{ 3 }

Cognitive Constraints on the Understanding and Acceptance of Evolution

Andrew Shtulman and Prassede Calabi

In September 2008, Electronic Arts released a computer game called "Spore," in which players control the evolution of a novel organism from microscopic cell to interstellar explorer. Billed as a potential teaching tool, the game has been sharply criticized by evolutionary biologists for its lack of scientific accuracy (Bohannon, 2008). In Spore, "evolution" proceeds by swapping "DNA points," earned through finding food and avoiding predators, for new body parts. The collection of selectable parts is determined not by variation in the population (all members of a given species are identical) but by the organism's size and intelligence—attributes that increase linearly, and deterministically, as the game proceeds. Death and reproduction occur in Spore but are unaffected by competition or selection, and are thus unrelated to the organism's "evolution." Death actually serves an opportunity to restart the game from wherever one's organism last spawned, and reproduction serves an opportunity to edit the organism, adding and subtracting body parts at will. Just as Spore makes no role for natural selection, it makes no role for common descent; every creature on the planet can trace its ancestry back to a different single-celled organism originally deposited by meteors.

While it is possible that evolutionary misconceptions embodied in Spore's game play reflect nothing more than confusion on the part of its designers, these misconceptions are not random or unique. Science education researchers have documented similar types of misconceptions among students of every stripe, from middle school students (Lawson & Thompson, 1988) to high school students (Banet & Ayuso, 2003; Settlage, 1994) to college undergraduates (Bishop & Anderson, 1990; Ferrari & Chi, 1998; Greene, 1990; Nehm & Reilly, 2007) to medical school students (Brumby, 1984) to preservice teachers (Crawford, Zembal-Saul, Munford, & Friedrichsen, 2005). These misconceptions include conflating mutation with adaptation, conflating species adaptation with individual adaptation, and preferring teleological explanations of adaptation to mechanistic ones.

From where do these misconceptions arise? Although different researchers have identified different sources (e.g., Jimenez, 1994; Southerland, Abrams, Cummins, &

Anzelmo, 2001), here we review evidence suggesting they derive, at least in part, from an early-emerging tendency to "essentialize" the biological world—that is, to assume that all members of a species share the same causal potential for growth, appearance, and behavior (see Gelman, 2003; Gelman & Rhodes, this volume). We also present evidence that these misconceptions foster skepticism toward evolution and must be addressed if science educators hope to increase public acceptance of evolution, especially in the United States, where only 40% of the population agrees with the statement "Human beings, as we know them, developed from earlier species of animals" (Miller, Scott, & Okamoto, 2006). The bulk of this chapter is devoted to describing the results of a teaching-intervention study in which college undergraduates' understanding and acceptance of evolution were assessed before and after an introductory course on evolution and behavior. We begin by describing the phenomenon of biological essentialism and its consequences for evolutionary reasoning, both in the history of evolutionary biology and in the practice of evolution education.

Biological Essentialism and Its Consequences

The concept of essentialism is well illustrated by Hans Christian Andersen's (1844) fairytale "The Ugly Duckling." The story begins with a mother duck sitting on a nest of eggs, waiting for her ducklings to hatch. One duckling hatches later than the others, and he is, to everyone's dismay, larger and "uglier" than his siblings. A neighboring duck suggests that he may be a turkey, but that suggestion is soon refuted by the fact that he can swim. This ability, paired with his unusual looks, makes him a target of ridicule from both the ducks and the turkeys. Frightened and upset, the duckling leaves his home in search of animals who will accept him as one of their own. During his journey, he meets geese, who reject him as an unsuitable mate; a tom cat, who rejects him for his inability to purr; and a hen, who rejects him for his inability to lay eggs. Finally, after months of travel, the duckling encounters a group of graceful white swans, who, to his surprise, accept him into their family. The reason, he soon discovers, is that he has grown into a graceful white swan himself. "To be born in a duck's nest, in a farmyard," writes Andersen, "is of no consequence to a bird, if it is hatched from a swan's egg" (p. 20).

Whatever moral the story was intended to convey, its plot is predicated on the assumption that an organism's properties—both current and potential—are determined by its species kind, which is, in turn, determined by its parentage. Much developmental research has shown that this assumption appears to be ubiquitous across cultures (Medin & Atran, 2004) and across ages (Gelman, 2003). Children and adults from all parts of the world tend to reason about an organism's outward appearance and behavior on the basis of an internal causal power, or "essence," inherited from the organism's parents and fixed at the organism's birth. This assumption serves us well in most situations, as an organism's species is, indeed, a reliable predictor of its properties. Knowing that an organism is a swan, for instance, allows us to make accurate predictions about how that organism should look (brown and fuzzy as an infant, white and sleek as an adult), where that organism should live (by water), what that organism should eat (vegetation), how that organism should reproduce (by laying eggs), and many other such properties.

Empirical evidence that young children are biological essentialists comes from at least three sources. First, many studies have shown that very young children privilege species kind over perceptual similarity when reasoning about the properties of novel organisms (Gelman & Markman, 1987; Gelman & Coley, 1990; Jaswal & Markman, 2007). In these studies, children are taught a novel property of a familiar organism (e.g., "This black-and-white cat can see in the dark") and asked whether certain novel organisms possess the same property. Some of the novel organisms are of the same species as the target organism but differ in appearance (e.g., a cat with different markings), and others share the same appearance but are of a different species (e.g., a skunk with identical markings). Children as young as age 2½ reliably project the target property to the former but not the latter, implying that they view species identity (conveyed via linguistic labels) as a better predictor of shared properties than mere appearance.

Second, preschoolers assume that an organism will retain its species identity across various natural and/or commonplace changes in appearance, like growing in size (Rosengren, Gelman, Kalish, & McCormick, 1991), growing in complexity (Hickling & Gelman, 1995), or donning a costume (DeVries, 1969). By age 7, children assume that an organism will retain its species identity even across drastic and/or unusual changes in appearance, like plastic surgery or chemical injections (Keil, 1989).

