Students' Mental Models of Evolutionary Causation: Natural Selection and Genetic Drift

Elizabeth P. Beggrow · Ross H. Nehm

Abstract In an effort to understand how to improve student learning about evolution, a focus of science education research has been to document and address students' naive ideas. Less research has investigated how students reason about alternative scientific models that attempt to explain the same phenomenon (e.g., which causal model best accounts for evolutionary change?). Within evolutionary biology, research has yet to explore how non-adaptive factors are situated within students' conceptual ecologies of evolutionary causation. Do students construct evolutionary explanations that include non-adaptive *and* adaptive factors? If so, how are non-adaptive factors structured within students' evolutionary explanations? We used clinical interviews and two paper and pencil instruments (one open-response and one multiple-choice) to investigate the use of non-adaptive and adaptive factors in undergraduate students' patterns of evolutionary reasoning. After instruction that included non-adaptive causal factors (e.g., genetic drift), we found them to be remarkably uncommon in students' explanatory models of evolutionary change in both written assessments and clinical interviews. However, consistent with many evolutionary biologists' explanations, when students used non-adaptive factors they were conceptualized as causal alternatives to selection. Interestingly, use of non-adaptive factors was not associated with greater understanding of natural selection in interviews or written assessments, or with fewer naive ideas of natural selection. Thus, reasoning using non-adaptive factors appears to be a distinct facet of evolutionary thinking. We propose a theoretical framework for an expert–novice continuum of evolutionary reasoning that incorporates both adaptive and non-adaptive factors, and can be used to inform instructional efficacy in evolutionary biology.

Keywords Undergraduates · Evolution · Natural selection · Non-adaptive factors · Genetic drift · Learning

Introduction

A central goal of science education reform is to refocus teaching, learning, and assessment on core concepts or "big ideas" (e.g., NRC 2001a, 2001b; AAAS 2011). One such big idea in the life sciences is biological evolution. Although evolution comprises the framework upon which the life sciences are structured (Dobzhansky 1973), student learning of the idea and its associated causal concepts remains extremely problematic for students at all levels of the educational hierarchy (Bishop and Anderson 1990; Nehm and Reilly 2007; Gregory 2009; Opfer et al. 2012). In an effort to understand how to improve student learning about evolution, a major thrust of science education research has been to focus on students' naive ideas (misconceptions) about evolution and develop pedagogies to initiate conceptual change (e.g., Demastes et al. 1995). Considerably less research has focused on how students reason about alternative *scientific* models seeking to explain the same phenomenon (e.g., of the possible causes, which most powerfully account for evolutionary change?). Indeed, within the field of science education, research has focused overwhelmingly on adaptive and selective explanations for evolutionary change. Research has yet to explore how non-adaptive factors are situated within students' conceptual ecologies[1] of evolutionary causation.

E. P. Beggrow (✉) · R. H. Nehm
The Ohio State University,
Columbus, OH, USA
e-mail: beggrow.7@osu.edu

---

[1] Conceptual ecologies refer to the interactive network of a person's epistemological commitments, beliefs and their different kinds of knowledge that collectively form the foundational, "organizing conceptions" of that person's worldview (Demastes et al. 1995, p. 638).

While intense debate about the causal factors accounting for the structure of the natural world persists in many science domains, and reflects deeply complex philosophical and epistemological tensions (e.g., Conway Morris 2003; Randall 2011), we wish to engage with this topic at a level that intersects with important questions about the teaching of evolution. Within the field of evolutionary biology, adaptive (e.g., natural selection) and non-adaptive (e.g., genetic drift) frameworks have been used to account for and explain (conceptually and methodologically) the processes that undergird evolutionary patterns in the living world, both past and present (Gould 2002). While the body of work in this area is immense, we wish to focus attention on two concepts that represent core perspectives on evolutionary causation and have direct relevance to undergraduate science educators: natural selection and genetic drift (Table 1).

Natural selection is considered by many evolutionary biologists to be the primary mechanism causing adaptive evolutionary change (Pigliucci and Müller 2010). Adaptive evolutionary change occurs when the frequency of a trait (genetic or phenotypic feature) increases *because* it confers a survival or reproductive advantage to the "individuals" (genes, bacteria, animals, etc.) possessing it (Dawkins 1976; Darwin 1859). Importantly, as noted above, natural selection is not the only process leading to evolutionary changes in the living world. Alternative mechanisms, collectively known as "non-adaptive" factors, include concepts such as genetic drift and developmental constraints, among others (Gould and Lewontin 1979; Nickels et al. 1996; Gould 2002). Non-adaptive change (e.g., genetic drift) occurs when the frequency of a trait increases or decreases in a population because of stochastic factors, regardless of whether the trait confers an advantage, disadvantage, or is neutral with respect to survival or reproduction (Freeman 2005).

As noted by Orr (1998, p. 2099), evolutionary biologists have long been seeking ways to determine whether

Table 1 Definitions of natural selection and genetic drift from glossaries of introductory college biology textbooks

Term	Definition	Textbook
Natural Selection	A process in which organisms with certain inherited characteristics are more likely to survive and reproduce than are organisms with other characteristics.	Campbell, N.A., Reece, J.B., Urry, L.A., Cain, M.L., Wasserman, S.A., Minorsky P.V., and R.B. Jackson. (2008). Biology. 8th ed. New York: Pearson Benjamin Cummings.
Natural Selection	The process by which individuals with certain heritable traits tend to produce more surviving offspring than do individuals without those traits, resulting in a change in the genetic makeup of the population. A major mechanism of evolution.	Freeman, S. (2005) Biological Science. 2nd ed. Upper Saddle River: Pearson Prentice Hall.
Natural Selection	The differential survival and/or reproduction of classes of entities that differ in one or more characteristics; the difference in survival and/or reproduction is not due to chance, and it must have the potential consequence of altering the proportions of the different entities to constitute natural selection. Thus natural selection is also definable as a partly or wholly deterministic difference in the contribution of different classes of entities to subsequent generations. Usually the differences are inherited. The entities may be alleles, genotypes or subsets of genotypes, populations, or in the broadest sense, species.	Futuyma, D.J. (1998). Evolutionary Biology. 3rd ed. Sunderland: Sinauer Associates, Inc.
Genetic Drift	A process in which chance events cause unpredictable fluctuations in allele frequencies from one generation to the next. Effects of genetic drift are most pronounced in small populations.	Campbell et al. (2008). Biology. 8th ed. New York: Pearson Benjamin Cummings.
Genetic Drift	Any change in allele frequencies due to random events. Causes allele frequencies to drift up and down randomly over time, and eventually can lead to the fixation or loss of alleles. Non-adaptive change (e.g., genetic drift) occurs when the frequency of a trait increases or decreases because of stochastic factors, regardless of whether the trait confers an advantage, disadvantage, or is neutral with respect to survival or reproduction	Freeman, S. (2005) Biological Science. 2nd ed. Upper Saddle River: Pearson Prentice Hall.
Genetic Drift	Random changes in the frequencies of two or more alleles or genotypes within a population	Futuyma, D.J. (1998). Evolutionary Biology. 3rd ed. Sunderland: Sinauer Associates, Inc.

differences in phenotype or genotype frequencies (e.g., proportions of short vs. long spines within a population) were caused by natural selection, genetic drift, or combinations of the two factors. Biologists often conceptualize selection and drift as alternative or competing evolutionary explanations; selection is deterministic, and drift is stochastic (Sober 1984, p. 110). Methodologically, several approaches have been developed by evolutionary biologists to empirically test the causal contributions of these two processes (for details, see Orr 1998). While many studies have modeled evolutionary change as being caused by either selection *or* drift, research has also looked at the collective contributions of both mechanisms to evolutionary change (i.e., selection *and* drift; Ackermann and Cheverud 2004; Parker and Maynard Smith 1990).

The relative importance of selection and drift has been controversial within evolutionary biology for some time. Evolution by natural selection was introduced through the work of Darwin (1859) and Wallace (1871) and gained ground after the "eclipse of Darwinism" in the early twentieth century (Bowler 1983). At this time, publications raised the question of the relative contributions of selection and random survival (e.g., Hagedoorn and Hagedoorn 1921; Fisher 1922). Fisher, for example, viewed deterministic processes as paramount and natural selection as the only important cause of evolutionary change (Provine 1971). Genetic drift was introduced primarily through Wright's studies of population genetics in the early 1930s, but a lack of convincing experimental evidence led some evolutionary biologists, including Fisher, to reject drift as an important cause of evolutionary change (Savage 1969; Provine 1971).

These earlier debates were revived with Kimura's (1968) finding of a high rate of neutral mutation; such empirical findings provided support for stochastic causal contributors (e.g., genetic drift) to evolutionary change. Nevertheless, many evolutionary biologists remained strong subscribers of what some have termed an "adaptationist approach" to evolution. For instance, Mayr argued that it is the goal of the evolutionary biologist to first try "to explain biological phenomena and processes as the product of natural selection" (1983, p. 326). Gould and Lewontin (1979) countered that Mayr's view is typical of what they caricatured as the "Panglossian paradigm." Gould and Lewontin refer to Voltaire's Dr. Pangloss of *Candide*, whose philosophy that "everything is for the best in this best of all possible worlds" leads to numerous just-so explanatory stories, such as "the nose has been formed to bear spectacles – thus we have spectacles" (Voltaire 2010, p. 2). The Panglossian paradigm serves as a metaphor for some biologists' adaptationist approach to evolution, where the presence of traits is explained through the useful function of those traits.

