5

Doubly Counterintuitive: Cognitive Obstacles to the Discovery and the Learning of Scientific Ideas and Why They Often Differ

Andrew Shtulman

Collectively, humans know more about how the world works than ever before. This knowledge is the hard-won achievement of innumerable scientists across innumerable years. Their labors include designing instruments of measurement, devising experimental protocols, recording and disseminating data, constructing theoretical accounts of those data, debating the merits of different theoretical accounts, and unifying insights across disparate paradigms or fields. The goal of these activities was to create ever-more accurate and ever-more coherent models of reality—models used to inform technological innovation and edify future generations.

Despite these collective advances in human knowledge, most individual humans know very little about science. Organizations like Gallup, the Pew Research Center, and the National Science Foundation have been polling the general public on their understanding of science for decades and have documented consistently low levels of scientific literacy. For instance, a research survey by the Pew Research Center (2015) found that only 65 percent of Americans believe that humans have evolved over time, compared with 98 percent of the members of the American Association for the Advancement of Science (AAAS); only 50 percent of Americans believe that climate change is due mostly to human activity, compared with 87 percent of AAAS members; and only 37 percent of Americans believe that genetically modified foods are safe to eat, compared with 88 percent of AAAS members.

Poor science education is one reason the average person knows little science, but it is not the only reason. Decades of research on science education have revealed that individuals exposed to extensive and comprehensive science instruction often fail to learn from it (e.g., Gregg et al., 2001; Kim and Pak, 2002; Libarkin and Anderson, 2005; Shtulman and Calabi, 2013). Students enter the science classroom with naïve, nonscientific ideas about how the world works, and they leave the classroom with those same ideas intact. Instruction is ineffective because science is deeply counterintuitive. Science defies our earliest and most accessible intuitions about how the world works, and those intuitions impede our ability to acquire more accurate models of the world. Learning science is of course possible, but the process is difficult and protracted (Carey, 2009; Shtulman, 2017; Vosniadou, 1994a).

This tension, between the advancement of science as a whole and the learning of science by individuals, has implications for the study of scientific knowledge. Scholars interested in the origin and character of scientific ideas are likely to learn different lessons from the professional activities of scientists than from the cognitive activities of science students. Our generalizations about scientific knowledge will differ depending on whom we take as the custodians of that knowledge and whose struggles we view as most informative. Students' knowledge cannot be written off as a corrupted or degraded version of scientists' knowledge, because the two forms of knowledge may embody different relations among scientific concepts or different relations between scientific concepts and empirical observations.

Here, I explore a question common both to the scientist's struggle to model reality and the student's: why are some scientific ideas particularly difficult to grasp? Atoms, germs, heat, inertia, heliocentrism, natural selection, continental drift: these ideas were slow to develop in the history of science and remain slow to develop in the minds of individuals, but the reasons for the historical delay are not necessarily the same as the reasons for the cognitive delay. Scientists and students have different explanatory goals, different empirical concerns, and different background assumptions, and I aim to show how these factors can render the same idea counterintuitive for different reasons. This comparison of scientists' and students' conceptual ecologies has implications not only for theories of scientific knowledge but also for the practice of teaching science to nonscientists.

Two caveats should be noted. First, my focus is on the content of scientific claims rather than the process of testing those claims. Much of the psychological research on scientific thought examines inquiry skills, like the ability to design informative experiments (Kuhn and Pease, 2008; Lorch et al., 2010), the ability to evaluate empirical data (Chinn and Brewer, 1998; Morris and Masnick, 2014), or the ability to coordinate data and theory (Bonawitz et al., 2012; Schauble, 1996). Scientists may engage in more sophisticated inquiry activities than students, but the focus of this chapter will be on the products of those activities—the concepts and theories informed by inquiry—and how those products are understood.

Second, my focus is on what makes scientific *ideas* counterintuitive for scientists and students, rather than how those ideas are constructed. The process of constructing scientific ideas often entails conceptual change, or knowledge restructuring at the level of individual concepts (Carey, 2009; Chi, 2005; Nersessian, 1989), and there is debate as to how closely conceptual change in the student mirrors conceptual change in the scientific community (see, for example, DiSessa, 2008; Kuhn, 1989). Differences in the process of conceptual change may yield differences in the outcome of that process, but the focus here will be on the latter. That is, I will focus on what makes scientific ideas difficult to grasp rather than on how we come to grasp them.

A common starting point: Intuitive theories

Humans are built to perceive the environment in ways that enhance survival, which do not always align with the categories of science (Carey and Spelke, 1996; Chi et al., 2012). These misalignments can take the form of omissions or commissions. The omissions are when we fail to perceive the entities or processes causally responsible for some phenomenon, whereas the commissions are when we mistakenly assume that a phenomenon is caused by entities or processes we can perceive.

Errors of omission are particularly common when reasoning about biology. Biological systems usually operate at too small a scale for us to observe firsthand. We cannot observe the functional relations among internal organs or the genetic underpinnings of heritable traits, so we gravitate toward generic explanations of metabolism and inheritance, such as *vitalism*, or the belief that organisms possess an internal life-force that maintains growth and health (Inagaki and Hatano, 2004; Morris, Taplin, and Gelman, 2000); and *essentialism*, or the belief that an organism's external properties are determined by an internal essence inherited at birth (Gelman, 2003; Johnson and Solomon, 1997).

Errors of commission, on the other hand, may be more common when reasoning about physics. Physical systems are multifaceted, and we observe only the facets that impact our interaction with the system. For instance, we perceive material objects in terms of their heft (felt weight) and bulk (visible size), not their actual weight and size (Smith, Carey, and Wiser, 1985; Smith, 2007), and we perceive gravity as pulling us down, not pulling us toward the center of the Earth (Blown and Bryce, 2013; Vosniadou and Brewer, 1994).

These perceptual biases lead humans down the wrong path when it comes to theorizing about the causes of natural phenomena, pushing us to draw distinctions that are not particularly meaningful from a scientific point of view (e.g., a distinction between motion and rest) and to overlook distinctions that *are* meaningful (e.g., a distinction between weight and density). What's more, they lead everyone down the wrong path, students and scientists alike. Students' preinstructional beliefs in many domains resemble the first theories to emerge in the history of science. For example, students' preinstructional beliefs about motion resemble the "impetus theory" of the Middle Ages more closely than Newtonian mechanics (McCloskey, 1983). Their beliefs about inheritance resemble Lamarck's theory of acquired characters more closely than a genetic theory (Springer and Keil, 1989). And their beliefs about astronomy resemble pre-Copernican models of the solar system more closely than post-Copernican ones (Vosniadou and Brewer, 1994).

