

Exploring Students' Descriptions of Mutation from a Cognitive Perspective Suggests How to Modify Instructional Approaches

FangFang Zhao^{1*} and Anita Schuchardt²

¹Department of Curriculum and Instruction and ²Department of Biology Teaching and Learning, University of Minnesota, Twin Cities, Minneapolis, MN 55455

ABSTRACT

Prior studies have shown that students have difficulty understanding the role of mutation in evolution and genetics. However, little is known about unifying themes underlying students' difficulty with mutation. In this study, we examined students' written explanations about mutation from a cognitive science perspective. According to one cognitive perspective, scientific phenomena can be perceived as entities or processes, and the miscategorization of processes as entities can lead to noncanonical ideas about scientific phenomena that are difficult to change. Students' incorrect categorization of processes as entities is well documented in physics but has not been studied in biology. Unlike other scientific phenomena that have been studied, the word "mutation" refers to both the process causing a change in the DNA and the entity, the altered DNA, making mutation a relevant concept for exploration and extension of this theory. In this study, we show that, even after instruction on mutation, the majority of students provided entity-focused descriptions of mutation in response to a question that prompted for a process-focused description in a lizard or a bacterial population. Students' noncanonical ideas about mutation occurred in both entity- and process-focused descriptions. Implications for conceptual understanding and instruction are discussed.

INTRODUCTION

In biological organisms, mutations (changes in DNA sequence) are the source of variation. Variation is a key concept in the biological sciences that impacts students' understanding of two core areas, evolution and genetics (Shaw *et al.*, 2008; American Association for the Advancement of Science [AAAS], 2011). However, few studies have been conducted on students' understanding of mutation. Existing studies have investigated students' misconceptions of mutation in the context of natural selection (Brumby, 1984; Abraham *et al.*, 2009), genetics (Shaw *et al.*, 2008; Smith and Knight, 2012), and more broadly as a source of variation (Prevost *et al.*, 2013; Bray Speth *et al.*, 2014). This study investigates students' understanding of the origin of mutations using a cognitive approach to examine how students describe mutation in their written responses to a short-answer question. The results provide insight into students' perspectives on mutation that might be underlying their expression of noncanonical ideas in multiple contexts.

Mutation Is a Fundamental Concept That Is Difficult for Students to Understand

The idea of mutation is fundamental in understanding biology. Mutation can be defined as a "technical term for some change in the individual gene" (Carlson, 2011, p. 1). The process of mutation can involve chromosomal rearrangements or the generation of small changes at the DNA level through errors in the action of DNA polymerase when replicating DNA. These errors either can occur spontaneously or can

Ross Nehm, *Monitoring Editor*

Submitted Nov 21, 2018; Revised May 2, 2019;
Accepted May 22, 2019

CBE Life Sci Educ September 1, 2019 18:ar45
DOI:10.1187/cbe.18-11-0225

*Address correspondence to: FangFang Zhao
(zhaox945@umn.edu).

© 2019 FF. Zhao and A. Schuchardt. CBE—Life Sciences Education © 2019 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

be induced by changes in the DNA structure caused by chemicals or X-rays. Mutation provides the source of variation essential for evolution to occur and is responsible for the cumulative changes that result in cancer. The field of genetics studies how mutations are passed from parents to offspring. A scientific understanding of mutation is therefore vital for biologists to understand evolution and genetics and also for the public to make sense of issues such as genetically modified foods and exposure to mutagens.

Research has shown that mutation is a difficult concept for students in multiple contexts. In genetics education, where an understanding of mutation is critical to understanding genetic diseases, existing research has shown that students often hold incorrect ideas about the nature and consequences of mutations (Shaw *et al.*, 2008; Smith and Knight, 2012). For example, on the Genetics Concept Assessment, students selected answers stating that “A change is a mutation only if it will produce a change in the amino acid sequence” and that “A frameshift mutation cannot lead to an early stop of translation” (Smith and Knight, 2012). Similarly, Shaw *et al.* (2008) found that students hold noncanonical ideas about the effect of a mutation on an organism. In essays responding to open-ended questions about genetics, students frequently wrote that one mutation always causes one disease, a deterministic idea that is not consistent with the expert perspective that is more probabilistic about the outcomes of mutations. Holding deterministic views about mutation may explain why some students have difficulty with concepts such as incomplete dominance or the idea that inheriting a mutation in a cancer-associated gene is not a death sentence but only increases the probability of eventually developing cancer. While instruction that aligns with promising practices in science education has been shown to decrease the frequency of students selecting the most common incorrect answers for questions on the nature or effects of mutations, the incorrect response rate is still greater than 25% (Smith and Knight, 2012).

Mutation is also fundamental to understanding evolution, a core concept for biology literacy, because mutation provides the variation that is acted upon by evolutionary processes (AAAS, 2011). Multiple studies have found that students have difficulty understanding the role of mutation in the process of evolution, either not recognizing the role of mutation as the source of organismal variation or believing mutation is caused by environmental change (Brumby, 1984; Abraham *et al.*, 2009; Gregory, 2009; Nehm and Ridgway, 2011; Prevost *et al.*, 2013; Bray Speth *et al.*, 2014). Even after explicit instruction, one-third of 182 students did not include mutation in their conceptual models for the origin of variation in the context of evolution, suggesting that failing to recognize the role of mutation in evolution is relatively robust to instruction (Bray Speth *et al.*, 2014). When students do recognize the role of mutation in producing variation, they often state that mutations will suddenly appear in response to environmental changes (Brumby, 1984; Abraham *et al.*, 2009; Gregory, 2009; Nehm and Ridgway, 2011). Even after instruction designed to target this noncanonical idea, 40% of students chose a response indicating that mutations were caused by a change in the environment (Clarke-Midura *et al.*, 2018).

Many of these studies explored students’ understanding of the role or effects of mutation within the context of genetics or evolution (Shaw *et al.*, 2008; Gregory, 2009; Smith and Knight,

2012; Bray Speth *et al.*, 2014). Only rarely has students’ understanding about the mechanism of mutation been investigated. One exception is Prevost *et al.*’s (2013) study, which asked students to describe in writing how new alleles arise in a population. This lexical analysis revealed that students tended to describe types of mutations instead of the mechanism producing mutations. One limitation of this study was that, in a lexical analysis, the removal of context from the analysis may have obscured the extent to which students understand the mutation mechanism.

Prevost *et al.* (2013) did not suggest an explanation for their results. Using theory derived from cognitive science, we hypothesize that students were treating mutations as an entity that existed independent of the process that produced it. We conducted an in-depth qualitative analysis of students’ written responses to investigate whether students have a tendency to think of mutation only as an entity and not as the process/mechanism that produces changes in DNA. Students’ focus on mutation as an entity may cause attributes associated with processes (e.g., steps, time, interactions between entities) to be ignored, accounting for some of the catalogued noncanonical ideas about mutation. In the following section, we explore the background of entity and process thinking from cognitive psychology.

The Cognitive Perspective of Viewing Scientific Phenomena as an Entity or a Process

Bringing a cognitive lens to discipline-based education research permits movement from descriptions of difficulties (misconceptions) to investigating patterns in learners’ processes of learning. The cognitive perspective emphasizes the nature of knowing and learning, which includes how concepts and theories are understood and descriptions of organized patterns in cognitive activities (Greeno *et al.*, 1996). The cognitive perspective has been used as a framework to explain students’ difficulties in understanding biological phenomena (Gregory, 2009; Coley and Tanner, 2012). The results of some studies have suggested that students’ incorrect ideas about scientific phenomena may arise from cognitive explanatory biases, such as teleological thinking (the phenomenon occurs to fulfill a goal) or anthropocentric thinking (organisms act as conscious agents; Nehm and Ridgway, 2011; Opfer *et al.*, 2012; Coley and Tanner, 2015). Opfer *et al.* (2012) proposed that students might resort to intuitive ideas if they lack a “scientifically normative explanation,” that is, mechanistic causes for phenomena. However, others have shown that students’ tendency to select cognitive construals is influenced by the context (Gouvea and Simon, 2018).

Other cognitive theorists have suggested that students’ difficulty in explaining scientific phenomena stems not from their tendency to ascribe an incorrect mechanistic cause but from their tendency to misclassify a phenomenon as an entity instead of a process (Chi *et al.*, 1994). Entities are matter or objects, while processes are events or activities. According to this theory, the attributes for entities and processes are mutually exclusive (Chi *et al.*, 1994; Chi, 2013). Entities “can be contained” (Chi, 2013), while processes cannot; processes are associated with the attribute of “occurring over time,” while entities are not.

