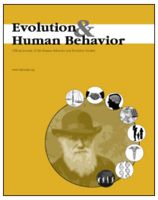
Evolution and Human Behavior xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org



A field guide for teaching evolution in the social sciences

Cristine H. Legare [a], John E. Opfer [b], Justin T.A. Busch [a], Andrew Shtulman [c]

[a] The University of Texas at Austin, United States
[b] The Ohio State University, United States
[c] Occidental College, United States

A R T I C L E I N F O

Article history:
Initial receipt 22 January 2017
12 January 2018
Final revision received 19 January 2018
Available online xxxx

Keywords:
Evolutionary social science
Educational assessment
Biological education
Cognitive biases
Evolution
Pedagogy
Science education

A B S T R A C T

The theory of evolution by natural selection has begun to revolutionize our understanding of perception, cognition, language, social behavior, and cultural practices. Despite the centrality of evolutionary theory to the social sciences, many students, teachers, and even scientists struggle to understand how natural selection works. Our goal is to provide a field guide for social scientists on teaching evolution, based on research in cognitive psychology, developmental psychology, and education. We synthesize what is known about the psychological obstacles to understanding evolution, methods for assessing evolution understanding, and pedagogical strategies for improving evolution understanding. We review what is known about teaching evolution about nonhuman species and then explore implications of these findings for the teaching of evolution about humans. By leveraging our knowledge of how to teach evolution in general, we hope to motivate and equip social scientists to begin teaching evolution in the context of their own field.

Contents

E-mail address: legare@austin.utexas.edu (C.H. Legare).

https://doi.org/10.1016/j.evolhumbehav.2018.01.002
1090-5138/

1. A field guide for teaching evolution

Social scientists are increasingly adopting an evolutionary perspective in how they study and describe human cognition and behavior (Boyd & Silk, 2015; Lewis, Al-Shawaf, Conroy-Beam, Asao, & Buss, 2017). The ability to teach evolution effectively should not be taken for granted. One reason is that few social science educators have deep knowledge of evolutionary biology. To our knowledge, Ph.D. programs in social science do not (yet) require coursework in evolution. Another reason is that evolution by natural selection is one of the most difficult scientific concepts for students to grasp (Rosengren, Brem, Evans, & Sinatra, 2012). Decades of research in cognitive psychology, developmental psychology, and science education have revealed that students regularly misunderstand what evolution is and how it occurs (Bean, Sinatra, & Schrader, 2010; Short & Hawley, 2014; Shtulman & Calabi, 2013; Shtulman & Schulz, 2008; Sinatra, Brem, & Evans, 2008; Sinatra, Southerland, McConaughy, & Demastes, 2003). Misunderstandings about the logic of evolutionary theory are rampant, which makes teaching it more difficult. For example, individuals who lack an understanding of evolution are less likely to accept it (Weisberg, Landrum, Metz, & Weisberg, in press). The objective of this paper is to provide social scientists with a field guide for teaching evolution. We synthesize what is known about the psychological obstacles to understanding evolution, methods for assessing evolution understanding, and pedagogical strategies for improving evolution understanding, with an eye toward informing the social science curriculum.

The field of evolutionary social science is quickly advancing, providing a more nuanced understanding of human cognition and behavior (Barrett, 2015; Bolhuis, Brown, Richardson, & Laland, 2011; Buss, 2015, 2016; Henrich, 2016; Wilson, 2007, 2015). We argue that integrating evolution more fully into the social science curriculum is long overdue. Our goal is to spur that integration by providing social scientists with a field guide on research on teaching evolution. First, we discuss obstacles to understanding evolution proper and then discuss how those obstacles might affect understanding the evolution of human cognition and behavior. Next, we discuss assessment of students' understanding and misunderstanding of evolution, as well as the possibility of adapting those assessments for use in the social sciences. Finally, we describe pedagogical techniques for teaching evolution in general and consider their strengths and weaknesses for teaching evolutionary social science. By leveraging our knowledge of how evolution can be taught successfully in a biological context, we hope to motivate and equip social scientists to begin teaching evolution in the context of their own field, addressing pedagogical questions specific to evolutionary social science along the way.

2. Obstacles to understanding evolution

Scientists overwhelmingly support the theory of evolution, with 98% agreeing that humans evolved over time whereas only 62% of the general U.S. population agrees with such a statement (Pew Research Center, 2014). The challenges associated with understanding evolution by natural selection are not exclusively the result of substantial popular resistance to scientific ideas on religious or other ideological grounds (Bloom & Skolnick Weisberg, 2007; Brem, Ranney, & Schindel, 2003; Evans, 2000a; Lombrozo, Shtulman, & Weisberg, 2006; Scott, 2004). Indeed, research shows that cultural factors such as religion and parental attitudes do not predict students' learning of natural selection (Barnes, Evans, Hazel, Brownell, & Nesse, 2017). Here we discuss the cognitive biases that pose substantial obstacles to understanding biological change (Evans, 2000b; Evans & Lane, 2011; Legare, Lane, & Evans, 2012; Shtulman, 2006; Sinatra et al., 2008). Among these are the essentialist tendency to view species as unchanging (Emmons & Kelemen, 2015; Evans, 2000a; Gelman, 2003; Herrmann, French, DeHart, & Rosengren, 2013; Mayr, 1982; Poling & Evans, 2002) and the teleological tendency to explain all kinds of natural phenomena by reference to purpose (Evans, 2001; Keil, 1992; Kelemen, 1999b). We also discuss the existential anxiety invoked by evolutionary theory and its implications fior accepting evolutionary explanations (Brem et al., 2003; Evans, 2000b; Evans, Legare, & Rosengren, 2011; Legare, Evans, Rosengren, & Harris, 2012; Legare & Visala, 2011; Tracy, Hart, & Martens, 2011).

2.1. Essentialism

Psychological essentialism is the belief that the members of a category (e.g., zebras) are united by a common essence, which determines the members' outwardly observable properties (e.g., their stripes, their hooves, their diet) (Gelman, 2003). Essentialist reasoning assumes that categories are stable (zebra babies grow into zebra adults) and immutable (once a zebra, always a zebra; Gelman & Rhodes, 2012, p. 8). Essentialist reasoning is largely incompatible with evolutionary theory. The idea that each species is undergirded by a separate, discrete essence is inconsistent with the idea that all extant life forms share a common ancestor (Mayr, 1982). Essentialist thinking about species likely reflects functional cognitive adaptations. The assumption that species are unchanging underlies many practical inferences in the biological world. Avoiding poisonous snakes or spiders, for example, requires no knowledge that modern snakes evolved from predecessor forms. Viewing them as having unchanging inherent properties that are hazardous to humans facilitates avoiding them. For all practical purposes, they are unchanging essences within human lifespans. Cognitive adaptations evolved to deal with problems that occurred in seconds, minutes, sometimes days, or occasionally months or years. We are less psychologically prepared to understand things that change gradually over hundreds of generations.

Essentialism also results in boundary intensification, which is incompatible with an evolutionary view of life. If species are perceived to be bounded, the relations among species can be difficult to discern, let alone the variation within a species (Shtulman & Schulz, 2008). To further compound the problem, essentialism is consistent with a need-based view of change, in which individual organisms develop traits based on their needs and then pass those traits to their offspring (Gelman & Rhodes, 2012; Ware & Gelman, 2014). It is true that *populations* of individuals do adapt to challenges of survival and reproduction, yet need-based explanations are insufficient for understanding population level variation and selection (Legare, Lane, & Evans, 2013).

2.2. Teleological reasoning

Evolution by selection involves two key components—blind chance variations (mutations), and selection by consequences. The first component is 'blind' in the key sense that it is not forward-looking, as in a watchmaker (or a God) designing something. In his autobiography, Charles Darwin states that he experienced "the extreme difficulty or rather impossibility of conceiving this immense and wonderful

Please cite this article as: Legare, C.H., et al., A field guide for teaching evolution in the social sciences, *Evolution and Human Behavior* (2018), https://doi.org/10.1016/j.evolhumbehav.2018.01.002

universe… as the result of blind chance" (Darwin, 1887, p. 92). Darwin struggled with the notion of blind chance because it contradicted the commonsense idea that everything exists for a purpose. Human-made objects—a guitar, for instance, exists for the purpose of making music—things like geological structures (rocks, mountains), weather patterns (wind, clouds), or whole organisms (animals, plants) do not exist for any external purpose. Young children, however, assume that anything and everything exists for a purpose. For instance, when children are asked to provide explanations for the properties of a non-living object—say, a rock—they inappropriately apply teleological reasoning, claiming, for instance, that rocks exist "so that animals could scratch on them when they got itchy." (Casler & Kelemen, 2008). This "promiscuous teleology" may emerge from a naïve theory of mind (Kelemen, 1999a), which attributes intentional origins to artifacts and is inappropriately applied to objects from the natural world (Evans, 2000a), or it may emerge through perceiving the interdependent relationships between species and assuming that these relationships were purposely forged (ojalehto, Waxman, & Medin, 2013).

Teleological thinking may reflect an important cognitive adaptation when applied to its proper domains. The component parts of organisms, such as the turtle's hard shell, the porcupine's sharp quills, and the skunk's noxious spray, do have purposes when this term is used to mean evolved functions (in these cases, specialized defenses against predators). Seeing functionality in the component parts of organisms can be useful in navigating the survival challenges posed by other species, as well as understanding their behavior (Opfer & Gelman, 2001). But teleological thinking poses challenges to accurately understanding evolution. Species did not evolve with any advanced foresight, but rather simply by natural selection favoring variants that successfully solved adaptive challenges better than other variants present in the population. Moreover, teleological thinking often assumes a designer with forward-looking goal-directed motivational properties, which contradicts the evolutionary logic of blind variation and selective retention.

