumentar la posibilidad de dinámicas caóticas. La teoría del caos está muy desarrollada en muchos sistemas físicos, su rigor avalado por estudios empíricos: cuando existe una dinámica caótica, las diferencias más mínimas entre sistemas pueden ampliarse y producir resultados radicalmente distintos. Si el aprendizaje en las aulas implica una dinámica caótica, es necesario un enfoque distinto de la mera acumulación de datos en múltiples ejemplos, como ocurre con los RCT, que son diseñados para lograr eliminar las pequeñas variaciones en favor de las medias estadísticas.

Esta es, pues, una versión refinada de nuestra pregunta para este artículo: '¿Cómo pueden contribuir a una investigación sobre el aprendizaje con rigor y validez perdurable los distintos episodios de estudiantes haciendo ciencia, con todas sus particularidades y complejidades?'

Objetivos y planes para este artículo

Esta pregunta surge en el contexto de nuestras investigaciones en curso para el estudio de 'las dinámicas del involucramiento y la persistencia de los estudiantes en la ciencia' (Hammer, Watkins, & Gouvea, 2015–2018). Abríamos este artículo con un ejemplo de estudiantes 'haciendo' ciencia, como caso paradigmático para centrar y motivar nuestra pregunta. En la sección siguiente describimos nuestro proyecto actual como ejemplo ilustrativo de nuestra metodología de investigación.

Discutimos nuestro enfoque en este proyecto y lo que percibimos como problemas pendientes de método y rigor investigador. A continuación analizamos investigaciones en el campo de la ecología, que ha lidiado con problemas similares de complejidad, en sus esfuerzos por comprender los sistemas biológicos. Por último, reflexionamos de nuevo sobre las implicaciones de nuestra revisión de las investigaciones en ecología y qué nos sugieren para nuestros trabajos de investigación sobre el aprendizaje.

Las dinámicas del involucramiento de los estudiantes en las ciencias

Una pregunta para la investigación y los desafíos para responder a ella

Nuestro proyecto en curso consiste en el estudio de las dinámicas en el involucramiento y la persistencia de los estudiantes mientras hacen ciencia. Con ello tratamos de contribuir al objetivo ya histórico de que los estudiantes aprendan a involucrarse en su propia construcción de sus comprensiones, en lugar de aprender ciencia como un cuerpo estático de conocimiento. A pesar del trabajo de educadores e investigadores para conseguir este objetivo, todavía es relativamente poco frecuente ver a estudiantes haciendo ciencia en las aulas (Duschl & Osborne, 2002; Lemke, 1990; Thompson et al., 2016). Existen trabajos en los que se trata de identificar los obstáculos o desafíos al involucramiento científico de los estudiantes (Chinn & Malhotra, 2002; Kuhn & Pease, 2008; Tang, Coffey, Elby, & Levin, 2010). En este proyecto tratamos de analizar momentos en que los estudiantes están involucrados para entender cómo llegan a ocurrir. Nuestro objetivo es ambicioso: abstraer patrones generales reconociendo al mismo tiempo las particularidades y complejidades de las dinámicas en los distintos casos.

En este artículo describimos los desafíos metodológicos a los que nos enfrentamos y las decisiones tomadas en tres fases de nuestro trabajo: la selección de los episodios para su estudio, el análisis de las dinámicas en cada uno de ellos y la comparación entre los casos para extraer patrones y temas.

Selección de los episodios para su estudio

El primer desafío al que nos enfrentamos es identificar los episodios objeto de estudio que, por supuesto, deben reflejar de algún modo qué es la ciencia. No existe una definición clara, ni por parte de los filósofos (Giere, 1988; Kitcher, 1993), ni de los educadores (Grandy & Duschl, 2007; Windschitl, Thompson, & Braaten, 2008) ni de los científicos (Mayr, 1998; Wilson, 1999). Cuando surge esta pregunta entre los físicos, la respuesta típica es admitir la derrota: 'Física es lo que hacen los físicos' o 'Física es el estudio de fenómenos físicos'. Parte del desafío es que lo que se considera ciencia y cómo se lleva a cabo cambia a través del tiempo: la teoría del caos y la mecánica cuántica, por ejemplo, han desafiado los supuestos fundacionales de cómo se desarrolla la física y su alcance. Para nosotros, el desafío consiste también en que estamos buscando *el surgimiento* de la ciencia en el pensamiento de los estudiantes.

En lugar de definir un conjunto de criterios específicos, nos basamos en un acuerdo intersubjetivo. Primero buscamos episodios en que los estudiantes parecen involucrados en comprender algún fenómeno natural, episodios que parecieran constituir ciencia. Después los presentamos como candidatos a un panel de colaboradores científicos e investigadores en el ámbito de la educación en ciencias. Los episodios que seleccionamos para estudio son aquellos en los que existe un consenso claro de que los datos reflejan a los estudiantes haciendo ciencia, aunque reconocemos que con este enfoque es posible que estemos excluyendo algunos datos cuyo estudio sería de gran valor.

En la primera revisión identificamos 15 candidatos que presentamos al panel de colaboradores, tomando datos de otros proyectos (especialmente de Goldberg, Hammer, Bendall, & Coffey, 2008–2011) y nuevos datos de asignaturas de grado en la Universidad de Tufts. El formato predominante de estos 15 episodios era el vídeo, complementado en muchos casos con escritos y dibujos, como en la libreta de Isaac². Ocho de los candidatos superaron con facilidad el consenso claro, y se dividían entre educación primaria y universitaria y pertenecían principalmente a las ciencias físicas. (En trabajos posteriores nos hemos centrado más en encontrar episodios relacionados con las ciencias de la vida y en niveles de educación secundaria y preuniversitaria.)

El siguiente desafío consiste en establecer los límites de los datos para el análisis. Se trata de una decisión importante: si establecemos unos límites muy restringidos alrededor de un evento particular, excluimos como datos para el análisis lo que ocurrió en la clase días o semanas antes, que con toda seguridad podrían haber ejercido cierta influencia. Por supuesto, nuestros datos excluyen lo que ocurre fuera del aula, como conversaciones o experimentos entre estudiantes. Reconociendo que nuestros datos no pueden estar ‘completos’, tomamos decisiones prácticas. En el caso de *Las ruedas de Isaac*, parecía bastante claro: el episodio comenzó con la participación de Isaac en la discusión y finalizó cuando el grupo pasó a otra tarea. En otros casos resulta más difícil decidir qué incluir y qué no, especialmente para casos con actividades que duran varios días.

Análisis de las dinámicas de un episodio

Diseñamos nuestro enfoque al análisis de datos cuidadosamente para tener en cuenta peculiaridades y complejidades, para considerar cada episodio con una mirada fresca y sin alegaciones predeterminadas. En la medida de lo posible, analizamos cada caso de manera independiente, evitando cualquier comparación explícita entre los casos y tratando de centrarnos en lo que surge de cada conjunto de datos. Además, distribuimos quién asume el rol principal en el análisis: los ocho análisis iniciales se dividieron entre cuatro miembros del equipo. El contenido de los análisis abarca tanto perspectivas teóricas como la envergadura del enfoque, incluidos los distintos pensamientos individuales (diSessa, 1993; Tannen, 1993) y las actividades e interacciones sociales (Jordan & Henderson, 1995; McDermott, Gospodinoff, & Aron, 1978).