In physics, several researchers have shown across multiple contexts that students have a tendency to misattribute processes to the entity category (Chi *et al.*, 1994; Reiner *et al.*, 2000;

Slotta and Chi, 2006; Lancor, 2014; Eshach *et al.*, 2018). This misattribution has been associated with what have been labeled robust misconceptions, noncanonical ideas that are very difficult to correct (Chi *et al.*, 1994; Reiner *et al.*, 2000; Slotta and Chi, 2006; Lancor, 2014; Eshach *et al.*, 2018). For example, in physics, students tend to incorrectly describe heat as a collection of objects instead of a process involving the movement of molecules. Thus, they will incorrectly talk about “hot molecules” (Chi, 2013, p. 57). According to Chi (2013), this type of miscategorization is problematic, because entities and processes do not share common attributes. Processes such as baking bread, heating, or electricity flowing through a circuit have an element of time associated with them. Entities such as bread, the incorrect “hot molecules,” or the electrons contained in the circuit do not occur over time (although they may continue to be in existence). Similarly, no process can have physical properties such as color, volume, or location (Chi, 2013). Chi (2013) claims that, when students do miscategorize processes as entities, robust misconceptions arise that are resistant to instruction, because it is not enough to point out the misconception. Instead, students have to change the way they categorize. For example, students tend to struggle with understanding heat transfer, and they often express the incorrect idea that “coldness from the ice flows into the water” (Chi, 2013, p. 51). Achieving long-lasting change requires more than simply changing this one idea. It requires a fundamental shift in the idea that heat transfer is something that occurs as the result of a transfer of a thing, “coldness” that contains “cold,” to a process-based concept that includes the movement and collision of molecules over time (Chi, 2013). The coding schemes to distinguish processes and entities employed by many of the researchers using this cognitive framework tend to be specific to the scientific phenomenon. Therefore, it is helpful to look elsewhere for a more general characterization of what constitutes evidence of process-based reasoning by students.

Characterizing Processes and Entities Using the Theory of Mechanism

In the science education literature, researchers have characterized students’ explanations of scientific phenomena in terms of their understanding of scientific mechanism (Russ *et al.*, 2008). As reviewed in Russ *et al.* (2008), scientific mechanism is often equated with scientific process. While disagreement exists on the nuances, most definitions of scientific mechanism agree that a scientific mechanism describes how entities within a phenomenon interact to produce outcomes associated with the phenomenon (Machamer *et al.*, 2000; Bechtel and Abrahamsen, 2005; Glennan, 2005; Illiari and Williamson, 2012). Relating this definition back to the idea of categorization of entities and processes, the entities would be defined as the objects in the phenomenon. In the example of heat transfer when hot and cold liquids are mixed, the entities would be the molecules in the hot and cold liquids. The mechanism is the heat-transfer process, which occurs through the interaction of molecules spatially and temporally. As the molecules move and bump into one another, they affect the motion of other molecules, either speeding them up or slowing them down. Students who refer to “hotness” as a thing to be transferred are missing the temporal and spatial interactions that are essential for understanding how one activity leads to the next to produce the end product

(medium liquid temperature) from the starting conditions (hot and cold liquid).

“Mechanistic reasoning involves more than noting which causes are associated with which effects” (Russ *et al.* 2008, p. 506). Russ *et al.* (2008) give the example of a student that responds to a teacher’s prompt by stating that “Gravity’s pulling the book down before the paper” (p. 518). According to the framework for mechanistic reasoning that they developed, this student has named an activity “pulling down” that is engaged in by gravity; the gravity is the cause for the effect of “pulling down.” In this phrasing, the mechanistic process of gravity is a thing that does the pulling. Without additional reasoning about either the chaining of events (the steps in a mechanism) or the interactions occurring between the entities involved in the mechanism of gravity (e.g., the paper and the earth), this statement shows a lack of mechanistic understanding of the scientific phenomenon of objects falling down.

Slotta *et al.* (1995) make a similar argument for the importance of context when they discuss how to categorize “wind” in two very similar example statements. They state that one example statement, “That sound you hear is just the wind moving through the canyon” (p. 385), should be categorized as an entity statement for “wind.” In contrast, Slotta *et al.* (1995) present another example statement about wind, “The wind in the canyon is just the air moving” (p. 385), as evidence of process thinking about wind. They explain that the different contextual aspects associated with students’ descriptions of what is moving results in the two different categorizations, but do not expand in a way that is generalizable to other phenomena.

However, by applying Russ *et al.*’s (2008) framework, it is possible to develop more generalizable parameters for distinguishing between entity and process statements. In the first statement (“wind moving through the canyon”), Russ *et al.* (2008) would probably have said that the student had named a cause (wind) for the effect (sound) but had not described a mechanism. On the other hand, the second statement, which was classified as a process by Slotta *et al.* (1995), contains indicators of two categories of Russ *et al.*’s (2008) framework that are associated with mechanistic descriptions: “identifying organization of entities” and “chaining.” The second statement specifies how two entities (air and the canyon) are organized with respect to one another: the wind in the canyon is caused by the air *moving* [in the canyon]. There is also an indication of chaining, because the statement implies that the air has to move in the canyon for wind to occur. These two features (identifying organization of entities and chaining) are absent for the process of wind from the entity-classified statement “the wind moving through the canyon,” which merely states that the wind is an entity that is moving through the canyon (although one could make the case that this is a process statement for sound).

The Dual Nature of Mutation as an Entity and a Process

Unlike many of the physical phenomena that have been studied, which can only be seen as either an entity (e.g., temperature) or a process (e.g., heating), the word “mutation” describes both an entity (the change in the DNA) and a process (how that change is produced). (Another example is adaptation, which can be defined both as a feature [an entity] that becomes more frequent in a population as a result of a selective process, and as

a process by which populations change over time in response to specific environmental pressures.) Both the entity-based and process-based perceptions of mutation have affordances and limitations depending on context. For example, an entity-based perception of mutations facilitates understanding inheritance, because it is necessary to know that mutations/alleles have fixed locations on chromosomes to understand recombination and the effect of the chromosome movements on sorting alleles into gametes. On the other hand, the perception of mutation as a dynamic process with attributes such as time, steps, and interactions is critical to understanding other biological processes. The process-based perspective of mutation is required to understand why the development of cancer is a multistep process and why a change in the environment cannot directly cause a change in the DNA in evolutionary processes.

Experts switch with ease between entity- and process-based associations (Gupta *et al.*, 2010, 2014; Chi, 2013). However, based on prior research in physics (Reiner *et al.*, 2000; Slotta and Chi, 2006; Lancor, 2014; Eshach *et al.*, 2018), it is likely that students tend to perceive mutations primarily as entities, subjugating the process that produced the mutations.

Adopting a cognitive lens and investigating how students perceive mutation rather than what students know about mutation will permit a broader instructional approach that will elevate students' scientific understanding across multiple contexts. Moreover, such research has the potential to add to cognitive understanding, because the entity-based perception of mutation is not treated as a misconceived idea as in prior studies (Brookes *et al.*, 2005; Slotta and Chi, 2006; Gupta *et al.*, 2010, 2014; Brookes and Etkina, 2015). This study uses a free-response question to elicit the ideas that are accessible to students about mutation, what students can access from their memories in that context (Tulving and Pearlstone, 1966). The prompt question was adapted from Prevost *et al.* (2013) and was intentionally left vague to elicit entity/process responses rather than to assess what students know about mutation or the mechanisms for mutation. We used qualitative analyses to answer the following research questions:

In written responses to a prompt asking students to explain how mutations originate in populations of organisms:

1. What topics do students discuss in their answers with respect to mutation?
2. To what extent do students describe mutation as an entity or as a process?
3. What types of noncanonical ideas are expressed?
4. How do students' noncanonical ideas about mutation expressed in their writing relate to their entity-/process-focused treatment of mutation?

METHODS

Context

Participants in this study were students enrolled in the first semester of an introductory course for biology majors at a large midwestern research university in the United States. Sixty-four students ($N = 64$; 90% of whom were freshmen) completed the assessment questions used in this study. Student pseudonyms were used. The course was cotaught by two experienced instructors with an average teaching experience of more than 15 years who have been trained in active-learning techniques.

This study was approved by the institutional review board at our university under IRB 1702E07721.

Analysis of the textbook (Brooker *et al.*, 2015) used by students in this study showed that, in the chapter that covers mutation, the 17 figures depict the types and effects of germ-line versus somatic cell mutations, the phenotypes for genetic disease caused by point mutations, and types of structural changes in the DNA caused by mutagens. All of these figures treated mutation as an entity, and none depicted the process of mutation.

Data Collection

The free-response question was administered during the 11th week of a 15-week-long course, after students had received instruction in mutation and DNA replication. This method was chosen to reveal what perspectives are accessible (readily retrieved in the moment) to students versus what might be available (previously memorized information that can be elicited if the right cue is provided; Tulving and Pearlstone, 1966). While multiple-choice questions tend to capture whether students know the correct answers in response to explicit or implicit cues in the prompts and distractors, free-response questions can elicit a more complex thinking or a higher-level cognitive processing (Martinez, 1999). Compared with the interview format, the use of written responses to questions avoids potential cueing by the interviewers and is less resource intensive for larger samples (Wiggins, 1993; National Research Council, 2001).