Third, numerous studies have shown that preschoolers assume an organism will retain its species identity across various changes in upbringing (Gelman & Wellman, 1991; Johnson & Solomon, 1997; Springer, 1996; Waxman, Medin, & Ross, 2007). In these studies, children are presented with Ugly Duckling–like scenarios, in which a baby animal is removed from its birth parents (e.g., cows) and raised by members of a different species (e.g., pigs). The children are then asked to predict which properties the baby animal would possess as an adult: the biological properties of its birth parents (e.g., a straight tail and a diet of grass) or those of its adopted parents (e.g., a curly tail and a diet of slop). Children of all ages tend to predict that the baby will grow to possess the biological properties of its birth parents. They also tend to justify their judgments with explicit appeals to the continuity of species identity (e.g., "It will eat grass because it's a cow, not a pig").

In sum, young children appear to construe the biological world in terms of hidden essences that give rise to its observable properties and causal regularities. Such biases are generally a useful constraint on biological induction, as they support the generally accurate projection of species-specific properties to individual organisms. In the absence of such biases, we would be unable to explain or predict the behavior of any organism we had not personally studied. Indeed, virtually all biological sciences operate on the assumption that information gleaned from observing a subset of a species is applicable to the species as a whole.

Yet, despite its utility for reasoning about the properties of individual organisms, biological essentialism has proven a major impediment for reasoning about population-level phenomena, such as evolution and natural selection. The problem, as articulated by historians of science like Gould (1996), Hull (1965), and Mayr (1982), is that essentialism leads one to treat species as discrete, homogeneous units, whose aggregate properties are true of all members of the species and only members of the species. This view correctly implies that any, and every, baby swan has the potential to grow into an adult swan—"swanness" is innate, discrete, and different from "duckness"—but it incorrectly implies that differences between swans are unimportant or inconsequential. Indeed, it positively obscures the fact that most baby swans will *not* survive to adulthood, let alone reproduce. Thus, easy understanding of the similarity among individuals within a species precludes easy understanding of differences *between* those individuals, especially differences that result in differential mortality and differential reproduction. As a result, students engaged in learning about evolution are likely to adopt what Mayr (2001) terms a "transformational" theory of evolution, or a theory in which evolution is (incorrectly) construed as the cross-generational transformation of a species' underlying essence, rather than what Mayr terms a "variational" theory of evolution, or a theory in which evolution is (correctly) construed as the selective propagation of within-species variation.

The effect of biological essentialism on the development of evolutionary biology was profound. According to Mayr (1982), Greek scholars had formulated the concept of descent with modification as early as 600 BC, but the mechanisms of evolution remained a mystery for another 25 centuries. Those who attempted to solve this mystery invariably fell prey to what Gould (1996) calls the "fallacy of reified variation," or the inclination "to abstract a single ideal or average as the essence of a system and to devalue or ignore variation among the individuals that constitute the full population" (p. 40). By focusing on the similarities among members of the same species rather than their differences, early evolutionary theorists ended up positing mechanisms of evolution that operated over individuals, not over populations—namely, mechanisms like the inheritance of acquired traits, the intrinsic properties of organic matter, or the law of acceleration of growth (see Bowler, 1983, for a review). Not until Darwin did evolutionary biologists begin eschewing species-wide similarities for within-species differences. Indeed, Darwin's recognition of the importance of intraspecific variation allowed him to combine three major insights: descent with modification, competition as a selective force (i.e., application of Malthus), and phylogeny as a "tree of life" (i.e., shared ancestry). The result was a qualitatively different view of evolution—"variationism"—which henceforth became the backbone of the biological sciences.

Because biological essentialism is seemingly universal (Medin & Atran, 2004), it is reasonable to suppose that, just as early evolutionary biologists were led astray by their essentialist intuitions, modern-day students of evolution are led astray as well, adopting transformational interpretations of evolutionary phenomena prior to learning correct, variational ones. Shtulman (2006) investigated this hypothesis by designing an in-depth assessment of evolutionary reasoning intended to distinguish between variational and transformational interpretations of six evolutionary phenomena: variation, inheritance, adaptation, domestication, speciation, and extinction. (Sample items from each section of the test are presented below.) The assessment was administered to 45 high school and college students enrolled in the Harvard Summer School and found that the majority (53%) held predominantly transformational views of evolution. Importantly, those who demonstrated transformational conceptions on one section of the assessment tended to do so on several others, implying that their responses were not isolated misconceptions but rather the by-products of a qualitatively different way of understanding species change and species adaptation. Indeed, a factor analysis of students' responses across the six different sections of the assessment revealed one, and only one, underlying factor. Students who scored high on this factor revealed a consistently variational understanding of evolution, whereas those who scored low revealed a consistently transformational understanding.

Shtulman and Schulz (2008) extended these findings by comparing adults' understanding of evolution, as measured by an abbreviated version of the Shtulman (2006) assessment tool, to their acceptance of within-species variation, as measured by a series of questions about whether certain biological properties can, and do, vary across different members of the same species (e.g., "Do all giraffes have spotted coats or just most giraffes?" "Do all ants have a tube-shaped heart or just most ants?"). As expected, adults who demonstrated a variational view of evolution were significantly more likely to accept within-species variation than those who demonstrated a transformational view. Indeed, the latter group of adults were no more likely to accept within-species variation than were 4-year-old children. Taken together, these findings suggest that deep-seated essentialist biases lead students to devalue within-species variation, and, as a result, fail to understand the mechanism of evolution that operates over such variation: natural selection.

Cognitive Constraints on Understanding

Having described the role of biological essentialism in evolutionary thought, we now turn to a detailed analysis of how biological essentialism constrains students' naive theories of evolution. This analysis is presented in the context of a study assessing the nature of those theories before and after a semester's worth of college-level instruction in evolutionary biology. Also assessed was the relation between students' understanding of how evolution works and their acceptance of various evolutionary claims.