Children may be more unrestrained than adults in their use of teleological reasoning, but adults also hold this cognitive bias (Rottman et al., 2017). Cross-cultural research has revealed that adults with minimal exposure to Western-style schooling express teleological explanations for the properties of natural objects about as often as American elementary school children who have yet to be exposed to extensive science education (Casler & Kelemen, 2008). Furthermore, adults suffering from Alzheimer's disease frequently endorse teleological explanations, explaining the existence of rain, for instance, by noting that rain provides water for animals to drink (Lombrozo, Kelemen, & Zaitchik, 2007). The authors suggest that the promiscuous teleology observed in children is not outgrown but rather "persists throughout life, reemerging when causal beliefs that might otherwise constrain it are limited or compromised" (Lombrozo et al., 2007, p. 1004; see also Kelemen, Rottman, & Seston, 2013; Shtulman & Harrington, 2016).

In sum, teleological reasoning is associated with three components with somewhat different implications for understanding evolution. First is that teleological reasoning is forward-looking. This is clearly inaccurate from an evolutionary perspective, because evolution has no foresight; it merely favors in each generation heritable qualities tributary to reproductive success. Second is that things have functions or purposes. This is right in some cases, for example, to say that a 'turtle has a shell to protect itself from predators' is a reasonable description of its evolved functionality; these are called 'adaptations.' Nonetheless, not all traits have evolved functions. And third is that functional explanations have a 'proper domain,' and teleological thinking often extends erroneously beyond a proper functional domain (e.g., to rocks and rain), which is also always incorrect.

2.3. Existential anxiety

Another challenge to achieving a comprehensive understanding of evolution is contemplating anxiety-provoking topics, like the violence inherent in nature (Zimmerman & Cuddington, 2007) or the extinction of one's own species (Legare & Visala, 2011). For example, thoughts about mortality have been shown to decrease endorsement of evolutionary explanations and increase support for intelligent design explanations, presumably because attitudes toward evolution are "shaped by unconscious psychological motives to maintain security and ward off existential angst through the cultivation of meaning and purpose" (Tracy et al., 2011, p. 12). Evolutionary theory also raises anxieties about human social relationships. Even those who endorse evolution believe that embracing it could have negative social consequences, including "increased selfishness and racism, decreased spirituality, and a decreased sense of purpose and self-determination" (Brem et al., 2003). In this vein, evolution constitutes a psychological threat, and psychologically threatening information tends to be processed in a biased, defensive manner (Hart, Shaver, & Goldenberg, 2005).

Other research suggests that people may adopt strategies to make evolutionary theory less existentially arousing. Research on explanatory coexistence indicates that people integrate evolutionary theory with religious explanations to explain multiple levels of causality (Legare et al., 2012). This type of integration can take three forms: *synthetic thinking*, in which evolutionary and religious explanations are combined (but not well integrated) into a causal chain; *target-dependent thinking*, in which evolutionary and religious explanations are used to account for distinct aspects of a given phenomenon and involve different kinds of causality; and *integrative thinking*, in which evolutionary and religious explanations are combined into a causal chain of proximate and distal causes in which God creates the conditions under which evolution can occur. By maintaining the existence of a supernatural creator, people are able to incorporate an empirically supported view of life without having to grapple with the existentially arousing topic of origins.

Another strategy for reducing the existential anxiety tied to evolution is to ascribe meaning to the process or products of evolutionary change (Rutjens, Van Der Pligt, & Van Harreveld, 2010). In one study (Tracy et al., 2011), participants read a passage by Carl Sagan arguing that purpose can be attained by embracing naturalism and "seeking to understand the natural origins of life." These participants were more likely to reject intelligent design theory and accept evolution than those who had not read the passage. Similarly, the same study reports that when mortality is salient, students majoring in the natural sciences are more likely to reject intelligent design theory because for them, "evolution is part of their understanding of the world and a source of meaning and purpose" (Tracy et al., 2011, p. 11).

2.4. Obstacles to understanding evolution in a social science context: Additional considerations

Evolutionary perspectives on cognition, behavior, and social organization have made substantial theoretical and empirical contributions to social science. At the same time, research on social scientists' attitudes toward evolution has uncovered several reasons why evolution is often excluded from the social science curriculum (Cabeza de Baca & Jordan, 2012; King & Cabeza de Baca, 2011; Perry & Mace, 2010; von Hippel & Buss, 2017). First, most social science educators receive little or no education in the fundamentals of evolutionary science and do not feel qualified to cover the concepts that a comprehensive understanding of evolution entails (e.g., variation, inheritance, selection, time, adaptation) (Evans, 2005). Others understand and accept the logic of evolution as applied to non-human species but have reservations about applying that logic to our own species. Still others may accept the logic of evolution as applied to human anatomy and human physiology but have reservations about applying that logic to human cognition or human social behavior, as if an ontological barrier exists at the neck, allowing evolutionary principles to apply to the human body but not the human brain and the psychological mechanisms housed in the brain (von Hippel & Buss, 2017).

Please cite this article as: Legare, C.H., et al., A field guide for teaching evolution in the social sciences, *Evolution and Human Behavior* (2018), https://doi.org/10.1016/j.evolhumbehav.2018.01.002

We suggest that there are two primary obstacles to integrating evolution into the social science curriculum. The first is the erroneous belief that learned behaviors are outside the scope of evolutionary explanation. Behavior can be both learned and evolved. Labeling something as learned does not, by itself, provide a satisfactory scientific explanation any more than labeling something as evolved does; it is simply the indisputable claim that environmental input changes the organism in some way. Learned and evolved are not competing explanations; rather, learning *requires* evolved psychological mechanisms—mechanisms which may be specific to a particular adaptive problem. Food aversion learning is an example of this dynamic. Clearly, there are specialized learning mechanisms to avoid eating toxic food (see Wertz & Wynn, 2014 for evidence of evolved aversion to plant consumption). Yet one is not born knowing which particular foods to avoid; this knowledge must be learned.

The second obstacle is that culture is seen as a competing explanation to evolution, most frequently when the trait in question varies across cultures or across development. Differences between groups are sometimes interpreted as evidence that culture alone shapes the human mind and that accounting for cultural variation obviates the need to seek evolutionary explanations. But truly satisfying cultural explanations identify the aspects of human cognition evoked by local social or ecological conditions. For example, cultures in which food resources show high variance evoke cooperative adaptations for group-wide sharing compared to those in which food variance is lower and more dependent on individual effort (Tooby & Cosmides, 1992). Understanding cultural variation, in short, requires understanding the evolved adaptations that are responsive to ecological and cultural input.

No research, to our knowledge, has explored the extent to which cultural learning is seen as distinct from—or opposed to—evolution and whether this impedes the teaching of evolution in a social science context. Research of this kind is needed to determine whether addressing the essentialist, teleological, and existential obstacles to understanding evolution is sufficient for teaching social science from an evolutionary perspective or whether an additional set of obstacles must be addressed as well.

An additional obstacle to learning evolution in the context of social science is a lack of understanding level of analysis (Buss, 1995). There are four levels to consider when testing evolutionary hypotheses. The first is evolutionary theory (e.g., natural selection and adaptation; modern genic selection). The second is *middle-level evolutionary theories* (e.g., Trivers's theory of parental investment). The third is *specific evolution-based hypotheses* (e.g., that derive from a middle-level theory; or based on an observation, such as higher child abuse in stepfamilies), and the fourth is *specific empirical predictions that test each hypothesis*. These distinct levels are often conflated. When scholars ask ‘What would falsify evolutionary theory?’ are they asking about levels 1, 2, or 3? Most scientific work does not test ‘evolutionary theory’ at level 1; most of the actual work is at levels 2, 3, and 4. A failure to appreciate these levels leads to inaccurate conclusions. For example, if an empirical finding falsifies a prediction (Level 4), based on a hypothesis (Level 3), it would not falsify ‘evolutionary theory’ in general. It would, however, call into question the level 3 hypothesis, and repeated level 3 failures would call into question the level 2 evolutionary theory. Relatedly, it is a mistake to assume there is one singular evolutionary hypothesis about any given phenomenon. In reality, there are competing evolutionary hypotheses, which is the normal state of science.

3. Strategies for teaching evolution

The obstacles outlined in the first section of the paper make the task of teaching and learning evolution a formidable challenge. Previous research on evolution understanding has documented misconceptions not only in novice biology students (e.g., Berti, Toneatti, & Rosati, 2010) but also in students who had taken multiple, college-level courses in biology, including college biology majors (Nehm & Reilly, 2007), medical school students (Brumby, 1984), pre-service biology teachers (Deniz, Donelly, & Yilmaz, 2008), and even doctoral students in biology (Gregory & Ellis, 2009). Despite years of intensive instruction, many students continue to harbor teleological and essentialist views of evolution that are logically incompatible with the principles of common ancestry and natural selection.

Shtulman and Calabi (2013) confirmed these findings in a longitudinal study of college students' understanding of evolution across a standard semester of biology instruction. Participants were recruited from six courses targeted to non-biology majors, and their understanding of evolution was assessed with Shtulman's (2006) 30-item instrument, described below. Prior to instruction, 235 of the 291 participants (or 81%) revealed more misconceptions about evolution than correct conceptions. Following instruction, 214 participants (or 74%) continued to reveal more misconceptions than correct conceptions. An analysis of individual response patterns revealed that, across courses, only 58 participants (or 20%) increased their evolution assessment score by a statistically reliable amount. The vast majority of participants (80%) left their courses with the same misconceptions they held upon entering those courses.

The objective of Shtulman and Calabi's (2013) study was not to test the efficacy of a particular intervention but to assess the effects of instruction in general. Their findings indicate that standard instruction is *not* effective. One reason that standard instruction is not effective is that complex concepts like natural selection and common ancestry are typically introduced in a single lecture or textbook chapter, with the rest of the course devoted to material predicated on these concepts but not illustrative of them. Moreover, concepts like common ancestry and natural selection are typically conveyed by definition rather than by more interactive forms of learning, such as inquiry, application, or analysis (Chi, 2009). As a result, students fail to develop a generative, mechanistic framework for understanding evolutionary phenomena and rely instead on a non-mechanistic (teleo-essentialist) framework for interpreting and encoding the subsequent course material.