The two questions in Table 1 were adapted from Prevost *et al.* (2013). Students randomly received one of the questions. We used two organisms here, because research has shown that assessment item features and contexts (such as organism type) can affect students' thinking (Nehm and Ha, 2011; Heredia *et al.*, 2012). Because research has shown that students tend to describe processes as entities (Slotta and Chi, 2006; Lancor, 2014; Eshach *et al.*, 2018), this question was phrased to theoretically bias toward a process-based description of mechanisms underlying mutation. The word "how" was used as to prompt for descriptions of a process. Prior studies have shown that 42% of primary and secondary students answered "how" questions about a scientific phenomenon with causal or mechanistic explanations and another 43% provided a reason for the phenomenon (Abrams *et al.*, 2001). The word "originate" was used, because theoretically this word should lead students into describing the molecular mechanisms *producing* mutations. Because we were also interested in whether students would have a dual perspective on mutations in their writing, the process cues were balanced by the use of the word "mutations" as a noun, an entity cue.

Data Analysis

All 64 participants responded to the assessment questions. Thirty students responded to the question on lizards. Thirty-four

TABLE 1. Free-response questions used to reveal student perspectives on mutation

Version	Questions
1	Explain how mutations originate in populations of lizards. (You can illustrate your answer with drawings.)
2	Explain how mutations originate in populations of bacteria. (You can illustrate your answer with drawings.)

responded to the question on bacteria. Student responses were analyzed in three phases. Phase 1 catalogued the topics contained in students' responses. Phase 2 analyzed whether the responses reflected an entity- or a process-focused treatment of mutation. Phase 3 characterized noncanonical ideas about mutation contained in students' responses.

Phase 1: Coding for Topics in Students' Descriptions. The first phase of coding broadly characterized the topics students mentioned with respect to mutation. The goal of this phase of coding was to determine whether students mentioned a cause for the origin of mutations.

Six codes were developed from phrases in students' responses using inductive coding (Glaser and Strauss, 2017): "DNA Replication," "Induction," "Other Cause," "Definition/Types," "Natural Selection," and "Reproduction." Examples of these codes are shown in Table 2. A single student response could be assigned multiple codes, because students could mention multiple topics with respect to mutation in their answers, and therefore, the unit of coding for phase 1 is a phrase within the entire response. Six responses included drawings along with the written answers. The drawings were analyzed using these six codes to capture the ideas the drawings conveyed.

Thirty responses ($n = 30$) were coded by FF.Z., who has a master's degree in botany and is pursuing a doctoral degree in science education, and independently by A.S., who has doctoral degrees in both biology and science education. The average Cohen's kappa was 0.94, indicating an almost perfect reliability (McHugh, 2012).

The six codes reflecting topics mentioned by students in their answers were grouped into three topic categories (Table 2): Cause, Description, and Propagation. The Cause category contains codes that mention a cause for mutation whether correct (i.e., DNA Replication or Induction) or incorrect (Other Cause; e.g., change in the environment). The Description category contains the Definition/Types code. In this category, students described what a mutation is or the types of mutations. The Propagation category contains codes that refer to how a mutation will spread through a population (Natural Selection) or be propagated from parent to offspring (Reproduction).

Phase 2: Coding for Defining Characteristic, Steps, and Interactions. The second phase of coding was to determine whether, within the entire response, students were treating mutation as an entity or a process. Student responses that did not contain a Cause code (DNA Replication, Induction, Other

TABLE 2. Descriptions and examples of six codes developed in the process of coding for response topics with the corresponding phrase bolded in the student responses

Category	Code	Description	Student response
Cause	DNA Replication	The student includes replication of the hereditary material of the organism as a way that mutation originates.	Mutations originate via errors in replication of DNA , or through horizontal transfer of DNA among individuals.—Jennifer
	Induction	The student claims that mutation is caused by agents in the environment, which include chemicals or radiation. The student may use "outside factors" or "induced" when talking about mutagens.	Mutation can occur in a variety of ways. Sometimes mutations occur spontaneously during DNA replication. If the lizard population is exposed to mutagens such as UV radiation or certain chemicals , this could also cause mutation.—Alan
	Other Cause	The student includes causes other than DNA replication or induction for how mutations arise, which may include transcription, transduction, or change in the environment.	There are different ways that mutations originate in bacteria, although I can really only think of a few. The insertion of viral DNA can cause mutations by changing the DNA sequence. Other ways to alter the DNA sequence, which causes mutations, is the addition or deletion of base pairs within DNA that can create a whole new amino acid sequence by changing one or more of the amino acids.—Sandy
Description	Definition/Types	Student talks about what a mutation is (e.g., a change in the sequence) or student lists types of mutation (e.g., insertions, deletion).	A lot of things can cause mutations. It could be a frameshift, missense, nonsense mutation. These can be insertions, deletions, or frameshift that can either change or not affect the amino acid sequence that will be transcribed and translated. These are then passed down from the parents' gametes.—Holly
Propagation	Inheritance	Student includes part of or the whole process of reproduction to describe how mutation gets passed down and preserved in the population.	It begins with a mutation in one organism and when it spreads either it survives in the environment or it dies. —Anne
	Natural Selection	The student talks about part or the whole process of how the mutation benefits or harms the organisms as related to natural selection. Or the student talks about how the environment interacts with the presence of a certain mutation.	Mutations originate slowly. Usually neutral mutations will arise that do not do anything for the individual's fitness. However, some lizards may accumulate an advantageous mutation which increases its fitness, allowing it to reproduce more. Because of this, there will be more lizards in the population with this mutation. —Kelly

TABLE 3. Descriptions and examples of three codes used to code for Entity- or Process-Focused phrases with Entity-Focused phrases italicized and Process-Focused phrases double-underlined in the student responses

Code	Description	Student response
Defining Characteristic (Entity-Focused)	The student describes the process as a definition statement and provides no explanation of the process.	Mutations originate <i>via errors in replication of DNA, or through horizontal transfer of DNA among individuals.</i> —Jennifer
Steps (Process-Focused)	The student attempts to describe steps in the process from a starting point to an ending point.	In the process of DNA replication/transcription, there are very rare events in <u>which an enzyme either makes a mistake</u> when replicating DNA by changing a nucleotide in a sequence or an RNA polymerase encoding RNA makes a transcript that slightly differs from the desired product. Any mistakes, such as these, <u>that are not corrected</u> by the cell permanently alter the genotype of that cell as well as any subsequent offspring made with genetic info from [the] altered organism.—Brendan
Interactions (Process-Focused)	The student attempts to describe the interaction among distinct objects or between objects and the outside factors to explain the dynamics in the process.	During DNA replication in a single gamete, a mutation may occur. Mutations arise from <u>DNA polymerase attaching an incorrect base pair</u> at any point during the replication process. If this mutated gamete is able to grow into maturity with this mutation and reproduce, the mutation will begin to spread through the population. If the mutation happens to provide an advantage in fitness, the mutation will continue to spread via higher rate of reproduction of the more fit mutated lizards.—Victor

Cause) were coded as treating mutation as an entity. These responses did not even name an activity that could result in the production of mutation. Thus, they did not even make a cause and effect association (Russ *et al.*, 2008).

However, for the responses that were coded as Cause, we needed to determine whether students were simply identifying a cause/effect association or whether they were providing a process-based description. Therefore, this phase of coding moved beyond coding which words were mentioned in students' responses to a coding scheme that captured context (Slotta *et al.*, 1995). Drawing on the distinctions between entity and process descriptions provided by Slotta *et al.* (1995) and the framework for mechanistic reasoning (Russ *et al.*, 2008), we developed a coding scheme that applied to this context but may have broader applications as well. Levels 6 (organization of entities) and 7 (chaining: backward and forward) of the framework for mechanistic reasoning (Russ *et al.*, 2008) indicate that students have moved beyond naming a cause or activity and, therefore, merely showing a knowledge of input–output relations (Slotta *et al.*, 1995; Russ *et al.*, 2008). Levels 6 and 7 indicate an acknowledgment of the mechanism/process that underlies this relationship: how an output was dynamically produced from the initial state. The framework for mechanistic reasoning was developed to characterize the discourse of young children in conversation with one another. However, this context involves written descriptions by undergraduate students in response to a prompt. In this static context, it can be difficult to detect evidence of chaining. Therefore, we chose to more specifically define chaining as “Steps”; the response shows evidence of sequential steps. The level 6 code, organization of entities, is vague and could describe spatial organization between entities that were not associated with one another, although that does not seem to be the way it was applied (Russ *et al.*, 2008). Therefore, this code was more specifically defined for this context as “Interactions,” evidence that two or more entities interacted with one another in some way. The Steps and Interactions codes also better encapsulate the mutation mechanism, which is a series of activities (binding, copying, repairing) carried out by multiple entities that interact with one

another (Machamer *et al.*, 2000). If the response containing a named causal activity did not show evidence of steps or interactions, then the response was coded as naming a “Defining Characteristic.” In these cases, the named activities are attributes (albeit causal) of the entity, but do not demonstrate a knowledge of a process or mechanism (Russ *et al.*, 2008).