The Panglossian paradigm demands deterministic explanations, thereby precluding any possibility of trait origins being a product of non-adaptive factors such as drift. Gould and Lewontin (1979, p. 587) warned of the dangers of such explanatory approaches: "One must not confuse the fact that a structure is used in some way…with the primary evolutionary reason for its existence and conformation." Likewise, Lynch (2007) viewed selection-centered views as reducing evolution to a form of "engineering," which he considered not only unnecessary, but also misleading. Other scientists, in contrast, have taken a holistic approach to evolutionary change that recognizes the relative contributions of both causal factors (e.g., Parker and Maynard Smith 1990; Ackermann and Cheverud 2004). In short, evolutionary biologists have debated how evolutionary change should be framed, with some taking an adaptationist approach, and others taking a more pluralistic approach to evolutionary causation.

The competing roles of selection and drift in evolutionary causation have also been a subject of debate for philosophers of biology, though the importance of both concepts is *not* disputed (Rosenberg and McShea 2008; Sober 1984). Discussions in the field of philosophy concern whether or not drift and selection are in fact two *distinct* concepts (e.g., Beatty 1984; Matthen and Ariew 2002; Millstein 2002; Rosenberg and McShea 2008; Sober 1984). Additionally, it has been argued that depending on what kind of question is being asked (e.g., evolutionary origin vs. evolutionary differences), explanations will involve collaborative or competing roles of selection and drift (Rosenberg and McShea 2008). Overall, in the philosophical and scientific communities, selection and drift are widely recognized by scientists and philosophers as two central evolutionary processes that should be considered when developing and testing explanations for trait differences between units (e.g., genomes, populations, species, etc.). Yet, Rosenberg and McShea (2008) point out that "recognizing the role of drift is not the same thing as agreeing on what it is, how it works, and what its relation to adaptation is" (p. 65). Thus, the primary debate concerns understanding what drift is, as well as the relative contributions of selection and drift.

Selection and Drift in Science Education Stochastic processes are acknowledged as potential contributors to evolutionary change by biologists and philosophers (Gould 2002; see above), and life science curricula appear to reflect such normative conceptualizations. For students to achieve more expert-like competency in evolutionary reasoning, then, they must consider both adaptive and non-adaptive processes as possible contributors to evolutionary causation. Alternative mechanisms to natural selection are notably present in undergraduate and graduate biology programs and most college biology textbooks (e.g., Campbell et al. 2008). Though non-adaptive causal mechanisms are present, coverage of these factors varies; indices of introductory biology textbooks indicate relatively low frequencies of topics related to "genetic drift" (Campbell et al. 2008: 3 pages/1267

content pages; Freeman 2005: 9 pages/1238 content pages). For evolution textbooks, the indices indicate slightly higher frequencies of topics related to "genetic drift" (Futuyma 1998: 32 pages/754 content pages; Ridley 2004; 47 pages/ 681 content pages). Reviews of collections of textbooks and laboratory manuals have also found non-adaptive factors to be present but minimally covered (Linhart 1997; Maret and Rissing 1998). In addition, considering the reliance on texts for teaching (Carpenter et al. 2006; Cuseo 2007), it is possible that most undergraduate science curricula may over-emphasize natural selection at the expense of non-adaptive factors. The possible consequences of a nearly exclusive focus on natural selection as the cause of evolutionary change raise the question as to whether biology education is reifying an adaptationist approach.

Although a number of studies have developed activities for teaching genetic drift in the classroom, or advocated for hands-on activities relating to genetic drift (e.g., Nickels et al. 1996; Maret and Rissing 1998; Staub 2002), remarkably little work has been done to explore how (or if) students think about non-adaptive factors in evolutionary change. We know that practicing evolutionary biologists attribute evolutionary change to selection, non-adaptive factors (e.g., genetic drift) or a combination of these two factors, but what do students' explanations of evolutionary causation look like? This question motivated our explorations of biology students' non-adaptive and adaptive evolutionary reasoning patterns.

Theoretical Framework

It is well established that students bring a variety of intuitive ideas to school that are in conflict with normative scientific perspectives (e.g., Wandersee et al. 1994). Students deal with these contradictions in a variety of ways: they may ignore the new information and continue using their previous frameworks, they may maintain both the new and old information in parallel, accessing each in specific contexts or situations, or they may construct a new conceptual framework that incorporates both the new and old knowledge (Fosnot 1996). Students use their existing conceptual frameworks to process new experiences (e.g., assimilation) or, when the students' current frameworks are inadequate in allowing them to make sense of new experiences, they must reorganize and/or replace them with new concepts (e.g., accommodation; Demastes et al. 1995; Posner et al. 1982; Sinatra et al. 2008).

Students' alternative conceptions in science, particularly those regarding natural selection, are well documented (e.g., Bishop and Anderson 1990; Settlage 1994; Demastes et al. 1995; Demastes et al. 1995; Ferrari and Chi 1998; Nehm and Reilly 2007; Nehm and Schonfeld 2007; Nehm 2009;

Gregory 2009). These alternative conceptions are both abundant and persistent, and efforts to target them through instruction have yielded varied results (e.g., Bishop and Anderson 1990; Demastes et al. 1995, 1996; Dagher and BouJaoude 1997; Ferrari and Chi 1998; Nehm and Reilly 2007; Sinatra et al. 2008). In an effort to understand how to improve student learning and thinking about evolution, most of the research in evolution education has focused on students' naive ideas and how to change them (e.g., Demastes et al. 1995; Demastes et al. 1995, 1996; Sinatra et al. 2008). Much less research has focused on how students incorporate alternative *scientific* models of evolutionary causation; that is, work has yet to explore how non-adaptive factors should be situated theoretically, or how non-adaptive factors such as genetic drift are incorporated into students' conceptual ecologies of evolution.

The Development of Evolutionary Reasoning: Theoretically Situating Non-adaptive Causation The National Research Council [NRC] (2001b) argues that an understanding of how students think and reason about domain-specific ideas should undergird the design of teaching, curriculum, and assessment. Such models should be based upon empirical evidence about how "students represent knowledge and develop competence in the domain" (2001b, p. 3). Taken collectively, the aforementioned studies about student thinking and alternative concepts of evolution and natural selection contribute to the development of a model of how students think about and learn evolutionary concepts. Despite important early work in this area (cf. Catley et al. 2005), non-adaptive factors remain to be incorporated within cognitive models of evolutionary competence and corresponding learning progressions.

Novice–expert studies, common in many areas of science education, provide crucial insights into notions of "competency" as well as what is meant by normative and accurate scientific understanding (NRC 2001b). When comparing novices with experts, research demonstrates that differences between novice students and expert scientists lie in a variety of factors, including metacognition, organization and categorization of knowledge, and presuppositions surrounding knowledge (e.g., Vosniadou 1999; Chi et al. 1981; Nehm and Ridgway 2011). Novice–expert studies have also compared knowledge representation and which concepts are used to build explanations to account for scientific phenomena (Keil and Wilson 2000). Unfortunately, only one study has explored novice–expert reasoning patterns in evolution, with a focus almost exclusively on natural selection (Nehm and Ridgway 2011).

While reasoning about evolutionary change is diverse and complex, expert–novice reasoning patterns, and associated benchmarks of competency, may be simplified and modeled along a continuum based on existing research (see Table 2). In this framework, novices are defined as those who tend to use

exclusively naive ideas, or both naive ideas and key concepts[2] of natural selection, in their explanations of evolutionary change (Nehm and Ridgway 2011). In contrast, "emerging experts" are those whose explanations include key concepts of natural selection, but not naive ideas (Nehm and Ridgway 2011). Though emerging experts provide accurate conceptions of selective causation, they may fail to incorporate possible non-adaptive factors, such as genetic drift, which are more common in experts' models of evolutionary change (Nehm and Ridgway 2011). Experts may attribute change to *either* natural selection *or* genetic drift (e.g., Orr 1998), or they may incorporate both selection *and* drift (or other stochastic, non-adaptive processes; e.g., Ackermann and Cheverud 2004). Considering that biologists employ either or both of these factors, it is reasonable to use these conceptual models as benchmarks for expert-like reasoning about evolution for life sciences students (Table 2, top row). While the goal of education is not to make *all* students scientific experts, domain-specific scientific literacy and competency, based on expert reasoning, is an appropriate expectation for undergraduate students, particularly those completing upper-level courses (Duit 2003; NRC 2001b).