El investigador líder en cada uno de los análisis redacta una primera versión de lo que ocurrió en cada episodio, citando evidencias que corroboran su relato. Por ejemplo, en *Las ruedas de Isaac*, la investigadora líder, Jennifer Radoff, describió que Scarlett hacía responsable a Isaac de la incipiente norma de explicaciones sobre mecanismos, y que su pregunta instó a Isaac a buscar una explicación sobre mecanismos. La presentación del análisis al grupo, que denominamos ‘trabajando en taller’, brinda la oportunidad a los demás compañeros a desafiar las afirmaciones, y en este caso a preguntar si existe evidencia suficiente de que Scarlett estaba buscando un mecanismo per se cuando rechazó la idea de que las ruedas pudiesen estar ‘cansadas’ o, por el contrario, consensuar que la evidencia es sólida que como respuesta, Isaac sugirió un mecanismo que pudiera explicar la desaceleración del coche.

Los análisis realizados constituyen un rico relato sobre lo acontecido en cada episodio con muchas alegaciones detalladas como aquellas sobre sus dinámicas, corroboradas por la evidencia en los datos. Estos relatos se extienden miles de palabras; 7,500 en el caso de *Las ruedas de Isaac*. Con frecuencia se indican distintas interpretaciones posibles que el equipo no ha podido dilucidar basándose en la evidencia disponible.

El último paso en cada caso consiste en elaborar una síntesis de ‘qué contribuyó’ al involucramiento y la persistencia del estudiante en las ciencias. Leemos cada uno de los casos e identificamos las principales características de las dinámicas. En *Las ruedas de Isaac*, incluyen el manejo de la actividad por parte de la maestra y su atención a las ideas y preguntas de los alumnos, los ricos recursos conceptuales de estos para poder pensar sobre el movimiento (e.g., saben que el coche se parará y tienen conocimientos sobre el rozamiento), sus respectivos papeles como agentes intelectuales y los tipos de conocimiento que estos tienen en cuenta (e.g., privilegiando las explicaciones sobre mecanismos, a diferencia de su trabajo anterior en esa unidad).

Comparación entre los casos

Una vez obtenida una cantidad suficiente de análisis, pasamos a mirar a través de ellos para identificar rasgos comunes, patrones que podamos abstraer más allá de las dinámicas particulares de los casos individuales. En primer lugar, partimos de las síntesis de lo que contribuyó para debatir distintas posibilidades y, de ahí, procedemos a seleccionar y ordenar las ideas que parecen tan apropiadas como importantes para estudiar como contribuciones potenciales a la literatura. El trabajo entonces consiste en definir estas ideas con mayor precisión, identificar literatura relevante y dar ejemplos y descripciones de lo que se considera evidencia en los datos.

Hasta ahora, hemos llevado a cabo este proceso una sola vez, con el conjunto inicial de los ocho análisis. Teníamos diversas ideas iniciales sobre patrones. En casi todos los casos aparecían demostraciones afectivas, es decir, de alegría o frustración; en varios casos no quedaba claro qué pregunta trataban de responder los estudiantes; mientras que en otros, su objetivo principal parecía ser formular una pregunta. Algunos casos incluían estudiantes que claramente identificaban

modelos (como en los dibujos de engranajes de Isaac). En todos los casos aparecían expresiones de confusión o incertidumbre por parte de algunos participantes, como la maestra Fargason y algunos estudiantes respecto a la propuesta de Isaac.

Algunos de estos patrones corroboraban claramente ciertas ideas ya existentes en la literatura. En casi todos los casos, por ejemplo, el profesor o maestra reconocía y se involucraba con las ideas o preguntas que no habían anticipado de los estudiantes (Robertson, Scherr, & Hammer, 2016). La maestra Fargason lo hizo durante todo el episodio, atendiendo a la idea de Isaac y después a las preguntas de Jourdan y Scarlett.

Otros patrones parecían ir en direcciones nuevas y trabajamos para desarrollarlos en forma de constructos más precisos. La ambigüedad en torno a las preguntas y el cuestionamiento de los estudiantes conectaba con dos ideas de la literatura. Una es el concepto de *problematización* (Engle & Conant, 2002; Engle, Lam, Meyer, & Nix, 2012), como aspecto del involucramiento disciplinar en las ciencias. La otra es la noción más general de *posicionamiento* (Harré & Van Langenhove, 1999), ‘una dinámica alternativa al concepto más estático de rol’ (Harré & Van Langenhove, 1999, p. 215). Alguien puede ser posicionado como ‘poderoso o impotente, seguro o tentativo, dominante o sumiso, definitivo o provisional, autorizado o no autorizado, etc’ (p. 217).

Definimos *problematización* como la búsqueda de una pregunta, y más precisamente como la actividad de ‘darse cuenta de un vacío en la comprensión, identificando y articulando su naturaleza específica y motivando a la comunidad sobre su existencia y significado’ (Phillips, Watkins, & Hammer, 2017). Entre otras evidencias de esta problematización se incluye la expresión de incomodidad por parte de los estudiantes sobre algo que no está bien o que falta, su expresión y re-expresión de distintas versiones de una misma pregunta y su indicación de que existe un vacío o una falta de consistencia que no acaban de comprender. Nos sorprendió la frecuencia con la que este aspecto aparecía como una parte importante de sus investigaciones, en seis de los primeros ocho casos³.

Y también identificamos el *posicionamiento de incompreensión* como posicionamiento social en el discurso de los participantes, el modo en que ‘revelan su incerteza o confusión sobre un fenómeno o una idea’ (Watkins, Hammer, et al., 2018, p. 576). Este aspecto también facilitó nuevos análisis, más exhaustivos, de los datos, en particular en busca de evidencia, por ejemplo, de muestras de ‘perplejidad, formulación de preguntas o indicaciones de su falta de conocimiento’ por parte de los estudiantes (Watkins, Hammer, et al., 2018, p. 577). En todos nuestros casos encontramos evidencia de al menos algunos de los participantes posicionándose a sí mismos en la incompreensión, a diferencia, quizás, de otros casos de argumentación entre estudiantes con distintos puntos de vista.

Dudas persistentes

Así es cómo estamos trabajando, nuestra respuesta actual a la pregunta ‘¿Cómo pueden contribuir a una investigación sobre el aprendizaje con rigor y validez perdurable los distintos episodios de estudiantes haciendo ciencia, con todas sus

particularidades y complejidades?’ Es la respuesta que podemos dar por el momento, para seguir trabajando, pero por supuesto no es la respuesta definitiva.

En esta sección reflexionamos sobre tres aspectos interconectados del proyecto, que seguimos considerando difíciles: (1) cómo delimitamos los datos objeto de estudio, (2) cómo analizamos las dinámicas complejas, y (3) cómo evaluamos la validez más general de los resultados.

Delimitación de los datos

Como ya hemos indicado, parecía claro comenzar nuestro análisis de *Las ruedas de Isaac* cuando la maestra pregunta a Isaac por primera vez, y finalizarlo cuando la clase pasa a otro tema. Aun así, seguía habiendo contrapartidas a esta delimitación que influían en nuestros resultados. Decidimos analizar solo los datos de esos 20 minutos, pero esos 20 minutos estaban integrados en un día completo, y en una unidad didáctica más amplia sobre el movimiento que había comenzado varias semanas antes. También formaban parte, de otro modo, de una clase en una escuela primaria de San Diego, en una comunidad particular; asimismo se integraban en una clase que formaba parte de un proyecto de investigación (Goldberg et al., 2008–2011) sobre ‘la enseñanza que responde a las ideas de los estudiantes’ (*responsive teaching*) (Robertson et al., 2016). Naturalmente, la inclusión de más datos nos hubiese brindado otra visión o perspectivas diferentes de las dinámicas observadas en esos 20 minutos.