The three quotes in Table 3 will be used to demonstrate how this coding scheme was applied. All three of the responses are presented in their entirety. Jennifer stated that “Mutations originate via errors in replication of DNA, or through horizontal transfer of DNA among individuals. She has named two activities, one partially correct (“DNA replication”) and one incorrect (“horizontal transfer”) that, according to her, cause mutations (“mutations originate via”). Mutations were treated as the effect of these causal activities, the end product. Moreover, mutations were equated with “errors in replication of DNA,” another entity (errors) that results from an activity (DNA replication). There is no evidence in this response of the dynamic mutational process that involves interactions between DNA and DNA polymerase resulting in errors that may then be repaired by the same enzyme or a different enzyme. Attributes of the entity have been defined (caused by *x* and *y*), but not attributes of the process (entities interacting in a spatiotemporal sequence to convert an initial state into an end state).

Contrast Jennifer's response to Brendan's response. Interestingly, Brendan did not mention the word “mutation.” Instead, he described a series of events involving interactions between more than one entity. First, he stated that there were “rare events in which an enzyme either makes a mistake when replicating DNA by changing a nucleotide in a sequence.”¹ Even this initial statement differs in a key way from Jennifer's statement that “mutations originate via errors in replication of DNA.” While Jennifer named mutations as the output of an activity, Brendan is describing a process that occurred over time when he proceeded: “Any

¹Brendan also incorrectly states that the initial event could be RNA polymerase making a mistake during transcription. We are not coding for correctness of answers, but indicators of steps or interactions.

mistakes, such as these [referring to the output of the first activity], that are not corrected.” Thus, he named a second activity, “correction,” that occurs after the first, describing a mutation process that has two sequential steps and therefore must occur over time—a key attribute of processes, but not entities.

The third response, by Victor, also differs from Jennifer’s response (albeit more subtly). Victor began by saying, “During DNA replication in a single gamete, a mutation may occur.” If Victor had stopped there, then this would have been equivalent to Jennifer’s naming of an activity that causes an output, a thing called “mutation.” However, Victor continued, “Mutations arise from DNA polymerase attaching an incorrect base pair at any point during the replication process.” In this part of the description, Victor described one entity (DNA polymerase) doing something (attaching) with another entity (base pair), describing an interaction between the two entities (DNA polymerase and base pair).

Using this coding scheme, all responses that named a causal activity for mutation, whether correct or incorrect (Cause category), were coded for Defining Characteristic, Steps, and Interactions by FF.Z. and independently by A.S. The average Cohen’s kappa was 0.93, indicating an almost perfect reliability (McHugh, 2012).

Categorizing Responses as an Entity-, Process-, or Dual-Focused Treatment of Mutation. Any causal response that was coded as Steps or Interactions was categorized as Process-Focused treatment of mutation, because we argue that these codes show evidence within the written description of attributes associated with processes. Any causal response that was coded as Defining Characteristic used the activity to define/characterize mutation as an entity that results from a particular activity and was thus classified as Entity-Focused. Because some responses were coded in phase 1 as Cause and either Propagation and/or Description, some responses showed evidence of both Entity-Focused treatment of mutation (from the Propagation or Description) and Process-Focused treatment of mutation (if the Cause was coded as containing Steps and Interactions). These responses were classified as Dual-Focused.

Coding for the Entity/Process Treatment of Propagation of Mutations. To determine whether students could apply a Process-Focused characterization to another mechanism, we re-examined responses that were coded as naming activities associated with Propagation of mutation using the same coding scheme used for Cause of mutations (described in Table 3). Therefore, all phrases of Propagation of mutation were coded to see whether students described the steps or interactions associated with reproduction (the process of how mutations are passed down from parents to the next generation) or evolution (the process of how mutations are spread in the population under natural selection), or simply provided defining characteristics. These responses will be described in detail in the *Results* section. All of the Propagation phrases ($n = 33$) were coded by FF.Z. and independently by A.S. The average Cohen’s kappa was 0.82, indicating a strong reliability (McHugh, 2012).

Phase 3: Coding for Noncanonical Ideas. All responses were analyzed for the presence of noncanonical ideas about mutations. Noncanonical ideas that appeared at least twice in students’ responses were captured and grouped into categories. Three categories were identified: Incorrect Cause, Confound Mechanisms, and Determinism. For example, Determinism denotes when a student elaborates on the effect of mutation, the student conveys the idea that one mutation will lead to a change in trait, or the student states that only beneficial mutations or nonharmful mutations can be passed on (as underlined by the wavy lines in Sally’s description, “If this mutation is silent it won’t impact anything, if it is detrimental to the survival, that bacteria will be less fit, if it is beneficial, it could be passed on to gametes proceeding [sic] it”). The description of these codes and examples are listed in Table 4. FF.Z. and a second coder with a doctoral degree in molecular biology coded all responses ($n = 64$) independently, and the average Cohen’s kappa was 0.85, indicating a strong reliability (McHugh, 2012).

RESULTS

The first section of the *Results* will discuss how students treat mutation in their responses: as an entity or as a process. We will

TABLE 4. Descriptions and examples of codes used for identifying noncanonical ideas with the noncanonical ideas highlighted using wavy underlining in the student responses

Code	Description	Student response
Incorrect Cause	The student indicates that mutations originate because of environmental change or evolutionary mechanisms.	Mutations can occur from random chance, <u>by gene flow, genetic drift...</u> –Kyle
Confound Mechanisms	The student describes a mechanism that produces mutation, and the student uses transcription or translation instead of DNA replication.	<u>In the process of DNA replication/transcription, there are very rare events in which an enzyme either makes a mistake when replicating DNA by changing a nucleotide in a sequence or an RNA polymerase encoding RNA makes a transcript that slightly differs from the desired product.</u> Any mistakes, such as these, that are not corrected by the cell permanently alter the genotype of that cell as well as any subsequent offspring made with genetic info from [the] altered organism. –Brendan
Determinism	When the student elaborates on the effect of mutation, he or she conveys the idea that one mutation will lead to a change in trait. Or the student states that only beneficial mutations or nonharmful mutations can be passed on.	Mutations can be random accidents that arose during transcription during binary fission, or if phages or other bacteria inject their DNA into the bacteria. They can also occur due to mutagens. If this mutation is silent it won’t impact anything, if it is detrimental to the survival, that bacteria will be less fit, <u>if it is beneficial, it could be passed on to gametes proceeding [sic] it.</u> –Sally

then present the types of noncanonical ideas that were observed and the association with the treatment of mutation. All example responses are presented in their entirety. Responses from both question contexts, the lizards and bacteria, were collapsed, because no significant differences in the patterns described below were identified according to the organism type (Supplemental Table A).

Phase 1: Most Students Named a Cause for Mutation

Out of 64 student responses, 48 mentioned a cause for mutation, suggesting that 75% of students recognized that the question was asking for a causal response. Twenty-five percent of student responses (16) only mentioned Description or Propagation without listing or describing any cause. Fifty percent of the 64 responses included the topic of Description, indicating that they contained a definition of mutation as a change in DNA or described the different types of mutations. Fifty-two percent of the 64 responses contained the topic of Propagation stating how mutations were passed down to the next generation or spread in the population. On average, each student response included 1.8 topic categories (i.e., Description, Propagation, or Cause).

Phase 2: Few Responses Reflected a Process-Focused Treatment of Mutation

Of the 64 student responses we analyzed, nine contained Process-Focused phrases with respect to mutation—phrases that described either steps or interactions in the mechanism of mutation (Table 3). Out of these nine responses, two contained only Process-Focused phrases and were classified as Process-Focused-only responses (Table 5). Seven contained both Process-Focused and Entity-Focused phrases and were classified as Dual-Focused responses (Table 5).

A Process-Focused-Only Treatment of Mutation. Only two out of 64 responses provided Process-Focused treatment of mutation without also containing an Entity-Focused description of mutation. Ethan's response to the prompt is

Mutations in a lizard population can arise spontaneously but this is less common due to proofreading and other counter measures. Additionally, if a lizard is exposed to chemicals, UV radiation, or any number of other mutagens, the number of mistakes in DNA transcription/translation increases and therefore there is less of a chance that all these get caught and either fixed or destroyed by the body.