Nonetheless, students have difficulty understanding and reasoning about evolution, and their naive ideas are well documented in the literature (e.g., Gregory 2009). Researchers have documented the persistence of naive ideas, while new, accurate scientific concepts are added to naive "knowledge" frameworks (Vosniadou et al. 2008; Kelemen and Rosset 2009; Nehm 2010). Indeed, students may assimilate scientific concepts learned in school into their pre-existing knowledge frameworks, unaware of any conflict between the two, thereby creating mixed or synthetic mental models of the phenomenon (Vosniadou et al. 2008; Nehm and Ha 2011). Based on prior studies, it is likely that most undergraduate students are concentrated at the bottom end of our expert–novice continuum for evolutionary reasoning (where reasoning about evolutionary change either involves only naive ideas and discarded historical concepts or comprises mixed models composed of both naive ideas and some accurate and/or discarded scientific concepts; see Table 2, bottom row).

Not all students neatly fit into these categories, however. Previous studies have shown that some undergraduate students are able to successfully reason about natural selection using accurate knowledge elements (key concepts) without employing any naive ideas (Nehm and Schonfeld 2008; Nehm and Ha 2011). Though prior work has not investigated non-adaptive reasoning specifically in relation to the aforementioned expert–

---

[2] Key concepts here refer to the generally agreed upon elements that constitute natural selection (e.g., presence and causes of variation and differential survival and reproduction; Nehm and Schonfeld, 2010). Naive ideas (or misconceptions; alternative conceptions) refer to explanatory units that are inaccurate based upon the context in which they are being used (Nehm et al. 2010b). For example, explaining that an animal or plant has a need is not inaccurate per se, but explaining evolutionary change as being caused by the animal or plant's need is inaccurate.

novice continuum, the students from these studies would be considered "emerging experts" rather than "experts," because despite their lack of naive ideas, they do not include non-adaptive processes (e.g., genetic drift) as possible mechanisms of evolutionary change in their explanations. Furthermore, the participants in these studies were enrolled in introductory biology courses for majors and, presumably, had limited exposure to non-adaptive mechanisms of evolutionary change. Whether these students needed more exposure to instruction about non-adaptive reasoning before integrating it into their mental frameworks of evolution remains to be determined. In short, this evolutionary competency framework is used to situate students' evolutionary reasoning sophistication in our study.

Research Questions

In this study we ask three questions: (1) Do students use non-adaptive factors to explain evolutionary change, and if so, does the frequency increase with increasing evolution coursework? (2) Is the use of non-adaptive reasoning patterns associated with greater knowledge of natural selection, or with fewer naive ideas? (3) What do students' explanatory models of evolutionary change look like when both adaptive and non-adaptive factors are included?

Sample

We gathered data from undergraduate biology majors at a large, public, Midwestern research university in the United States. Fifty-five students from two groups were studied. The first group was from early in a college biology program (second semester introductory biology) and the second was from late in the program (an advanced organismal biology class with a prerequisite of an upper-division evolution class). The first group was exposed to basic evolution content (including non-adaptive factors such as genetic drift), and evolution was also considered by the course instructor to be a "key theme." For the first group, Campbell and colleagues' (2008) introductory textbook, *Biology*, was recommended but not required reading. Two lab exercises (out of eight) and nine lecture topics (out of 17) covered evolutionary concepts. Of the assigned readings in the recommended book, four pages consisting of approximately fifteen paragraphs covered genetic drift and gene flow (non-adaptive processes) and an entire chapter and approximately 25 paragraphs covered natural selection (Campbell et al. 2008).

The second group had extensive exposure to evolution (including selective and non-adaptive factors) in the introductory biology course, the advanced evolution course, and in the advanced organismal biology course. The pre-requisite evolution course required Ridley's (2004) *Evolution* textbook, and the entire course was devoted to micro- and macroevolutionary topics. "Neutral

Table 2 A theoretical framework for novice-expert evolutionary reasoning about natural selection and genetic drift

Level of Expertise	Reasoning Model	Example	Reference
Expert	Adaptive+Non-adaptive Both natural selection and genetic drift collectively explain patterns of evolutionary change.	"Although the initial divergence of Homo from the australopiths may have involved selection, divergence after this time (at least in the facial characters analyzed) could have occurred through random processes alone."	(Ackermann and Cheverud 2004, p. 17951)
		"This result suggests that both random and to a lesser extent nonrandom processes played an important role in the diversification of this morphologically diverse group; it does not necessarily mean that both played a role across all parts of the group."	(Ackermann and Cheverud 2004, p. 17949)
	Adaptive vs. Non-adaptive Either natural selection or genetic drift leads to evolutionary change.	"QTL data do provide information on the roles of natural selection vs. genetic drift in phenotypic evolution."	(Orr 1998, p. 2102)
Emerging Expert	Adaptive (key concepts only) Only natural selection explains evolutionary change.	"A mutation may have taken place that allowed a locust to be immune to DDT, this trait was then passed on. These immune locust were the only (ones) that survived and reproduced. Over time, the mutated trait became common of the locust species 'migratoria'."	(Nehm and Ha 2011; supplementary materials) See also this paper.
Novice	Mixed/Synthetic (naive ideas and key concepts) Naive ideas and natural selection explain evolutionary change.	"Due to the fact that animals continually ate the "broken bush" species, the species developed a poison that would fight off predators. This poison worked, and more and more plants decided to use such a survival strategy. Only the strong survived and reproduced, which were the plant species containing poison."	(Nehm and Ha 2011; supplementary materials) See also this paper.
	Pure Naive (naive ideas) Only Naive ideas explain evolutionary change.	"Flightless bird species could have originated from other bird species that can fly because they did not have a specific need for flight. Since they didn't need and/or use their wings for flight, a selective pressure may have worked on them to cause their wings to become flightless."	(Nehm and Ha 2011; supplementary materials) See also this paper.

theory and genetic drift" was the topic for one lecture (out of 18) and there was an associated class on "selection, variation and drift" spread over two recitation periods (out of 16). The corresponding assigned chapters were titled "Random events in population genetics," "Natural selection and random drift," and "Adaptive explanation" (Ridley 2004). Non-adaptive factors like genetic drift and constraint were covered much more extensively in this text compared to the introductory biology textbook. The advanced organismal course did not specifically cover any evolutionary causal factors (e.g., selection or drift), nor did the assigned reading from the required textbook, *Mammalogy* by Feldhamer et al. (2007). This lack of specific coverage of natural selection and genetic drift was expected, as the evolution course is a prerequisite for the advanced organismal biology course, and thus students are expected to have an understanding of evolutionary causal factors when they begin the advanced course.

For brevity, we will refer to these two samples of participants as "majors" and "advanced majors." The majors sample consisted of 28 students (43% male, 57% female) with an average age of 20.4 years. The advanced majors sample comprised 27 students (56% male and 44% female) with an average age of 21.9 years. The majority of students in both samples were White, non-Hispanic.

Methods

We used three methods to gather data on evolutionary reasoning patterns in the participants: (1) clinical oral interviews, (2) the open-response ACORNS assessment (Nehm et al. 2012), and (3) the multiple-choice CINS test (Anderson et al. 2002). Despite displaying psychometric problems (Battisti et al. 2010; Nehm and Schonfeld 2008, 2010), the CINS is recognized as an instrument capable of generating valid inferences about general levels of students' evolutionary knowledge. Each item of the CINS has one correct response option for

each question; therefore, the total score of the CINS instrument ranged from 0 to 20. While the original CINS paper suggests that it is a test only of natural selection knowledge, in fact it includes some questions about speciation, which is widely recognized as a macroevolutionary concept (Futuyma 2009). For our sample of biology students, the reliability of CINS scores (measured with Cronbach's alpha) was 0.7.

The second instrument that we used was the newly developed open-response ACORNS (Assessing COntextual Reasoning about Natural Selection; Nehm et al. 2012). We used four isomorphic ACORNS items, standardized by familiarity: (1) How would biologists explain how a living mouse species without claws evolved from an ancestral mouse species that had claws? (2) How would biologists explain how a living lily species without petals evolved from an ancestral lily species that had petals? (3) How would biologists explain how a living snail species with teeth evolved from an ancestral snail species that lacked teeth? (4) How would biologists explain how a living grape species with tendrils evolved from an ancestral grape species that lacked tendrils?

The ACORNS is a test of both microevolutionary and macroevolutionary knowledge because it prompts students to explain the causes responsible for between-species (i.e., macroevolutionary) change from a biologist's perspective. To score students' ACORNS responses, we utilized the published rubrics of Nehm and colleagues (Nehm et al. 2010a). This scoring rubric includes seven key conceptions and six naive ideas. This scoring rubric identifies three core key concepts as necessary and sufficient to explain evolutionary patterns using the natural selection model: (1) the presence and causes of variation (mutation, recombination, sex), (2) the heritability of variation, and (3) the differential reproduction and/or survival of individuals. It also includes four other key concepts that are widely accepted additional elements for explaining evolutionary patterns of change by natural selection: (4) hyper-fecundity or "overproduction" of offspring, (5) limited resources, (6) competition, and (7) a change in the distribution of produced phenotypic/genotypic variation across generations. The scoring rubric also includes naive ideas (e.g., "needs and/or goals cause evolutionary change," that "pressures" applied to organisms can "push" them to change, and that the disuse of phenotypic features proximally produces evolutionary loss). Key concept scores for each item ranged from 0 to 7, and naive idea scores for each item ranged from 0 to 6. The ACORNS responses were scored to consensus by two raters: a Ph.D. student in biology education and an evolutionary biologist. ACORNS reliabilities (measured using Cronbach's alpha) were 0.8 for key concepts, 0.6 for naive ideas, and 0.8 for reasoning about non-adaptive factors (for detailed scoring examples, see below).