The beliefs of today's students resemble those of yesterday's scientists in both form and function. Physics students, for instance, make essentially the same predictions about motion that Medieval physicists made, and they provide essentially the same explanations (Eckstein and Kozhevnikov, 1997; McCloskey, 1983). Across tasks and contexts, their beliefs about motion are generally as coherent as Medieval physicists', which is one reason psychologists terms those beliefs *theories*. Another reason is that the beliefs facilitate the same cognitive activities as scientific theories: explaining past events, predicting future events, intervening on present events, and reasoning about counterfactual events (Gelman and Legare, 2011; Gopnik and Wellman, 2012; Shtulman, 2017).

Intuitive theories are the starting point for how humans represent and understand the natural world, emerging early in life and in similar forms across cultures (Shtulman, 2017). They shape the foundations of everyday reasoning, including the foundations of scientific inquiry. However, the historical pathway from intuitive theories to scientific theories is often quite different from the developmental pathway. Scientists revise their theories through iterative cycles of data collection and data interpretation, whereas nonscientists typically maintain the same intuitive theory until confronted with a scientific alternative. Students are the beneficiaries of a vast effort to vet empirical ideas without doing any of the vetting. But the vetting, I will argue, may change scientists' understanding of the role and value of the vetted product.

Divergent paths: Discovering vs. learning scientific truths

Scientific truths can be difficult to discover for different reasons than they are difficult to learn. Here, I will sketch some ways in which the motivations and assumptions of the discoverers (scientists) differ from those of the learners (students) and how those differences can shape the cognitive obstacles to embracing a scientific truth. The contrast I draw between discoverers and learners is not meant to be holistic, in the sense that some people are discoverers and others are learners, but rather concept-specific. Discoverers are those who first formulate a scientific concept, and learners are those who are introduced to the concept secondhand.

Divergent explanatory goals

When scientists construct a new theory, their primary goal is to account for an existing body of data, but other goals are pursued as well. Scientists try to maximize explanatory scope, minimize auxiliary assumptions, generate new hypotheses, avoid internal inconsistencies, and be consistent with established theories in related domains (Laudan et al., 1986; Thagard, 1978). Nonscientists care about these additional considerations when they are directly asked to compare two theories (Koslowski et al., 2008; Samarapungavan, 1992), but there is little evidence that these considerations inform the construction of intuitive theories. Intuitive theories are a response to everyday phenomena—motion, heat, weather, illness, growth—and are constructed primarily to account for how we perceive those phenomena. Higher-order considerations like generativity, parsimony, and breadth may implicitly guide the construction of intuitive theories but do not seem to be engaged explicitly (DiSessa, 1993; Kuhn, 1989).

Consider the domain of matter. The idea that objects are composed of microscopic particles—atoms—was debated within the scientific community for hundreds of years (Toulmin and Goodfield, 1962) and is not fully embraced by children until the second decade of life (Smith, 2007). Early chemists agreed that material objects were composed of more fundamental elements, but they disagreed about what those elements were, what shape they took, how they interacted, whether they could be transmuted from one type to another, and whether they were the constituents of all entities or only inorganic entities. Early chemists also disagreed about the relations between matter and space, matter and motion, matter and sensation, and matter and mind. The disagreements were moral as well as empirical. Early skeptics of atomism saw bleak implications in the claim that reality consists of nothing more than atoms and void, fearing it implied a lack of purpose and design.

The goals of early chemists thus extended beyond the domain of matter into the domains of biology, psychology, and ethics. The goals of nonchemists, on the other hand, are much simpler: to account for everyday material transformations like sinking and floating, shrinking and expanding, freezing and burning. These transformations occur at a macroscopic level, but they are constrained by processes operating at a microscopic level, and nonchemists have difficulty relating the two levels, preferring to interpret material transformations as directed processes rather than emergent phenomena (Chi et al., 2012). Thus, while early chemists were reluctant to embrace atomic theory for metaphysical reasons, nonchemists are reluctant to do simply because they cannot see microscopic particles or fathom how such particles could give rise to outcomes they can see (Chi, 2005).

Another example of divergent explanatory goals can be seen in the domain of motion. A guiding principle for sixteenth- and seventeenth-century physicists, like Kepler and Newton, was to account for terrestrial motion (e.g., an apple falling from a tree) and celestial motion (e.g., a moon orbiting a planet) with the same laws (Horton, 1988). Beginning with the Greeks, terrestrial motion and celestial motion were treated as separate phenomena and explained by separate principles. Even terrestrial motion was subdivided into separate phenomena—flinging vs. falling vs. spinning—and explained by separate principles. Physicists like Kepler and Newton sought to unify all motion with a single mechanics.

Nonphysicists, on the other hand, are concerned not with explaining motion in general but with explaining particular instances of motion: the trajectory of a ball off a bat, the speed of a sled down a hill, the arc of water in a drinking fountain (DiSessa, 1993). The nonphysicist wants to know where these things are going and how quickly. Little effort is spent comparing one instance of motion to another, leading to discrepant predictions. A ball rolled off a cliff is expected to move forward as it falls, but a ball dropped from a plane is expected to fall straight down. Both have horizontal velocity and thus both would move forward as they fell. However, nonphysicists attribute a force to the first ball—the "force of motion"—but attribute no force to second (Kaiser, Proffitt, and McCloskey, 1985). Nonphysicists are hard-pressed to see the similarity between a ball that is dropped and a ball that is rolled, whereas many early physicists saw the similarity but were hard-pressed to explain it.

Divergent empirical concerns

Just as scientists approach theory construction with a wider range of goals than nonscientists, they are also aware of a wider range of phenomena to be covered by those theories, including anomalies. Anomalies are observations that cannot be explained by a field's prevailing theory, such as the observation that Uranus does not follow a perfectly elliptical path (leading to the discovery of Neptune) or the observation that some metals gain weight when burned (leading to the discovery of oxygen). They play an important role in scientific innovation (Kuhn, 1962), but nonscientists know nothing of them. Nonscientists struggle to account for everyday observations, whereas scientists struggle to account for both everyday observations and anomalies—the latter typically discovered through careful, systematic observation.