His response provided evidence for a Process-Focused treatment of mutation, because multiple steps (which occur over time) were included in his description. The steps were 1) exposure to mutagens (“if a lizard is exposed to chemicals, UV radiation, or any number of other mutagens”), 2) mistakes in DNA

TABLE 5. Distribution of student responses under Process-Focused, Entity-Focused, and Dual-Focused categories (N = 64)

Treatment of mutation	Number of students	Percentage of students
Process-Focused only	2	3.1
Entity-Focused only	55	86
Dual-Focused	7	10.9

occur (“the number of mistakes in DNA transcription/translation increases”), 3) mistakes get caught (“therefore there is less of a chance that all these get caught”), and 4) mistakes are “either fixed or destroyed.” Although Ethan confounded DNA replication with DNA transcription/translation, we are not coding for correctness of answers, but for indicators of steps or interactions.

A Dual-Focused Treatment of Mutation. Seven out of 64 responses included both phrases that indicated a focus on process and phrases that indicated a focus on describing mutation as an entity. Mary stated,

Mutations originate in a population of bacteria when the polymerase molecule makes a mistake when extending the second strand of complementary DNA. This mistake can be inserting the wrong base pair. This might happen because there are 4 nitrogenous base pairs, 2 pyrimidines and 2 primines; Adenine pairs with Thymine while Guanine pairs with Cytosine. The polymerase could accidentally pair the other pyrimidine with the primine. This may be harmful during translation of the mRNA of the gene. This substitution mutation can be silent in that the codon still codes for the same amino acid. It could code for another amino acid which may not have a huge effect on the protein other than changing the protein shape slightly making it more difficult for the protein to function properly. Or it could code for a stop codon which would terminate translation early and completely throw off protein function.

Mary's response was classified as Process-Focused because it was coded as showing evidence of Interactions. Mary included two sections that were indicative of polymerase interacting with base pairs (indicated by the underlined sections). The most explicit statement occurs when she stated, “The polymerase could accidentally pair the other pyrimidine with the primine.” Although not coded as explicitly showing evidence of Steps, this response contains a suggestion of time associated with the formation of mutation with the use of the phrase “when polymerase molecule makes a mistake when extending.” Mary also included descriptions of one type of mutation (a “substitution mutation”) that depicted mutation as an entity to be categorized and was coded as Description. Therefore, Mary's response can be categorized as a Dual-Focused treatment of mutation.

Three of these seven students whose responses reflected a Dual-Focused treatment of mutation described the propagation of mutation through the population after describing the mechanism of mutation.

The Majority of Student Responses Reflected an Entity-Focused Treatment of Mutation

Student responses that contained only Entity-Focused phrases were much more common than those that contained both. Eighty-six percent (55) students treated mutation only as an entity in their responses. Table 6 shows the distribution of coded topic phrases that were mentioned under each of the response categories. For example, 55 out of 64 student responses treated mutation as an entity, either by describing the Cause of mutation (DNA Replication or Induction or Other Cause) as a defining statement or by only including the Description or Propagation of mutations. As summarized in Table 6, out of 55 Entity-Focused responses, 39 student responses named a causal activity that was coded as describing a Defining Characteristic

TABLE 6. Distribution of coded topic phrases under response categories of Entity-Focused, Process-Focused, and Dual-Focused treatments of mutation^a

Entity/Process-Focused (phase 2 codes) Response category (64 student responses)	Topics mentioned (phase 1 codes)		
	Cause	Description	Propagation
Entity-Focused (55)	Cause as a Defining Characteristic (39)	Yes (16)	Yes (10) None (6)
		None (23)	Yes (9) None (14)
	No Cause (16)	Yes (12)	Yes (7) None (5)
		None (4)	Yes (4) None (0)
Process-Focused (2)	Cause with Steps or Interactions (2)	Yes (0)	—
		None (2)	Yes (0) None (2)
Dual-Focused (7)	Cause with Steps or Interactions (7)	Yes (4)	Yes (0) None (4)
		None (3)	Yes (3) None (0)
Phrases total	48	32	33

^aOn average, each student response included 1.8 topic phrases.

of mutation Cause and thus were classified as Entity-Focused. The remaining 16 responses did not include any Cause phrases. Twelve of these responses were coded as providing a Description (categorizing or defining mutation) and 11 were coded as Propagation (discussing how mutations are inherited).

Entity-Focused Responses That Mentioned a Cause for Mutation. Out of 48 student responses that mentioned Cause, 39 responses treated the cause as a Defining Characteristic and mutation as an entity—something to be defined. In the response that follows, Emily defined a mutation as an entity that can be “caused by something (e.g., UV radiation),” noting the cause (UV radiation) and the effect (mutation) without delineating the dynamics of the process (the Steps or Interactions). Emily proceeded to characterize the types of mutations (“addition of a base, deletion of a base”) and the possible effects, treating mutation as the end product to be categorized. Because Emily’s response provided no evidence of the steps or interactions (the dynamic activity) associated with a process, even though she named a causal activity, her response was coded as Entity-Focused:

Mutations can either be random or can be caused by something (e.g., UV radiation). There are a few types of mutations that may occur such as addition of a base, deletion of a base, or it may switch the base (not complimentary [sic]). It can have different effects on expression depending on the type of mutation and the location of the mutation, for example a mutation in the 3rd spot of an amino acid group will most likely not change the expression of the gene.

Entity-Focused Responses That Did Not Mention a Cause for Mutation. Because the question asked students to explain how mutations originate, not how they are propagated or how they are defined, it was expected that all students would mention a cause for mutation. However, 16 students did not even mention

a cause for mutation (Table 6). Instead, these students defined a mutation and/or listed types of mutations (Description phrases) and/or wrote about how mutations were transferred from parents to offspring or propagated in a population (Propagation phrases). All of these responses were classified as Entity-Focused treatment of mutation. For example, Kelly wrote,

Mutations originate slowly. Usually neutral mutations will arise that do not do anything for the individual’s fitness. However, some lizards may accumulate an advantageous mutation which increases its [sic] fitness, allowing it to reproduce more. Because of this, there will be more lizards in the population with this mutation.

This response described one type of mutation, “neutral mutation” and the Propagation of mutation “some lizards may accumulate an advantageous mutation which increases its [sic] fitness, allowing it to reproduced more.” Both of these phrases treat mutations as a *thing* that can be classified into bins or can be propagated in the population. There is some evidence of time associated with the phrase “Mutations originate slowly,” but even in this phrase, mutations are treated as something that appears over time, not a process that is causing something to appear. Therefore, Kelly’s response is characterized as Entity-Focused.

The Majority of Student Responses Included a Process-Focused Treatment of Propagation of Mutation

Surprisingly, considering that the prompt asked about origination of mutations, not propagation, more than half (33 of 64) of students’ responses included Propagation phrases that characterized how mutations were transferred from parent to offspring either in a single mating or in populations (Table 6). Although mutation was treated as an entity (something that can be passed on) in the Propagation phrases, the mechanism of Propagation has the potential to be described as a process or an entity.

We wanted to know whether students could adopt a process focus when describing a biological phenomenon, as opposed to whether they adopted a process focus when writing about mutation. Therefore, in the 33 responses that contained Propagation phrases, student writing about propagation was coded using the same three codes used to distinguish between an Entity- and Process-Focused treatment of mutation (Defining Characteristic, Steps, Interactions).

Holly wrote,

A lot of things can cause mutations. It could be a frameshift, missense, or sense mutation. These can be insertions, deletions, or frameshift that can either change or not affect the amino acid sequence that will be transcribed and translated. These are then passed down from the parents' gametes.

Holly included the topic of Propagation, stating that mutations are passed down in the process of reproduction. This reflected an Entity-Focused treatment of Propagation, because the phrase provided a naming activity, "passed down," without any dynamics (Interaction or Steps) in the process of Propagation. On the other hand, Ali's response shows evidence of a Process-Focused treatment of Propagation. He wrote,

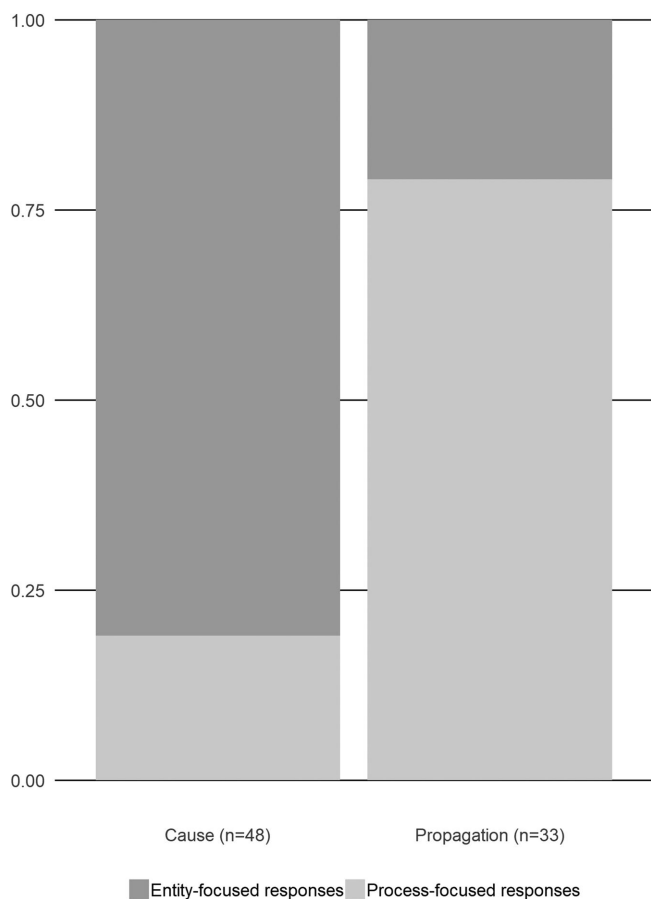


FIGURE 1. Comparison of descriptions of Propagation of mutation and Cause of mutation shows that Propagation descriptions were Process-Focused, while Cause descriptions were Entity-Focused.