The third approach that we used to explore students' evolutionary reasoning was clinical oral interviews. All 55 students were recruited as volunteers by the interviewer by e-mail

as well as at the beginning and end of various class periods and laboratory sessions, and these participants reflected the performance distribution in the overall sample. All participants were offered USD $20 for their participation. Approximately 16 hours of interviews were audio recorded and transcribed. The majors participated in more than ten hours of oral questioning (mean 19 minutes/student; range of 12–34 minutes). The advanced majors participated in more than six hours of oral questioning (mean 15 minutes/student; range 8–24 minutes).

The interview protocol was comprised of two ACORNS items (identical to those on the written instrument) and two novel, isomorphic items (i.e., taxa and traits of comparable familiarity; see Nehm et al. 2012). The two novel isomorphic items were included in the interview to minimize a potential testing effect (i.e., higher scores because of prior attempts to solve the same problem). While answering, students were prompted by the interviewer to elaborate on what they had said or to clarify what they meant by the words that they used. Follow-up questions included prompts such as "Can you tell me more about X?" "Can you explain what you mean when you use the word X?" and "Can you tell me a little bit more about how X would happen, in general terms?" Interviews were analyzed by two raters and scored 0 for the absence of non-adaptive factors, key concepts, or naive ideas, and 1 for the presence of non-adaptive factors, key concepts, or naive ideas. Holistic competency scores (−1, 0, +1) were also assigned to each student following Nehm and Schonfeld (2008). An evolutionary biologist and a biology education Ph.D. student assigned all oral interviews an overall competency score. Initial inter-rater reliabilities were 0.75 for oral interview scoring, and all discrepancies were subsequently resolved by deliberation. Consensus scores were used in all subsequent analyses.

We calculated descriptive statistics for the two samples (majors and advanced majors) and compared participant performance for all measured variables between the two groups using t tests. This information is useful for aligning our student sample with previously studied student samples in evolution education that used the same instruments (e.g., CINS). We also calculated Pearson correlation coefficients to examine putative interrelationships among measured variables from the multiple-choice CINS test, the open-response ACORNS test, and the clinical interviews. Variables included (1) the number of correct CINS scores, (2) the number of written key concepts of natural selection documented in the ACORNS, (3) the number of written naive ideas identified in the ACORNS, (4) the number of written non-adaptive evolutionary factors in the ACORNS, (5) the number of mentions (or "naming") of non-adaptive evolutionary factors in the clinical interviews, (6) the number of key concepts of natural selection in the interviews, and (7)

the naive ideas mentioned in the interviews. Pearson correlation coefficients were calculated in PASW v. 18.

Finally, we used qualitative methods to examine the structure of students' reasoning patterns. We examined transcripts of the interviews for (1) patterns of how students incorporated non-adaptive causal factors, such as genetic drift, into their explanations and (2) whether they conceptualized non-adaptive factors to be an alternative to, or synergistic with, selection. The purpose of including oral interviews was not only to validate inferences derived from students' written answers, but also to create a holistic snapshot of students' evolutionary reasoning patterns.

Results

Associations Among Evolutionary Reasoning Elements Pearson correlation analyses indicated that non-adaptive factors scores from the interview were significantly correlated with non-adaptive factors scores from the open-response ACORNS, but not with any scores from the multiple-choice CINS (Table 3). Specifically, both "mentioning" and "scientifically explaining" non-adaptive factors scores from the interviews displayed strong and significant associations with ACORNS non-adaptive factors scores ($r=0.86$, $p<0.01$ and $r=0.84$, $p<0.01$, respectively). Higher non-adaptive "explaining" scores were *not* significantly associated with greater key concept scores for the ACORNS ($r=0.04$, n.s.) or with higher CINS scores ($r=0.08$, n.s.). Thus, reasoning using non-adaptive factors appears to be a somewhat distinct reasoning pattern from selective reasoning.

Explanatory Elements Used by Majors and Advanced Majors We compared the types and frequencies of conceptual elements used in students' explanations of evolutionary change (Fig. 1). Though not significant, we found slight increases in students' accurate knowledge elements (key

concepts, CINS scores) between the samples of majors and advanced majors using both interview data and ACORNS data (CINS: $t_{47}=1.33$, $p>0.05$; ACORNS key concept: $t_{53}=0.63$, $p>0.05$; Interviews key concept: $t_{53}=1.31$, $p>0.05$). The types and frequencies of naive ideas also did not differ appreciably between the two groups (ACORNS: $t_{53}=0.88$, $p>0.05$; Interviews: $t_{53}=0.24$, $p>0.05$). In contrast, use of non-adaptive factors was much more frequent with the advanced majors, as measured by both the clinical interviews and the open-response ACORNS test (the CINS does not include non-adaptive options; Fig. 1). However, use of non-adaptive factors was only significant for those students who "mentioned" non-adaptive factors in their interview responses ($t_{28}=2.25$, $p<0.05$, $d=0.61$; Cohen 1988) and not for those whose non-adaptive factors were "scientific."

Explanatory Elements Used in the ACORNS and Clinical Interviews The key concepts of *variability*, *differential survival*, and *limited resources* were the three most frequently used in the ACORNS responses and in the clinical interviews (Fig. 2). However, the relative frequencies of these concepts differed for each assessment. In the interviews, *variability* was the most frequent, followed by *limited resources* and then *differential survival*. Overall, key concepts were used more often in the interviews than in the written ACORNS instrument. In the written ACORNS assessment, *differential survival* was used most frequently, followed by *variability* and then *limited resources*. Relative proportions of the other four key concepts were similar between written and oral responses, and in order of relative frequencies were *heritability*, *change in the frequency of a variant in a population*, *competition*, and *hyper-fecundity*. *Hyper-fecundity* was the least frequently applied key concept; it was never used in the ACORNS and was only used once in the interviews. Overall, key concepts were greater in number in the interviews than in the ACORNS, which is expected given the greater time allotted to oral questioning. Students spent almost twice as much time

Table 3 Correlations among scores derived from the CINS, ACORNS, and clinical interviews

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) CINS	1.000							
(2) ACORNS KC	0.426**	1.000						
(3) Interview KC	0.401**	0.566**	1.000					
(4) ACORNS MIS	−0.355**	−0.377**	−0.521**	1.000				
(5) Interview MIS	−0.158	−0.213	−0.354**	0.550**	1.000			
(6) ACORNS NAF	0.017	−0.038	0.046	−0.169	−0.043	1.000		
(7) Mentioning NAF	0.002	−0.089	0.046	−0.158	−0.121	0.861**	1.000	
(8) Scientific NAF	0.079	0.039	0.060	−0.157	−0.068	0.835**	0.851**	1.000

$N=55$

NAF non-adaptive factors, *KC* key concepts, *MIS* naive Ideas
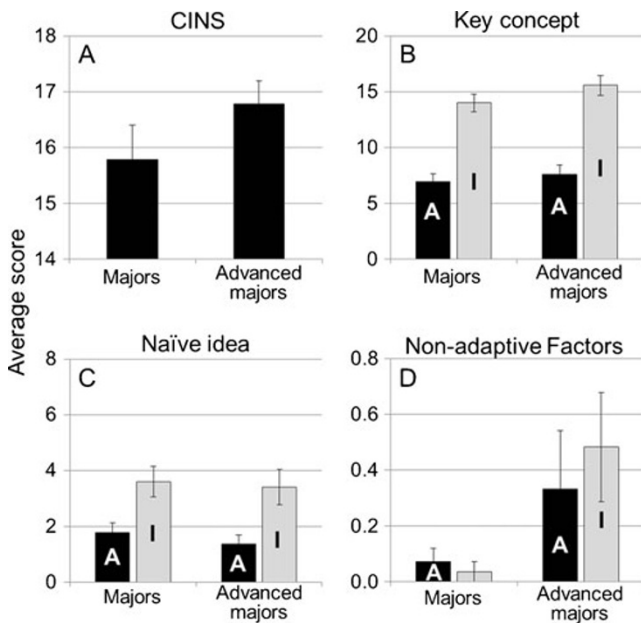
*$p<0.05$; **$p<0.01$