Otras de las decisiones que tomamos tenían que ver con los detalles de los datos que íbamos a tener en cuenta. A modo de ejemplo, el vídeo nos hubiese permitido analizar la ubicación física de los niños en el aula, quién se sentaba junto a quién, sus posturas, etc. Todos los profesores saben bien que la disposición de los asientos puede influir en las dinámicas de la clase, pero estos aspectos no formaron parte de nuestro análisis. El vídeo es un medio excesivamente rico y no sería posible analizar todos los aspectos de los datos, incluso limitando el análisis a los 20 minutos de este episodio.

Con estos dos enfoques, mirando hacia afuera del episodio, y también hacia adentro, cómo delimitar los datos de estudio es un aspecto tan relevante como poco claro. Por el momento, hacemos lo que podemos, caso a caso, y en cada uno de ellos nos preguntamos si los datos que seleccionamos respaldan un análisis productivo para nuestra investigación.

Identificación de fenómenos y entidades

Sea cual sea la delimitación de los datos de los episodios, todos presentan múltiples aspectos que interactúan en dinámicas complejas. Esto representa desafíos muy relacionados para realizar inferencias sobre el caso objeto de estudio, y también para conjeturar qué puede resultar útil en otros casos.

Durante el taller sobre el análisis de *Las ruedas de Isaac*, por ejemplo, consideramos la inferencia sobre la pregunta de Scarlett ‘¿Cómo pueden cansarse las ruedas?’, que refleja su intuición sobre una norma epistemológica emergente,

de que en esta actividad, las explicaciones deberían incluir mecanismos tangibles. No obstante, tal vez Scarlett objetaba a la idea de que un objeto inanimado pudiese ‘cansarse’. Aunque no entraba en los límites de los datos seleccionados para este caso, sabíamos que la clase había debatido recientemente si el coche de juguete tenía ‘libre albedrío’. Esa discusión centró la atención de los alumnos en la diferencia entre seres vivos y objetos inanimados; los coches no tienen ‘libre albedrío’. Por tanto, podríamos analizar la pregunta de Scarlett centrándonos en su intervención en ese momento, ‘la pregunta de Scarlett’, o podríamos analizarla como parte de una dinámica que se inició antes de ese episodio y que incluyó a toda la clase.

Aquí, como en el resto de los casos, nos enfrentamos a cuestiones de adecuación de escala, de tiempo y de relevancia del sistema; cómo diseccionar y circunscribir los fenómenos y las unidades de análisis. Es decir, tenemos que seleccionar la granularidad más apropiada de tiempo y de la entidad o el sistema. Podríamos centrarnos en Scarlett como individuo, y en el momento particular en el que formula su pregunta; o podríamos centrarnos en el sistema formado por Jourdan, la maestra Fargason, Isaac y Scarlett durante los minutos en los que consideran la pregunta sobre qué hace que el coche desacelere; o también podríamos centrarnos en la clase en su conjunto, en los patrones de conversación y de atención durante varios días (si ampliamos los límites de los datos para el análisis).

Evaluación de la validez de los resultados

Como ya hemos descrito, primero trabajamos manteniendo los casos separados, tratando de generar y corroborar ideas sobre qué ocurrió que surgen de los datos de cada episodio. No obstante, la validez de una afirmación sobre las dinámicas de un caso particular no la avala necesariamente como resultado valioso; tal vez sea idiosincrásica de ese momento. Obviamente, en ese caso era relevante que Isaac pensara en los engranajes cuando trataba de darle sentido al rodamiento; pero no está claro que podía considerarse algo de valor generalizable para cualquier estudiante de tercer grado.

Nuestra búsqueda de patrones a través de los distintos casos nos exigía adoptar un nivel de abstracción por encima de las particularidades. Para dar un seguimiento a esas ideas, era necesario desarrollarlas y refinarlas, identificar evidencias, iterar los análisis y el desarrollo de un marco como se expone en los informes de investigación cualitativa (Chi, 1997; Glaser & Strauss, 1967). También evaluamos las ideas mediante lecturas y vínculos con la literatura, considerando cómo aluden a lo ya conocido o a las creencias más extendidas. De este modo podríamos corroborar nuestras afirmaciones de patrones a través de los casos, de la problematización y el posicionamiento de incomprensión. Cada una de ellas, podríamos afirmar, contribuye a la literatura relevante actual.

Al mismo tiempo, no propondríamos que estas ideas tienen una validez general como ‘leyes de pensamiento científico’. Ciertamente, es fácil pensar en ejemplos de estudiantes haciendo ciencia que no expresan *problematizaciones* ni

posicionamiento de incomprensión, como ocurre en episodios de argumentación entre estudiantes que se muestran seguros y convencidos de sus respectivas posiciones (e.g., Engle & Conant, 2002; Hammer, 1996).

Comparación con las ciencias naturales

En un escrito publicado en la revista *American Journal of Physics* en 1996, David planteó la cuestión de comparar las ideas en la investigación educativa con las ideas en la investigación física:

Tendemos a asumir que, para ser útil, la investigación debe producir resultados inequívocos y de validez demostrable, así como principios y métodos fiables y en términos lo suficientemente precisos como para no ser sensibles a la interpretación y el juicio. Estas asunciones podrían ser apropiadas si la educación ya fuese una ciencia aplicada — esperamos que los físicos faciliten a los ingenieros resultados precisos y fiables sobre los que basar sus diseños — pero no lo es. Describimos que los estudiantes albergan ‘ideas equivocadas’, pero no podemos presentar ese relato con la precisión y confianza con las que podemos describir, por ejemplo, las propiedades del hidrógeno.

Es posible que estas suposiciones no lleguen nunca a ser apropiadas para la educación, que las diferencias epistemológicas y fundamentales entre el estudio de la cognición humana y el estudio del mundo físico impidan un formalismo en la educación parecido al de la física. También es posible que la educación, como la física, logre alcanzar en algún momento un formalismo productivo. (Hammer, 1996, p. 1,316)

Allí se comparaba la investigación educativa con la física; partimos de lo que sabemos. Reflexionando sobre el trabajo y los desafíos de la investigación en nuestro proyecto actual, Julia percibió más conexiones con la biología. Sospechamos que esta puede ser una comparación más útil.

Sin duda, en psicología cognitiva existe una larga historia de vinculación con la biología. Piaget reconoció la inteligencia como un fenómeno biológico (Piaget, 1950/2005). Su modelo de desarrollo, incluidos los conceptos de asimilación y acomodación, se basaba en sus propios estudios previos de la adaptación y la autorregulación. Aquí perseguimos una conexión distinta, en particular con la ecología. En última instancia, esperamos que nuestras consideraciones coincidirán con las de Piaget, especialmente a la luz de los avances más recientes en los modelos de desarrollo de sistemas dinámicos (Thelen & Smith, 1994), construidos sobre la base de las ideas previas de autorregulación. Por ahora, nos centramos en la ecología.

En la investigación educativa también existe una larga historia de referencias a la ecología. Algunos autores se han pronunciado a favor de los modelos ecológicos para la comprensión conceptual (Southerland, Johnston, & Sowell, 2006; Strike & Posner, 1992) y el desarrollo humano (Bronfenbrenner, 1977; O'Connor & McCartney, 2007). Hemos recibido una influencia más directa del ‘enfoque de sistemas dinámicos’ de Thelen y Smith (1994), que los autores

describen mediante discusiones tanto de sistemas físicos como biológicos. Este último podría ser más apto, como discutimos más adelante, pero aun así, existe un solapamiento evidente con anteriores caracterizaciones de modelos ecológicos. De hecho, los desafíos epistemológicos y metodológicos que hemos descrito anteriormente reflejan nuestra perspectiva teórica ecológica sobre las dinámicas de la actividad en el aula y del aprendizaje de los estudiantes.