If it was a helpful mutation, then natural selection will benefit this lizard + it will be more fit than the other lizards + will have more offspring. These offspring will have that mutation + will also be more fit, so they will have also have more offspring + pass the mutation onto them. Eventually after generations, the mutation will be prominent in the population.

This response included the sequence of activities as steps for the process of Propagation, starting from the lizard with a helpful mutation, then the lizard will become more fit, then it will have more offspring, then these offspring will have more offspring, until the termination condition "be prominent in the population." Ali described multiple steps in the process for how mutations can be propagated in a population, thus this response was categorized as a Process-Focused treatment of Propagation.

As shown in Figure 1, Propagation phrases were mentioned in 33 student responses, and Cause phrases were mentioned in 48 student responses. The first bar shows the proportion of Entity-Focused responses (21%, gray) and Process-Focused responses (79%, black) for Propagation of mutation; the second bar shows the proportion of Entity-Focused responses (83%, gray) and Process-Focused responses (17%, black) for Cause of mutation. The results show that the majority of responses described Propagation of mutations as a process (black portion in first bar in Figure 1), while the majority of responses described Cause for mutations as an entity (gray portion in second bar in Figure 1).

Alex's response included an Entity-Focused treatment of mutation and a Process-Focused treatment of Propagation of mutations:

Due to either *mistakes in gene replication or a mutagen substance in germlines or an embryo*, a new allele arises. Rarely that allele is beneficial and few individuals that have it can reproduce more frequently and are more fit than others to survive. Natural selection acts on the new allele and select[s] for that and against the older one. After a long time and if conditions have remained stable the new allele is dominant in the population.

The Cause phrase defined "mistakes in gene replication or a mutagen substance" (something) as the cause for mutation (italicized phrase). He has named an activity and ascribed a cause ("mistakes in gene replication or a mutagen substance in germlines or an embryo") to an effect ("a new allele arises") but has not provided a mechanistic characterization of mutation that could be coded as Interactions or Steps. In contrast, the depiction of Propagation includes multiple steps: 1) "Allele is beneficial," 2) "Few individuals that have it can reproduce more frequently," 3) "Natural selection acts on the new allele and select[s] for that and against the older one," and 4) "After a long time and if conditions have remained stable the new allele is dominant in the population." The phrase "after a long time" (while not coded) is also indicative of a Process-Focused treatment of Propagation, because time is a key attribute of processes.

Noncanonical Ideas Occur in Both Process- and Entity-Focused Responses

Because it has been proposed that students' process/entity perspective is associated with noncanonical thinking (Chi, 2013),

TABLE 7. Noncanonical ideas about mutation organized by process/entity treatment of mutation

Entity/Process treatment of mutation	Noncanonical ideas		
	Confound Mechanism	Determinism	Incorrect Cause
Dual-Focused responses ($n = 7$)	2	1	
Process-Focused responses ($n = 2$)	1		
Entity-Focused responses ($n = 55$)	3	5	19

responses were also coded for the presence of noncanonical ideas. Three main types of noncanonical ideas were identified after coding the 64 student responses: Incorrect Cause (including a cause other than DNA Replication or Induction to explain how mutations arise), Confound Mechanisms (confounding transcription or translation with DNA replication), Determinism (mutation must lead to change in protein or phenotype; only beneficial or nonharmful mutations can be passed down). Twenty-eight students had responses with noncanonical ideas that fit into these three categories (only two students had more than one noncanonical idea). Table 7 shows the association between the process/entity treatment of mutation in the response and the type of noncanonical idea.

Most of the students (19 out of 28) with a noncanonical idea mentioned Incorrect Cause for mutation, with or without mentioning the correct causes for mutation of DNA replication or induction. Four students included the idea that mutations occur to adapt to the environment. Four students who answered the bacteria question included transduction as the cause for mutation. Eleven students each mentioned a random cause such as “mishap in the cell formation,” “random cellular metabolic processes,” “crossing over,” “gene flow,” “horizontal transfer.” All of these 19 responses provided an Entity-Focused description of mutation; no step or interaction was mentioned to explain the origin of the mutation.

Six responses confounded transcription/translation with DNA replication. These were categorized separately from Incorrect Cause, because it was difficult to ascertain whether this was a naming difficulty or whether students really did not understand how mutations are generated. Out of these six responses, three responses were Entity-Focused descriptions, one was a Process-Focused description, and two reflected a Dual-Focused treatment of mutation. For example, Luke wrote, “Most mutations occur through a random set of ‘mistakes’ that occur through DNA replication, translation and/or transcription,” using the correct cause “DNA replication” and incorrect cause “translation and/or transcription” to define mutation. Brendan’s response reflecting a Dual-Focused treatment of mutation stated,

In the process of DNA replication/transcription, there are very rare events in which an enzyme either makes a mistake when replicating DNA by changing a nucleotide in a sequence or an RNA polymerase encoding RNA makes a transcript that slightly differs from the desired product. Any mistakes, such as these, that are not corrected by the cell permanently alter the genotype of that cell as well as any subsequent offspring made with genetic info from altered organism.

Brendan included the steps 1) an enzyme makes a mistake and 2) mistakes are not corrected, indicating a Process-Focused

treatment of mutation (double-lined phrases). He also described the mutation being spread to “subsequent offspring,” treating mutation as an entity that can be passed down to offspring. Therefore, this whole response was coded as a Dual-Focused treatment of mutation. However, the interaction of RNA polymerase with RNA was incorrectly listed as being involved in mutation (wavy-lined phrase), which is a noncanonical idea confounding transcription with DNA replication.

Six responses, five of which are Entity-Focused descriptions of mutation, included a form of Determinism for two types of ideas. Four out of those six responses stated that only beneficial or nonharmful mutations can be passed on. For example, Sally included the phrases “If this mutation is silent it won’t impact anything, if it is detrimental to the survival, that bacteria will be less fit, if it is beneficial, it could be passed on to gametes proceeding [sic] it.” Sally described three effects of mutation—silent, detrimental, and beneficial—but indicated that only beneficial mutations can be passed down, which is a deterministic idea. Two responses included the idea that a mutation would definitely cause a change in phenotype, as Kay wrote,

Mutations originate in bacteria when part of the DNA sequence is changed. This can be the addition, removal or replacement of nucleotides. When the nucleotides are changed, the DNA can not code for what it was originally intended for. If it codes something different, for example, spots on an animal’s fur, this is considered a mutation. In bacteria, these mutations are then expressed in it, such as the function or form of the bacteria.

The definition and types of mutations were listed (“the addition, removal or replacement of nucleotides”), while Kay mentioned that any mutation will lead to a change in phenotypes (wavy-lined phrase), which is a deterministic idea.

DISCUSSION

This study identified three ways students treated mutation in response to the prompt: “Explain how mutations originate in populations of lizards/bacteria”: Entity-Focused only, Process-Focused only, and Dual-Focused (both Entity- and Process-Focused). In contrast to previous studies on how surface features (such as organism type) affect students’ thinking (Nehm and Ha, 2011; Heredia *et al.*, 2012; Prevost *et al.*, 2013), students in our study tended to treat mutation as an entity for both organism types. There was a slightly greater proportion of responses for the lizard that were Process- or Dual-Focused (7/30 for lizards vs. 2/34 for bacteria). However, this tendency did not rise to the level of significance. Given the paucity of responses containing a process perspective, a larger sample would be necessary to draw firm conclusions about the effect of organism type on students’ responses.