Fig. 1 *A* ACORNS and *I* interview. **a** Average scores of both majors and advanced majors for the CINS, **b** key concepts in the written ACORNS assessment and interviews **c** naive ideas used in the ACORNS and interviews, and **d** Non-adaptive ideas used in the ACORNS and interviews. Though advanced majors show a slight increase in key concepts used and a slight decrease in naive ideas used, this trend is non-significant. However, advanced majors do use significantly more non-adaptive ideas in their evolutionary explanations compared to majors

answering interview questions as they did answering ACORNS items (ACORNS: *M*=10.9, SD=7.2 minutes; Interview: *M*=21.9, SD=7.7 minutes (Fig. 2). Overall, however, there was good correspondence between measures derived from the two methods.
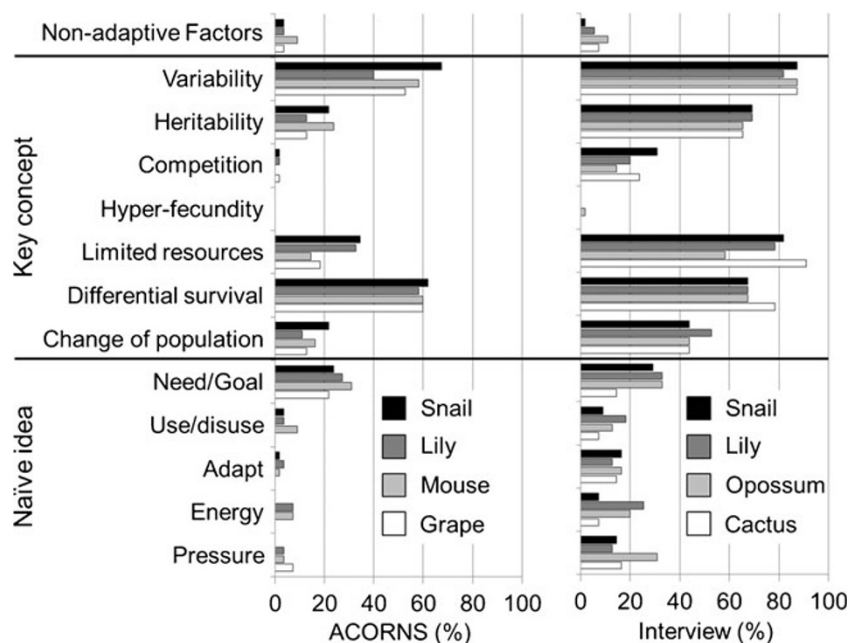
Similar to the key concept patterns that we documented, naive ideas were more abundant in interview responses than in the ACORNS responses (Fig. 2). In both oral interviews and in written responses, however, naive ideas were much more variable across items than key concepts. The most frequent naive idea used in both the interviews and in the ACORNS was *needs/goals*. The naive ideas *use/disuse*, *energy*, and *pressure* were about equally common in ACORNS responses. *Adapt* was the least common naive idea found in the ACORNS responses. In the interviews, the second most common naive idea was *pressure*, while *adapt* and *energy* were equally the next most frequent. *Use/disuse* was the least common naive idea used during interviews. Two notable patterns were noted relating to non-adaptive factors in both the ACORNS and in the interviews. First, non-adaptive factors were used in the first item more often than in the other items, and second, they were applied inconsistently across items (Fig. 2, top row). In short, the interviews tended to elicit greater frequencies of ideas, but not different ideas, than those revealed in the ACORNS. Non-adaptive factors were rarely used regardless of the method of detection used, and naive ideas were less common than key concepts.

Holistic Reasoning Patterns Interview data revealed that all of the students who incorporated non-adaptive factors into their explanations of evolutionary change (eight out of 55) presented non-adaptive factors in the form of "genetic drift," "bottleneck," or "founder's effect" (See Table 4). Moreover, in all cases students represented a non-adaptive factor as an alternative to natural selection. This finding is in line with how many evolutionary biologists conceptualize the two

Fig. 2 Diversity and abundance of explanatory elements used in the ACORNS and the clinical interviews. Both key concepts and naive ideas were more abundant in interviews than in the ACORNS. Non-adaptive factors were used more often in the first item than subsequent items

Table 4 Selected excerpts from students who used non-adaptive factors in clinical oral interviews

Participant	Non-adaptive terms used	Example use of term	Scientific NAF Score for item excerpted (0–1 scale)	Overall interview score (−1, 0 or +1)
A	Bottleneck, founder's effect	"It could be the bottleneck effect or the founder's effect, like the ancestor species could have had like a nat-, disaster, like a, calamity, so only a few of the, uh, individuals survived and those individuals like maybe had like a dwarf gene, or like a, like a gimp tail I guess. And then eventually it started being more prevalent in the, in the new species that started from there."	0.5	0
B	Bottleneck	"Or, um, if there were opossums with shorter tails and something happened to like bottleneck them or something, and it didn't have, uh, an advantage or disadvantage they could just go away randomly."	0.5	1
C	Founder's effect	(defining the word) "Um, it's like when, uh, part of the population gets isolated and, uh, um, it's like alleles get uh, a rare allele becomes dominant."	0	1
D	Founder effect, bottleneck	"Um, I guess that would be a founder effect I think, or a bottleneck. Um, so the species without petals not necessarily that it was favored but because it was in greater variety or greater, uh, it was in greater number than the ones without, that had the larger petals, it eventually, uh, I guess it was eventually derived to a species without petals, in the living species today."	0	1
E	Genetic drift	"I always forget the definition of genetic drift. Um, it's kinda just like, uh, a swing towards one extreme instead of where it was before, but I guess, so I guess that kind of takes away from what I was going towards, I guess, kinda contradicted myself."	0	1
F	Bottleneck, founder's effect, genetic drift	"I guess with genetic drift what would happen is, um, like a random grouping of like, let's say out of a hundred only four of these just happen to have, like, this mutation…so over time, like, their genes were the ones that were being passed down randomly, um, into the population, and it just happened to be that as time went on, it was their genes that caused the population, like, not really caused the population but, um, is what the population formed into."	0	1
G	Gene drift, genetic drift	"you've got your random genetic drift option where there was variation in the presence or absence of teeth in the first species, and then randomly, you know, maybe due to natural disasters, or just certain members got lucky in breeding, you'd have the fixture of the lack, or the presence of teeth."	1	1
H	Genetic drift	"I think the way genetic drift influence it, is that through time and passing on of traits and genes, that it was affected, oh gosh, I feel like I'm rambling and I'm not really saying anything. (long pause) I can't really think of what I want to say with genetic drift."	−1	0

concepts, where evolutionary change is attributed to either natural selection *or* genetic drift (e.g., Orr 1998).

While at first glance the observation that students discussed genetic drift as an alternative to natural selection was indicative of expert-like thinking (see the "Introduction" section), follow-up questions during the interviews often revealed that many students poorly understood the concepts of genetic drift and/or natural selection. Indeed, only one student (Participant G from the advanced majors sample) accurately and consistently used both genetic drift and natural selection as possible mechanisms for evolutionary change across the four interview items (see below).

> Participant G: Another way would be more like genetic drift oriented, where, the presence or absence of a tail doesn't matter very much. So, maybe it has no

fitness ramifications in survivability and it just kinda [sic] fluctuates based on just randomness, you know, statistics, until eventually you've got all the opossums don't have tails. That sort of random stuff does happen, sometimes, and it brings genes towards fixture that way. But, I would say, cause it, I would say most biologists would go with either natural selection or gene drift to explain the loss of tail over time in opossums.

A second student from the advanced majors presented "genetic drift" as one possible solution for the fourth interview item about cactus spines (see the "Methods" section).

> Participant E: I guess that could be similar with either, like a genetic mutation or maybe a genetic drift and, uh, just, could also have to do with, uh, like being an

anti-predatory defense to, uh, protect it since it's in a harsh environment already, they kinda [sic] have to guard themselves from anything that's going to get water out of them.

Interviewer: You said that it could be a genetic mutation or genetic drift, can you explain what you mean by genetic drift?

Participant E: I always forget the definition of genetic drift. Um, it's kinda [sic] just like, uh, a swing towards one extreme instead of where it was before, but I guess, so I guess that kind of takes away from what I was going towards, I guess, kinda [sic] contradicted myself.

Interviewer: What's swinging? What is it that is swinging towards one extreme?

Participant E: Uh, just kinda [sic] like, the genetic makeup or the, uh, actual structures that are present are kinda [sic] more of a shift from one to another based on the pressures that they're getting from the environment or from, from other species around them.

Interviewer: When you say structures, do you mean, um, what do you mean by that?

Participant E: Um, sorta [sic] like the actual spines, in this case, or uh, the the [sic] cells within, that are used to maintain the water and the moisture and hold it.

Interviewer: And how, how does that process of swinging towards one of the extremes or another come about?

Participant E: Um, I guess that would be more, type of a, kind of, more of natural selection, where they're, where if they don't change then they're just going to die out, so without, if the, if they didn't develop the spines, there'd be no way for them to completely protect themselves, they would, without spines they'd probably evolved in another way and came up with some sort of poison or something else that would have deterred predators and stuff from eating them.

While participant E mentions the term "genetic drift," his explanation for this term was scientifically inaccurate (he confuses genetic drift with directional selection, and he was therefore given a scientific non-adaptive factors score of "0"). His answer demonstrates that a student's use of a scientific term is not always indicative of scientific understanding of a concept (cf. Rector et al. 2012). This situation also occurred with students' use of the terms "bottleneck" and "founder effect."

Overall, clinical interviews corroborated our statistical findings that knowledge of non-adaptive factors is not associated with greater understanding of natural selection, or

with fewer naive evolutionary ideas. Similarly, in clinical interviews and in ACORNS responses, most students inconsistently applied non-adaptive factors, or lacked what has been termed "knowledge coherence" (Kampourakis and Zogza 2009).