Lo que no hemos visto en llamamientos anteriores en favor de modelos ecológicos en la investigación sobre la educación es una consideración directa de la investigación en ecología. ¿Cómo gestionan los ecólogos las particularidades y complejidades de su trabajo? ¿Cómo conceptualizan y producen una investigación rigurosa de validez perdurable? A continuación nos ocupamos de estas preguntas y, después, volveremos a reflexionar sobre cómo la investigación en el ámbito de la ecología puede contribuir a nuestro trabajo.

La investigación en el ámbito de la ecología

Los ecólogos estudian sistemas complejos formados por poblaciones de especies en evolución que interactúan en entornos cambiantes. Por supuesto, la ecología es un campo relativamente joven en comparación con la física, pero también con otras subdisciplinas de la biología como la genética y la evolución. A medida que este campo de estudio ha ido creciendo, ha tenido lugar una discusión formal en la literatura dirigida a comprender y desarrollar una epistemología productiva de la ecología. Los desafíos que plantea su estudio son, en muchos aspectos, similares a los que nos encontramos en educación, en particular por su complejidad a través de las distintas escalas de los fenómenos estudiados.

En ecología es importante distinguir entre las ideas de lo *complicado* e *intricado* — ambos aplicables a los sistemas biológicos — de lo *complejo*. Podríamos asemejar un ecosistema a una maquinaria *complicada*, *intricada*, como un reloj, que podría estudiarse examinando su estructura y las propiedades de sus partes y cómo estas se combinan. De hecho, en los libros de texto de biología, muchos sistemas biológicos se describen en términos de sus partes y fases: mitosis, redes tróficas o el ciclo de Krebs.

Pero la analogía del reloj esconde a algunos aspectos de la *complejidad* del ecosistema. Como apuntaba el conocido ecólogo Charles Elton, para que un reloj se asemeje a un ecosistema, los engranajes tendrían que poder desarrollarse y evolucionar con el tiempo. Los componentes no solo estarían conectados; sus conexiones, mediante procesos análogos a la depredación y la competencia, provocarían cambios en los componentes. Por último, una parte del sistema ‘conserva el derecho a desvincularse, migrar y establecerse en otro reloj’ (Elton, cit. Hardy, 1968, p. 5). Los sistemas estudiados por los ecólogos son mucho más complejos que un reloj, porque sus componentes y las interacciones entre estos cambian y evolucionan con el tiempo, y se mueven a través de entornos heterogéneos. En consecuencia, los ecosistemas son muy variables en el tiempo y en el espacio. Cualquier ecosistema particular es un producto de la geografía y

el clima locales, así como de las historias particulares de los organismos que llegaron al lugar y evolucionaron en él.

Con toda esta complejidad y variabilidad, los ecólogos se han esforzado por generar un conocimiento riguroso y perdurable. Con la física como referente, algunos investigadores han buscado estos el valor en principios imperecederos que podríamos calificar como ‘leyes ecológicas’ (Lange, 2005; Lawton, 1999; Murray, 1992) o modelos matemáticos sencillos y elegantes (May, 1974, 1976). Otros han afirmado que una teoría ecológica general no tiene ninguna utilidad práctica y que la ecología debería abrazar su estatus como una ‘ciencia de estudios de casos’ (Hansson, 2003; Schrader-Freschette & McCoy, 1993). Así pues, los ecólogos han estado peleando con cuestiones similares a las nuestras y podríamos aprender de sus esfuerzos por conceptualizar y re-conceptualizar su disciplina.

En lo que queda de esta sección, hacemos una digresión de nuestra investigación para revisar algunos temas relacionados al progreso en ecología: los objetivos dejan de ser tanto las leyes generales y se convierten cada vez más en constructos de nivel medio de posibles mecanismos; los métodos tienen más en cuenta las dinámicas y escalas múltiples y tienden a invitar la variación y la complejidad. En la sección siguiente volvemos a reflexionar sobre cómo estas tendencias en el campo de la ecología pueden ayudarnos en nuestra investigación en el campo de la educación.

Abandonando las leyes

La búsqueda de ‘leyes de la ecología’ estaba motivada por las comparaciones con las leyes de la física: principios generales y abstractos, aplicables a través de los sistemas. En esta búsqueda, los ecólogos trataban de ‘no dejarse distraer por los equivalentes biológicos de la fricción’ (Murray, 1992, p. 596). En física, el esfuerzo por dejar de lado la fricción hizo posible la formulación de las leyes de Newton y parte de su éxito fue su capacidad, después, de explicar la fricción como un tipo de fuerza. Para investir la ecología de rigor, algunas voces defendieron la necesidad de dar pasos similares y dejar de lado la variabilidad y las fluctuaciones intranscendentes, con la esperanza de facilitar la formulación de leyes básicas.

Algunos ecólogos adoptaron ese camino y han propuesto algunas leyes como posibles candidatas. Un ejemplo es la ley de crecimiento exponencial: ‘Una población crecerá (o disminuirá) exponencialmente siempre que el entorno habitado por todos los individuos de la población permanezca constante’ (Turchin, 2001, p. 18). Como en el caso de las leyes de Newton, esta ley puede escribirse en forma de ecuación; surge de sencillas leyes físicas de las interacciones entre organismos, leyes sencillas de nacimiento y muerte y es una idealización en tanto que ignora aspectos de sistemas reales como las limitaciones de recursos y de espacio. Como ocurre con las leyes de Newton, los sistemas reales se desvían de la idealización; en poblaciones reales, los recursos y el espacio se agotan rápidamente, por lo que ninguna población natural crece exponencialmente durante mucho tiempo. Sin embargo, a diferencia de las leyes de Newton, las modificaciones necesarias para

ajustar la ecuación básica en cualquier población real fuera del laboratorio la hacen prácticamente inservible. La ley básica, aunque general, tiene un poder explicativo o predictivo limitado para los sistemas reales.

Otras leyes potenciales de la ecología corrieron una suerte similar y, con el tiempo, la búsqueda de leyes generales ha ido menguando. Muchos ecólogos han reconocido que, aunque fuera posible articular leyes de ecología, era poco probable estas que fuesen de utilidad (e.g., Hansson, 2003; Lange, 2005; O'Hara, 2005; Simberloff, 2004).

Otros fueron más reticentes a desistir de las leyes. Lawton (1999) hizo su célebre alusión a la investigación en ecología de las comunidades como 'un desastre' por el número de factores interrelacionados que influían en las dinámicas. 'La contingencia es abrumadoramente complicada en escalas intermedias características de la ecología de las comunidades, en las que se sitúa un gran número de casos y poco más que una generalización débil y difusa' (Lawton, 1999, p. 177). Para Lawton (1999), la implicación era que la ecología debía avanzar hacia una investigación a niveles que facilitarían el descubrimiento de 'patrones, leyes y normas generales' (p. 177).

Balance entre conocimiento global y local: una teoría intermedia

Simberloff (2004) discrepaba con Lawton en dos cosas. En primer lugar, defendía que la comprensión de la ecología a nivel de las comunidades tiene un gran valor para la sociedad, por ejemplo, para la gestión de piscifactorías o de los bosques. Estaba de acuerdo en que identificar los factores que influyen en las dinámicas de una comunidad particular no permitiría predecir cuál de esos factores sería importante en otra comunidad distinta. No obstante, el segundo punto de desencuentro era que Simberloff consideraba un error presumir que las leyes generales son la única forma de conocimiento útil. Por el contrario, la ecología ha ido avanzando significativamente mediante el desarrollo de mecanismos causales y enfoques locales para probar su incidencia en los sistemas.