In the domain of heat, an anomalous finding that spurred scientific innovation was Black's discovery that heat and temperature are dissociable at a phase change (Fox, 1971; Wiser and Carey, 1983). Black observed that adding heat to a mixture of ice and water did not raise its temperature but rather increased the proportion of water to ice. Only after all the ice had melted did the water's temperature rise. Black observed the same pattern for a mixture of boiling water and steam; adding heat to this mixture did not raise its temperature until all the water had turned to steam. Chemists before Black had assumed that thermometers measure heat, not temperature, and they had no explanation for how heat could be added to a physical system without a concurrent change in temperature. Black posited a new substance—caloric—to explain his findings. Caloric was believed to pool inside substances at their melting point or boiling point, changing the substance's chemical composition but not its temperature.

Caloric is a fiction; the correct explanation for why heating a substance does not increase its temperature at a phase change is that the added energy is spent breaking molecular bonds. This explanation requires thinking of heat as kinetic energy at the molecular level. Black did not think of heat this way, and neither do nonchemists today. But nonchemists are generally unaware of the thermal dynamics of phase change and are thus uncompelled to explain them (Wiser and Amin, 2001).

For nonchemists, the phenomena most in need of explanation are the physical sensations of warmth and cold. Warmth and cold are reified either as properties intrinsic to matter or as substances that flow in and out of matter (Erickson, 1979; Reiner et al., 2000). This view leads to a conflation of heat and temperature, as well as the misconception that heat and cold are distinct kinds of substances. This set of beliefs overlaps with Black's theory of caloric in some regards but not others. It leads non-chemists to construe heat as a kind of substance, as Black did, but it also leads them to conflate heat with temperature, which Black did not.

An additional example of the divergent concerns of scientists and students comes from astronomy. For centuries, the Earth was believed to be at the center of the universe, but this model of the universe could not easily account for early observations of planetary motion (Toulmin and Goodfield, 1961). Viewed from Earth, the planets appear to move backward for several weeks of their orbit. Early astronomers like Hipparchus and Ptolemy accounted for this anomaly by positing a complicated system of epicycles, or small circles traversed by each planet along their larger circle around the Earth.

Eventually, this convoluted, geocentric model was replaced with a simpler, heliocentric model, but the heliocentric model remains counterintuitive to nonastronomers. Only around 75 percent of Americans accept that the Earth revolves around the sun; the remaining 25 percent believe that the sun revolves around the Earth (National Science Board, 2014). Nonastronomers are reluctant to embrace heliocentrism not because they are committed to epicycles but because they perceive the sun as rising and setting, and they do not perceive the Earth as rotating or revolving (Harlow et al., 2011; Vosniadou and Brewer, 1994). Accounting for the sun's apparent motion is the chief concern of nonastronomers, who neither perceive nor consider the motions of other planets.

Divergent background assumptions

Scientific ideas are constrained by a host of background assumptions about how the world works in general. The assumptions that make an idea counterintuitive to a scientist, steeped in particular methodological and theoretical traditions, may be quite different from those that make the same idea counterintuitive to a nonscientist. The sticking points for scientists may not be non-issues for nonscientists and vice versa.

Consider the claim that the Earth's continents move. On first blush, nonscientists view this claim as absurd. They see the Earth as essentially an inert chunk of rock, solid and eternal (Libarkin et al., 2005; Marques and Thompson, 1997). Accepting that the continents move requires reconceptualizing the Earth itself, from a static object characterized by small, inconsequential changes (like eroding mountains and shifting coastlines) to a dynamic system characterized by large, continual change (like sinking landmasses and colliding plates).

Geologists, on the other hand, were initially resistant to this idea for different reasons. They did not view the Earth as static. They knew, by the time that Wegener proposed his theory of continental drift, that the Earth had begun its existence in a molten state, that its interior was still hotter and more fluid than its exterior, that its oceans were once vaster, and that its mountains were once flatter. What made geologists of the early twentieth century skeptical of Wegener's theory is that they were unable to reconcile the theory with its implications. They were willing to concede that the Earth's crust could crack or fold, but they were unwilling to concede that it could rearrange itself into new configurations, for they knew of no mechanism that would allow whole continents to move (Oreskes, 1999; Gould, 1992).

Tellingly, a key piece of evidence that convinced geologists that the continents move was the discovery of magnetic stripes on the seafloor. These stripes indicate that currents of molten rock deep within the Earth's interior have changed the magnetic properties of the Earth's crust, as that crust forms anew at the boundaries of tectonic plates (Oreskes, 1999). Nongeologists do not know of the existence of magnetic stripes, let alone appreciate their implications.

The domain of illness provides another example. Many, if not most, of the illnesses that plague humanity are caused by microbial infection. Microbes, or germs, cannot be perceived, nor can they be tracked in their transmission from one host to another, so their discovery took centuries. The perceptual obstacles to identifying germs were compounded by conceptual ones. How could something alive be too small to be seen? How could one living thing survive and reproduce inside another? Nonbiologists continue to be puzzled by such dilemmas. They now know about the existence of germs—even preschoolers know that germs make a person sick (Kalish, 1996) and that germs spread on contact with an infected individual (Blacker and LoBue, 2016)—but they do not conceive of illness as the biological consequence of a parasite hijacking the host's resources to further its own survival and reproduction (Au et al., 2008).

The historical discovery of germs was also hampered by considerations that nonbiologists do not entertain: that the body contains a supply of internal fluids, or "humors," whose balance is critical for health and vitality (Lederberg, 2000; Thagard, 1999). Beginning with Hippocrates, early physicians analyzed disease as the interplay between blood (the sanguine humor), phlegm (the phlegmatic humor), yellow bile (the choleric humor), and black bile (the melancholic humor). Too much blood was thought to cause headaches; too much phlegm, epilepsy; too much yellow bile, fevers; too much black bile, depression. The prescribed cures were to relieve the body of the excess humor by inducing vomiting, defecation, or bleeding.

This framework made the notion of microbial infection even more problematic. Early biologists were willing to accept that humors could become imbalanced by external factors—notably, bad air or "miasma"—but the true cause of illness was the imbalance, not the imbalancer. External factors were not construed as sources of contagion. Once again, a telling sign of the difference between scientists and nonscientists' acceptance of the correct theory comes from a discovery that only scientists would find convincing: the discovery that fermentation of wine requires a living organism—yeast—which consumes the sugar in grapes and excretes alcohol as waste. Yeast was an existence proof for nineteenth-century biologists of how a foreign microbe could alter the functioning of a biological system, but most nonbiologists remain unaware that yeast is alive (Songer and Mintzes, 1994), let alone the correspondence between yeast's role in fermentation and a pathogen's role in human illness.