Evolutionary Reasoning Competencies in the Two Groups We found that on the spectrum of novice to expert reasoning about evolutionary change, the majority of our students (both majors and advanced majors) fell into the *novice* category (Table 2, Fig. 3). Recall that the novice category was characterized by naive or naive+scientific reasoning. Based upon the written explanations, only one student from the majors held a purely *naive* model, while no students from the advanced majors group held such models. In the interviews, none of the students from either group were found to exhibit purely naive models. Most students exhibited mixed models comprised of both naive ideas and key concepts (Fig. 3). Written responses indicated that 16 majors and 14 advanced majors held mixed models, whereas interview responses indicated that both groups had 19 students exhibiting mixed models. Interestingly, a small number of students from each group displayed mixed models while also incorporating non-adaptive reasoning into their explanations (ACORNS: two from the majors, two from the advanced majors; Interviews: one from the majors, four from the advanced majors) (Fig. 3). There were a
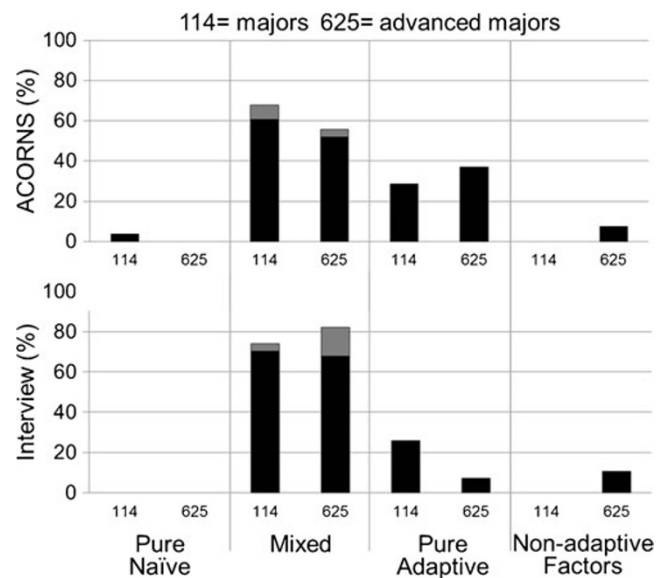

Fig. 3 Percentage of students who fall into each category of expertise. The majority of students used mixed models, though a small portion of these also included non-adaptive factors into their explanations (*shaded gray*). A large portion of students also held pure adaptive models and would therefore be considered emerging experts. Only advanced majors reached expert-like levels of reasoning and these students used adaptive vs. non-adaptive models. No students held adaptive+non-adaptive models of evolutionary reasoning

number of students categorized as *emerging experts* based on their written explanations using purely adaptive models (eight majors, ten advanced majors). During the interviews, in contrast, seven majors exhibited purely adaptive models, although only two advanced majors used key concepts exclusively (Fig. 3). Nevertheless, a small number of advanced majors fell into the *expert* category (that is, students explaining evolutionary change using both adaptive and non-adaptive conceptual models; ACORNS: two students; Interviews: three students). No students from the majors reached the *expert* level.

It was challenging to unambiguously situate a few students along our novice–expert continuum. For instance, one student from the advanced majors displayed a complex mixed model (naive ideas+key concept+non-adaptive factors) in the interview (See Supplementary Materials, Participant F), but not in the written assessment. Based on her ACORNS responses, we placed Participant F in the adaptive vs. non-adaptive model of reasoning even though she had provided explanations using non-adaptive factors exclusively. She was placed in this category because she had used key concepts (*variability* and *change of population*) in her explanations of genetic drift. Her response demonstrated that some key concepts were not specific to natural selection; in fact, key concepts such as *variability* and *change of population* are necessary for explaining evolutionary change by stochastic processes (e.g., genetic drift) as well as deterministic processes (e.g., natural selection). Accordingly, this student was labeled as having an expert-like model of evolutionary reasoning. However, placing students along a continuum of evolutionary reasoning competency was straightforward in most cases. Classifying our two student samples revealed that, regardless of the assessment method or amount of biology coursework completed, and despite direct instruction of genetic drift, most students failed to reach expert-like levels of reasoning about evolutionary change.

Discussion

While science education research has produced a large body of work investigating the learning, teaching and assessment of natural selection (e.g., Bishop and Anderson 1990; Demastes et al. 1995; Settlage 1994; Nehm and Schonfeld 2008; Nehm and Ha 2011), strikingly few studies have focused on students' thinking about non-adaptive evolutionary factors such as genetic drift, despite its important role in experts' empirical tests and theoretical models of evolutionary change (e.g., Ackermann and Cheverud 2004; Lande 1976; Orr 1998; Parker and Maynard Smith 1990). This gap in the literature motivated our exploration of evolutionary reasoning using non-adaptive factors in undergraduate students, and whether such patterns differed between groups of introductory and advanced biology students exposed to varying degrees of non-adaptive content. We used three different methodologies to explore students' evolutionary reasoning: the multiple-choice CINS, which ignores non-adaptive factors but measures understanding of natural selection and speciation; the ACORNS, a constructed-response format test that captures students' evolutionary explanations across contextual features; and extended clinical oral interviews. We used these three different methods to rigorously and holistically determine how, and to what degree, students used stochastic and non-deterministic factors to explain evolutionary change, and how their explanatory models were constructed.

Students' Use of Non-adaptive Factors Given Gould and Lewontin's (1979) widely cited criticisms about biologists' over-reliance on exclusively adaptive factors to explain evolutionary change, and contemporary biologists' use of hypothesis tests that explore the relative contributions of non-adaptive mechanisms such as genetic drift in evolutionary change, the goal for biology students should be to understand and apply the mechanisms that the field of evolutionary biology currently uses to account for organismal diversity through time and space (Gould 2002). In short, students' competency in evolutionary reasoning should be measured by their ability to consider both drift and selection as possible causal mechanisms in their explanations for trait change. What our study reveals is that the vast majority of our participants have not reached this competency benchmark.

Student use of non-adaptive factors in their reasoning is not associated with greater knowledge of natural selection or with fewer naive ideas about evolution. Most students in the majors and advanced major samples used mixed models of evolutionary reasoning, suggesting that despite increased instruction in evolution, most students failed to progress in their understanding of evolutionary change. Instead, their responses imply that they added new concepts (e.g., *genetic drift* or *differential survival*) into their existing naive explanatory frameworks. Regardless of the amount of biology coursework that students completed, they expressed both key concepts and naive ideas at comparatively the same frequency when cued to reason about evolutionary change (Fig. 3). Nevertheless, advanced majors' relatively greater use of non-adaptive factors suggests that increased exposure to instruction about non-adaptive processes, such as genetic drift, is in fact associated with increases in students' use of non-adaptive factors in their evolutionary explanations, (albeit not at desired magnitudes; see Fig. 1).

The Structure of Students' Explanatory Models of Evolutionary Change When students did use non-adaptive factors in their explanatory models of evolution, they either used them within a mixed model of reasoning, or they used them

as competing mechanisms of evolutionary change (i.e., in an expert-like model of adaptive vs. non-adaptive reasoning; Fig. 3). Additionally, most students did not consistently apply non-adaptive factors across items (Fig. 2). It could be that the association of this concept within students' evolutionary reasoning framework is not theory-like (Vosniadou et al. 2008), that it was selectively cued across items differing in surface features (Nehm et al. 2012; Nehm and Ha 2011) or that students simply prefer one scientific model over another (Bachelard 1968; Duit 2003). Indeed, only two of the three students from the advanced majors group who did use non-adaptive factors did so consistently and expressed an adaptive vs. non-adaptive model of reasoning about evolutionary change. This corroborates previous work suggesting that expert-like evolutionary reasoning models include more stable associations of concepts within conceptual reasoning frameworks (Nehm and Ha 2011; Nehm and Ridgway 2011).

Those students who reached an expert-like model of evolutionary causation adopted models of adaptive vs. non-adaptive change, but none of them expressed integrative causal models (i.e., the top level of Table 2). For example, Participant G (see excerpt above) was a student from the advanced majors group who included both natural selection and non-adaptive factors across all four items in both the ACORNS written assessment and in the interview. Participant G employed what we term an expert-like model of reasoning, although one that frames adaptive causation as *competing* with stochastic causation of evolutionary change. Another student, Participant F, provided only non-adaptive factors in the ACORNS explanations, yet also discussed non-adaptive factors and selection as possible alternative causal mechanisms of evolutionary change in the interviews. It is interesting that the written format of the ACORNS elicited non-adaptive factors only, while the interviews elicited both possible mechanisms. However, this seems to be an anomaly and not significant enough to warrant any adjustments to our framework of novice–expert reasoning about evolutionary change (Fig. 2). Regardless of assessment format, this student reached expert levels of reasoning.