While standard instruction has proven ineffective, other forms of instruction have proven more successful. Studies demonstrating the effectiveness of such instruction are reviewed below in terms of their key pedagogical innovation. Each study is accompanied by a measure of the difference between students' pre- and post-instructional scores on the study's chosen assessment of evolution understanding (computed as Cohen's d), for comparison's sake.

3.1. Refutation of pre-instructional misconceptions

One reason students fail to learn evolutionary principles from standard instruction is that they enter their biology classes with a host of misconceptions that are never explicitly addressed or refuted. Interventions that target those misconceptions have proven effective at instilling proper conceptions. Bishop and Anderson (1990), for instance, designed a curriculum in which correct conceptions were introduced only after participants (college undergraduates) completed activities that highlighted the inadequacy of their prior conceptions. Before the curriculum, around 25% of Bishop and Anderson's participants demonstrated a correct understanding of the assessment material. After the curriculum, 50% did—a significant, though far from complete, gain in conceptual understanding (Cohen's $d = 1.34$). Demastes, Settlage, and Good (1995) extended Bishop and Anderson's findings by supplementing their curriculum with additional inquiry-based activities, leading to a larger (32%) increase in conceptual understanding from pretest to posttest (Cohen's $d = 1.50$).

Adopting a slightly different approach, Jensen and Finley (1995) confronted college students' pre-instructional misconceptions with a curriculum that traced the history of evolutionary thought from Lamarck to Darwin to the modern synthesis. Participants (college undergraduates) were taught Lamarck's theory of evolution, evidence against Lamarck's theory, Darwin's theory, and evidence in support of Darwin's theory. Lamarck's theory was taught prior to Darwin's on the assumption that Lamarck's theory would resemble the inaccurate views of evolution students brought with them to the class. Jensen and Finley measured participants' understanding of evolution using a combination of multiple-choice questions and essay questions, covering seven evolutionary concepts. Participants' assessment scores increased by 22% from pretest to posttest (Cohen's $d = 1.52$). Articulating and confronting students' misconceptions prior to introducing correct conceptions may help students avoid assimilating the correct conceptions into a teleo-essentialist framework.

3.2. Extended illustration of evolutionary principles

Because evolutionary principles are foreign to everyday experience, students may need additional support connecting those principles to their prior knowledge and prior experiences. One way to forge these connections is by illustrating evolutionary principles with realistic, in-depth case studies of evolutionary change. Adopting this approach, Spiegel et al. (2012) introduced participants to four evolutionary concepts (variation, inheritance, selection, and time) illustrated in each of seven case studies of evolution (e.g., sexual selection in Hawaiian flies, the coevolution of ants and fungi, the rapid evolution of HIV). Participants were visitors to a natural history museum, and their understanding of evolution was assessed before and after the intervention with a series of Likert-scale items in which participants rated their agreement with several explanations of adaptations, some correct and some incorrect. While the intervention had little effect on participants' pre-instructional misconceptions, it still yielded significantly higher endorsement of correct conceptions (Cohen's $d = 0.48$).

Kelemen, Emmons, Schillaci, and Ganea, (2014) also taught participants evolutionary principles in the context of case studies, but they focused on a single, extended example of evolutionary change rather than multiple, brief examples. The participants in their study were five to eight-year-old children, who were taught evolution in the context of a picture book. The picture book introduced participants to a fictional animal—the elephant-like "pilosa"—and illustrated how pilosas evolved from having predominantly thick trunks to predominantly thin trunks. Participants were shown that pilosas varied in their trunk thickness (variation), that pilosas with thin trunks were better able to access food than pilosas with thick trunks (resource limitation), that pilosas who ate more food lived longer (differential survival), that pilosas who lived longer had more babies (differential reproduction), that the babies inherited their parents' thin trunks (inheritance), and that this process led to an increase in the proportion of pilosas with thin trunks over multiple generations (population change).

Participants' understanding of this causal sequence was assessed by their ability to incorporate the illustrated principles into their explanations of adaptation for animals other than pilosas before and after the intervention. Before the intervention, most children (84%) exhibited no understanding of evolution; after the intervention, approximately half exhibited some level of understanding, though the sophistication of that understanding varied (Cohen's $d = 0.53$). These results have been replicated by Shtulman, Neal, and Lindquist (2016), who were equally successful at teaching elementary-school-aged children the logic of natural selection (Cohen's $d = 1.63$).

In a third variant of the case-study method, Heddy and Sinatra (2013) taught evolution to college undergraduates using a "teaching for transformative experiences" curriculum. This curriculum emphasized three facets of learning: active use of a concept, expansion of perception, and experiential value. Each facet was instantiated in multiple, student-generated examples, ranging from the predatory behavior of polar bears to the extinction of dodo birds. Participants' understanding of evolution was measured with a 14-item, multiple-choice assessment covering six evolutionary concepts (variation, inheritance, adaptation, domestication, speciation, and extinction). On this measure, participants increased their score by an average of 6.3 points from pretest to posttest (Cohen's $d = 1.52$).

3.3. Collaborative problem solving

Another empirically successful approach to teaching evolution is involving students in joint problem-solving activities. In one study, Nehm and Reilly (2007) involved college biology majors in activities where they analyzed data or provided explanations relevant to the concepts of variation, inheritance, genetics, biomechanics, biodiversity, ecology, speciation, extinction, common ancestry, and natural selection. Participants' understanding of evolution was assessed by asking them to define evolutionary terminology and explain specific instances of adaptation, which were coded for evidence of several key concepts. Across assessment items, participants incorporated approximately two more key concepts into their explanations and definitions after instruction compared to before (Cohen's $d = 0.85$).

Asterhan and Schwarz (2007) achieved comparable success by involving college undergraduates in a collaborative argumentation task. Participants generated explanations for two instances of adaptation before and after collaborating with a partner on explaining a different set of adaptations. Participants were instructed to generate their own explanations first and then critically evaluate their partners' explanations. Emphasis was placed on providing evidence in support of one's preferred explanation and counterevidence against unfavorable explanations. On average, participants cited 2.4 evolutionary principles at pretest and 4.3 evolutionary principles at posttest—a significant improvement (Cohen's $d = 0.92$). Collaboration was effective in this instance for potentially many reasons: participants were required to articulate their pre-instructional beliefs; they were required to justify those beliefs; they were confronted with evidence against those beliefs; and they were exposed to alternative, potentially more accurate beliefs. In this way, collaboration yields many of the same benefits as those yielded by the instructional strategies reviewed above but does so in a more ecologically valid—and socially motivating—context (see also Shtulman & Checa, 2012).

3.4. Teaching evolution in a social science context: Additional considerations

The studies reviewed above indicate that, although evolutionary concepts are difficult to convey with standard instruction (e.g., a single lecture, a single textbook chapter), they can be conveyed more effectively with interventions that target pre-instructional misconceptions, that provide extended illustration of correct conceptions, or that encourage debate among students with varying levels of understanding. This research, which was conducted in the context of a biology curriculum, certainly has applications to teaching evolution in other contexts, but the specifics of those contexts may matter as well. In a social science context, the target organisms are humans, rather than nonhumans, and the target traits are often psychological, rather than physical. Additionally, social science students may be predisposed to view evolutionary accounts of human behavior as incompatible with sociocultural accounts, as noted above. Below we consider four questions pertaining to the teaching of evolution in a social science context, the concerns they raise, and possible avenues for further investigating those concerns.

3.4.1. Which evolutionary principles are most critical to cover?

There are three big evolutionary ideas that can inform the instruction of social science material: natural selection, adaptationism, and common ancestry. Each idea constrains students' understanding of the target material in different ways and can potentially be taught independent of the others. Indeed, teaching all three would be difficult given that time spent on evolution is time *not* spent on the primary social science content. It is thus an open question as to which evolutionary ideas are most critical to cover and for what purpose.

Natural selection, or the differential survival and reproduction of some organisms in a population relative to others by virtue of differences in heritable traits, is the primary mechanism of evolution. It is also important to understand modern genic selection, the foundation of modern evolutionary theory–that is, genes that have effects that increase their own replicative success relative to competing genes. Teaching social science students about natural selection would provide them with a mechanistic understanding of the origin of human traits and, accordingly, a means of discriminating empirically-testable evolutionary claims from untestable ones and theoretically-plausible evolutionary claims from implausible ones. But the research reviewed earlier in the paper indicates that teaching students about natural selection is not trivial. It requires covering several other concepts—e.g., trait variation, trait inheritance, resource limitation—and social scientists would need to devote sufficient time to covering the entire suite.

Adaptationism is the analysis of existing biological forms in terms of their survival-enhancing or reproduction-enhancing functions using criteria such as reliability, efficiency, and specificity of design for a particular function (Williams, 1966). This type of analysis does not inherently require an understanding of the mechanisms of evolution or the relations among different biological kinds and may thus be the easiest principle to teach on its own. The benefit of doing so would be to help students look beyond proximal sources of human behavior (e.g., experience, instruction, enculturation) and identify distal sources, namely, selection pressures in ancestral environments. The danger of doing so, however, is encouraging students to take an adaptationist stance without proper consideration of the mechanisms that could—or could not—have given rise to the hypothesized adaptation (e.g., explaining grammatical differences between English and Mandarin in terms of evolved differences between English speakers and Mandarin speakers). Adaptationism, if presented in isolation, may also lead to confusion regarding the origin of non-adaptive traits (e.g., the human tailbone) or maladaptive traits (e.g., the blind spot in the vertebrate eye). The logic of adaptationism should ideally be presented using criteria for invoking adaptation (Tooby & Cosmides, 1992; Williams, 1966), as well as factors that constrain adaptations and lead to non-optimal design, such as developmental constraints, lack of available genetic variation, tradeoffs with other adaptations, and time lags (Dawkins, 1982).