Case study: Divergent paths to understanding evolution

Evolution by natural selection is a prime example of a theory that was counterintuitive to early scientists for different reasons than it is counterintuitive to science students. Here, I will outline key differences in the motivations and assumptions of early evolutionary theorists and those who learn about evolution secondhand. I focus on differences in background assumptions, as these differences are perhaps the most important factor in this domain, theoretically and pedagogically.

Divergent explanatory goals

The idea that species change over time was entertained as early as antiquity, but it was not widely investigated until the eighteenth and nineteenth centuries

(Bowler, 1992; Mayr, 1982). Biologists of that period had amassed a large database of specimens and sought a naturalistic explanation for the origin of species and their adaptation to particular environments. The traditional explanations were *divine creation*, the idea that species were created in their present form by a divine power; and *spontaneous generation*, the idea that species emerged from the Earth whole-cloth. Neither provided a generative framework for empirical inquiry.

Biologists dissatisfied with creationism and spontaneous generation revisited the idea that the Earth's species had not always existed but were instead the descendants of some smaller number of ancestral species. It was agreed that the ancestral species spread and diversified, but it was debated whether this happened because of environmental circumstances or because of some inherent property of living things. For eighteenth- and nineteenth-century biologists, evolution was an accepted possibility—and a preferred alternative to supernatural explanations—but the mechanism was a mystery.

For nonbiologists, on the other hand, evolution itself is an unlikely supposition (Blancke et al., 2012), as it is not readily observed nor inferred. Casual observation of plants and animals conveys no impression that they have changed over time. The biological world appears to be as static and eternal as the Earth itself. The pressing biological questions for nonbiologists are not where species come from and why they are adapted to their environment but which species are safe to interact with and which should be avoided (Barrett and Broesch, 2012; Wertz and Wynn, 2014). The very idea of evolution has to be suggested by others; it is not intuited as a possibility (Shtulman, Neal, and Lindquist, 2016).

Another reason evolution is viewed as irrelevant to everyday biological concerns is that nonbiologists already have an explanation for adaptation and speciation: divine creation (Heddy and Nadelson, 2012; Newport, 2010). Divine creation is endorsed by individuals of varying ages and upbringings, including children raised in secular households (Evans, 2001). When elementary schoolers are asked where the first bear came from or where the first lizard came from, they usually say that God created them, even when their own parents say that bears and lizards evolved from earlier forms of life. Divine creation embodies a form of causation we are all familiar with—intentional design—and it is thus preferred to a more complicated explanation like evolution. Evolution may provide a naturalistic account of the origins of life—a primary desiderata for biologists—but nonbiologists are generally unperturbed by the supernatural aspects of creationism. Most people consider supernatural causes to be as just as plausible as natural ones (Legare and Shtulman, 2018).

Divergent empirical concerns

Eighteenth- and nineteenth-century biologists knew of a wide range of empirical phenomena that current students do not. They knew of extinct species, through their fossilized remains, and wondered how those species are related to extant species. They knew of *analogous* traits, or traits with similar functions but dissimilar structures, such as bat wings and bird wings, and wondered how those traits emerged seemingly independently. They knew of *homologous* traits, or traits that have taken on new functions or lost their old functions, such as the blind mole rat's eye or the human tailbone, and wondered whether those traits are the remnants of a shared body plan. These facts constrained early theories of evolution and even suggested possible mechanisms. For instance, widespread homologies across species motivated Cope's theory of accelerated growth, or the theory that evolution results from the acceleration and compression of universal stages of embryonic growth, with new stages added on top of old ones (Bowler, 1992).

Nonbiologists are generally unaware of these facts. They may know of fossils and shared traits, but they do not necessarily see these phenomena as evidence of evolution (Evans et al., 2010). The primary evidence of evolution for a nonbiologist is public discourse about evolution and public representations of evolution. The discourse includes claims about common ancestry (e.g., that humans and chimps share 98 percent of their DNA), claims about adaptation (e.g., that white fur is an adaptation to Arctic climates), and the controversy over teaching evolution in school. The public representations include evolutionary trees, nature documentaries, cartoons, and even video games (e.g., *Spore*, *SimEarth*, *Pokémon Go*). Nonbiologists learn about evolution not through observation but through culturally transmitted information, and the challenge for nonbiologists is interpreting this information, which is often vague or misleading.

Evolutionary trees are a prime example of misleading information. Evolutionary trees depict *speciation*, or the emergence of new species. Speciation is inherently a branching process, of one species diverging from another, but it is often depicted as a linear process, of one species giving rise to another, by the evolutionary trees in textbooks and science museums (Catley and Novick, 2008; MacDonald and Wiley, 2012). The nodes in these trees are labeled with extinct species, implying that they gave rise to the extant species along the trees' tips, which is highly unlikely given the ubiquity of extinction. Other problematic features of evolutionary trees include varying the thickness of a tree's branches without explanation, varying the endpoints of a tree's branches without explanation, segregating "higher" organisms from "lower" organisms, and placing humans on the top-most branch of a vertically arrayed tree or the right-most branch of a horizontally arrayed tree (Catley and Novick, 2008; MacDonald and Wiley, 2012; Shtulman and Checa, 2012).

Evolutionary trees, and other popular depictions of evolution, thus present significant interpretive challenges to nonbiologists. Whereas early evolutionary theorists struggled to interpret varied traces of evolution in the fossil record and the zoological record, nonbiologists struggle to interpret ambiguous or misleading representations of evolution in the public record.

Divergent background assumptions

Darwin's discovery of the principle of natural selection revolutionized the biological sciences. While Darwin was one of many biologists trying to understand speciation and adaptation from a naturalistic point of view, he was one of the first to realize that evolution proceeds via selection over a population. Darwin's predecessors and contemporaries had posited many mechanisms of their own—the inheritance of acquired characters (Lamarck's mechanism), the law of accelerated growth (Cope's mechanism), the inherent properties of organic matter (Eimer's mechanism)—but all such mechanisms operated indiscriminately, propelling evolution in each and every lineage of living things. Darwin, on the other hand, realized that evolution is an emergent property of the selective survival of only some lineages within a population.

From where did this insight arise? The history of science suggests that three events were critical: (1) Darwin's journey to the Galapagos, which opened his eyes to the ubiquity of variation within a species (Lack, 1947/1983); (2) Darwin's reading of Malthus's *Essay on the Principle of Population*, which opened his eyes to resource limitation and its role in inciting competition within a population (Millman and Smith, 1997); and (3) Darwin's reading of Lyell's *Principles of Geology*, which opened his eyes to the transformative power of incremental change over vast periods of time (Gruber, 1981).