THE PARTICIPANTS

The participants were 45 college undergraduates enrolled in a one-semester course on behavior and evolution at a large, public northeastern university. The course was taught three times, with about 15 students each semester. Because there was no effect of semester scheduling on any of the variables reported below, we collapsed the three samples into one. Most participants were psychology majors who had enrolled in the course to fulfill an upper-division requirement of the major. All had taken at least one high school or college-level biology course prior to the course in which they were currently enrolled, and some had taken as many as three. Although it is unclear how much of that coursework entailed evolutionary concepts/phenomena, most participants (76%) claimed to have taken a class or read a book that sufficiently explained the concept of natural selection prior to instruction.

THE TEACHING INTERVENTION

The main objective of the teaching intervention was to help participants derive, for themselves, the concepts of evolution and natural selection from basic principles of biology, natural history, and population dynamics. The intervention was based on Mayr's (1982) analysis of how Darwin initially derived the concepts of evolution and natural selection from four basic phenomena (superfecundity, resource limitation, trait variation, and trait heritability) and two intermediate inferences (differential survival and differential reproduction). Five clusters of activities produced opportunities for participants to reproduce this chain of inferences, either by generating their own data or by analyzing preexistent data from real populations of organisms. These activities also produced opportunities for participants to confront, articulate, and question their nongenetic, nonvariational assumptions about species adaptation.

The opening activity was designed to introduce the concept of superfecundity (the potential inherent in all species to grow exponentially) and to set the stage for the other three phenomena. That activity began with the instructor posing the question, "Why is the earth not covered in dogs?" Participants were asked to estimate the number of offspring a single pair of dogs would produce over six generations of breeding if each pair in each generation produced ten offspring per year (the so-called "overlapping generations model"). After making an estimation, participants calculated and graphed the actual number of offspring produced. This value typically exceeded their estimates by several orders of magnitude and raised questions that served as entry points for discussing the other phenomena and inferences. For instance, the follow-up question "Why do so few individuals survive?" raised issues of resource limitation, intraspecies competition, predation, disease, and bad weather. The follow-up question "Who dies?" raised issues of trait variability, differential reproduction, and chance (for more on the intervention, see Calabi, 1998, 2005).

This way of teaching evolution contrasts with typical instruction in that it allows students to derive relevant outcomes from first principles and real data, while also accounting for their preinstructional misconceptions. This approach has proven successful in a variety of domains (e.g., Moss & Case, 1999; Smith, 2007; Wiser & Amin, 2001), though it has not been used in the domain of evolution (to our knowledge). Our hope was that, by helping students derive the relevant concepts themselves, they would more likely understand those concepts and would also more likely accept them as valid, thereby side-stepping the typical "dualistic" outcome of science education, in which students maintain their intuitive beliefs alongside those explicitly required by the instructor (Bloom & Weisberg, 2007).

THE COMPREHENSION ASSESSMENT

At the beginning and end of the 15-week semester, participants' were administered a 30-question assessment of their understanding of variation, inheritance, adaptation, domestication, speciation, and extinction. The same questions were used at both pretest and posttest and were never discussed during the teaching intervention itself. Each question was designed to differentiate inaccurate, transformational interpretations of the target phenomenon from accurate, variational ones, and participants were instructed to answer each question based on their best understanding of evolution regardless of whether they believe that evolution actually occurs. This assessment can be found, in its entirety, in the appendix of Shtulman (2006). Below, we discuss sample questions from each section of assessment to illustrate the nature of the instrument as a whole.

Variation

The evolution of the peppered moth, *Biston betularia*, was used as a vehicle for eliciting participants' reasoning about the prevalence and importance of within-species variation. On the first question of this section, participants were told that nineteenth-century England underwent an industrial revolution with the unfortunate side effect of covering the English countryside in soot and ash and that during the same time period England's native moth species, *Biston betularia*, became, on average, darker in color. Participants were then asked to speculate how a change in the moths' environment might have brought about a change in the moths' color. Responses that referenced individual differences in fitness (e.g., "predators lunched on the lighter ones, leaving the darker ones to reproduce") were coded as variational, and responses that referenced the needs of the population as a whole (e.g., "the moths needed to blend into their environment in order to survive") were coded as transformational. Prior to instruction, 44% of participants provided variational responses, 44% provided transformational responses, and 12% provided ambiguous responses (e.g., "evolution"). Following instruction, 51% provided variational responses, 38% provided transformational responses, and 11% provided ambiguous responses.

Inheritance

A fictitious species of woodpeckers was used as a vehicle for eliciting participants' reasoning about the heritability of various traits. Participants were told that two such woodpeckers migrated to a windier environment and, in consequence, developed stronger wing muscles. Participants were then asked to decide whether offspring of these two woodpeckers would be born with (a) stronger wing muscles than the parents had at birth, (b) weaker wing muscles than the parents had at birth, or (c) either stronger wing muscles or weaker wing muscles, neither being more likely. Participants who chose (c) and justified their response by referencing the randomness of mutations (e.g., "phenotypic differences occur by random chance") or the phenotype-genotype distinction (e.g., "things that develop or are learned during the lifetime of an animal cannot be passed down to its offspring") were coded as having provided a variational response. Participants who chose (a) and justified their responses by referencing the necessity of adaptation (e.g., "the offspring need to have strong wing muscles to survive in a windier environment") were coded as having provided a transformational response. At pretest, 38% of participants provided variational responses, 42% provided transformational responses, and 20% provided ambiguous responses. At posttest, 53% provided variational responses, 33% provided transformational responses, and 14% provided ambiguous responses.

Adaptation

Five analogical-reasoning questions were used to assess participants' interpretation of the mechanism of adaptation. On each question, participants were shown four explanations for why a group of individuals had improved their performance along some particular dimension and asked to select the explanation that was most analogous to Darwin's explanation for the adaptation of species. For example, participants were shown the following four explanations for why a youth basketball team had won more games in the current season than in the previous season: (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 season; or (d) each team member practiced harder this season than he did last season. Whereas choice (c) attributes the improvement to changes in group membership (a variational analogy), choices (a), (b), and (d) attribute the improvement to the transformation of each group member (a transformational analogy). Prior to instruction, 40% of participants chose the variational analogy, and 60% chose one of the three transformational analogies. Following instruction, 53% chose the variational analogy and 47% chose one of the three transformational analogies.