Common ancestry is the idea that the lineages of any two biological kinds (organisms, species, families, etc.) can be traced back to the same ancestor at some point in their evolutionary past and that these two kinds share many of the traits present in that ancestor. Teaching social science students about common ancestry would allow them to appreciate how experiments with nonhuman organisms are relevant to humans (e.g., how fear-conditioning experiments in mice relate to human fear responses or how spatial-navigation experiments with birds relate to human spatial memory). It might also foster an appreciation of the larger historical context in which evolution unfolds. Depth perception, for instance, is a trait that would be valuable not just to humans but to any organism with eyes. That said, common ancestry may be less relevant in courses where nonhuman comparisons are either nonapplicable or unknown.

3.4.2. What are good case studies for teaching the evolution of human cognition and behavior?

Whichever evolutionary principles a social science instructor choses to emphasize, there are further questions about which case studies are well suited to illustrate those principles. One question is whether the target trait is unique to humans (e.g., bipedalism, language, cooking, prolonged tolerance to lactose) or is shared with other organisms as well, and if it is shared with other organisms, how widely. One could potentially pick a trait shared with most other primates (e.g., coalitional psychology, dominance hierarchies), a trait shared with most other mammals (e.g., mating strategies, parenting strategies), or a trait shared with most other animals (e.g., depth perception, spatial navigation).

One could also pick a trait shared only by animals that occupy the same ecological niche or face the same survival problems, even though those traits are likely to have arisen independently in the target organisms (e.g., tool use in octopuses and humans, prolonged childhood in corvids and humans). Focusing exclusively on human-specific traits may help emphasize the fact that human behavior has evolutionary consequences and that human traits have been shaped by evolutionary pressures. But doing so may understate the phylogenetic context of those traits—i.e., that humans share most of their traits with other organisms (common ancestry) and that human-specific traits are just variants of the traits possessed by those other organisms (descent with modification).

Even among human traits, there are questions of how widespread the exemplar traits should be. One could focus either on traits common to all humans (e.g., social cognition) or on traits that vary (e.g., higher levels of physical aggression and risk taking in males). The latter, concerning human *differences*, raise larger issues that may be desirable to address in the course at hand. For instance, how might those differences have originated if not by evolution? Which among the competing modes of evolutionary analysis furnish the greatest heuristic and predictive value? Are individual differences in proclivity toward cooperation caused by different cultural environments activating or deactivated cooperative adaptations? Or a history of selection that favors heritable proclivities toward cooperation in some ecologies more than others? Most students are not prepared to answer such questions on their own, which makes broaching the topic of evolved differences in human traits both an opportunity and a challenge (Buss, 2009a).

Orthogonal to the question of how widespread a trait is (either among humans or among animals more generally) are questions about its function. Some traits are more patently adaptive than others, and this difference may have consequences for learning. Patently adaptive traits may be useful in developing form-function reasoning, whereas other traits—e.g., vestigial traits, maladaptive traits, exaptive (byproduct-like) traits—may be more useful in developing an appreciation of the historical contingencies of evolutionary change or the ubiquity of common ancestry (see Buss, 2009b). Traits also vary in the degree to which they are under conscious control. Traits over which humans appear to exhibit control (e.g., mating strategies, parenting strategies) may be useful in highlighting how proximal influences on human decision-making, operative in the immediate environment, relate to more distal ones, operative over evolutionary history. On the other hand, more reflexive traits (e.g., yawning, shivering, sneezing, blushing) may be useful in highlighting “hard” constraints on human biology—constraints that clearly were not learned through culture or experience.

Finally, traits also vary in their moral valence, and selecting morally reprehensible traits (e.g., dispositions toward infanticide, dispositions toward rape) could have strong consequences on students' engagement with the material. On one hand, such traits may turn students away from an evolutionary perspective if they perceive that perspective as justifying immoral behavior, but on the other hand, they may capture students' attention and motivate critical analysis. Such traits are also useful in highlighting the difference between descriptive, empirical claims about human nature (what humans *tend* to do) and prescriptive, value-laden claims about human nature (what humans *ought* to do).

3.4.3. Do the principles or cases need to be covered in a particular order?

Some evolutionary principles could be introduced through nonhuman examples before moving to humans. Doing so would constitute a “backdoor” approach to confronting preconceptions about human uniqueness, in that students would learn the evolutionary logic of a particular trait in nonhuman organisms (e.g., mating strategies of peafowl) before considering similar traits in humans. This approach could backfire, however, if the similarities are compartmentalized (e.g., if students

Please cite this article as: Legare, C.H., et al., A field guide for teaching evolution in the social sciences, *Evolution and Human Behavior* (2018), https://doi.org/10.1016/j.evolhumbehav.2018.01.002

are willing to concede cross-species similarities in mating strategies but nothing else) or if they are rejected altogether. This approach could also backfire if the point of the example was to highlight *differences* between humans and nonhuman organisms, as students might come to associate differences in outcome (e.g., differences between humans' and chimpanzees' tolerance for unfamiliar conspecifics) with differences in process (e.g., differences in how humans and chimpanzees acquired their social instincts—one by culture and the other by evolution). One compromise would be to highlight examples of *coevolution* between humans and other species (e.g., horses, dogs, lice, bacteria) on the assumption that such examples would make it difficult to accept evolutionary explanations for the nonhuman trait but reject evolutionary explanations for the corresponding trait in humans.

Questions of how to order'teaching examples arise even within the realm of exclusively human examples. For example, should the evolution of human universals be discussed before the evolution of human differences? Should the evolution of physical traits be discussed before the evolution of psychological ones? The evolution of psychological traits tends to be more controversial and thus subject to greater scrutiny regarding evidence and evidential standards—a set of issues that may be too complex for an introductory course but desirably complex for an advanced course (Buss, 2009a). One way to negotiate this complexity would be to start with a less controversial trait and then move to a more controversial one—e.g., starting with a human universal (e.g., the evolution of parasite avoidance) and then moving to a corresponding human difference (e.g., the evolution of gender-specific mating preferences related to fertility cues) or starting with a physical trait (e.g., the evolution of the hippocampus) and then moving to a corresponding psychological trait (e.g., the evolution of spatial navigation). A downside to this strategy, however, is that it might signal to students that the latter trait in each pair is less subject to evolutionary considerations.

There are, of course, sequencing constraints imposed by the material itself. It may be wise to explain natural selection before explaining other types of selection (e.g., sexual selection, kin selection, artificial selection), to explain the origin of adaptive traits before explaining the origin of other types of traits (e.g., exapted traits, epigenetic traits, vestigial traits), and to explain common ancestry before explaining other macro-evolutionary phenomena (e.g., interspecies homologies, gene-culture coevolution). On the other hand, introducing students to these “advanced” concepts may help clarify the meaning of more basic concepts. For example, introducing students to the idea that taste preferences for fat and sugar can be either adaptive or maladaptive, depending on the abundance of fat and sugar in one's environment, may help clarify the meaning of “adaptive.” Either way, every case of evolutionary change touches on several, interrelated concepts, and it remains an open question how to sequence those cases without presupposing unfamiliar concepts or raising uninformed objections, particularly with respect to cases of human evolution.

3.4.4. Should evolution be treated as an independent perspective or as a unifying framework?

In the last few decades, many social scientists have begun to embrace evolution as a valid perspective for analyzing social science phenomena, but treating evolution merely as a “perspective” may not be enough. Evolution is a unifying framework for the biological sciences, and many social scientists (e.g., the authors of this paper) have argued that it should be a unifying framework for the social sciences as well. But teaching evolution as a unifying framework requires a complete overhaul of the standard curriculum in many social science disciplines. Two examples of such an overhaul are Pinker's (1997) “How the Mind Works” and Gray and Bjorklund's (2014) “Psychology,” both of which introduce social science phenomena in terms of the selection pressures and survival problems that give rise to those phenomena. Most introductory social science texts do not go this far. Instead, they treat evolution as one of several equally valid and equally important perspectives. For instance, Myers's (2012) highly popular “Psychology” presents evolution as one of seven perspectives—along with neuroscience, behavioral genetics, psychodynamic theory, behaviorism, cognitivism, social/cultural systems—and identifies the evolution perspective mainly with the analysis of human differences (e.g., gender differences in mating preferences, gender differences in aggression).

To be fair, the “perspectives” approach is an accurate reflection of the methodological diversity of social science research. Few social scientists directly test evolutionary claims, even if they value an evolutionary approach in general, and introductory texts tend to emphasize the different kinds of methods that social scientists employ (e.g., developmental methods, neuroscientific methods). But there is no intrinsic reason why the material has to be organized this way, and doing so may have negative implications for learning. In particular, the ‘perspectives’ approach may imply that many human traits are *not* evolved and that evolutionary explanations for human traits are mutually exclusive from non-evolutionary ones. On the other hand, a “framework” approach may be difficult to achieve for many social science topics, given the current lack of data in support of specific evolutionary accounts. Ultimately, it remains an empirical question which approach is more pedagogically useful. Whereas the perspectives approach may gain traction because it is seen as less ideologically threatening, the framework approach makes it clear that evolutionary explanations are not “optional”—that all traits have a basis in evolution and can be analyzed accordingly.

4. Methods for assessing evolution understanding

In the previous section, we reviewed research on how to teach evolution, but what do students actually know about evolution? We address this question next, reviewing research on methods for assessing students' understanding of evolution in general and considerations for adapting those assessments to a social science context.

4.1. Assessment characteristics

Current measures of evolution understanding vary considerably in how their items are formatted (e.g., closed-response items versus open-response items), how scoring occurs, which evolutionary concepts are included, and which exemplars are used to illustrate those concepts. Each of these features is important because each is associated with varying levels of competence among novice reasoners.