These events instilled in Darwin an appreciation of intraspecific variation, intraspecific competition, and geologic time, respectively. All were important to Darwin's theorizing, but one concept in particular—intraspecific variation—has been implicated as his most important insight. Philosophers of biology commonly argue that what set Darwin's theory apart from his peers' was that Darwin's was population-based whereas those of his peers were typological. Darwin treated species as continuums of variation whereas his peers treated species as discrete, homogenous types (Gould, 1996; Hull, 1965; Mayr, 1982; Sober, 1994).

Students of biology today have difficulty understanding the same three concepts that proved critical to Darwin's theorizing. Students view variation between species as pervasive and adaptive but variation within species as minimal and nonadaptive (Nettle, 2010; Shtulman and Schulz, 2008). They claim, for instance, that most traits appear in duplicate form across the entire species and that it is unlikely a member of the species could be born with a different version of the trait. Students also hold overly simplistic views of the relations among organisms within an ecosystem—views that downplay competition for resources between species, let alone within species (Özkan, Tekkaya, and Geban, 2004; Zimmerman and Cuddington, 2007). Most believe that stable ecosystems are characterized by ample food, water, and shelter, and that all inhabitants of the ecosystem are able to survive and reproduce. Lastly, students underestimate the duration of geological events by several orders of magnitude (Lee et al., 2011; Trend, 2001). They date the origin of mammals hundreds of millions of years too close to present day and the origin of life billions of years too close to present day.

A psychological question motivated by the history and philosophy of biology is whether understanding evolution by natural selection requires an understanding of all three concepts—intraspecific variation, intraspecific competition, and geologic time—or whether one concept in particular is most critical, namely, intraspecific variation. In my lab, we explored this question directly, surveying students' understanding of variation, competition, and time in relation to their understanding of evolution (Shtulman, 2014). The students were recruited from introductory psychology courses, and they reported having taken an average of 1.2 college-level biology courses. Some were biology majors, but most were not.

We assessed students' understanding of evolution using a battery of questions designed specifically to differentiate population-based reasoning from typological reasoning (Shtulman, 2006). The questions covered six evolutionary phenomena—variation, inheritance, adaptation, domestication, speciation, and extinction—and solicited a combination of closed-ended and open-ended responses.

Here is a sample question regarding adaptation: "A youth basketball team scores more points per game this season than they did the previous season.

Which explanation for this change is most analogous to Darwin's explanation for the adaptation of species? (a) Each returning team member grew taller over the summer; (b) Any athlete who participates in a sport for more than one season will improve at that sport; (c) More people tried out for the same number of spots this year; (d) On average, each team member practiced harder this season." The correct answer is (c), as it is the only answer that evokes selection, but most survey respondents chose one of the other answers, which evoke mechanisms that operate on the group as a whole. And those who chose (a), (b), or (d) as most analogous to Darwin's explanation for adaptation typically chose (c) as *least* analogous, further indicating that they do not see selection as relevant to evolution.

Scores on this survey, in its entirety, could range from −30 to +30, with negative scores indicating typological reasoning and positive scores indicating population-based reasoning. In actuality, they ranged from −25 to +24, with an average score of −2.3.

To measure students' understanding of intraspecific variation, we adapted a task from Shtulman and Schulz (2008). Participants were asked whether each of three traits—a behavioral trait, an external anatomical trait, and an internal anatomical trait—could vary for each of six animals. Half the animals were mammals (giraffes, pandas, kangaroos) and half were insects (grasshoppers, ants, bees). One trial pertained to kangaroos having two stomachs. For this trial, participants were told, "It is commonly observed that kangaroos have two stomachs," and they were then asked (1) "Do you think all kangaroos have two stomachs or just most kangaroos?" and (2) "Could a kangaroo be born with a different number of stomachs?" Across species and traits, participants judged traits actually variable (question 1) 47 percent of the time and potentially variable (question 2) 61 percent of the time.

To measure participants' understanding of intraspecific competition, we presented participants with sixteen behaviors and asked them to indicate which of six animals exhibit that behavior. The behaviors came in four types: cooperation within a species (e.g., nursing the offspring of an unrelated member of the same species), cooperation between species (e.g., sharing a nest or burrow with an animal from a different species), competition within a species (e.g., eating another member of the species), and competition between species (e.g., tricking an animal from a different species into raising one's young). We paired the properties with unfamiliar animals, such as plover birds and bluestreak wrasse, so that participants would be unlikely to know the correct answers and would have to guess. In reality, half the animals exhibited the target behavior and half did not. Overall, participants estimated that cooperative behaviors are more common than competitive ones, and this asymmetry was larger for intraspecific behaviors (where the average difference was 10 percent) than for the interspecific behaviors (where the average difference was only 2 percent).

To measure participants' understanding of geologic time, we adapted a task from Lee et al. (2011). Participants were presented with eighteen historic or geologic events and were asked to estimate how much time had passed since the event occurred. They registered their estimate by selecting one of ten time periods, beginning with "between 100 and 1000 years ago" and ending with "between 100,000,000,000 and 1,000,000,000,000 years ago." The events included the time since Rome was founded, the time since the extinction of dinosaurs, the time since the Earth was formed, and the time since the Milky Way galaxy was formed. Consistent with the findings of Lee et al. (2011), participants systematically overestimated how much time had passed for events occurring less than 10,000 years ago and systematically underestimated how much time had passed for events occurring more than 10,000 years ago.

In sum, participants underestimated the prevalence of intraspecific variation, the prevalence of competition relative to cooperation (especially within a species), and the duration of geologic events. Still, participants varied in their accuracy on each task, and we ran a regression analysis to determine whether understanding each target concept relates to understanding evolution. We regressed scores on our measure of evolution understanding against scores on the intraspecific variation task (the proportion of traits judged potentially variable), scores on the intraspecific competition task (the proportion of behaviors accurately attributed), and scores on the geologic time task (the proportion of events accurately time-stamped). We used a stepwise regression, in which the predictor variables are entered into the regression model by the amount of variance they explain. The first predictor entered was intraspecific competition, which explained 15 percent of the variance in evolution understanding. The second was intraspecific variation, which explained an additional 4 percent. And the third was geologic time, which explained an additional 2 percent. All predictors were significant.