El conocimiento sobre los mecanismos posibles (a diferencia de los necesarios) representa un nivel intermedio de construcción teórica; no una ley universal, pero sí de valor general. Un ejemplo clásico es la revisión realizada por Dayton (1984) de las investigaciones sobre la diversidad de las especies en los ecosistemas marinos. Su trabajo demostró cómo y cuándo se pueden aplicar posibles mecanismos como la competencia, la depredación, las alteraciones ambientales o la creación de nichos.

Por ejemplo, en la zona 'intermareal rocosa' — comunidades de especies que habitan las zonas rocosas y piscinas naturales de las zonas afectadas por las mareas — la competencia por el espacio domina las dinámicas de la diversidad de especies. Muchos organismos como los percebes y los mejillones se incrustan entre las rocas, excluyéndose entre ellos. La competencia entre múltiples especies evita que una sola domine sobre las demás. Actuando de manera paralela, la perturbación (de las olas o de los predadores) expone nuevos espacios libres de roca que facilita la reposición de especies que, a su vez, mantiene la diversidad permanente de la comunidad.

En el mar ártico, donde todas las especies son relativamente escasas, el índice global de interacción entre especies es también muy bajo y, por lo tanto, ni la exclusión competitiva ni la perturbación frecuente son esenciales para mantener la diversidad. Por el contrario, en estas comunidades el dominio de una especie determinada como los erizos de mar puede mantenerse bajo control a través de un depredador especializado muy poco frecuente pero muy eficaz, la estrella de mar. La ausencia de este único depredador reduciría drásticamente la diversidad global de la comunidad. En otras comunidades como, por ejemplo, el lecho marino, procesos de retroalimentación positiva pueden incrementar la diversidad de los organismos: los gusanos perforan las formaciones de coral, por ejemplo, creando nuevos hábitats que pueden ser ocupados por otros organismos.

Dayton y otros ecólogos construyeron este conjunto de conocimientos mediante el estudio intensivo de sistemas particulares: las esponjas en el estrecho de McMurdo en la Antártida (Dayton, 1972), los corales del monte marino Jasper en el Pacífico (Genin, Dayton, Lonsdale, & Spiess, 1986) y los bosques de algas de las costas centrales de California (Dayton, Currie, Gerrodette, Keller, Rosenthal, & Tresca, 1984). Los resultados de estos estudios son detalladas descripciones de posibles mecanismos y cuándo y por qué pueden ocurrir. Algunos de ellos, como Simberloff (2004), resultaron de gran ayuda a la comprensión de los ecosistemas locales por parte de los ecólogos.

Atención a las dinámicas a través de distintas escalas

El cambio de enfoque hacia posibles mecanismos y cuándo podrían darse centra la atención en el alcance y la escala de los efectos.

Por ejemplo, una de las principales contribuciones en ecología ha sido precisar la dependencia de la escala en la relación entre las diversas especies y la productividad del ecosistema; el volumen de biomasa producido por unidad de superficie (Purvis & Hector, 2000). En una escala global, la productividad está relacionada positivamente con la diversidad: una selva tropical muy densa produce mucha más biomasa por unidad de superficie que un bosque de un clima templado. Y, sin embargo, en la escala de la comunidad individual, la diversidad puede hacer disminuir la productividad (Fraser et al., 2015; Gaston, 2000). Un número elevado de especies crea más competencia, tanto por el espacio como a través de la producción de partículas químicas inhibitorias y otras defensas. Las plantas tienen que desviar parte de sus recursos a esta competencia, en lugar de invertirlos en su propio crecimiento, lo que produce un decremento global de la producción. El patrón se invierte de nuevo a nivel de las especies individuales. Si los individuos son demasiado similares, la producción puede decrecer cuando la población se ve atacada por predadores herbívoros o por una enfermedad. A este nivel, la diversidad genética puede, de nuevo, incrementar la productividad global (Hughes, Inouye, Johnson, Underwood, & Vellend, 2008).

La escala ha sido un factor crítico para entender la supervivencia y la extinción de las especies. Una población observada en una escala reducida en el tiempo y en el espacio puede extinguirse, causando alarma entre los conservacionistas. No obstante, una población particular puede entenderse como parte de una red más

amplia de poblaciones vinculadas por la migración; una ‘meta población’ (Hanski, 1998). Desde este punto de vista, la permanencia de las especies puede seguir siendo posible siempre que la población extinta sea sustituida a través del nacimiento o el crecimiento de una población en otro espacio. Si los esfuerzos de conservación se centran en salvar poblaciones individuales en lugar de manejar meta poblaciones, los recursos pueden ser adjudicados en la escala equivocada sin que surjan apenas efecto.

Aceptación del estudio de la variación y la complejidad

Dada la gran complejidad de los ecosistemas, los ecólogos han tenido que formular hipótesis simplificadas para poder estudiarlas. Estas hipótesis parecían necesarias para generar un conocimiento general que pudiese superar las variaciones inherentes. Sin embargo, a medida que este campo de estudio ha ido madurando, los investigadores han empezado a revisar la validez de muchas de estas hipótesis y a aceptar el estudio de la variación y la complejidad.

Un cambio reciente hace referencia a cómo gestionar la variación intra-especies. Los ecólogos saben que la variación individual es omnipresente y, por supuesto, esa variación es central en la selección natural y la evolución. Aun así, en cuestiones de dinámicas de población y de comunidad, durante décadas se ha considerado aceptable tratar a los individuos de una población como unidades relativamente intercambiables. Investigaciones recientes han cuestionado este supuesto (Bolnick et al., 2011, 2003; Hughes et al., 2008; Sih, Bell, & Johnson, 2004; Violle et al., 2012).

En el estudio del comportamiento animal, por ejemplo, los estudios clásicos informaban sobre el comportamiento promedio de una especie, obviando cualquier diferencia a nivel individual. Nuevas investigaciones centradas en medir y estudiar el impacto de las variaciones individuales han descubierto que estas variaciones pueden ser importantes para las poblaciones y las comunidades.

Por ejemplo, no todos los depredadores individuales en una especie particular se comportan exactamente de la misma manera ante una presa. Algunos son sistemáticamente más ‘atrevidos’, atacando sin vacilar, mientras que otros se muestran más ‘tímidos’, con ataques menos frecuentes y vigorosos (Sih et al., 2004). La composición específica de tipos individuales, no solo el número de depredadores, puede tener un impacto en las dinámicas de masa y de energía del sistema más amplio. Los investigadores documentaron este efecto de manera experimental en comunidades de arácnidos (Pruitt, Bolnick, Sih, Dirienzo, & Pinter-Wollman, 2016). Primero midieron los comportamientos de arañas individuales de varias especies y, después, combinaron distintos individuos para crear comunidades con la misma composición de especies pero con mayor o menor variación en comportamiento de los individuos. Observaron que las comunidades en las que la variación individual era mayor ganaban más masa y mostraban menor tendencia a disolverse con el tiempo. Si hubiesen tratado a todos los individuos como si estos mostrasen el mismo comportamiento promedio, se

hubiese impedido a los ecólogos comprender por qué algunas comunidades perduran y otras se desploman.