Domestication

The domestication of corn from Teosinte, a wild grass native to Central America, was used as a vehicle for assessing participants' interpretation of the role of human intervention in the domestication process. On one set of questions, participants were asked to rank six factors in order of their relevance to the domestication of corn: (a) the degree of similarity among plants of the same generation, (b) the average amount of time each plant was exposed to direct sunlight, (c) the preferences of those who decided which kernels to plant, (d) the fertility of the soil in which the kernels were planted, (e) the average rainfall per year, and (f) the percentage of each crop used to breed the next generation. Whereas factors (a), (c), and (f) are relevant to the modification of an entire species, factors (b), (d), and (e) are relevant only to the modification of individual organisms. At pretest, 53% of participants assigned higher rankings to the species-relevant factors than to the organism-relevant factors (a variational response), and 47% did the reverse (a transformational response). At posttest, 87% assigned higher rankings to the species-relevant factors than to the organism-relevant factors, and 13% did the reverse.

Speciation

Primate evolution was used as a vehicle for eliciting participants' reasoning about common ancestry and species individuation. On one question, participants were shown a list of nine species—lemurs, elephants, salamanders, sparrows, bees, jellyfish, algae, daffodils, and brontosauruses—and asked to place a check next to each species that shares a common ancestor with humans. Common descent is perfectly consistent with variationism, as variationists interpret divergence as a product of prolonged geographic isolation or assortative mating. It is not, however, consistent with transformationism, for transformationists assume that all members of the same species share a common essence unaffected by geographical constraints and thus impervious to divergence. Transformationists must therefore assume either that all extant species evolved independently of one another or that some extant species are actually "preevolved" versions of others (e.g., viewing chimpanzees as a parent species, not a sister species, to humans). Consistent with this view, 56% of participants claimed, at pretest, that humans share a common ancestor with fewer than half of the nine species; only 36% claimed that humans share a common ancestor with all nine. At posttest, 47% claimed that humans share a common ancestor with fewer than half of the nine species, and 40% claimed that humans share a common ancestor with all nine.

Extinction

The evolution of bacteria was used as a vehicle for eliciting participants' beliefs about the prevalence of extinction. On the first question of the section, participants were asked to decide whether the number of extinct bacteria species is (a) greater than the number of living bacteria species, (b) smaller than the number of living bacteria species, or (c) either greater or smaller than the number of living bacteria species, neither option being more likely. Participants who chose (a) and justified their response by referencing the unlikelihood of adaptation (e.g., "evolution is about trial and error; many fail and few succeed") or the scope of a geological timescale (e.g., "the extinct species have accumulated over billions of years") were coded as having provided a variational response. Participants who chose (b) or (c) and justified their responses by referencing the likelihood of adaptation (e.g., "although some bacteria went extinct, most adapted to their environment") were coded as having provided a transformational response. Prior to instruction, 40% of participants provided variational responses, 40% provided transformational responses, and 20% provided ambiguous responses. Following instruction, 44% provided variational responses, 33% provided transformational responses, and 22% provided ambiguous responses.

DIFFERENCES FROM PRETEST TO POSTTEST

Participants' responses were analyzed quantitatively by scoring them along a three-point scale. Responses consistent with variationism were scored +1; responses consistent with transformationism were scored −1; and responses consistent with both theories (i.e., ambiguous responses) were scored 0. Summed across 30 questions, participants' assessment scores could range from −30 to +30. In actuality, they ranged from −24 to +27 at pretest and −22 to +28 at posttest. The actual distribution of participants' scores is displayed in Figure 3.1. At pretest, most participants (47%) scored between −30 and −11, and only a small minority (24%) scored between +11 and +30. At posttest, most participants (44%) scored between −10 and +10, with approximately the same number (36%) scoring between +11 and +30. The increase in the proportion of participants who scored between +11 and +30 was not statistically significant, but the *decrease* in the proportion of participants who scored between −11 and −30 was. Thus, the intervention appeared to be more effective at eliminating strong transformational reasoning than at fostering strong variational reasoning.

Overall, the mean pretest score was −3.5 ($SD = 14.8$), and the mean posttest score was 2.3 ($SD = 13.3$). This difference was highly significant, yet the difference itself is ambiguous as to the nature of participants' conceptual progress. On one hand, participants who began the course with a predominantly transformational view of

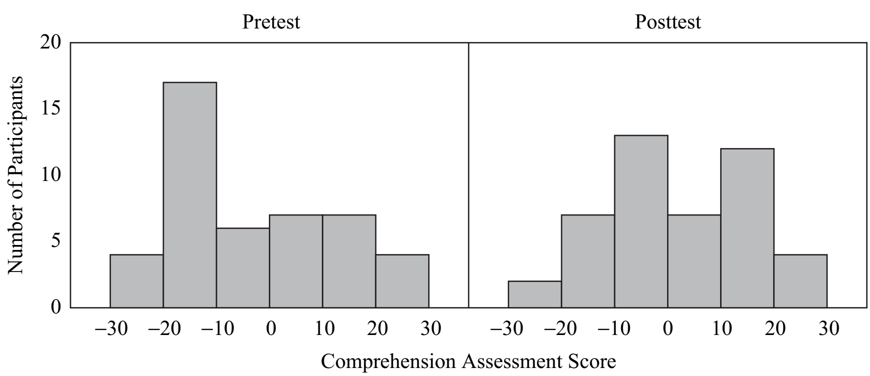

FIGURE 3.1 *Frequency distributions of participants' comprehension assessment scores at pretest and posttest (range = −30 to +30). Negative scores are indicative of transformational reasoning and positive scores are indicative of variational reasoning.*

evolution might have ended the course with significantly fewer transformational conceptions and significantly more variational conceptions. On the other, participants who began the course with a predominantly transformational view might have ended the course more confused, producing significantly more ambiguous responses (scored 0) but not significantly more variational responses (scored +1). A closer analysis of participants' responses revealed that the former scenario was actually more typical than the latter. Prior to instruction, participants provided an average of 11.2 variational responses, 15.4 transformational responses, and 3.4 ambiguous responses. Following instruction, they provided an average of 14.2 variational responses, 13.8 transformational responses, and 2.8 ambiguous responses. The increase in variational responses from pretest to posttest was statistically significant, but the decrease in ambiguous responses was not. Thus, the decrease in participants' transformational reasoning was accompanied not by confusion but by a corresponding increase in variational reasoning.