These results confirm the general finding that theory development in the history of science often parallels conceptual development in the individual. Just as Darwin's discovery of natural selection appears to have been based on the conceptual foundations of intraspecific variation, intraspecific competition, and geologic time, students' understanding of natural selection is based on the same foundations. That said, the relative contributions of these foundations were strikingly different. Appreciating within-species competition explained nearly four times as much variance as appreciating within-species variation and nearly eight times as much as appreciating geologic time. Thus, the focus on variation in the philosophy of science does not align with the psychology of evolution understanding. Recognizing that conspecifics compete for resources appears to be more critical to learning about evolution than recognizing that conspecifics vary in their traits.

From an empirical point of view, it's debatable whether organisms are truly more competitive than cooperative—that is, whether nature is better characterized as a "peaceable kingdom" or as "red in tooth and claw" (see De Waal, 2006). Regardless, the latter appears to foster a more accurate, population-based view of evolution. Indeed, what predicted participants' understanding of evolution was not their recognition of competition in general but their recognition of competition within a species. The better participants appreciated that members of the same species compete for resources, the better they understood the logic of natural selection and its consequences for phenomena as diverse as speciation and extinction.

Implications for understanding and improving scientific knowledge

Scientists' pathways to scientific truths are often quite different than students' pathways to the same truths. For instance, in trying to understand where species came from and why they are adapted to their environment, early biologists struggled with (a) the need to account for these phenomena within a naturalistic framework; (b) the need to account for a wide diversity of relevant data, from fossils to analogous traits to homologous traits; and (c) the deep-seated assumption that species are homogenous "types" rather than continuums of a variation. Biology students, on the other hand, struggle with (a) finding value in a naturalistic explanation for phenomena they can already explain by divine creation, (b) interpreting ambiguous or misleading information about evolution conveyed through public discourse and public representations, and (c) conceiving of species as more competitive than cooperative—that is, recognizing that conspecifics compete for food, shelter, and mates. Differences like these are common in other domains of knowledge as well (noted earlier), and they likely have implications for the acquisition and representation of scientific concepts.

One implication is that different forms of cultural input can catalyze the same theory change—from an intuitive theory to a scientific theory—but the output of that process might not be the same, even if the starting point is the same. Humans typically converge on the same intuitive theories, despite living in different environments or in different time periods (Eckstein and Kozhevnikov, 1997; McCloskey, 1983; Vosniadou, 1994b; Wiser and Carey, 1983). Our innate ideas about objects, agents, and organisms furnish us with shared expectations about how those entities will behave, and those expectations are further refined through shared experiences (Carey, 2009; Spelke, 2000). For instance, early-emerging expectations about contact causality and free fall lay the groundwork for an intuitive theory of motion that varies little from one country to the next, whether it be China, Mexico, Israel, Turkey, Ukraine, or the Philippines (Shtulman, 2017).

Science can reshape and restructure our theories to a point where they are no longer intuitive, but it's an open question whether the now-counterintuitive theories are equally counterintuitive for those who discovered them as for those who learned them secondhand. The steps involved in deriving a scientific theory, via data and inference, may be critical to integrating that theory with the expectations and experiences that predated it. On the other hand, deriving a scientific theory could lead one to quarantine the theory, viewing it as relevant to controlled, lab-based observations but irrelevant to observations from everyday life. Scientific innovation is cultural innovation writ large, and there is much we still do not understand about how knowledge obtained through culture is combined with knowledge obtained through experience.

From a practical point of view, comparing scientists' and students' understanding of the same ideas can lead to more effective science education. One of the hallmarks of intuitive theories is their resistance to counterevidence and counterinstruction. Intuitive theories of evolution, for instance, have been documented in individuals of all levels of education, including college biology majors (Nehm and Reilly, 2007), medical-school students (Brumby, 1984), preservice biology teachers (Deniz, Donelly, and Yilmaz, 2008), and even graduate students in the biological sciences (Gregory and Ellis, 2009). Understanding evolution does not increase linearly with exposure to evolutionary ideas; a whole semester of college-level biology typically has no impact on a student's ability to grasp the logic of natural selection (Shtulman and Calabi, 2013).

One reason that instruction may fail to facilitate conceptual change is that it targets the wrong preconceptions. Instruction that follows the sequence of findings that led to the discovery of a scientific idea may miss the mark for nonscientists, who hold a different set of assumptions and struggle with a different set of concerns. Curricula informed by the history of science have proven effective in some domains (see, e.g., Wandersee, 1986), but they may not be effective in all domains or for all students. Additional research comparing the conceptual ecologies of scientists and students is needed to determine whether students should be led to scientific truths along the same path they were discovered or along different paths.

Conclusion

History repeats itself. Students of science today face many of the same difficulties in understanding scientific ideas as the scientists who discovered those ideas. The first theory of a domain explicitly articulated by scientists often resembles the intuitive theories implicitly constructed by nonscientists. That said, there is more than one pathway from intuitive theories to scientific theories, and the pathways taken by scientists may differ systematically from those taken by students. Here, I have outlined three factors that lead scientists and students down different paths: scientists and students hold different explanatory goals; they know of different empirical phenomena; and their theorizing is constrained by different background assumptions. These factors may alter how scientific ideas are mentally represented, either in relation to the world or in relation to each other, and they point to the need for additional research comparing the concepts and theories of professional scientists to those of science students. Scientific knowledge is instantiated in many forms—papers, models, technologies, the records of early scientists, the minds of science students—and all forms can shed light on the structure and origin of such knowledge, particularly if analyzed together.

References

Au, T. K. F., Chan, C. K., Chan, T. K., Cheung, M. W., Ho, J. Y., and Ip, G. W. (2008). Folkbiology meets microbiology: A study of conceptual and behavioral change. *Cognitive Psychology*, 57: 1–19.

Barrett, H. C. and Broesch, J. (2012). Prepared social learning about dangerous animals in children. *Evolution and Human Behavior*, 33: 499–508.

Blacker, K. A. and LoBue, V. (2016). Behavioral avoidance of contagion in children. *Journal of Experimental Child Psychology*, 143: 162–70.

Blancke, S., De Smedt, J., De Cruz, H., Boudry, M., and Braeckman, J. (2012). The implications of the cognitive sciences for the relation between religion and science education: The case of evolutionary theory. *Science & Education*, 21: 1167–84.

Blown, E. J. and Bryce, T. G. K. (2013). Thought-experiments about gravity in the history of science and in research into children's thinking. *Science & Education*, 22: 419–81.

Bonawitz, E. B., van Schijndel, T. J., Friel, D., and Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, 64: 215–34.