Un segundo ejemplo son los enfoques matemáticos en el modelado de los tamaños de población que, en algunos casos, puede fluctuar drásticamente de un año a otro. Los ecólogos teóricos se preguntaban si podían hacerse modelos de estas fluctuaciones utilizando la teoría del caos (Hastings, Hom, Ellner, Turchin, & Godfray, 1993; May, 1974, 1976) que, como ya hemos apuntado, puede explicar con rigor matemático el efecto amplificador de pequeñas diferencias para producir resultados radicalmente distintos. Los ecólogos esperaban que, si eran capaces de identificar las marcas del caos en las poblaciones naturales, podrían utilizar las herramientas matemáticas de esta teoría para comprender las dinámicas de los sistemas (Hastings et al., 1993; May, 1974, 1976). Pero estas marcas han resultado difíciles de identificar, excepto en experimentos de laboratorio cuidadosamente controlados o en sistemas naturales muy sencillos (Beninca, Ballantine, Ellner, & Huisman, 2015; Bjørnstad, 2015; Constantino, Cushing, Dennis, & Desharnais, 1995). Incluso algunos de los ecólogos más optimistas han expresado desde entonces que las dinámicas previsibles, aunque sean caóticas, tienden a ser muy poco habituales en la mayoría de los ecosistemas (Hastings, 2001).

Por eso los problemas de cómo abordar la variabilidad y la complejidad siguen siendo temas recurrentes en la investigación: es propio de la investigación ecológica mostrarse interesado en el estudio de estas dinámicas; ya no es aceptable simplemente utilizar un promedio.

De vuelta a la investigación sobre educación

Comenzamos este artículo con *Las ruedas de Isaac* y utilizamos el episodio para plantear la pregunta: ‘¿Cómo pueden contribuir a una investigación sobre el aprendizaje con rigor y validez perdurable los distintos episodios de estudiantes haciendo ciencia, con todas sus particularidades y complejidades?’ Después, pasamos a exponer la metodología utilizada en nuestro proyecto en curso, *Las dinámicas del involucramiento de los estudiantes y su persistencia en la ciencia*, utilizado para ilustrar las posibilidades y los desafíos de la investigación sobre la educación.

A partir de nuestra breve visita virtual a los ecólogos, observamos sus esfuerzos y discusiones sobre cómo gestionar las idiosincrasias de los fenómenos, discusiones que incluyen decidir si ‘seguir dedicando tanto tiempo y esfuerzo a los estudios tradicionales sobre la ecología de las comunidades’ (Lawton, 1999; citado en Simberloff, 2004; p. 787). Las semejanzas entre sus desafíos y los nuestros son especialmente llamativas dadas las ventajas relativas de la investigación sobre ecología frente a la investigación sobre educación. Muchos de los resultados que los ecólogos están interesados en comprender — poblaciones, productividad de biomasa, biodiversidad, etc. — son mucho más fáciles de reconocer y cuantificar que, pongamos por caso, la comprensión conceptual, la motivación o el involucramiento en prácticas disciplinarias.

Pero estas diferencias entre ambas disciplinas hacen la comparación particularmente útil. Por un lado, dada la relativa facilidad de cuantificar en ecología, y el escaso valor de la teoría del caos para su trabajo, nos sentimos menos inclinados a adoptarla en nuestro trabajo. Más globalmente, la tangibilidad de los objetivos del estudio de la ecología la hacen particularmente útil como base analógica para reflexionar sobre educación. Las cuestiones que hemos revisado sobre el progreso epistemológico y metodológico en ecología — abandonando las leyes universales para enfatizar los posibles mecanismos de nivel intermedio, estudiando distintas escalas y adoptando la variación — parecen todas relevantes para nuestras preguntas.

En esta sección final del artículo, reflexionamos de nuevo sobre nuestro proyecto como ejemplo ilustrativo de investigación educativa, regresando a las dudas persistentes que mencionábamos anteriormente. ¿Cómo delimitamos los datos de estudio? ¿Cómo identificamos fenómenos y entidades en los análisis de datos? ¿Qué esperamos lograr con nuestros resultados, más allá de los episodios particulares?

Delimitación de los datos

Hemos descrito nuestros enfoques y dificultades para delimitar los datos objeto de análisis, así como para seleccionar, dentro de esos límites, los detalles a considerar. En la ecología existen cuestiones similares, relacionadas con las escalas de los fenómenos y los sistemas.

En líneas generales, los esfuerzos de los ecólogos por estudiar la estabilidad de una comunidad ecológica, por ejemplo, parecen análogos a nuestros esfuerzos por estudiar las dinámicas del involucramiento de los estudiantes durante una clase. Las decisiones sobre la escala del sistema, sus elementos y el intervalo de tiempo tienen implicaciones en la identificación de los fenómenos y las entidades. Los ecólogos tienen que decidir en qué escala temporal tienen sentido las mediciones de estabilidad; tienen que decidir si incluyen las interacciones con comunidades próximas. Podría ayudarnos observar más allá de los 20 minutos que hemos seleccionado para el estudio para observar la influencia de una idea surgida en una discusión previa o buscar el origen y el transcurso de las ideas en otros contextos.

A grandes rasgos, la selección de los detalles objeto de estudio por parte de los ecólogos, tales como el seguimiento de arañas individuales en el tiempo, se asemeja a nuestra decisión de los elementos en los que centrar la atención en un episodio. Aquí también las decisiones iniciales tienen implicaciones: las ‘personalidades’ individuales de las arañas solo podían emerger en estudios que prestasen atención a esos detalles.

Como ya hemos comentado, comenzamos los análisis de los estudios de caso tomando decisiones sobre la selección de los datos objeto de estudio, pero la investigación del ámbito ecológico nos incita a replantearnos este enfoque. Tal vez deberíamos incluir iteraciones en nuestra manera de delimitar los datos de un modo similar a la iteración en el refinamiento y aplicación de sistemas de codificación (e.g.,

Chi, 1997; Glaser & Strauss, 1967). Hasta cierto punto, examinamos y reexaminamos los datos dentro de los límites establecidos, a medida que los miembros del grupo observan y debaten la relevancia de ciertos aspectos que el investigador líder no había considerado en la fase inicial.

Sin embargo, no se ajustaron los límites de los episodios en iteración con los análisis. Sospechamos, por ejemplo, que una discusión anterior sobre ‘el libre albedrío’ podría haber influido en el cuestionamiento de Isaac por parte de Scarlett, ‘¿Cómo pueden cansarse las ruedas?’ Tal vez hubiese tenido sentido ampliar la delimitación de los datos para incluir esa discusión. Por supuesto, esto aumentaría significativamente el trabajo de análisis del caso que, a primera vista, parecía de fácil delimitación, pero el esfuerzo adicional podría producir resultados más sólidos. En otros episodios, no hay una opción inmediatamente evidente, y nos ha resultado difícil establecer los límites; la iteración podría limitar las repercusiones de esas decisiones iniciales.

Identificación de fenómenos y entidades

Para la identificación de fenómenos y entidades nos enfrentamos a dos dificultades relacionadas: una es la disección de dinámicas complejas en cuanto a la envergadura de los datos en el análisis de un episodio particular y, la otra, determinar qué aspectos de ese episodio pueden tener relevancia general.

La primera dificultad está relacionada con las escalas, tanto temporales como sistémicas, de lo que entendemos por fenómenos en un episodio concreto y está estrechamente relacionada con la cuestión de cómo delimitar los datos. En uno de los ejemplos anteriores, la ampliación del alcance de los datos brindó apoyo al concepto de metapoblación (Hanski, 1998) que venían desarrollando los ecólogos. Similarmente, nuestra ampliación en la envergadura de los datos podría apoyar nuestra consideración de la posibilidad de que la pregunta de Scarlett formase parte de un fenómeno más amplio sobre su resistencia al pensamiento antropomórfico, derivado de un debate anterior sobre el ‘libre albedrío’. Y, por tanto, podríamos considerar si atribuir esa preocupación a Scarlett como individuo o a la clase, dado que la clase había creado un significado y una norma compartidos de explicaciones sobre mecanismos. Tal vez ambas escalas son relevantes, análogas a la identificación, por parte de los ecólogos, de la relación reflexiva entre las personalidades individuales de las arañas y la resiliencia de las comunidades de arañas (Pruitt et al., 2016).