Participants' pretest and posttest scores are broken down by section in Figure 3.2. These data show that the effects of instruction were widespread, as participants increased their score on all six sections (though the increase on the Extinction section was not statistically reliable). Interestingly, there was a strong correlation between the mean pretest score for each section and its corresponding pre-post gain ($r = 0.73$), implying that participants made greater conceptual progress on the sections they understood better from the start.

Postinstructional gains in understanding were widespread not only across sections but across participants as well. Seventy-six percent of participants increased their score by at least 1 point, 49% increased their score by at least 5 points, and 27% increased their score by at least 10 points. Overall, there was a negative correlation between a participant's pretest score and his/her pre-post gain ($r = -0.41$)

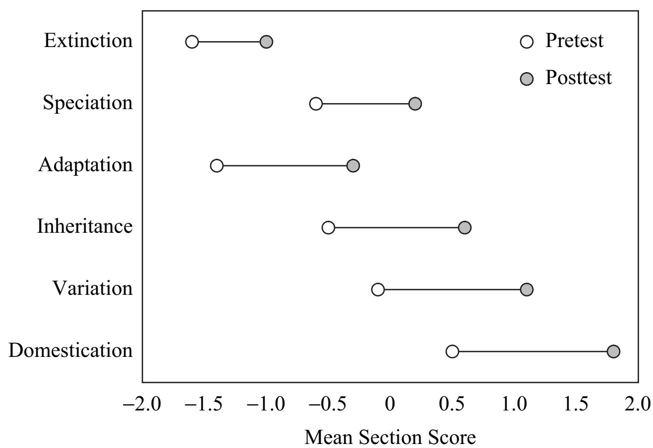

FIGURE 3.2 *Mean scores on the individual sections of the comprehension assessment (range = −5 to +5), ordered from smallest pre-post gain to largest. Pre-post gains were statistically significant for all sections except Extinction.*

such that participants with low pretest scores gained more points (or lost fewer points) than participants with high pretest scores. This correlation was not due to a ceiling effect among participants with high pretest scores, as all but one participant with a positive pretest score could have improved his or her score by at least five points. Rather, instruction appears to have been more effective for participants who entered the classroom with moderate to strong transformational misconceptions.

WITHIN-PARTICIPANT CONSISTENCY

One of the hallmarks of conceptual change is the degree to which an individual's beliefs about various domain-specific phenomena cohere both before and after the change (e.g., Au, Chan, Chan, et al., 2008; Smith, Solomon, & Carey, 2005; Vosniadou & Brewer, 1992). To measure the coherence in participants' beliefs, we looked for within-participant consistency across different sections of the same assessment. The average correlation among participants' section scores was high at both pretest ($r = 0.44$) and posttest ($r = 0.39$), with nearly all such correlations proving statistically reliable. Furthermore, a factor analysis of participants' scores on the six different sections revealed one, and only one, factor capable of explaining the majority of variance in those scores at both pretest (54%) and posttest (51%). Thus, participants' understanding of a diversity of evolutionary phenomena was well described by a single factor; those who scored high on this factor demonstrated consistently variational reasoning, whereas those who scored low demonstrated consistently transformational reasoning (replicating Shtulman, 2006).

That said, participants who began the course with strong transformational beliefs tended to end the course with a less coherent view of evolution, as evidenced both by a significant decrease in number of participants who scored between –11 and –30 on the comprehension assessment (observable in Figure 3.1) and by a significant decrease in the strength of the intercorrelations among participants' section scores. These participants apparently held "mixed" or "synthetic" theories of evolution, reasoning about some phenomena on the basis of transformational principles and others on the basis of variational principles. This outcome, though perhaps less than ideal from an instructional point of view, mirrors that documented in domains like cosmology (Vosniadou & Brewer, 1992; Samarapungavan, Vosniadou, & Brewer, 1996) and physiology (Astuti & Harris, 2008; Legare & Gelman, 2008), where knowledge derived from intuition frequently conflicts with that derived from testimony.

Cognitive Constraints on Acceptance

Both before and after the instruction participants were asked to rate their agreement with five statements of belief: (1) "Species have changed over time"; (2) "The species in existence today have not always existed"; (3) "Natural selection is the best explanation for how species adapt to their environment"; (4) "Natural selection is the best explanation for the origin of new species"; and (5) "The origin of human beings does not require a different explanation than the origin of other species." Participants used a 5-point scale, with 1 indicating "strongly disagree;" 2, "disagree;" 3, "neutral;" 4, "agree," and 5, "strongly agree." The ordering of the statements was determined by their controversiality, as public opinion polls have shown that Americans are more accepting of microevolution than macroevolution and more accepting of nonhuman evolution than human evolution (Scott, 2005).

Because the focus of our study was on measuring understanding, not acceptance, our instrument for measuring acceptance was less comprehensive than those developed by other researchers (e.g., Rutledge & Warden, 1999). Nevertheless, participants' agreement ratings were highly correlated across the five statements of belief at both pretest ($r = 0.51$) and posttest ($r = 0.49$), implying that the various ratings reflected a single attitude or disposition toward the endorsement of evolutionary claims. Before instruction, participants' ratings averaged 4.2 across the five statements of belief ($SD = 0.7$); after instruction, they averaged 4.4 ($SD = 0.6$). This increase was statistically significant, though the magnitude of change (0.2) was small. Closer inspection of the data revealed 14 participants whose preinstructional agreement ratings were at ceiling and could not therefore have increased. Removing those participants from the sample yielded a mean pre-post difference of 0.4, which is equivalent to a 10% increase in agreement ratings. The magnitude of this increase is virtually identical to that documented by Ingram and Nelson (2006), which is noteworthy given that these authors used a different instrument for measuring acceptance, a different curriculum for teaching evolution, and a different participant sample (upper-level biology majors).