Bowler, P. J. (1992). *The Eclipse of Darwinism: Anti-Darwinian Evolution Theories in the Decades around 1900*. Baltimore: John Hopkins University Press.

Brumby, M. N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education*, 68: 493–503.

Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.

Carey, S. and Spelke, E. (1996). Science and core knowledge. *Philosophy of Science*, 63: 515–33.

Catley, K. M. and Novick, L. R. (2008). Seeing the wood for the trees: An analysis of evolutionary diagrams in biology textbooks. *BioScience*, 58: 976–87.

Chi, M. T. H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *The Journal of the Learning Sciences*, 14: 161–99.

Chi, M. T. H., Roscoe, R. D., Slotta, J. D., Roy, M., and Chase, C. C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science*, 36: 1–61.

Chinn, C. A. and Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching,* 35: 623–54.

De Waal, F. (2006). Morally evolved: Primate social instincts, human morality, and the rise and fall of "Veneer Theory". In S. Macedo and J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 1–58). Princeton: Princeton University Press.

Deniz, H., Donelly, L. A., and Yilmaz, I. (2008). Exploring the factors related to acceptance of evolutionary theory among Turkish preservice biology teachers. *Journal of Research in Science Teaching*, 45: 420–43.

DiSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10: 105–225.

DiSessa, A. A. (2008). A bird's-eye view of the "pieces" vs. "coherence" controversy (from the "pieces" side of the fence). In S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 35–60). New York: Routledge.

Eckstein, S. G. and Kozhevnikov, M. (1997). Parallelism in the development of children's ideas and the historical development of projectile motion theories. *International Journal of Science Education*, 19: 1057–73.

Erickson, G. L. (1979). Children's conceptions of heat and temperature. *Science Education*, 63: 221–30.

Evans, E. M. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology*, 42: 217–66.

Evans, E. M., Spiegel, A. N., Gram, W., Frazier, B. N., Tare, M., Thompson, S., and Diamond, J. (2010). A conceptual guide to natural history museum visitors' understanding of evolution. *Journal of Research in Science Teaching*, 47: 326–53.

Fox, R. (1971). *The Caloric Theory of Gases: From Lavoisier to Regnault*. Oxford: Clarendon.

Gelman, S. A. (2003). *The Essential Child*. Oxford: Oxford University Press.

Gelman, S. A. and Legare, C. H. (2011). Concepts and folk theories. *Annual Review of Anthropology*, 40: 379–98.

Gopnik, A. and Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, 138: 1085–108.

Gould, S. J. (1992). *Ever Since Darwin: Reflections in Natural History*. New York: Norton.

Gould, S. J. (1996). *Full House: The Spread of Excellence from Plato to Darwin*. New York: Three Rivers Press.

Gregg, V. R., Winer, G. A., Cottrell, J. E., Hedman, K. E., and Fournier, J. S. (2001). The persistence of a misconception about vision after educational interventions. *Psychonomic Bulletin & Review*, 8: 622–26.

Gregory, T. R. and Ellis, C. A. J. (2009). Conceptions of evolution among science graduate students. *Bioscience*, 59: 792–99.

Gruber, H. E. (1981). *Darwin on Man: A Psychological Study of Scientific Creativity*. Chicago: University of Chicago Press.

Harlow, D. B., Swanson, L. H., Nylund-Gibson, K., and Truxler, A. (2011). Using latent class analysis to analyze children's responses to the question, "What is a day?" *Science Education*, 95: 477–96.

Heddy, B. C. and Nadelson, L. S. (2012). A global perspective of the variables associated with acceptance of evolution. *Evolution: Education and Outreach*, 5: 412–18.

Holton, G. J. (1988). *Thematic Origins of Scientific Thought: Kepler to Einstein*. Boston: Harvard University Press.

Hull, D. (1965). The effect of essentialism on taxonomy: 2000 years of stasis. *British Journal for the Philosophy of Science,* 15–16: 314–26.

Inagaki, K. and Hatano, G. (2004). Vitalistic causality in young children's naïve biology. *Trends in Cognitive Sciences*, 8: 356–62.

Johnson, S. C. and Solomon, G. E. (1997). Why dogs have puppies and cats have kittens: The role of birth in young children's understanding of biological origins. *Child Development*, 68: 404–19.

Kaiser, M. K., Proffitt, D. R., and McCloskey, M. (1985). The development of beliefs about falling objects. *Perception & Psychophvsics*, 38: 533–39.

Kalish, C. W. (1996). Preschoolers' understanding of germs as invisible mechanisms. *Cognitive Development*, 11: 83–106.

Kim, E. and Pak, S. J. (2002). Students do not overcome conceptual difficulties after solving 1000 traditional problems. *American Journal of Physics*, 70: 759–65.

Koslowski, B., Marasia, J., Chelenza, M., and Dublin, R. (2008). Information becomes evidence when an explanation can incorporate it into a causal framework. *Cognitive Development*, 23: 472–87.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96: 674–89.

Kuhn, D. and Pease, M. (2008). What needs to develop in the development of inquiry skills? *Cognition and Instruction*, 26: 512–59.

Lack, D. (1947/1983). *Darwin's Finches*. Cambridge: Cambridge University Press.

Laudan, L., Donovan, A., Laudan, R., Barker, P., Brown, H., Leplin, J., Thagard, P., and Wykstra, S. (1986). Scientific change: Philosophical models and historical research. *Synthese*, 69: 141–223.

Lederberg, J. (2000). Infectious history. *Science*, 288: 287–93.

Lee, H. S., Liu, O. L., Price, C. A., and Kendall, A. L. (2011). College students' temporal-magnitude recognition ability associated with durations of scientific changes. *Journal of Research in Science Teaching*, 48: 317–35.

Legare, C. H. and Shtulman, A. (2018). Explanatory pluralism across cultures and development. In J. Proust and M. Fortier (Eds.), *Interdisciplinary Approaches to Metacognitive Diversity* (pp. 415–32). Oxford: Oxford University Press.

Libarkin, J. C. and Anderson, S. W. (2005). Assessment of learning in entry-level geoscience courses: Results from the Geoscience Concept Inventory. *Journal of Geoscience Education*, 53: 394–401.

Libarkin, J. C., Anderson, S. W., Dahl, J., Beilfuss, M., Boone, W., and Kurdziel, J. P. (2005). Qualitative analysis of college students' ideas about the earth: Interviews and open-ended questionnaires. *Journal of Geoscience Education*, 53: 17–26.