4.1.1. Item format: closed- vs. open-response

One of the primary ways that tests of evolution differ is in their use of closed-response items (such as multiple-choice questions) relative to open-response items (such as essay questions). Closed-response items are popular for their efficiency and reliability. When developed through iterative rounds of pilot work, they can have as much construct validity as open-response items (Rodriguez, 2003). Despite these benefits, some researchers argue that closed-response items are unable to assess the depth of students' knowledge or their ability to synthesize information (Martinez, 1999; Popham, 2010). They also argue that closed-response items are poor predictors of real-world scientific reasoning (Nehm & Ha, 2011; NRC 2001).

Open-response tests, in contrast, permit students to express ideas comprised of both correct and incorrect elements, which is a common feature of students' progression to mastery of a domain (Alonzo & Gotwals, 2012; Chi, 2006; Nehm & Ridgway, 2011; Vosniadou & Brewer, 1992). Consequently, open-response tests typically elicit more graded levels of cognitive activity during problem solving than closed-response tests (Martinez, 1999). Perhaps for this reason, performance on open-response tests of evolution has been found to have greater correspondence to clinical interviews than closed-response tests (Nehm & Schonfeld, 2008). Due to advances in automatic scoring, rule-based text analysis and supervised machine learning have been used to code open-

Please cite this article as: Legare, C.H., et al., A field guide for teaching evolution in the social sciences, *Evolution and Human Behavior* (2018), https://doi.org/10.1016/j.evolhumbehav.2018.01.002

response items on evolution tests automatically (Moharreri, Ha, & Nehm, 2014).

4.1.2. Concepts inventoried: normative vs. non-normative

An important characteristic that varies among tests involves the specific concepts of evolutionary theory that are inventoried, as well as the types of misconceptions that are assessed. For example, almost all tests index whether students grasp three quantitative facts that are arguably necessary and sufficient for evolutionary change to occur—that a population varies in the presence or absence of a trait (*variability*), that the probability of the trait being passed from parent to offspring is greater than zero (*heritability*), and that individuals possessing the trait have a higher probability of reproductive success than other individuals (*differential selection*). Beyond these three causal, normative concepts, some tests additionally index non-causal normative concepts, such as *competition*, *hyperfecundity*, *scarcity of resources*, *evolution as change in phenotypic frequency*, and the scientific concept of *adaptation*.

In addition to assessing normative ideas about evolution by natural selection, tests also assess a number of common misconceptions, such as *essentialism* and *teleology*. These misconceptions can be assessed either directly or indirectly. For example, essentialist reasoning has been assessed directly by whether students underestimate the degree of variability in a trait on a closed-response item (Shtulman & Schulz, 2008) or indirectly by whether they use generic labels in their explanations for evolutionary change, for example, “the elm” in “the elm mutated to have long-winged seeds” (Opfer, Nehm, & Ha, 2012). Likewise, teleology has been assessed directly by whether students endorse teleological explanations in cases where they should not (Kelemen et al., 2013) or indirectly by whether students use “need-based” reasoning to explain the origin of traits (Southerland, Abrams, Cummins, & Anzelmo, 2001). Additional misconceptions that have been coded include *intentionality biases*, misuse of *pressure* or *force* as a cause of new traits appearing in individuals, and *use*/*disuse* as a cause of the gain or loss of traits.

4.1.3. Exemplars inventoried: breadth vs. depth

Another important characteristic of assessments is the *breadth* of exemplars used to illustrate or probe understanding. The crux of biology education, after all, is to foster effective evolutionary reasoning across all branches on the tree of life, not just for a few disparate twigs. Clough and Driver (1986) were among the first to explore item context effects in novice's evolutionary explanations, documenting fairly substantial consistency in the use of normative ideas, but not misconceptions. More recently, isomorphic open-response items that differed only in taxon (e.g., plant vs. animal), exemplar familiarity (e.g., penguin vs. prosimian), and type of evolutionary change (e.g., gain vs. loss) have been shown to elicit markedly different scores in novice students'—but not experts'—evolutionary knowledge and misconceptions (Nehm & Ha, 2011; Opfer et al., 2012; see also Heredia, Furtak, & Morrison, 2016). These studies highlight the importance of providing a wide range of exemplars (across taxa, familiarity, and gain/loss) to assess progress toward expert levels of evolutionary reasoning.

4.2. Current assessments: design choices, reliability, and validity

In this section, we review the characteristics of currently published assessments of evolutionary thinking. Our review is arranged chronologically to highlight how the development of new assessment instruments have attempted to build on the foundational work provided by their predecessors.

Bishop and Anderson (1985) provided the first standardized test of evolution understanding and the only test for nearly two decades. The format of their test used a hybrid open- and closed-response format. Two open-response items ask students to describe how a biologist would explain the evolution of a trait in a population, with one item concerning the gain in a trait (i.e., running speed in cheetahs) and another concerning the loss of a trait (i.e., sight in cave salamanders). The closed-response items include a Likert-style graded response and a follow-up open-response. For example, when asked “If a population of ducks were forced to live in an environment where water for swimming was not available,” students grade the likelihood of two different scenarios (“Many ducks would die because their feet were poorly adapted to this environment” vs. “The ducks would gradually develop non-webbed feet”) and are asked to provide an open-ended explanation for their responses.

As the first standardized test of evolution, the Bishop and Anderson (1985) assessment was highly successful. It was rapidly adopted in multiple training studies that sought to improve evolution understanding. Their test also revealed that understanding how evolution occurs is not the same as believing that evolution does occur. Indeed, belief in evolution had little impact on students' use of natural selection as a tool for understanding evolutionary change. As Bishop and Anderson (1990) observed, “Most students who believed in the truth of evolution apparently based their beliefs more on acceptance of the power and prestige of science than on an understanding of the reasoning that had led scientists to their conclusions.”

Despite its initial success, several undesirable features of the assessment became clear over the next twenty years. Researchers have reported inconsistent inter-item reliability scores (Asterhan & Schwarz, 2007; Bishop & Anderson, 1990; Demastes et al., 1995; Jensen & Finley, 1995), failure of the closed-response items to correlate very highly with open-response items (Nehm & Schonfeld, 2008), and failure of closed-response items to predict third assessments (such as oral interviews). As a result, the original Bishop and Anderson (1985) test fell out of favor among researchers, who developed novel tests aimed at avoiding these weaknesses.

Anderson, Fisher, and Norman (2002) introduced a novel assessment that was more comprehensive in scope, easier to grade, and subject to multiple tests of reliability and validity. The Conceptual Inventory of Natural Selection (or CINS) is a 20-item, multiple-choice test. For each question, students are asked to select among four options, with one option indexing the correct use of an important concept from evolutionary theory and at least one option indexing a common misconception related to that concept. The concepts covered are *variation*, *origin of variation*, *heritability*, *differential survival*, *hyperfecundity*, *population equilibrium*, *scarcity of resources*, *competition*, *evolution as change in phenotypic frequency*, and the *origin of species*. Anderson et al. (2002) reported acceptable levels of internal reliability, face validation by content experts, and acceptable ability to discriminate between low and high achievers for 14 of the 20 items. Similar results were reported by Nehm and Schonfeld (2008), who also reported that the ability of the CINS to index normative concepts tended to be greater than that of the original Bishop and Anderson (1985) open-response items, whereas the ability of the CINS to index misconceptions was lower than that of the original Bishop and Anderson (1985) assessment. Additionally, the clustering of both key concepts and misconceptions did not correspond to any hypothesized pattern of reasoning. As a result, future assessments have tried to address wider patterns or themes in students' misconceptions of evolution.

One such assessment was developed by Shtulman (2006). This assessment sought to distinguish between two theoretically meaningful patterns of evolutionary reasoning—the *variational* reasoning of Darwin, in which adaptation is the result of a non-random selection of randomly generated variation, versus the *transformationist* reasoning of the biological essentialist, in which adaptation is the result of a transformation in a species' underlying nature or "essence." Put differently, variational reasoning (correctly) interprets evolution as the selective propagation of within-species variation, whereas transformational reasoning (incorrectly) interprets evolution as the holistic transformation of an entire species, akin to metamorphosis. The purpose of Shtulman's assessment was to detect the coherence among these two, logically distinct views of evolution—coherence undetected in previous assessments. Indeed,

Please cite this article as: Legare, C.H., et al., A field guide for teaching evolution in the social sciences, *Evolution and Human Behavior* (2018), https://doi.org/10.1016/j.evolhumbehav.2018.01.002

cluster analysis of item responses supported this view, suggesting three major types of reasoning: variational, pre-variational, and transformational. This discovery is important because it suggests a generative framework for assessing a students' place in the progression from novice to expert, and a subset of the assessment has been used successfully to track changes in a recent training study (Heddy & Sinatra, 2013).

Shtulman's assessment presents students with scenarios organized around six key evolutionary phenomena, including *variation*, *inheritance*, *adaptation*, *domestication*, *speciation*, and *extinction*. Scoring of open-response items required trained coders, whereas closed-response items did not. The 30 items probe understanding of evolutionary phenomena in both plant and animal populations, including bacteria, corn, moths, woodpeckers, chimpanzees, and humans, as well as drawing analogies to population level changes in non-biological traits (e.g., SAT scores at a private school). Evaluations of the reliability and validity of Shtulman's assessment are limited, but expert evolutionary scientists were unequivocally identified as variationists by the test and respondents identified as transformationists were more likely than variationists to use evolutionarily inappropriate language (e.g., teleological language) in their responses.

Evans et al. (2010) published an assessment that, like Bishop and Anderson (1990), presented participants (museum visitors) with an open-response test that asked participants to explain how a trait evolved from an ancestral population to a novel population. Of interest was the participants' use of central concepts in evolutionary theory. Like Shtulman's (2006) test, the Evans et al. (2010) battery explicitly targeted thematic patterns of misconceptions, such as essentialist and teleological reasoning.