We found that, even after instruction in DNA replication and mutation, the majority of students treated mutation as an entity in response to a prompt that calls for an explanation of the process of mutation in both the lizard and bacterial populations. It is worth noting that 48 students (75%) interpreted the question as intended, at least naming a cause for an effect. However, this study shows that most (39/48) of these students do not go beyond naming an activity that produces an effect, which is not sufficient to show the mechanistic reasoning (Machamer *et al.*, 2000; Russ *et al.*, 2008) that we argue is associated with Process-Focused descriptions based on parallels with Slotta *et al.* (1995). This extends prior studies that found that students tended to describe physical phenomena as entities (Reiner *et al.*, 2000; Wisner and Amin, 2001; Slotta and Chi, 2006; Gupta *et al.*, 2014; Lancor, 2014; Eshach *et al.*, 2018). In these studies, the description of physical phenomena as entities instead of processes was often incorrect (Reiner *et al.*, 2000; Wisner and Amin, 2001). Our findings show that, even when an entity- and process-focused characterization of the scientific phenomenon are both valid, students tend to provide an entity-focused description.

These results suggest that theoretical discussions of why students use an entity perspective to explain scientific phenomena needs to move beyond targeting incorrect categorization of phenomena. It has been proposed that incorrect categorization of scientific phenomena as entities instead of processes may be associated with students' noncanonical ideas about the phenomena (Chi, 2013). Therefore, this study also examined the association between entity- and process-focused categorization of mutation and noncanonical ideas about mutation. We found that noncanonical ideas were associated with approximately half of Entity-Focused responses and approximately half of Process-Focused responses. Therefore, with respect to phenomena for which either an Entity-Focused or Process-Focused treatment is correct, there does not seem to be an association between a particular perspective and the presence of noncanonical ideas.

Consequences of an Entity-Focused versus a Process-Focused Perspective of Mutation

Although students tend to describe mutation as an entity regardless of the context, experts in biology maintain dual perspectives and switch fluidly between the two depending on the context, in the same way that experts in physics switch between Entity-Focused metaphors and Process-Focused explanations (Brookes *et al.*, 2005; Brookes and Etkina, 2015; Gupta *et al.*, 2010, 2014). In physics, straddling these two mind-sets is believed to be productive and essential to understanding and communicating concepts such as energy and gravity (Gupta *et al.*, 2010, 2014). Gupta *et al.* (2010, 2014) have suggested that entity-focused descriptions for nonmaterial concepts can be used to better communicate with peers when discussing process-focused concepts (Gupta *et al.*, 2010) and that this entity-focused view can help learners focus on the sense-making process rather than on using the "right" words (Gupta *et al.*, 2014). Researchers have argued against the use of entity perspectives of processes even as analogy or metaphor, because this treatment could lead to noncanonical ideas that will be hard to displace (Slotta and Chi, 2006; Slotta, 2011). The data presented here suggest that noncanonical ideas are associated

with responses containing only Entity-Focused perspectives and with responses containing Dual-Focused and Process-Focused perspectives. Moreover, the majority of students with a Process-Focused perspective are capable of switching between an Entity- and Process-Focused perspective within the same short response. Furthermore, as we discuss later, there are affordances to students using both perspectives when considering mutation. We suggest, therefore, that it may be appropriate for instructors to leverage students' existing entity-focused perceptions of a phenomenon, particularly in cases in which both perspectives are valid, with the goal of switching perspectives to suit the context. Additional research needs to be done to determine the affordances and constraints of such an approach when an entity-focused perspective is not legitimate.

In biology, the word "mutation" has both terms entity-focused attributes (fixed location) and process-focused attributes (steps and interactions) depending on the context. An entity-focused view of mutation (e.g., with a fixed location on the chromosome) may help learners focus on sense-making of the process of inheritance of mutation, while the process-focused view of mutation initiates mechanistic thinking and the sense-making of the effects that mutation can have in the processes of natural selection and the development of genetic diseases or cancer. If students stick to an entity-focused view of mutation, they may struggle with concepts requiring an understanding of process-focused attributes of mutation. For example, students may have difficulty associating the idea that it takes time for mutations to form and accumulate in cancer (accounting for the benefits of early detection). Additionally, they may have difficulty recognizing that an interaction between the environment and DNA would be necessary for an organism to mutate in response to the environment, accounting for the persistence of this often-heard noncanonical idea (Brumby, 1984; Abraham *et al.*, 2009; Gregory, 2009; Nehm and Ridgway, 2011; Bray Speth *et al.*, 2014). Others have attributed the persistence of this idea to students' difficulty in understanding the randomness of mutation (Abraham *et al.*, 2009; Bray Speth *et al.*, 2014) or to the presence of cognitive biases, such as intentionality or teleology (Brumby, 1984; Gregory, 2009; Nehm and Ridgway, 2011). However, an alternative explanation is that students may not be recognizing that mutation is a process that has attributes of time, steps, and interactions. Therefore, they would not consider that most environmental contexts would not physically be able to cause changes in the DNA. They would also be unlikely to associate the idea of randomness with mutation, because they would not be considering mutation as a process that has this attribute. Instead, they would be thinking of mutation as an entity that already exists.

Noncanonical Ideas Appeared in Both Process-Focused and Entity-Focused Treatment of Mutation. Chi *et al.* (1994) have argued that misconceiving processes as entities can lead to robust misconceptions, because the attributes associated with these two categories are mutually exclusive. Thus, students would incorrectly assign inappropriate entity-focused attributes to the process. In the contexts studied by Chi and others, the fundamental error is the incorrect assignment of a process to an entity category (Slotta and Chi, 2006; Chi, 2013). Therefore, it has been argued that correcting students' categorization will correct these robust misconceptions (Chi, 2013).

This study, however, found that students who had process descriptions and those who had entity descriptions both expressed noncanonical ideas. For example, the noncanonical idea of describing errors in transcription as responsible for mutations was present in all categories of descriptions (Dual-Focused, Process-Focused-only, Entity-Focused-only treatment of mutation). It could be that this noncanonical idea is of the type Chi (2013) calls “false beliefs” or “flawed mental models” and may not be the robust type that is associated with confounding entities and processes. This type of noncanonical idea can be corrected relatively easily by explicit or implicit refutation, for example, showing students the correct belief or model and asking students to compare and contrast it with their noncanonical ideas (Chi, 2013).

Effect of an Absence of a Process-Focused Perspective on Students’ Understanding of Phenomena. The most common noncanonical idea in this study is ascribing an Incorrect cause (19 responses out of 64). All 19 responses show an Entity-Focused treatment of mutation. The wide variety of incorrect causes students specified as generating mutations aligns with previous studies that showed students had difficulty in understanding mutation (Prevost *et al.*, 2013; Bray Speth *et al.*, 2014). Because only nine Process-Focused descriptions were identified in our study as opposed to 55 Entity-Focused descriptions, future research on a sample with more Process-Focused responses is needed to determine whether a Process-Focused view of mutation inoculates students against noncanonical ideas associated with incorrect causes.

Students Are Capable of Process-Focused Explanations of Scientific Phenomena

To our knowledge, this is the only scientific phenomenon with a dual nature for which student treatment of the phenomenon has been investigated using the cognitive theory of entity/process classification. In previous studies, the entity-focused perspective of investigated scientific phenomenon is often an incorrect understanding (Reiner *et al.*, 2000; Wisner and Amin, 2001; Slotta and Chi, 2006; Gupta *et al.*, 2014; Lancor, 2014; Eshach *et al.*, 2018). In this study, an Entity-Focused perspective is a choice that students make that reflects how they access their knowledge about mutation. Our results show that, in response to a question that should have elicited a description of the process of mutation, most students defaulted to an Entity-Focused description of mutation, but one that is not incorrect. Their responses included phrases that define mutation as altered DNA or errors in DNA replication, or describe types of mutations, or show the propagation of mutations in the population. Most students did not include phrases that describe the processes of mutation. They did not describe the interactions or steps in the process of mutation, even though they were answering this prompt after instruction in DNA replication and mutation. This indicates that mutation as a process was not accessible to them when they answered this prompt, although it is possible that they may show that this information is available to them in other contexts (Tulving and Pearlstone, 1966).

Many students were able to access information about interactions and steps of another scientific phenomenon in their descriptions. Taking advantage of students’ spontaneous and unexpected production of descriptions of propagation of mutations (either

from parent to offspring or within populations), we investigated whether students produced Entity- or Process-Focused descriptions of propagation. Many students who produce an Entity-Focused response about mutation are capable of Process-Focused descriptions of propagation. Almost twice as many descriptions of the process of propagating a mutation in a population are Process-Focused as opposed to Entity-Focused. This implies that students are not inherently “process thinkers” or “entity thinkers.” Instead, it is necessary to look at the instructional context surrounding mutation as opposed to propagation.