Lorch, R. F., Lorch, E. P., Calderhead, W. J., Dunlap, E. E., Hodell, E. C., and Freer, B. D. (2010). Learning the control of variables strategy in higher and lower achieving classrooms: Contributions of explicit instruction and experimentation. *Journal of Educational Psychology*, 102: 90–101.

MacDonald, T. and Wiley, E. O. (2012). Communicating phylogeny: Evolutionary tree diagrams in museums. *Evolution: Education and Outreach*, 5: 14–28.

Marques, L. and Thompson, D. (1997). Misconceptions and conceptual changes concerning continental drift and plate tectonics among Portuguese students aged 16–17. *Research in Science & Technological Education*, 15: 195–222.

Mayr, E. (1982). *The Growth of Biological thought: Diversity, Evolution, and Inheritance*. Cambridge: Harvard University Press.

McCloskey, M. (1983). Naïve theories of motion. In D. Gentner and A. Stevens (Eds.), *Mental Models* (pp. 299–324). Hillsdale: Erlbaum.

Millman, A. B. and Smith, C. L. (1997). Darwin's use of analogical reasoning in theory construction. *Metaphor and Symbol*, 12: 159–87.

Morris, B. J. and Masnick, A. M. (2014). Comparing data sets: Implicit summaries of the statistical properties of number sets. *Cognitive Science*, 39: 156–70.

Morris, S. C., Taplin, J. E., and Gelman, S. A. (2000). Vitalism in naïve biological thinking. *Developmental Psychology*, 36: 582–95.

National Science Board (2014). *Science and Engineering Indicators*. Arlington: National Science Foundation.

Nehm, R. H. and Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, 57: 263–72.

Nersessian, N. J. (1989). Conceptual change in science and in science education. *Synthese*, 80: 163–83.

Nettle, D. (2010). Understanding of evolution may be improved by thinking about people. *Evolutionary Psychology*, 8: 205–28.

Newport, F. (2010). Four in 10 Americans believe in strict creationism. Washington: Gallup Organization.

Oreskes, N. (1999). *The Rejection of Continental Drift: Theory and Method in American Earth Science*. Oxford: Oxford University Press.

Özkan, Ö., Tekkaya, C., and Geban, Ö. (2004). Facilitating conceptual change in students' understanding of ecological concepts. *Journal of Science Education and Technology*, 13: 95–105.

Pew Research Center (2015). *Public and Scientists' Views on Science and Society*. Washington: Author.

Reiner, M., Slotta, J. D., Chi, M. T. H., and Resnick, L. B. (2000). Naïve physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction*, 18: 1–34.

Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition*, 45: 1–32.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, 32: 102–19.

Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, 52: 170–94.

Shtulman, A. (2014). *Using the History of Science to Identify Conceptual Prerequisites to Understanding Evolution*. Poster presented at the 40th meeting of the Society for Philosophy and Psychology, Vancouver, Canada.

Shtulman, A. (2017). *Scienceblind: Why Our Intuitive Theories about the World Are So Often Wrong*. New York: Basic Books.

Shtulman, A. and Calabi, P. (2013). Tuition vs. intuition: Effects of instruction on naïve theories of evolution. *Merrill-Palmer Quarterly*, 59: 141–67.

Shtulman, A. and Checa, I. (2012). Parent–child conversations about evolution in the context of an interactive museum display. *International Electronic Journal of Elementary Education*, 5: 27–46.

Shtulman, A., Neal, C., and Lindquist, G. (2016). Children's ability to learn evolutionary explanations for biological adaptation. *Early Education and Development*, 27: 1222–36.

Shtulman, A. and Schulz, L. (2008). The relationship between essentialist beliefs and evolutionary reasoning. *Cognitive Science*, 32: 1049–62.

Smith, C. L. (2007). Bootstrapping processes in the development of students' commonsense matter theories: Using analogical mappings, thought experiments, and learning to measure to promote conceptual restructuring. *Cognition and Instruction*, 25: 337–98.

Smith, C., Carey, S., and Wiser, M. (1985). On differentiation: A case study of the development of the concepts of size, weight, and density. *Cognition*, 21: 177–237.

Sober, E. (Ed.) (1994). *Conceptual Issues in Evolutionary Biology*. Boston: MIT Press.

Songer, C. J. and Mintzes, J. J. (1994). Understanding cellular respiration: An analysis of conceptual change in college biology. *Journal of Research in Science Teaching*, 31: 621–37.

Spelke, E. S. (2000). Core knowledge. *American Psychologist*, 55: 1233–43.

Springer, K. and Keil, F. C. (1989). On the development of biologically specific beliefs: The case of inheritance. *Child Development*, 60: 637–48.

Thagard, P. R. (1978). The best explanation: Criteria for theory choice. *The Journal of Philosophy*, 75: 76–92.

Thagard, P. (1999). *How Scientists Explain Disease*. Princeton: Princeton University Press.

Toulmin, S. and Goodfield, J. (1961). *The Fabric of the Heaves*. Chicago: University of Chicago Press.

Toulmin, S. and Goodfield, J. (1962). *The Architecture of Matter*. Chicago: University of Chicago Press.

Trend, R. (2001). Deep time framework: A preliminary study of U.K. primary teachers' conception of geological time and perceptions of geoscience. *Journal of Research in Science Teaching,* 38: 191–221.

Vosniadou, S. (1994a). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4: 45–69.

Vosniadou, S. (1994b). Universal and culture-specific properties of children's mental models of the earth. In L. A. Hirschfeld and S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 412–30). Cambridge: Cambridge University Press.

Vosniadou, S. and Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science*, 18: 123–83.

Wandersee, J. H. (1986). Can the history of science help science educators anticipate students' misconceptions? *Journal of Research in Science Teaching*, 23: 581–97.

Wertz, A. E. and Wynn, K. (2014). Selective social learning of plant edibility in 6- and 18-month-old infants. *Psychological Science*, 25: 874–82.

Wiser, M. and Amin, T. (2001). “Is heat hot?” Inducing conceptual change by integrating everyday and scientific perspectives on thermal phenomena. *Learning and Instruction*, 11: 331–55.

Wiser, M. and Carey, S. (1983). When heat and temperature were one. In D. Gentner and A. L. Stevens (Eds.), *Mental Models* (pp. 267–97). Hillsdale: Erlbaum.

Zimmerman, C. and Cuddington, K. (2007). Ambiguous, circular and polysemous: Students’ definitions of the “balance of nature” metaphor. *Public Understanding of Science*, 16: 393–406.