Implications for Teaching and Learning Our study demonstrates that current approaches to teaching genetic drift and other non-deterministic processes may not be effective at helping the majority of students build expert-like models of evolutionary causation. Students' models of evolutionary change in our sample were overwhelmingly composed of a combination of naive ideas and key concepts of natural selection (Fig. 2). Not only is this the case for majors in introductory courses; it is also true of advanced majors who have successfully completed an entire course in evolutionary biology. Exposure to genetic drift does not appear to be sufficient for inducing students to accommodate non-adaptive factors into their mental models of evolutionary change and build

expert-like explanatory models (i.e., Table 2, top row). This finding suggests that teachers should move beyond definitions or simulations of genetic drift (e.g., Staub 2002) and illustrate to students how genetic drift is conceptually structured within models of evolutionary causation. Teachers could present cases of how evolutionary biologists currently use both genetic drift and natural selection to test alternative hypotheses and build explanatory models of evolutionary change. Such examples may facilitate more advanced perspectives on non-adaptive and adaptive causal factors in evolutionary biology (e.g., Orr 1998; Ackermann and Cheverud 2004).

Our findings also suggest that evolutionary reasoning using non-adaptive factors and scientifically accurate selective reasoning are distinct. This raises the question of whether the lack of an association among non-adaptive factors, key concepts, and naive ideas is a product of the way these evolutionary concepts are taught, or if it is an intrinsic way of thinking about evolution. For instance, perhaps students fail to use non-adaptive factors in their evolutionary explanations because they view the lack of emphasis on drift in the classroom as indicative of its relatively low level of importance in evolutionary change. On the other hand, students may fail to use non-adaptive factors in their evolutionary explanations because the concept is more difficult to accommodate into existing cognitive frameworks compared to selective reasoning concepts. Regardless, based on students' response patterns, it appears that they have not sufficiently accommodated non-adaptive factors into their evolutionary reasoning frameworks, even after advanced instruction in evolution. Rather, interviews and open-response assessments clearly indicate that most students at introductory and advanced levels of instruction assimilate non-adaptive concepts with key concepts and naive ideas of selective reasoning into unstable (i.e., non-coherent), naive models of evolutionary causation (cf. Kampourakis and Zogza 2009; Nehm and Ha 2011). It is important to note, however, that student response patterns may not always mirror cognitive processes, and our interpretations are constrained by the methods we used to uncover student thinking and the sample that we studied.

The finding that *no* students used adaptive+non-adaptive causation models in their explanations of evolutionary change raises the question of whether this, too, is a product of teaching experiences or a default approach to thinking about evolution. Non-adaptive factors such as genetic drift often receive minor instructional focus and are presented as an alternative model to natural selection (e.g., Linhart 1997). It is possible that standard approaches prevent students from conceptualizing cases in which stochastic processes work in tandem with selective processes to generate patterns of genotypic and phenotypic change. Such integrated causal models (adaptive +non-adaptive) are perhaps more complex, and thereby may be beyond the grasp of students still struggling to restructure common naive ideas. Indeed, it may require more

comprehensive evolution instruction using extensive theoretical and experimental examples. However, it is important to note that some evolution experts continue to use an adaptive vs. non-adaptive model of evolutionary change in their research programs (e.g., Orr 1998), and so it is possible that an adaptive+non-adaptive model is simply a less common, alternative framework for experts within evolutionary biology. Further work is needed to explore this issue.

Overall, our study highlights the fact that non-adaptive factors receive very little attention in the science education or evolution education literature. Future research exploring novice and expert understanding and application of non-adaptive factors would help to reveal what a more expert-like model of evolutionary reasoning would look like and could inform the design of evolution learning progressions (cf. Catley et al. 2005).

Implications for Assessment Assessments are central to helping teachers foster meaningful science learning (NRC 2001b). However, it is imperative that those assessments meet quality control criteria established by the educational measurement community (AERA et al. 1999). Measurement instruments, among other things, must comprehensively assess all facets of a well-defined construct (Neumann et al. 2011). In the domain of evolutionary biology, natural selection and genetic drift are the two most important causal factors that biologists use to explain evolutionary change (Orr 1998). Thus, to measure evolutionary thinking, instruments must provide the opportunity for students to explain change using natural selection, genetic drift, or combinations of drift+selection. Currently, there are no such instruments.

One important consideration for helping teachers to understand students' thinking about evolution is to employ instruments that are capable of capturing progress in students' conceptual growth. As we argue, measuring progress in evolutionary reasoning requires the consideration of both non-adaptive and adaptive factors in evolutionary causation, in addition to common naive ideas about evolutionary change (Table 2). Many widely used instruments, nevertheless, ignore the possibility that non-adaptive factors (such as genetic drift) could contribute to patterns of evolutionary change that the instrument scenarios present (e.g., Bird beak evolution in Anderson et al.'s CINS instrument). Thus, teachers across the educational hierarchy must develop and deploy instruments that include the measurement of reasoning using non-adaptive factors in formative and summative assessments of evolutionary thinking.

Multiple-choice assessments are often a popular choice among instructors considering the time and expertise needed to develop and grade open-ended items. However, multiple-choice items rarely provide mixed model options (that is, not just right or wrong options), despite the fact that such reasoning models may be very common in samples (as we found in

our study; Nehm 2011). If we want our assessments to fulfill their purposes, we must consider the large proportion of students who have mixed models for explaining evolutionary change and revise our multiple choice assessments to reflect this well established finding. No multiple-choice evolution instruments to our knowledge allow for the measurement of mixed models.

Study Limitations and Implications for Future Work on Evolutionary Reasoning Using Non-adaptive Factors Although we were able to describe *how* students reason about evolutionary change and their use of non-adaptive factors, we were not able to determine *why* students did not incorporate non-adaptive factors into their models of evolutionary causation. Because all students in our sample were assigned readings that covered genetic drift, and had the opportunity to listen to lectures that included the topic of genetic drift, the lack of non-adaptive factors in students' explanations could be due to poor teaching of the topic. Additionally, it is unclear whether teachers incorporated developmental constraints, or other lesser-known non-adaptive factors (see Gould 2002), into their teaching in all the biology classes or if it was only found in the evolution course. It is possible that other student groups exposed to instruction on drift in a more integrated and sophisticated manner would show different patterns than those observed with our sample. Thus, our findings may not generalize to other student samples. Future studies should explore the relationship between instruction and evolutionary reasoning using non-adaptive factors. Such a focus would help determine whether specifically addressing non-adaptive factors (beyond genetic drift) through instruction or teaching non-adaptive factors in different ways, influences students' use of non-adaptive factors in their evolutionary explanations.

References

Ackermann RR, Cheverud JM. Detecting genetic drift versus selection in human evolution. Proc Natl Acad Sci USA. 2004;101:17946–51.

American Association for the Advancement of Science: Vision and Change in Undergraduate Biology Education, Washington, DC: http://visionandchange.org (2011).

American Educational Research Association: American Psychological Association, and National Council on Measurement in Education (AERA, APA, and NCME). Standards for educational and psychological testing. Washington, DC: AERA; 1999.

Anderson DL, Fisher KM, Norman GJ. Development and evaluation of the conceptual inventory of natural selection. J Res Sci Teach. 2002;39:952–78.

Bachelard G. The philosophy of no. A philosophy of the new scientific mind. New York: The Orion Press; 1968.

Battisti BT, Hanegan N, Sudweeks R, Cates R. Using item response theory to conduct a distracter analysis on Conceptual Inventory of Natural Selection. Int J Sci Math Educ. 2010;8:845–68.

Beatty J. Chance and natural selection. Philos Sci. 1984;51(2):183–211.

Bishop BA, Anderson CW. Student conceptions of natural selection and its role in evolution. J Res Sci Teach. 1990;27(5):415–27.

Bowler P. The eclipse of darwinism: anti-darwinian evolution theories in the decades around 1900. The Johns Hopkins University Press: Baltimore. 1983.

Campbell NA, Reece JB, Urry LA, Cain ML, Wasserman SA, Minorsky PV, Jackson RB. Biology. 8th ed. New York: Pearson Benjamin Cummings; 2008.

Carpenter P, Bullock A, Potter J. Textbooks in teaching and learning: the views of students and their teachers. Brookes e-J Lear Teach. 2006;2:1–10.

Catley K, Lehrer R, Reiser B. Tracing a prospective learning progression for developing understanding of evolution. Paper commissioned by the National Academies Committee on Test Design for K-12 Science Achievement. Washington, DC: National Academies. http://www7.nationalacademies.org/bota/Test_Design_K-12_Science.html (2005).

Chi MTH, Feltovich PJ, Glaser R. Categorization and representation of physics problems by experts and novices. Cogn Sci. 1981;5(2):121–52.

Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale: Lawrence Erlbaum; 1988.

Conway Morris S. Life's solution: inevitable humans in a Lonely Universe. Cambridge: Cambridge University Press; 2003.

Cuseo J. The empirical case against large class size: adverse effects on the teaching, learning, and retention of first-year students. J Fac Dev. 2007;21(1):5–21.

Dagher ZR, BouJaoude S. Scientific views and religious beliefs of college students: the case of biological evolution. J Res Sci Teach. 1997;34(5):429–45.

Darwin C. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. London: John Murray; 1859.

Dawkins R. The selfish gene. Oxford: Oxford University Press; 1976.