La comparación de nuestro trabajo con el de los ecólogos ha puesto de relieve estas cuestiones relacionadas con la creación de modelos. Ha renovado nuestro interés en las interpretaciones de las distintas escalas de dinámicas basadas en los datos (Conlin, Gupta, & Hammer, 2010; Conlin & Hammer, 2016). En líneas generales, corrobora el llamamiento de las ciencias del aprendizaje en favor de una investigación que contemple distintas escalas temporales y sistémicas (e.g., Jacobson et al., 2016). Cobb y sus colegas han desarrollado una metodología para analizar las relaciones reflexivas entre niveles individuales y sociales en la actividad y la práctica matemáticas (Cobb, Stephan, McClain, & Gravemeijer,

2001; McClain & Cobb, 2001; Yackel & Cobb, 1996). Saxe y sus colegas han propuesto un enfoque al estudio del desarrollo de ideas matemáticas que coordina interacciones ‘momento a momento’ con patrones de cambios en el tiempo tanto en los individuos como en las comunidades (Saxe & Esmonde, 2005; Saxe et al., 2009).

Philip y sus colegas han estudiado las intersecciones entre dinámicas individuales y ‘momento a momento’ del aprendizaje con estructuras sistémicas de poder y opresión (Philip, 2011; Philip, Gupta, Elby, & Turpen, 2018). Su trabajo nos inspira a estudiar cómo las interacciones entre dinámicas sistémicas más amplias de poder en torno a la raza, las clases sociales y el género podrían influir en las dinámicas locales de la *problematización* de los estudiantes y su *posicionamiento*.

La segunda faceta de la dificultad de identificar fenómenos y entidades tiene que ver con la relevancia más allá de los episodios particulares, y está estrechamente relacionada con la cuestión de cómo evaluar el valor y la validez de los resultados. El objetivo epistémico de la validez general podría informar en cómo decidimos enfocar nuestros análisis, para considerar si es plausible que un fenómeno que observamos en un caso podría aparecer en otros. Si esperamos que, en última instancia, nuestra decisión sobre la validez implique superar una prueba según un criterio de referencia, como el RCT, deberíamos elegir objetos de estudio que permitan realizar ese tipo de pruebas.

En cambio, hemos diseñado nuestro enfoque específicamente para centrarnos en los datos disponibles, caso por caso, y para evitar las comparaciones entre los distintos casos hasta que tuviésemos un conjunto de casos completos. Se trata de un intento de aplicar rigor cualitativo, de resistir la observación de los datos de ciertas maneras que ya anticipamos. La comparación de nuestro trabajo con el de los ecólogos nos anima a continuar en esta línea, a estudiar episodios concretos por lo que cada uno de ellos muestra, aplazando la cuestión de la generalización.

Los ecólogos han reconocido la importancia de las idiosincrasias, como la llegada de un depredador individual poco frecuente, una migración lejana poco probable o un individuo anormalmente agresivo en una comunidad. Parte de su productividad académica está relacionada con la abstracción a partir de particularidades: la importancia de la variación entre los individuos de una comunidad inspira a contemplar esas mismas posibilidades en otras. En nuestro proyecto, el fenómeno de la *problematización* y del *posicionamiento de incomprensión* aparecieron con mayor claridad en casos individuales, para después surgir como patrones de estudio cuando realizamos comparaciones entre distintos casos.

Evaluación de la validez de los resultados

Los ecólogos han alcanzado un consenso general sobre la ausencia de leyes generales, en particular en las ‘escalas intermedias’ de la ecología de las comunidades, en las que ‘la contingencia es abrumadoramente complicada’ (Lawton, 1999, p. 177). En lugar de abandonar la investigación en estas escalas, la mayoría de los ecólogos decidieron, como Simberloff (2004), desistir de las leyes como

objetivos epistémicos de la investigación. En cambio, tratan de encontrar y producir resultados seminales en forma de constructos de nivel intermedio de *mecanismos posibles* que ocurren en algunos sistemas. No se trata de resultados en forma de leyes fiables, pero pueden tener una validez perdurable como verdades contingentes que facilitan la comprensión de ciertos sistemas particulares. El conocimiento de posibles dinámicas ayuda a los investigadores o gestores (de piscifactorías o bosques) en la construcción de modelos locales y particulares.

Así entendemos la posible validez general de nuestros resultados, como posibles aspectos de los estudiantes haciendo ciencia. No esperamos que la *problematización* o el *posicionamiento de incomprensión* sea evidente en todos los episodios, pero estamos seguros de que ambos aspectos ocurren en algunos de ellos. Esperamos que nuestros resultados ayuden a otros investigadores o educadores a comprender y fomentar lo que observan en sus estudiantes.

En este artículo hay margen para incluir distintos tipos de poder predictivo. En el sentido tradicional, los investigadores podrían diseñar estudios explícitamente para fomentar la problematización y el posicionamiento de incomprensión, y comprobar las predicciones de que estas condiciones producen con mayor frecuencia momentos de estudiantes haciendo ciencia.

Pero, reflexionando sobre el giro de la ecología hacia la búsqueda de mecanismos posibles y de adopción de la variación y la complejidad, nos hace pensar en el poder predictivo a escala local: ¿Podríamos evaluar los resultados como útiles si ayudan a los investigadores o educadores a comprender mejor a sus estudiantes, con los aspectos particulares de sus situaciones? ¿Podríamos preguntar qué utilidad tienen los resultados en la construcción de modelos locales y particulares que generen ideas predictivas sobre cómo contribuir al aprendizaje de estudiantes particulares?

Al mismo tiempo, nuestra revisión de la investigación en ecología ha incrementado tanto nuestro escepticismo respecto a criterios fijos para evaluar la validez de los resultados como nuestra paciencia para evaluar su aplicabilidad general. Podemos hacer ciertas afirmaciones con gran confianza sobre casos particulares y, por supuesto, podemos tener en cuenta el posible valor de estas afirmaciones para la literatura cuando pensamos en su divulgación, pero quizá no sabremos hasta mucho más tarde si tendrán ese valor. El modo en que evaluamos su validez tiene que variar en función de la naturaleza de las afirmaciones. Como Cartwright declaró (2007) cuando cuestionó el estatus de los RCT, no existe un 'estándar de oro' para evaluar la validez.

Búsqueda permanente

Obviamente, no hemos llegado a una respuesta clara y definitiva a nuestra pregunta. La referencia al ámbito de la ecología nos ha mostrado una vez más que la ciencia en sí misma evoluciona y se adapta; en sus objetivos, valores, virtudes y procesos (Chinn et al., 2011). Lo hemos observado como parte del desafío planteado en nuestro proyecto, de disponer de una definición clara de la

ciencia para guiarnos en nuestros esfuerzos por identificarla en nuestros estudiantes. Lo mismo debe ocurrir en la investigación sobre el aprendizaje.

Nuestro recorrido por la ecología ha corroborado el valor de los estudios de caso sobre el aprendizaje, incluida la búsqueda de patrones, pero también simplemente para reconocer el uso posible de un ejemplo individual. Estamos experimentando con la iteración como enfoque para delimitar los datos y para identificar fenómenos y entidades. Nos ha ayudado a reconocer otros modos en que los resultados pueden tener validez duradera. Por ejemplo, del mismo modo que el trabajo de Dayton en la identificación de posibles mecanismos para mantener la diversidad de las especies en ecosistemas marinos tiene validez para los conservacionistas, identificar diversas dinámicas de clase posibles para fomentar o inhibir a los estudiantes hacer ciencia podría contribuir a las comprensiones y las decisiones pedagógicas de los educadores. Nuestra revisión de la ecología también nos ha alentado a diseñar investigaciones posteriores para analizar cuándo y en qué condiciones ocurren estas posibles dinámicas.