The number of participants who selected "agree" or "strongly agree" for each statement of belief are displayed in Figure 3.3, at both pretest and posttest. Two effects are observable. First, the number of participants who agreed with each statement decreased from statement 1 (about species change) to statement 5 (about human evolution), as predicted by the controversiality of statement content. Second,

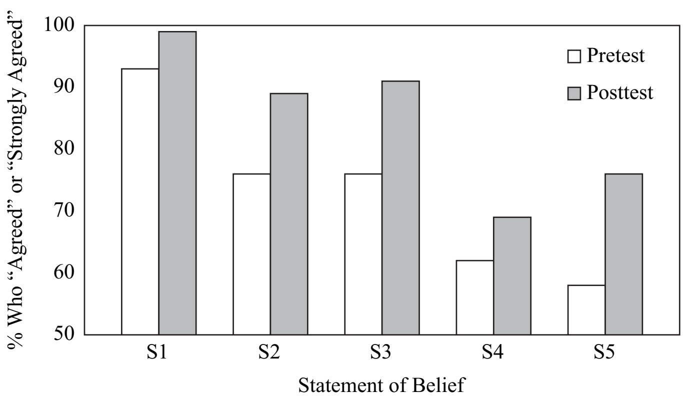

FIGURE 3.3 *Percent of participants who "agreed" or "strongly agreed" with the five statements of belief at pretest and posttest. Pre-post gains in percent agreement were statistically significant for S2, S3, and S5.*

the number of participants who agreed with each statement increased as a function of instruction (though the increase was statistically reliable only for statements 2, 3, and 5). The fact that instruction increased acceptance of statement 5 (“The origin of human beings does not require a different explanation than the origin of other species”) is particularly noteworthy, as it is this claim that Americans are least likely to endorse (Miller et al., 2006; Newport, 2004).

Although most participants agreed with most statements, there was still sufficient variation in participants’ agreement ratings (summed across the five statements) to compare them to their comprehension assessment scores. In contrast to previous studies that have found no correlation between understanding evolution and accepting evolution (Bishop & Anderson, 1990; Brem, Ranney, & Schindel, 2003; Demastes, Settlage, & Good, 1995; Lawson & Worsnop, 1992; Sinatra, Southerland, McConaughy, & Demastes, 2003), the present study found strong correlations between these two measures at both pretest ($r = 0.55$) and posttest ($r = 0.46$). In other words, variationists were more likely than transformationists to endorse the five statements of belief at both assessment periods. These correlations may have gone undetected in prior studies due to differences in how understanding was measured, how acceptance was measured, or both. They may also have gone undetected due to insufficient variation in those measures within the particular populations under investigation. That said, the present study is not the first, or the only, study to have documented correlations between understanding and acceptance. Similar findings have been obtained by Nadelson and Sinatra (2009), Nehm, Kim, and Sheppard (2009), and Rutledge and Warden (2000), all of which used different measures of understanding and different measures of acceptance than those used here.

Theoretical and Pedagogical Implications

Consistent with previous research (Shtulman, 2006; Shtulman & Schulz, 2008), participants in the present study demonstrated pervasive, preinstructional misconceptions of a transformational nature. Some of these misconceptions were corrected by instruction, and some were not. Although pre-post gains in assessment scores were modest in size, they were frequent in occurrence. A full 76% of participants increased their score by 1 or more points, 49% increased their score by 5 or more points, and 27% increased their score by 10 or more points. This rate of change is unprecedented in the evolution education literature (e.g., Bishop & Anderson, 1990; Demastes et al., 1995; Jensen & Finley, 1995), which implies that teaching interventions targeted at transformational misconceptions may be more successful than those that trace historical changes in evolutionary thought (Jensen & Finley, 1995) or those that focus strictly on natural selection (Demastes et al., 1995). This finding complements findings from other science education studies demonstrating that students’ misconceptions must be adequately addressed before they can be replaced by new, accurate conceptions (e.g., Moss & Case, 1999; Slotta & Chi, 2006; Smith, 2007; Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001; Wiser & Amin, 2001).

One intriguing aspect of the change in participants' understanding of evolution from before instruction to after is that participants made significant conceptual progress in five of the six areas tested. This finding was unexpected, as the teaching intervention focused primarily on microevolution and only briefly touched on such macroevolutionary topics as speciation and domestication. Apparently, participants were able to apply some of their newfound understanding of microevolutionary principles to their prior (inaccurate) understanding of macroevolutionary phenomena. This spontaneous transfer of information from one set of beliefs to another may have been a byproduct of their interrelatedness. In other words, the coherence in participants' preinstructional beliefs may have actually facilitated their revision (see Au et al., 2008, and Slaughter & Lyons, 2003, for a similar pattern of results).

Another intriguing aspect of the change in participants' understanding of evolution is that it was accompanied by a change in their acceptance of evolutionary claims—namely, increased understanding led to increased acceptance. This finding implies that Americans' skepticism toward evolution is rooted, at least in part, in a misunderstanding of what evolution is. Rather than construe evolution as the selective propagation of within-species variation, many appear to construe evolution as the uniform adaptation of all individuals within a species. This construal is not only incorrect but is also highly problematic for appreciating how biological phenomena bear on evolutionary claims and how evolutionary claims make sense of biological phenomena.