Demastes SS, Good RG, Peebles P. Students' conceptual ecologies and the process of conceptual change in evolution. Sci Educ. 1995a;79(6):637–66.

Demastes SS, Settlage J, Good RG. Students' conceptions of natural selection and its role in evolution: cases of replication and comparison. J Res Sci Teach. 1995b;32(5):535–50.

Demastes SS, Good RG, Peebles P. Patterns of conceptual change in evolution. J Res Sci Teach. 1996;33(4):407–31.

Dobzhansky T. Nothing in biology makes sense except in the light of evolution. Am Biol Teach. 1973;35(3):125–9.

Duit R. Conceptual change: a powerful framework for improving science teaching and learning. Int J Sci Educ. 2003;25(6):671–88.

Feldhamer GA, Drickamer LC, Vessey SH, Merritt JF, Krajewski CW. Mammalogy: Adaptation, diversity, ecology. 3rd edition. Baltimore, MD: Johns Hopkins University Press. 2007.

Ferrari M, Chi MTH. The nature of naïve explanations of natural selection. Int J Sci Educ. 1998;20(10):1231–56.

Fisher RA. On the dominance ratio. P Roy Soc Edinb B. 1922;42:321–41.

Fosnot CT. Constructivism: a psychological theory of learning. In: Fosnot CT, editor. Constructivism: theory, perspectives and practice. New York: Teachers College Press; 1996. p. 8–33.

Freeman S. Biological science. 2nd ed. Upper Saddle River: Pearson Prentice Hall; 2005.

Futuyma DJ. Evolutionary biology. 3rd ed. Sunderland: Sinauer Associates, Inc; 1998.

Futuyma DJ. Evolution. 2nd ed. Sunderland: Sinauer Associates, Inc; 2009.

Gould SJ. The structure of evolutionary theory. Cambridge: Belknap Press; 2002.

Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. P Roy Soc Lond B Bio. 1979;205:581–98.

Gregory TR. Understanding natural selection: essential concepts and common misconceptions. Evol Educ Outreach. 2009;2(2):156–75.

Hagedoorn AL, Hagedoorn AC. The relative value of the processes causing evolution. The Hague: Martinus Nijhoff; 1921.

Kampourakis K, Zogza V. Preliminary evolutionary explanations: a basic framework for conceptual change and explanatory coherence in evolution. Sci Educ. 2009;18(10):1313–40.

Keil FC, Wilson RA. Explanation and cognition. Cambridge: MIT Press; 2000.

Kelemen D, Rosset E. The human function compunction: teleological explanation in adults. Cognition. 2009;111(1):138–43.

Kimura M. Evolutionary rate at the molecular level. Nature. 1968;217:624–6.

Lande R. Natural selection and random genetic drift in phenotypic evolution. Evolution. 1976;30(2):314–34.

Linhart YB. The teaching of evolution: we need to do better. Bioscience. 1997;47(6):385–91.

Lynch M. The evolution of genetic networks by non-adaptive processes. Nat Rev Genet. 2007;8:803–13.

Maret TJ, Rissing SW. Exploring genetic drift & natural selection through a simulation activity. Am Biol Teach. 1998;60(9):681–3.

Matthen M, Ariew A. Two ways of thinking about fitness and natural selection. J Philos. 2002;99(2):55–83.

Mayr E. How to carry out the adaptationist program? Am Nat. 1983;121:324–34.

Millstein RL. Are random drift and natural selection conceptually distinct? Biol Philos. 2002;17:33–53.

National Research Council. Investigating the influence of standards: a framework for research in mathematics, science, and technology education. Washington, DC: National Academy Press; 2001a.

National Research Council. Knowing what students know. Washington, DC: National Academy Press; 2001b.

Nehm RH. Understanding undergraduates' problem solving processes. J Biol Microbiol Ed. 2010;1(2):119–21.

Nehm RH, Ha M. Item feature effects in evolution assessment. J Res Sci Teach. 2011;48(3):237–56.

Nehm RH, Reilly L. Biology majors' knowledge and misconceptions of natural selection. Bioscience. 2007;57(3):263–72.

Nehm RH, Ridgway J. What do experts and novices "see" in evolutionary problems? Evol Educ Outreach. 2011;4(4):666–79.

Nehm RH, Schonfeld IS. Does increasing biology teacher knowledge of evolution and the nature of science lead to greater preference for the teaching of evolution in schools? J Sci Teac Educ. 2007;18(5):699–723.

Nehm RH, Schonfeld IS. Measuring knowledge of natural selection: a comparison of the CINS, an open-response instrument, and an oral interview. J Res Sci Teach. 2008;45(10):1131–60.

Nehm RH, Schonfeld IS. The future of natural selection knowledge measurement. A reply to Anderson et al. J Res Sci Teach. 2010;47(3):258–362.

Nehm RH, Kim SY, Sheppard K. Academic preparation in biology and advocacy for teaching evolution: biology versus non-biology teachers. Sci Educ. 2009;93(6):1122–46.

Nehm RH, Ha M, Rector M, Opfer J, Perrin L, Ridgway J, Mollohan K. Scoring guide for the open response instrument (ORI) and evolutionary gain and loss test (EGALT). Technical Report of National Science Foundation REESE Project 0909999. http://evolutionassessment.org/Evo_Assessment/Scoring.html(2010a). Accessed 18 February 2012.

Nehm RH, Rector MA, Ha M. "Force-Talk" in evolutionary explanation: metaphors and misconceptions. Evol Educ Outreach. 2010b;3:605–13.

Nehm RH, Beggrow E, Opfer J, Ha M. Reasoning about natural selection: diagnosing contextual competency using the ACORNS instrument. Am Biol Teach. 2012;74(2):92–8.

Neumann I, Neumann K, Nehm RH. Evaluating instrument quality in science education: Rasch-based analyses of a nature of science test. Int J Sci Educ. 2011;33(10):1373–405.

Nickels MK, Nelson CE, Beard J. Better biology teaching by emphasizing evolution & the nature of science. Am Biol Teach. 1996;58 (6):332–6.

Opfer JE, Nehm RH, Ha M. (2012) Cognitive foundations for science assessment design: knowing what students know about evolution. J Res Sci Teach (in press).

Orr HA. Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. Genetics. 1998;149: 2099–104.

Parker GA, Maynard Smith J. Optimality theory in evolutionary biology. Nature. 1990;348:27–33.

Pigliucci M, Müller GB. Elements of an extended evolutionary synthesis. In: Pigliucci M, Müller GB, editors. Evolution, the extended synthesis. Cambridge: MIT Press; 2010. p. 3–17.

Posner GJ, Strike KA, Hewson PW, Gertzog WA. Accommodation of a scientific conception: towards a theory of conceptual change. Sci Educ. 1982;67(4):489–508.

Provine WB. The origins of theoretical population genetics. Chicago: The University of Chicago Press; 1971.

Randall L. Knocking on heaven's door: How physics and scientific thinking illuminate the universe and the modern world. New York: Ecco Publishers; 2011. p. 442.

Rector MA, Nehm R, Pearl D (2012) Learning the language of evolution: Lexical ambiguity and word meaning in student explanations. Res Sci Educ (in press).

Ridley M. Evolution. 3rd ed. Malden: Blackwell Publishing; 2004.

Rosenberg A, McShea DW. Further problems of Darwinism: constraint, drift, function. In philosophy of biology: a contemporary introduction. New York: Routledge; 2008. p. 65–95.

Savage JM. Evolution. 2nd ed. New York: Holt, Rinehart and Winston, Inc; 1969.

Settlage J. Conceptions of natural selection: a snapshot of the sense-making process. J Res Sci Teach. 1994;31(5):449–57.

Sinatra GM, Brem SK, Evans EM. Changing minds? Implications of conceptual change for teaching and learning about biological evolution. Evol Educ Outreach. 2008;1:189–95.

Sober E. The nature of selection: evolutionary theory in philosophical focus. Cambridge: MIT Press; 1984.

Staub NL. Teaching evolutionary mechanisms: genetic drift and M&M's®. Bioscience. 2002;52(4):373–7.

Voltaire: Candide. Cricket House Books, LLC. http://books.google.com/books?id=x4Te1GEczYC&pg=PA86&dq=candide+cricket+house&hl=en&sa=X&ei=0ADQT9TtGYbz0gGRurzQDQ&ved=0CDYQ6AEwAA#v=onepage&q&f=false (2010). Accessed 6 June 2012.

Vosniadou S. Conceptual change research: state of the art and future directions. In: Schnotz W, Vosniadou S, Carretero M, editors. New perspectives on conceptual change. Pergamon: Oxford; 1999. p. 3–14.

Vosniadou S, Vamvakoussi X, Skopeliti I. The framework theory approach to the problem of conceptual change. In: Vosniadou S, editor. International handbook of research on conceptual change. New York: Routledge; 2008. p. 3–34.

Wallace AR. Contributions to the theory of natural selection. 2nd ed. New York: Macmillan and Co; 1871.

Wandersee JH, Mintzes JJ, Novak JD. Research on alternative conceptions in science. In: Gabel DL, editor. Handbook of research on science teaching and learning. New York: Macmillan Publishing Company; 1994. p. 177–210.