The Evans et al. (2010) scale had a similar format, asking how a new species arose from an ancestral species. Scoring required trained coders. Explanations of evolutionary change were coded for use of *any* valid evolutionary concept, as well as eight specific concepts (*variation*, *heritability*, *selection*, *common descent*, *time*, *chance*, *sexual selection*, and *ecological pressure*). Misconceptions coded included naturalistic errors (e.g., essentialism) as well as creationist misconceptions, such as *deity intervention*, *intelligent design*, *rejection of geological time*, and *rejection of common descent*. Finally, the test probed a wide variety of biological populations, including viruses, diatoms, ants, flies, finches, whales, and humans, thereby assessing patterns of evolutionary reasoning that cut across diverse life forms.

The Assessing Contextual Reasoning about Natural Selection, or ACORNS, tool was developed by a cross-disciplinary team, including an evolutionary biologist, cognitive scientist, science education researcher, statistician, and computer scientist (Opfer et al., 2012). The goal of the ACORNS design team was to respond to the problems discovered in the CINS and Bishop and Anderson (1985) battery, as well as to incorporate discoveries from the cognitive sciences made by Shtulman (2006) and Evans et al. (2010). Unlike previous assessments, the ACORNS instrument was designed to fix the syntax of open-response questions about evolutionary change while letting the item characteristics vary according to the user's needs. Specifically, all questions in the ACORNS instrument identify a particular species with a particular trait and then ask how biologists would explain the evolution of that species from an ancestral form that did not possess the trait. For example, "A species of labiatae (plants) is known to have pulegone. How would biologists explain how this labiatae species with pulegone evolved from an ancestral species that had no pulegone." This template permits open-ended responses to be machine scored via EvoGrader (Moharreri et al., 2014) and a potentially infinite number of items.

A subset of items from the ACORNS was introduced by Opfer et al. (2012). These four open-response items were coded for normative use of evolutionary concepts (*variation*, *heritability*, *differential survival*, *competition*, *hyperfecundity*, *limited resources*, and *change of population*) and for use of major misconceptions identified by Shtulman (2006) and Evans et al. (2010) (*essentialism*, *teleology*, and *intentionality bias*). The face validity of items has been confirmed by domain experts, including M.A. biology students, Ph.D. biology students, and tenured professors of biology (Nehm & Ridgway, 2011). Additionally, scores on ACORNS items predict later course grades for undergraduate students in both evolutionary and non-evolutionary biology courses after controlling for overall GPA (Opfer et al., 2012). Further, interrater reliability of human coders is high to very high, with the reliability of the automated coder obtaining nearly "flawless" scores (Moharreri et al., 2014). Thus, the latest assessments of evolution understanding have successfully built on a large research base in the biological, cognitive, and education sciences to produce a rich, qualitative assessment with excellent psychometric properties.

Complementing these measures of evolution understanding is the Evolutionary Attitudes and Literacy Survey (EALS), a multi-dimensional survey designed to assess attitudes about the relevance of evolutionary theory (Hawley, Short, McCune, Osman, & Little, 2011). There are two forms of the EALS, the long form (104 items) and the short form (62 items). Both have been validated using Confirmatory Factor Analysis (Short & Hawley, 2012). Unlike many previous assessments, the EALS relies exclusively on closed-response questions. The main purpose of the EALS is to diagnose socio-political factors that affect people's endorsement or rejection of evolution, including religious ideation, political ideation, knowledge of science, trust in science, and the endorsement of intelligent design fallacies.

4.3. Assessing evolution understanding in a social science context; additional considerations

Current methods for assessing students' understanding of evolution are grounded in the evolution of non-human animals and non-mentalistic traits. These tools could be used productively in the social science classroom, but social scientists may need their own assessment tools—tools capable of gauging students' understanding of the evolved nature of human cognition and human behavior. Below we consider four considerations that might shape the construction of such tools. Social scientists who embrace evolution in their teaching have likely pondered these considerations already and developed assessment tools for their own instruction, but research is needed to construct a shared set of tools, which have the demonstrated validity and reliability of the tools reviewed above.

4.3.1. What knowledge should the assessment measure?

Perhaps the most basic knowledge one could assess is *conceptual* knowledge—knowledge of the mechanisms of evolution (e.g., variation, inheritance, selection) and knowledge of the products of those mechanisms (e.g., adaptation, exaptation, homologous traits, analogous traits, vestigial traits, trait loss). Assessments of this kind exist for nonhuman animals and could be used as models for developing human-based ones. In this vein, Nettle (2010) has developed a human-based assessment by recasting standard questions about nonhuman evolution (e.g., how might light-colored fossa have evolved from dark-colored fossa?) in terms of humans (e.g., how might light-haired humans have evolved from dark-haired humans?). He found that students score around 10% higher on the human-based assessment than on the nonhuman version, possibly because students hold more accurate conceptions of variation and inheritance for humans than for nonhuman animals. Nettle's assessment does not, however, cover the evolution of human cognition or human behavior, and it remains an open question how to probe such topics authentically and systematically.

Another type of knowledge that social scientists could assess is *methodological* knowledge of how the origins of human traits can be approached from an evolutionary perspective. One facet of this knowledge is constructing plausible evolutionary hypotheses of human traits, which entails understanding the kinds of selection pressures early humans faced, as well as the pathways by with such pressures gave rise to between-species and within-species differences. Another facet of this knowledge is devising empirical tests of evolutionary hypotheses, which entails identifying data sources capable of discriminating among competing hypotheses, as well as analyzing distributional data for changes in the frequency of particular traits across populations or across time.

A third type of knowledge that social scientists might want to assess is knowledge of when and how to apply evolutionary principles to behavioral phenomena, or evolutionarily-informed *habits of mind*. Habits of mind that are most relevant to social science include (1) the ability to identify implausible evolutionary accounts of human traits, such as accounts that presuppose nonexistent selection pressures or infeasible means of adapting to a selection pressure; (2) the ability to respond to non-evolutionary accounts of human traits, such as creationist or sociocultural accounts untenable with a biological perspective; and (3) the ability to identify naturalistic fallacies–claims that properties which are natural are morally permissible and properties that are unnatural are not. Assessments of students' habits of mind are perhaps most revealing of whether students will use their knowledge of evolutionary social science beyond the classroom, when there are no requirements to do so.

4.3.2. What content should the assessment cover?

Decisions about what content to cover are invariably tied to decisions about what knowledge to assess. If conceptual knowledge is of interest, then one could use "why" questions—questions about why humans possess traits that other animals do not or why certain groups of humans possess traits that other groups do not. Examples include: Why are humans bipedal? Why do humans use tools? Why do humans have language? Why do humans cook? Why do humans exhibit prolonged tolerance to lactose? Why do humans exhibit prolonged childhoods? Why do humans have unusually large brains?

Such questions could be posed in either an open-response format or a closed-response format, and they could cover traits that "call out" for an evolutionary explanation (e.g., bipedalism, differences in skin color) or traits that could be explained from a non-evolutionary perspective just as easily—or more easily—than from an evolutionary one (e.g., tool use, language use). The latter potentially constitute a more stringent test of students' evolutionary reasoning, as they would have to inhibit non-evolutionary explanations in order to provide evolutionary ones. Another possible variant is to describe traits that are currently infrequent in humans (e.g., immunity to the AIDS virus) and ask how such traits could become more frequent over time. Such questions would require students to reason predictively rather than postdictively, which is a more resource-demanding task (Fischhoff, 1975).

To assess students' methodological knowledge of evolutionary social science, at least two approaches are possible. One approach is to ask students to go beyond explaining the origin or frequency of a trait and identify an empirical means of testing those explanations. Another approach is to give students data relevant to a pre-specified evolutionary claim and ask them to determine whether that claim is, in fact, supported by the data. Students could, for example, evaluate the claim that children are more likely to be killed by a stepparent than by a biological parent by analyzing a dataset on the circumstances surrounding real-life cases of infanticide. Depending on the nature of the claim, students could evaluate data that either support the claim or contradict it, and depending on the nature of the data, students could evaluate claims that are either consistent with their beliefs about human nature or inconsistent with them. Discerning that a claim is contradicted by data is more difficult than discerning that it is supported by data, and discerning that a counterintuitive claim is supported by data is more difficult than discerning that an intuitive claim is supported (Koehler, 1993; Kuhn, 1989).

Finally, to assess students' evolutionarily-informed habits of mind, one could present students with misconceptions and evaluate how they respond to those misconceptions—i.e., whether they can identify the misconceptions as wrong or, better yet, whether they can articulate precisely why they are wrong. One could present misconceptions that relate either to the nature of evolutionary change (e.g., how many generations are needed to effect change given different selection gradients and different levels of adaptation complexity?) or to the implications of evolutionary findings (e.g., discovering a psychological adaptation for aggression in specific circumstances does not imply that aggression is inevitable or that social programs to reduce its occurrence will be ineffective). Such questions would require students to differentiate correct and incorrect applications of evolutionary theory, which is a reasonable expectation for students in upper-level classes but may not be for students in introductory classes. Such questions require not just a correct understanding of evolutionary principles but also a meta-level appreciation of the difference between correct and incorrect applications of those principles.

4.3.3. What format should the assessment take?

A popular format for knowledge assessment in the sciences is a "concept inventory" (see, for example, Hestenes, Wells, & Swackhamer, 1992; Libarkin & Anderson, 2005). At least one concept inventory exists for evolution as applied to nonhumans (Anderson et al., 2002, discussed above) but none exists for evolution as applied to humans.