Pero, por encima de todo, nos anima a seguir participando en esta conversación entre disciplinas. Tenemos proyectado reclutar ecólogos en nuestro grupo de estudio multidisciplinar para leer y discutir sobre metodología. Parte de esa lectura incluirá una revisión de literatura sobre educación, como los trabajos de Cobb y Saxe y sus colegas, citados anteriormente, la investigación basada en el diseño (Barab & Squire, 2004; Brown, 1992) y muchos otros trabajos de académicos que están estudiando temas similares de escala y *situacionalidad* (Bang, 2015; Maxwell, 2004; Sandoval, 2013; Tabak, 2004). Somos conscientes de que sus trabajos incluyen muchos de los aspectos que hemos visto durante nuestra revisión de la investigación en el ámbito de la ecología; será un verdadero placer volver a leer sus trabajos con interés renovado.

Notas

1. La clase formaba parte de un proyecto de investigación sobre la indagación científica centrado en la 'enseñanza que responde a las ideas de los estudiantes' (Goldberg et al., 2008–2011). Tanto la grabación de video como la descripción del episodio *Las ruedas de Isaac* pueden consultarse en studentsdoingscience.tufts.edu. Véase también Hammer y Radoff (2014).
2. El vídeo ofrece una gran riqueza de datos, pero no es estrictamente necesario. El requisito es que los datos tengan la riqueza suficiente para transmitir al panel una idea de lo que los estudiantes hacían en términos de ciencia.
3. *Las ruedas de Isaac* era una de las dos excepciones en las que no identificamos la problematización como aspecto relevante. No obstante, esta resultaba evidente, en particular en el trabajo de Jourdan, Scarlett y Jamir, con la ayuda de la maestra Fargason, de su intento por identificar y convencer a Isaac de que había un vacío en su razonamiento sobre la desaceleración del coche. Este episodio también sugiere que Isaac había llevado a cabo cierta problematización por sí mismo con anterioridad, dada su necesidad de explicar cómo las ruedas permiten al coche moverse sin arrastrarse.

Acknowledgements / Agradecimientos

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. / *Agradecemos a Infancia y Aprendizaje y a Bárbara Brizuela que nos invitaran a presentar este artículo, así como sus valiosos comentarios. Nuestro inmenso agradecimiento también a la Gordon and Betty Moore Foundation y a Janet Coffey por su financiación del proyecto (ayudas GBMF3475 y 3475.01). Gracias también a Dan Levin por sus comentarios en un borrador de este artículo y, por supuesto, a todos los estudiantes y colegas que han trabajado con nosotros en este proyecto.*

Disclosure statement

No potential conflict of interest was reported by the authors./ Los autores no han referido ningún potencial conflicto de interés en relación con este artículo.

References / Referencias

- Bang, M. (2015). Culture, learning, and development and the natural world: The influences of situative perspectives. *Educational Psychologist, 50*, 220–233.
- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*, 1–14.
- Bauer, N. (2015). *How to do things with pornography*. Cambridge, MA: 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*, 6252–6253.
- Blanchette, I., & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory & Cognition, 28*, 108–124.
- Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., ... Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology and Evolution, 26*, 183–192.
- Bolnick, D. I., Svanbäck, R., Fordyce, J. A., Yang, L. H., Davis, J. M., Hulseley, C. D., & Forister, M. L. (2003). The ecology of individuals: Incidence and implications of individual specialization. *The American Naturalist, 161*, 1–28.
- 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*, 513–531.
- Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 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*, 499–531.
- 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*, 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*, 141–167.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*, 175–218.
- Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences, 10*, 113–163.
- Conlin, L. D., Gupta, A., & Hammer, D. (2010). 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. (2015). 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, NY; 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. (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.
- 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. K., Currie, V., Gerrodette, T., Keller, B. D., Rosenthal, R., & Tresca, D. V. (1984). Patch Dynamics and Stability of Some California Kelp Communities. *Ecological Monographs, 54*, 253–289.
- diSessa, A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior, 10*, 117–160.
- diSessa, A. A. (1993). Towards an epistemology of physics. *Cognition and Instruction, 10*, 105–225.
- Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38*, 39–72.
- 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*, 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*, 215–231.
- 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.
- Giere, R. N. (1988). *Explaining science: A cognitive approach*. Chicago, IL: 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, 141–166.
- 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, 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., & 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. Retrieved from 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, 424–431.
- 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.
- Hansson, L. (2003). Why ecology fails at application: Should we consider variability more than regularity? *Oikos*, 100, 624–627.
- Hardy, A. (1968). Foreword: Charles Elton's influence in ecology. *Journal of Animal Ecology*, 37, 3–8.
- Harré, R., & van Langenhove, L. (1999). *Positioning theory: Moral contexts of intentional action*. Oxford; Malden, MA: Blackwell.
- Hastings, A. (2001). Transient dynamics and persistence of ecological systems. *Ecology Letters*, 4, 215–220.
- 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, 1–33.
- Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, 11, 609–623.
- 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, 210–218.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4, 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, 512–559.
- Lange, M. (2005). Ecological laws: What would they be and why would they matter? *Oikos*, 110, 394–403.
- Lawton, J. H. (1999). Are there general laws in ecology? *Oikos*, 84, 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, 142–154.

- Levy, S. T., & Wilensky, U. (2008). Inventing a “Mid level” to make ends meet: Reasoning between the levels of complexity. *Cognition and Instruction, 26*, 1–47.
- Lineback, J. E. (2015). The redirection: An indicator of how teachers respond to student thinking. *Journal of the Learning Sciences, 24*, 419–460.
- Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist, 39*, 57–68.
- Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3–11.
- 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*, 236.
- McCloskey, M. (1983). Naive theories of motion. In D. Gentner, & A. L. Stevens (Eds.), *Mental models* (pp. 299–324). Hillsdale, NJ: Erlbaum.
- McDermott, R., Gospodinoff, K., & Aron, J. (1978). Criteria for an ethnographically adequate description 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*, 50–110.
- Murray, B. (1992). Research methods in physics and biology. *Oikos, 64*, 594–596.
- 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*, 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*, 318–335.
- O’Hara, R. (2005). The anarchist’s guide to ecological theory. Or, we don’t need no stinkin’ laws. *Oikos, 110*, 390–393.
- 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*, 297–329.
- 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*, 183–223.
- 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*. New York, NY: 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.
- 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, NY: 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*, 499–525.
- Sandoval, W. (2013). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*, 18–36.
- 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*, 171–225.
- 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 (Eds.), *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*, 129–184.
- Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: An ecological and evolutionary overview. *Trends in Ecology and Evolution*, *19*, 372–378.
- Simberloff, D. (2004). Community ecology: Is it time to move on? (An American society of naturalists presidential address). *The American Naturalist*, *163*, 787–799.
- Southerland, S. A., Johnston, A., & Sowell, S. (2006). Describing teachers' conceptual ecologies for the nature of science. *Science Education*, *90*, 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*, 797–811.
- 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, NY: State University of New York Press.
- Svoboda, J., & Passmore, C. (2013). The strategies of modeling in biology education. *Science & Education*, *22*, 119–142.
- Tabak, I. (2004). Reconstructing context: Negotiating the tension between exogenous and endogenous educational design. *Educational Psychologist*, *39*, 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*, 29–47.
- 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, MA: 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*, 17–26.

- 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, 244–252.
- 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.
- 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, 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, NY: 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, 941–967.
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27, 458.