As an illustration, consider the recent discovery that humans share over 80% of their genes with mice (Waterston et al., 2002). This discovery is easily assimilated by a variationist, who sees species as continuums of variation related by common ancestry, but is not easily assimilated by a transformationist, who sees species as discrete entities characterized by unique, nonoverlapping essences. A transformationist must either recast mice as the evolutionary forbearers of humans (as done by Russell [2002], the *San Francisco Chronicle* reporter who asserted that "scientists have found a wealth of common chemistry between human beings and our tiny, four-legged ancestors") or downplay the importance of genes in determining a species' identity (as done by McKie [2001], the *London Observer* reporter who asserted that "environmental influences are vastly more powerful [than genetic influences] in shaping the way humans act").

This example highlights a particular means by which understanding might influence acceptance: evidential reasoning (Chinn & Brewer, 2001; Kuhn, 1991; Sa, Kelley, Ho, & Stanovich, 2005). Changes in how students understand a particular theory can lead to changes in how they evaluate data relevant to the theory, which, in turn, can lead to changes in how well they think the theory is supported by evidence. Because much of the evidence for evolution cannot be interpreted, let alone appreciated, without an understanding of natural selection, we suspect that critics of evolution are incapable of engaging with the very evidence they find "unconvincing." That said, we did not assess our participants' ability to evaluate novel evolutionary data or novel evolutionary claims, so it remains an empirical question whether the relationship between understanding and acceptance is indeed mediated by evidential reasoning. It has been shown, however, that acceptance of evolution is significantly correlated with an understanding of the nature of science, even when controlling for general interest in science and past science education (Lombrozo, Thanukos, & Weisberg, 2008).

In a similar vein, it should be noted that our findings, while implicating understanding evolution as an important influence on accepting evolution, do not implicate understanding as the *only* influence. Religious commitments are certainly an important influence as well (see Brem, et al., 2003; Evans, 2001; Miller et al., 2006; Poling & Evans, 2004). The influence of such commitments was evident in the present study from participants' agreement ratings for two statements of belief regarding explicitly religious matters: (1) "I believe in the existence of God" and (2) "I believe in the existence of souls." These ratings did not change as a function of instruction and, when averaged together, were negatively correlated with participants' mean agreement ratings for the five statements about evolution, both before instruction ($r = -.33$) and after ($r = -.32$). They were also negatively correlated with participants' assessment scores before instruction ($r = -.25$) and after ($r = -.33$), indicating that, throughout the duration of the study, religious participants were less likely than nonreligious participants to understand evolutionary concepts and accept them as valid.

Clearly, multiple factors influence an individual's acceptance of evolution. An understanding of evolution is, however, the main factor that science educators are charged with changing. Although different researchers hold different opinions on the question of whether science educators should advocate for evolution rather than merely explain it (Smith, Siegel, & McInerney, 1995), our own opinion is that fostering an acceptance of evolution is crucial to the long-term advancement of scientific literacy and scientific reasoning. Accordingly, we see the correlation between understanding evolution and accepting evolution as highly informative to those goals. Although more research needs to be done to determine why and how this relationship obtains, one straightforward implication of our findings is that improving evolution education in the United States could help to increase the U.S. public's acceptance of evolution to a level more typical of other first-world nations.

References

Andersen, H. C. (1844/1981). The ugly duckling. In L. Owens (Ed. & Trans.), *The complete Hans Christian Andersen fairy tales* (pp. 15–20). New York: Avenel Books.

Astuti, R., & Harris, P. L. (2008). Understanding mortality and the life of the ancestors in rural Madagascar. *Cognitive Science, 32,* 713–740.

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

Banet, E., & Ayuso, G. E. (2003). Teaching of biological inheritance and evolution of living beings in secondary schools. *International Journal for Science Education, 25,* 373–407.

Bishop, B., & Anderson, C. A. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching, 27,* 415–427.

Bohannon, J. (2008). Flunking spore. *Science, 322,* 531.

Bloom, P., & Weisberg, D. S. (2007). Childhood origins of adult resistance to science. *Science, 316,* 996–997.

Bowler, P. J. (1983). *The eclipse of Darwinism: Anti-Darwinian theories in the decades around 1900*. Baltimore: Johns Hopkins University Press.

Brem, S. K., Ranney, M., & Schindel, J. (2003). Perceived consequences of evolution: College students perceive negative personal and social impact in evolutionary theory. *Science Education, 87,* 181–206.

Brumby, M. N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education, 68*(4), 493–503.

Calabi, P. (1998). *Ecology: A systems approach*. Dubuque, IA: Kendall Hunt.

Calabi, P. (2005). *Teaching evolution by natural selection*. Unpublished manuscript.

Crawford, B. A., Zembal-Saul, C., Munford, D., & Friedrichsen, P. (2005). Confronting prospective teachers' ideas of evolution and scientific inquiry using technology and inquiry-based tasks. *Journal of Research in Science Teaching, 42,* 613–637.

Chinn, C. A., & Brewer, W. F. (2001). Models of data: A theory of how people evaluate data. *Cognition and Instruction, 19,* 323–393.

Demastes, S. S., Settlage, J., & Good, R. J. (1995). Students' conceptions of natural selection and its role in evolution: Cases of replication and comparison. *Journal of Research in Science Teaching, 32,* 535–550.

DeVries, R. (1969). Constancy of generic identity in the years three to six. *Society for Research in Child Development Monographs, 34* (Whole No. 127).

Evans, E. M. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology, 42,* 217–266.

Ferrari, M., & Chi, M. T. H. (1998). The nature of naïve explanations of natural selection. *International Journal of Science Education, 20,* 1231–1256.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26*(5), 796–804.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories over appearances. *Child Development, 58,* 1532–1541.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understanding of the nonobvious. *Cognition, 38,* 213–244.

Gould, S. J. (1996). *Full house: The spread of excellence from Plato to Darwin*. New York: Three Rivers Press.

Greene, E. D. (1990). The logic of university students' understanding of evolution. *Journal of Research in Science Teaching, 27,* 875–885.

Hickling, A. K., & Gelman, S. A. (1995). How does your garden grow? Early conceptualization of seeds and their place in the plant growth cycle. *Child Development, 66,* 856–876.

Hull, D. (1965). The effect of essentialism on taxonomy: 2000 years of stasis. *British Journal for the Philosophy of Science, 15–16,* 314–326.