Concept inventories are easy to administer and easy to score, but they have several drawbacks, as noted above. One way to address these problems is to intermix open-response questions with the closed-response ones, possibly in the form of justifications (e.g., "Why do you think so?"). Another way is to collect agreement ratings for all answers to each multiple-choice question. Three response profiles are then possible: (1) the "novice" profile of agreeing with incorrect answers and disagreeing with correct ones; (2) the "mixed" or "undifferentiated" profile of agreeing with both correct and incorrect answers; and (3) the "expert" profile of agreeing with correct answers and disagreeing with incorrect ones. Prior research suggests that, with instruction, students are more likely to move from the first profile to the second before moving all the way to the third (Schneider & Hardy, 2013; Shtulman & Calabi, 2013)—a pattern not readily detectable with concept inventories administered in a standard, forced-choice fashion.

4.3.4. Against what criteria should the assessment be validated?

The most straightforward means of validating a knowledge assessment of evolutionary social science would be to confirm that social scientists themselves score high on the assessment. Of course, not all social scientists find value in evolutionary approaches; the proportion of social scientists who deem evolution as applicable to their area of research is, in fact, alarmingly small (King & Cabeza de Baca, 2011; Perry & Mace, 2010; von Hippel & Buss, 2017). Even among social scientists who value evolutionary approaches, not all are practiced at adopting such approaches. It is thus unclear who should count as an expert for the sake of validation.

Even if the pool of potential experts was narrowed to social scientists who take an evolutionary approach to their own work, this group still disagrees about which evolutionary explanations—or classes of explanation—are appropriate for explaining human nature (Barrett, Pollet, & Stulp, 2014; Burke, 2014), including questions pertaining to cultural differences, racial differences, gender differences, developmental differences, and phylogenetic differences. This disagreement also spans questions pertaining to human similarities, as those similarities could represent either homologous traits (arising from shared genetic foundations) or analogous traits (arising from shared experiences but different genetic foundations). There are legitimate scientific disagreements about these issues, which is all part of healthy science. Developing a knowledge assessment of evolutionary social science will likely require collaboration among several groups of scientists, so as to ensure that the assessment is not biased in favor of the theoretical or methodological commitments of any one particular group.

Please cite this article as: Legare, C.H., et al., A field guide for teaching evolution in the social sciences, *Evolution and Human Behavior* (2018), https://doi.org/10.1016/j.evolhumbehav.2018.01.002

5. Conclusion

The theory of evolution is the unifying framework for the biological sciences and, as argued in this paper, should be the unifying framework for the social sciences. To achieve this however, a major revision of the social science educational model is required. Evolution is an exceedingly difficult theory to understand and is counterintuitive due to multiple cognitive biases. Essentialist biases lead people to believe that species are stable and immutable. Teleological biases lead us to believe that all species and their component parts exist for a reason. Existential anxieties lead us to believe that accepting evolution threatens our sense of meaning and purpose. To counteract these psychological disadvantages, there is a need for incoporating evolution education into social science disciplines, which utilizes the strategies outlined in this paper. Begin by countering misconceptions about evolution before attempting to instill accurate beliefs. Move on to forging deeper understanding by illustrating evolutionary principles with detailed, real-world case studies of evolution taking place, such as prolonged tolerance to lactose in humans. In addition, instructors should utilize collaborative problem-solving activities that necessitate students working together to analyze data and generate their own evolutionary explanations. These educational strategies all show empirical promise to enhance students' understanding of evolution. Finally, teaching evolution in the social sciences requires devising accurate assessments of student comprehension. This chapter highlights a number of measures tested in the literature. We argue that educators should adopt a three-pronged approach to assessment by evaluating students' conceptual knowledge, methodological knowledge, and evolutionary "habits of mind."

Humans are biological beings, and, as such, they have evolved. Evolution has shaped the human body, but it also shaped the human brain and the human mind. Many social scientists initially ignored the evolutionary origins of human nature, but a new generation of social scientists see this practice as a lost opportunity at best, and a critical omission at worst. In this paper, we have outlined a path for moving forward on the issue of teaching evolution in the social sciences—a path informed by our understanding of teaching evolution in general. We urge the community of social scientists to embark on that path, both by investigating different approaches to teaching and assessing evolution understanding and by actually incorporating evolution into the social science curriculum. If "nothing in biology makes sense except in the light of evolution," as Dobzhansky (1973) famously argued, then understanding how biological beings think and behave must also require the light of evolution.

References

Alonzo, A. C., & Gotwals, A. W. (Eds.). (2012). *Learning progressions in science: Current challenges and future directions*. Rotterdam: Sense Publishers.

Anderson, D. L., Fisher, K. M., & Norman, G. L. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, *39*, 952–978.

Asterhan, C. S., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology*, *99*, 626–639.

Barnes, M. E., Evans, E. M., Hazel, A., Brownell, S. E., & Nesse, R. M. (2017). Teleological reasoning, not acceptance of evolution, impacts students' ability to learn natural selection. *Evolution: Education and Outreach*, *10*(7), 1–12.

Barrett, H. C. (2015). *The shape of thought: How mental adaptations evolve*. New York: Oxford University Press.

Barrett, L., Pollet, T. V., & Stulp, G. (2014). From computers to cultivation: Reconceptualizing evolutionary psychology. *Frontiers in Psychology*, *5*(867).

Bean, T. E., Sinatra, G. M., & Schrader, P. G. (2010). SPORE: Spawning evolutionary misconceptions? *Journal of Science Education and Technology*, *19*(5), 409–414.

Berti, A. E., Toneatti, L., & Rosati, V. (2010). Children's conceptions about the origin of species: A study of Italian children's conceptions with and without instruction. *The Journal of the Learning Sciences*, *19*, 506–538.

Bishop, B. A., & Anderson, C. W. (1985). Evolution by natural selection: A teaching module. (Occasional Paper no. 91). Retrieved from http://files.eric.ed.gov/fulltext/ED272383.pdf.

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

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

Bolhuis, J. J., Brown, G. R., Richardson, R. C., & Laland, K. N. (2011). Darwin in mind: New opportunities for evolutionary psychology. *PLoS Biology*, *9*(7), 1–8.

Boyd, R., & Silk, J. (2015). *How humans evolved* (7th Ed). W.W. Norton & Company Inc.

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*, 493–503.

Burke, D. (2014). Why isn't everyone an evolutionary psychologist? *Frontiers in Psychology*, *5*(910).

Buss, D. M. (1995). Evolutionary psychology: A new paradigm for psychological science. *Psychological Inquiry*, *6*(1), 1–30.

Buss, D. M. (2009a). How can evolutionary psychology successfully explain personality and individual differences? *Perspectives on Psychological Science*, *4*, 359–366.

Buss, D. M. (2009b). The great struggles of life: Darwin and the emergence of evolutionary psychology. *American Psychologist*, *64*, 140–148.

Buss, D. M. (2015). *Evolutionary psychology: The new science of mind* (5th edition). New York: Pearson.

Buss, D. (2016). *The handbook of evolutionary psychology* (2nd ed.). Hoboken, NJ: Wiley.

Cabeza De Baca, T., & Jordan, A. C. (2012). To know is not to love: Cognitive and affective barriers toward the adoption of evolutionary theory. *Personality and Individual Differences*, *53*, 681–686.

Casler, K., & Kelemen, D. (2008). Developmental continuity in teleo-functional explanation: Reasoning about nature among Romanian Romani adults. *Journal of Cognition and Development*, *9*, 340–362.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21–30). New York: Cambridge University Press.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105.

Clough, E. E., & Driver, R. (1986). A study of consistency in the use of students' conceptual frameworks across different task contexts. *Science Education*, *70*, 473–496.

Darwin, C. R. (1887). *The life and letters of Charles Darwin*. London: John Murray.

Dawkins, R. (1982). *The extended phenotype*. New York: Oxford University Press.

Demastes, S. S., Settlage, J., & Good, R. (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.

Deniz, H., Donelly, L. A., & Yilmaz, I. (2008). Exploring the factors related to acceptance of evolutionary theory among Turkish preservice biology teachers. *Journal of Research in Science Teaching*, *45*, 420–443.

Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher*, *35*, 125–129.

Emmons, N. A., & Kelemen, D. A. (2015). Young children's acceptance of within-species variation: Implications for essentialism and teaching evolution. *Journal of Experimental Child Psychology*, *139*, 148–160.

Evans, E. M. (2000a). The emergence of beliefs about the origins of species in school-age children. *Merrill-Palmer Quarterly*, *46*, 221–254.

Evans, E. M. (2000b). Beyond scopes: Why creationism is here to stay. In K. Rosengren, C. Johnson, & P. Harris (Eds.), *Imagining the impossible: Magical, scientific and religious thinking in children* (pp. 305–331). Cambridge: Cambridge University Press.

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

Evans, E. M. (2005). Teaching and learning about evolution. In J. Diamond (Ed.), *The virus and the whale: Explore evolution in creatures small and large* (pp. 25–37). Arlington, VA: NSTA Press.

Evans, E. M., & Lane, J. D. (2011). Contradictory or complementary? Creationist and evolutionist explanations of origins. *Human Development*, *54*, 144–159.

Evans, E. M., Legare, C., & Rosengren, K. (2011). Engaging multiple epistemologies: Implications for science education. In M. Ferrari, & R. Taylor (Eds.), *Epistemology and science education: Understanding the evolution vs. intelligent design controversy* (pp. 111–139). New York, NY: Routledge.

Evans, E. M., Spiegel, A., Gram, W., Frazier, B. F., Tare, M., Thompson, S., & Diamond, J. (2010). A conceptual guide to natural history museum visitors' understanding evolution. *Journal of Research in Science Teaching*, *47*, 326–353.

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

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

Gelman, S. A., & Rhodes, M. (2012). Two-thousand years of stasis. How psychological essentialism impedes evolutionary understanding. In K. S. Rosengren, S. Brem, E. M. Evans, & G. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution* (pp. 3–21). Cambridge, UK: Oxford University Press.

Gray, P., & Bjorklund, D. F. (2014). *Psychology* (7th edition). New York: Worth.

Gregory, T. R., & Ellis, C. A. J. (2009). Conceptions of evolution among science graduate students. *Bioscience*, *59*, 792–799.