Ingram, E. L., & Nelson, C. E. (2006). Relationship between achievement and students' acceptance of evolution or creation in an upper-level evolution course. *Journal of Research in Science Teaching, 43,* 7–24.

Jaswal, V. K., & Markman, E. M. (2007). Looks aren't everything: 24-month-olds' willingness to accept unexpected labels. *Journal of Cognition and Development, 8,* 93–111.

Jensen, M. S., & Finley, F. N. (1995). Teaching evolution using historical arguments in a conceptual change strategy. *Science Education, 79,* 147–166.

Jimenez, A. M. P. (1994). Teaching evolution and natural selection: A look at textbooks and teachers. *Journal of Research in Science Teaching, 31,* 519–535.

Johnson, S. C., & Solomon, G. E. A. (1997). Why dogs have puppies and cats have kittens: The role of birth in young children's understanding of biological origins. *Child Development, 68*(3), 404–419.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.

Lawson, A. E., & Thompson, L. D. (1988). Formal reasoning ability and misconceptions concerning genetics and natural selection. *Journal of Research in Science Teaching, 25,* 733–746.

Lawson, A. E., & Worsnop, W. A. (1992). Learning about evolution and rejecting a belief in special creation: Effects of reasoning skill, prior knowledge, prior belief and religious commitment. *Journal of Research in Science Teaching, 29,* 143–166.

Legare, C. H., & Gelman, S. A. (2008). Bewitchment, biology, or both: The co-existence of natural and supernatural explanatory frameworks across development. *Cognitive Science, 32,* 607–642.

Lombrozo, T., Thanukos, A., & Weisberg, M. (2008). The importance of understanding the nature of science for accepting evolution. *Evolution: Education and Outreach, 1,* 290–298.

Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*. Cambridge, MA: Harvard University Press.

Mayr, E. (2001). *What evolution is*. New York: Basic Books.

McKie, R. (2001, February 11). Revealed: The secret of human behavior; Environment, not genes, key to our life. *London Observer*, p. 1.

Medin, D., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review, 111*(4), 960–983.

Miller, J. D., Scott, E. C., & Okamoto, S. (2006). Public acceptance of evolution. *Science, 313,* 765–766.

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Science Education, 30,* 122–147.

Nadelson, L. S., & Sinatra, G. M. (2009). Educational professionals' knowledge and acceptance of evolution. *Evolutionary Psychology, 7,* 490–516.

Nehm, R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience, 57,* 263–272.

Newport, F. (2004). *Third of Americans say evidence has supported Darwin's evolution theory*. Princeton, NJ: Gallup Organization.

Poling, D. A., & Evans, E. M. (2004). Religious belief, scientific expertise, and folk ecology. *Journal of Cognition and Culture, 4,* 485–524.

Rosengren, K. S., Gelman, S. A., Kalish, C., & McCormick, M. (1991). As time goes by: Children's early understanding of biological growth. *Child Development, 62,* 1302–1320.

Russell, S. (2002, December 5). Of mice and men: Striking similarities at the DNA level could aid research. *San Francisco Chronicle*, p. A1.

Rutledge, M. L., & Warden, M. A. (1999). The development and validation of the measure of acceptance of the theory of evolution instrument. *School Science and Mathematics, 99,* 13–18.

Sa, W. C., Kelley, C. N., Ho, C., & Stanovich, K. E. (2005). Thinking about personal theories: Individual differences in the coordination of theory and data. *Personality and Individual Differences, 38,* 1149–1161.

Scott, E. (2005). *Evolution vs. creationism: An introduction*. Berkeley: University of California Press.

Samarapungavan, A., Vosniadou, S., & Brewer, W. F. (1996). Mental models of the earth, sun, and moon: Indian children's cosmologies. *Cognitive Development, 11,* 491–521.

Settlage, J. (1994). Conceptions of natural selection: A snapshot of the sense-making process. *Journal of Research in Science Teaching, 31,* 449–457.

Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology, 52,* 170–194.

Shtulman, A., & Schulz, L. (2008). The relationship between essentialist beliefs and evolutionary reasoning. *Cognitive Science, 32,* 1049–1062.

Sinatra, G. M., Southerland, S. A., McConaughy, F., & Desmastes, J. W. (2003). Intentions and beliefs in students' understanding and acceptance of biological evolution. *Journal of Research in Science Teaching, 40,* 510–528.

Slaughter, V., & Lyons, M. (2003). Learning about life and death in early childhood. *Cognitive Psychology, 43,* 1–30.

Slotta, J. D., & Chi, M. T. H. (2006). Helping students understand challenging topics in science through ontology training. *Cognition and Instruction, 24,* 261–289.

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,* 337–398.

Smith, C., Solomon, G. E. A., & Carey, S. (2005). Never getting to zero: Elementary school students' understanding of the infinite divisibility of number and matter. *Cognitive Psychology, 51,* 101–140.

Smith, M. U., Siegel, H., & McInerney, J. D. (1995). Foundational issues in evolution education. *Science and Education, 4,* 23–46.

Southerland, S. A., Abrams, E., Cummins, C. L., & Anzelmo, J. (2001). Understanding students' explanations of biological phenomena: Conceptual frameworks or p-prims? *Science Education, 85,* 328–348.

Springer, K. (1996). Young children's understanding of a biological basis for parent-offspring relations. *Child Development, 67,* 2841.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24,* 535–585.

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, F. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction, 11,* 381–419.

Waterston, R. H., Lindblad-Toh, K., Birney, E., et al. (222 coauthors). (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature, 420,* 20–562.

Waxman, S., Medin, D., & Ross, N. (2007). Folkbiological reasoning from a cross-cultural developmental perspective: Early essentialist notions are shaped by cultural beliefs. *Developmental Psychology, 43*(2), 294–308.

Wiser, M., & Amin, T. (2001). Is heat hot? Inducing conceptual change by integrating everyday and scientific perspectives on thermal phenomena. *Learning and Instruction, 11,* 331–355.