Hart, J., Shaver, P. R., & Goldenberg, J. L. (2005). Attachment, self-esteem, worldviews, and terror management: Evidence for a tripartite security system. *Journal of Personality and Social Psychology*, *88*, 999–1013.

Hawley, P. H., Short, S. D., McCune, L. A., Osman, M. R., & Little, T. D. (2011). What's the matter with Kansas?: The development and confirmation of the Evolutionary Attitudes and Literacy Survey (EALS). *Evolution: Education and Outreach*, *4*, 117–132.

Please cite this article as: Legare, C.H., et al., A field guide for teaching evolution in the social sciences, *Evolution and Human Behavior* (2018), https://doi.org/10.1016/j.evolhumbehav.2018.01.002

Heddy, B. C., & Sinatra, G. M. (2013). Transforming misconceptions: Using transformative experience to promote positive affect and conceptual change in students learning about biological evolution. *Science Education*, *97*, 723–744.

Henrich, J. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton, NJ: Princeton University Press.

Heredia, S. C., Furtak, E. M., & Morrison, D. (2016). Exploring the influence of plant and animal item contexts on student response patterns to natural selection multiple choice items. *Evolution: Education and Outreach*, *9*(10), 1–11.

Herrmann, P. A., French, J. A., DeHart, G. B., & Rosengren, K. S. (2013). Essentialist reasoning and knowledge effects on biological reasoning in young children. *Merrill-Palmer Quarterly*, *59*, 198–220.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*, 141–158.

von Hippel, W., & Buss, D. M. (2017). Do ideologically driven scientific agendas impede the understanding and acceptance of evolutionary psychology? In L. Jussim, & J. Crawford (Eds.), *The politics of social psychology* (pp. 7–25). New York: Frontiers in Psychology.

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

Keil, F. C. (1992). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Kelemen, D. (1999a). Functions, goals and intentions: Children's teleological reasoning about objects. *Trends in Cognitive Sciences*(12), 461–468.

Kelemen, D. (1999b). The scope of teleological thinking in preschool children. *Cognition*, *70*, 241–272.

Kelemen, D., Emmons, N. A., Schillaci, R. S., & Ganea, P. A. (2014). Young children can be taught basic natural selection using a picture-storybook intervention. *Psychological Science*, *25*, 893–902.

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General*, *142*, 1074–1083.

King, A. C., & Cabeza de Baca, T. (2011). The stagnancy of family studies in modern academia: Resistances toward the integration of evolutionary theory. *Evolution and Education Outreach*, *4*, 64–74.

Koehler, J. J. (1993). The influence of prior beliefs on scientific judgments of evidence quality. *Organizational Behavior and Human Decision Processes*, *56*, 28–55.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, *96*, 674–689.

Legare, C. H., Evans, E. M., Rosengren, K., & Harris, P. (2012). The coexistence of natural and supernatural explanations across cultures and development. *Child Development*, *83*, 779–793.

Legare, C. H., Lane, J., & Evans, E. M. (2013). Anthropomorphizing science: How does it affect the development of evolutionary concepts? *Merrill-Palmer Quarterly*, *59*, 168–197.

Legare, C. H., & Visala, A. (2011). Between religion and science: Integrating psychological and philosophical accounts of explanatory coexistence. *Human Development*, *54*, 169–184.

Lewis, D. M. G., Al-Shawaf, L., Conroy-Beam, D., Asao, K., & Buss, D. M. (2017). Evolutionary psychology: A how-to guide. *American Psychologist*, *72*(4), 353–373.

Libarkin, J. C., & 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.

Lombrozo, T., Kelemen, D., & Zaitchik, D. (2007). Inferring design evidence of a preference for teleological explanations in patients with Alzheimer's Disease. *Psychological Science*, *18*, 999–1006.

Lombrozo, T., Shtulman, A., & Weisberg, M. (2006). The intelligent design controversy: Lessons from psychology and education. *Trends in Cognitive Sciences*, *10*, 56–57.

Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist*, *34*(4), 207–218. https://doi.org/10.1207/s15326985ep3404_2.

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

Moharreri, K., Ha, M., & Nehm, R. H. (2014). EvoGrader: An online formative assessment tool for automatically evaluating written evolutionary explanations. *Evolution: Education and Outreach*, *7*(15).

Myers, D. G. (2012). *Psychology (10th edition in modules)*. New York: Worth.

National Research Council (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Nehm, R. H., & Ha, M. (2011). Item feature effects in evolution assessment. *Journal of Research in Science Teaching*, *48*, 237–256.

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

Nehm, R. H., & Ridgway, J. (2011). What do experts and novices "see" in evolutionary problems? *Evolution: Education and Outreach*, *4*, 666–679.

Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, *45*, 1131–1160.

Nettle, D. (2010). Understanding of evolution may be improved by thinking about people. *Evolutionary Psychology*, *8*, 205–228.

Opfer, J. E., & Gelman, S. A. (2001). Children's and adults' models for predicting teleological action: The development of a biology-based model. *Child Development*, *72*(5), 1367–1381.

Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, *49*, 744–777.

Perry, G., & Mace, R. (2010). The lack of acceptance of evolutionary approaches to human behavior. *Journal of Evolutionary Psychology*, *8*, 105–125.

Pew Research Center (2014). *Religious landscape study*. Washington, DC: Author.

Pinker, S. (1997). *How the mind works*. New York: Norton & Company.

Poling, D. A., & Evans, E. M. (2002). Why do birds of a feather flock together? Developmental change in the use of multiple explanations: Intention, teleology and essentialism. *British Journal of Developmental Psychology*, *20*, 89–112.

Popham, W. J. (2010). *Everything school leaders need to know about assessment*. Thousand Oaks, CA: Corwin Press.

Rodriguez, M. C. (2003). Construct equivalence of multiple-choice and constructed-response items: A random effects synthesis of correlations. *Journal of Educational Measurement*, *40*(2), 163–184.

Rosengren, K. S., Brem, S., Evans, E. M., & Sinatra, G. (Eds.). (2012). *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. Cambridge, UK: Oxford University Press.

Rottman, J., Zhu, L., Wang, W., Schillaci, R. S., Clark, K. J., & Kelemen, D. (2017). Cultural influences on the teleological stance: Evidence from China. *Religion, Brain & Behavior*, *17*(1), 17–26.

Rutjens, B. T., Van Der Pligt, J., & Van Harreveld, F. (2010). Deus or Darwin: Randomness and belief in theories about the origin of life. *Journal of Experimental Social Psychology*, *46*, 1078–1080.

Schneider, M., & Hardy, I. (2013). Profiles of inconsistent knowledge in children's pathways of conceptual change. *Developmental Psychology*, *49*, 1639–1649.

Scott, E. C. (2004). *Evolution Vs. Creationism: An introduction*. Berkeley: University of California Press.

Short, S. D., & Hawley, P. H. (2012). Evolutionary attitudes and literacy survey (EALS): Development and validation of a short form. *Evolution: Education and Outreach*, *5*(3), 419–428.

Short, S. D., & Hawley, P. H. (2014). The effects of evolutionary education: Examining attitudes toward and knowledge of evolution in college courses. *Evolutionary Psychology*, *13*(1), 67–88.

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

Shtulman, A., & Calabi, P. (2013). Tuition vs. intuition: Effects of instruction on naïve theories of evolution. *Merrill-Palmer Quarterly*, *59*, 141–167.

Shtulman, A., & 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., & Harrington, K. (2016). Tensions between science and intuition across the lifespan. *Topics in Cognitive Science*, *8*, 118–137.

Shtulman, A., Neal, C., & Lindquist, G. (2016). Children's ability to learn evolutionary explanations for biological adaptation. *Early Education and Development*, *27*, 1222–1236.

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

Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evolution: Education Outreach*, *1*, 189–195.

Sinatra, G. M., Southerland, S. A., McConaughy, F., & Demastes, J. (2003). Intentions and beliefs in students' understanding and acceptance of biological evolution. *Journal of Research in Science Teaching*, *40*(5), 510–528.

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.

Spiegel, A. N., Evans, E. M., Frazier, B., Hazel, A., Tare, M., Gram, W., & Diamond, J. (2012). Changing museum visitors' conceptions of evolution. *Evolution: Education and Outreach*, *5*, 43–61.

Tooby, J., & Cosmides, L. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.

Tracy, J. L., Hart, J., & Martens, J. P. (2011). Death and science: The existential underpinnings of belief in intelligent design and discomfort with evolution. *PLoS One*, *6*, e17349.

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

Ware, E. A., & Gelman, S. A. (2014). You get what you need: An examination of purpose-based inheritance reasoning in undergraduates, preschoolers, and biological experts. *Cognitive Science*, *38*, 197–243.

ojalehto, b., Waxman, S. R., & Medin, D. L. (2013). Teleological reasoning about nature: Intentional design or relational perspectives? *Trends in Cognitive Sciences*, *17*, 166–171.

Weisberg, D. S., Landrum, A. R., Metz, S. E., & Weisberg, M. (in press). No missing link: Knowledge predicts acceptance of evolution in the United States. Biosciences.

Wertz, A. E., & Wynn, K. (2014). Selective social learning of plant edibility in 6- and 18-month-old infants. *Psychological Science*, *24*, 874–882.

Williams, G. C. (1966). Natural selection, the costs of reproduction, and a refinement of Lack's principle. *The American Naturalist*, *100*(916), 687–690.

Wilson, D. S. (2007). *Evolution for everyone: How Darwin's theory can change the way we think about our lives*. New York, NY: Bantam Dell.

Wilson, D. S. (2015). *Does altruism exist? Culture, genes, and the welfare of others*. New Haven, CT: Yale University Press.

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

Please cite this article as: Legare, C.H., et al., A field guide for teaching evolution in the social sciences, *Evolution and Human Behavior* (2018), https://doi.org/10.1016/j.evolhumbehav.2018.01.002