Instructional Context Favors Entity-Focused Treatment of Mutation

Students’ tendency to describe mutation as an entity could be the result of the tradition in biology instruction of a treatment of mutation that emphasizes the entity-focused perspective. The textbook used in this class indicates bias toward depicting mutations as entities to be characterized instead of processes with steps and interactions of objects. A similar pattern was observed with another commonly used textbook for college biology majors, *Biological Sciences* (Freeman *et al.*, 2017). Similarly, in physics, physical phenomena tend to be treated as entities in textbooks, as more than 80% of the clauses containing “heat” in two of three mainstream textbooks for introductory physics are describing “heat” as an entity, suggesting that instruction may play a role in students’ entity-focused view of physical phenomena as well (Brookes *et al.*, 2005).

Exposure to the entity-focused perspective of scientific phenomena may be reinforced by the coverage of science within the classroom. In the classroom in this study, a review of notes and classroom artifacts showed that the instructors spent much more time covering mutation as an entity (either embedded within a process) or describing the types of mutations than they did on explaining the process. This may reflect instructors’ prior training (e.g., evolution has a more entity-focused treatment of mutation than molecular biology) and/or a traditional view of how mutations should be taught. The instruction of mutation reported in previous studies also focused on the treatment of mutation as an entity (Smith and Knight, 2012; Clarke-Midura *et al.*, 2018). On the other hand, instructors may perceive that the process-focused perspective of mutation should be taught in an advanced biology class rather than an introductory class. Bray Speth *et al.* (2014) mentioned that mutation was taught in an introductory biology class as the origin of variation after instruction of evolution, because “a detailed overview of different molecular types of mutation and of DNA repair mechanisms was beyond the scope of the course” (p. 531).

More broadly, students may also have been conditioned by prior science instruction to view science as a collection of facts rather than a set of explanations of how and why things happen (McComas *et al.*, 1998; Sandoval, 2003; Danielak *et al.*, 2014; Russ, 2018). For example, researchers have shown that educators tend to reward students’ knowledge of the correct terminology rather than students’ efforts in making sense of the phenomenon (Russ *et al.*, 2009; Russ, 2018; Talanquer *et al.*, 2015). This conditioning may explain why others have found that students struggle with constructing causal explanations for scientific phenomena (Abrams *et al.*, 2001; Haskel-Ittah and Yarden, 2018).

Implications for Instruction

This study shows that students tend to describe mutation as an entity, not including the process-focused attributes of mutation when answering how mutations originate. Because a process-focused perspective can benefit students' sense-making of other biological phenomena, we suggest that instruction should emphasize the process-focused attributes of mutation, for example, the steps and interactions, while minimizing the instruction of categorization or definition of mutations. In addition, our findings suggest that instructors need to be aware that the characteristics of mutation that are accessible to novices may not include process-focused attributes. Therefore, instructors should clarify when they use the word "mutation" whether they are referring to the process or the entity. Such steps may help to improve communication and minimize the development of non-canonical ideas.

Future Directions

This study was designed to explore how students described mutation, not what they could understand about the mechanism of mutation. Moreover, it was designed to explore what ideas about mutation were accessible to students, not what was available to them. Thus, it remains unclear from this work whether students' perspectives are influenced by their understanding of the mechanism of mutation and whether or not students *can* associate process attributes to mutation (vs. whether they *do*). While we chose this approach because we wanted to understand what students might be accessing when they hear or see the word "mutation," additional studies are needed to explore these questions. However, our results show that, when instructors use the word "mutation" in class in a process context (as in our written prompt), it is likely that students are not associating process attributes with the concept, which may be contributing to noncanonical ideas that appear in diverse contexts.

ACKNOWLEDGMENTS

We thank the members in the Schuchardt-Warfa research group who provided valuable feedback throughout this study: Abdi Warfa, Petra Kranzfelder, Catherine Kirkpatrick, Jenna Hicks, Marcos Garcia-Ojeda, Jessica Dewey, Linh Chau, and Jennifer Bankers-Fulbright. We also thank Gillian Roehrig for critical comments on an earlier draft of this article. We appreciate the critical feedback from two anonymous reviewers. We also thank the instructor and the students who participated in this study.

REFERENCES

- Abraham, J. K., Meir, E., Perry, J., Herron, J. C., Maruca, S., & Stal, D. (2009). Addressing undergraduate student misconceptions about natural selection with an interactive simulated laboratory. *Evolution: Education and Outreach*, 2(3), 393–404.
- Abrams, E., Southerland, S., & Cummins, C. (2001). The how's and why's of biological change: How learners neglect physical mechanisms in their search for meaning. *International Journal of Science Education*, 23(12), 1271–1281.
- American Association for the Advancement of Science (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC.
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421–441.
- Bray Speth, E., Shaw, N., Momsen, J., Reinagel, A., Le, P., Taqieddin, R., & Long, T. (2014). Introductory biology students' conceptual models and explanations of the origin of variation. *CBE—Life Sciences Education*, 13(3), 529–539.
- Brooker, R. J., Widmaier, E. P., Graham, L. E., & Stiling, P. D. (2015). *Biology*. New York: McGraw-Hill Education.
- Brookes, D., Horton, G., Van Heuvelen, A., & Etkina, E. (2005). Concerning scientific discourse about heat. *AIP Conference Proceedings*, 790(1), 149–152.
- Brookes, D. T., & Etkina, E. (2015). The importance of language in students' reasoning about heat in thermodynamic processes. *International Journal of Science Education*, 37(5–6), 759–779.
- Brumby, M. N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education*, 68(4), 493–503.
- Carlson, E. A. (2011). *Mutation: The history of an idea from Darwin to genomics*. New York: Cold Spring Harbor Laboratory Press.
- Chi, M. T. (2013). Two kinds and four sub-types of misconceived knowledge, ways to change it, and the learning outcomes. In Vosniadou, S. (Ed.), *International Handbook of Research on Conceptual Change* (pp. 61–82). New York: Routledge.
- Chi, M. T., Slotta, J. D., & De Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, 4(1), 27–43.
- Clarke-Midura, J., Pope, D. S., Maruca, S., Abraham, J. K., & Meir, E. (2018). Iterative design of a simulation-based module for teaching evolution by natural selection. *Evolution: Education and Outreach*, 11, 1–17.
- Coley, J. D., & Tanner, K. D. (2012). Common origins of diverse misconceptions: Cognitive principles and the development of biology thinking. *CBE—Life Sciences Education*, 11(3), 209–215.
- Coley, J. D., & Tanner, K. (2015). Relations between intuitive biological thinking and biological misconceptions in biology majors and nonmajors. *CBE—Life Sciences Education*, 14(1), ar8.
- Danielak, B. A., Gupta, A., & Elby, A. (2014). Marginalized identities of sense-makers: Reframing engineering student retention. *Journal of Engineering Education*, 103(1), 8–44.
- Eshach, H., Lin, T. C., & Tsai, C. C. (2018). Misconception of sound and conceptual change: A cross-sectional study on students' materialistic thinking of sound. *Journal of Research in Science Teaching*, 55(5), 664–684.
- Freeman, S., Quillin, K., Allison, L., Black, M., Podgorski, G., & Taylor, E. (2017). *Biological sciences*. New York: Pearson Education.
- Glaser, B. G., & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. New York: Routledge.
- Glennan, S. (2005). Modeling mechanisms. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 443–464.
- Gouvea, J. S., & Simon, M. R. (2018). Challenging cognitive construals: A dynamic alternative to stable misconceptions. *CBE—Life Sciences Education*, 17(2), ar34.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. *Handbook of Educational Psychology*, 77, 15–46.
- Gregory, T. R. (2009). Understanding natural selection: Essential concepts and common misconceptions. *Evolution: Education and Outreach*, 2(2), 156.
- Gupta, A., Elby, A., & Conlin, L. D. (2014). How substance-based ontologies for gravity can be productive: A case study. *Physical Review Special Topics—Physics Education Research*, 10(1), 010113.
- Gupta, A., Hammer, D., & Redish, E. F. (2010). The case for dynamic models of learners' ontologies in physics. *Journal of the Learning Sciences*, 19(3), 285–321.
- Haskel-Ittah, M., & Yarden, A. (2018). Students' conception of genetic phenomena and its effect on their ability to understand the underlying mechanism. *CBE—Life Sciences Education*, 17(3), ar36.
- Heredia, S., Furtak, E. M., & Morrison, D. (2012, April). Item context: How organisms used to frame natural selection items influence student response choices. Paper presented at: Proceedings of the National Association for Research in Science Teaching (NARST) annual conference (Indianapolis, IN).
- Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2(1), 119–135.

- Lancor, R. A. (2014). Using student-generated analogies to investigate conceptions of energy: A multidisciplinary study. *International Journal of Science Education*, 36(1), 1–23.
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1–25.
- Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist*, 34(4), 207–218.
- McComas, W. F., Clough, M. P., & Almazroa, H. (1998). The role and character of the nature of science in science education. In McComas, W. F. (Ed.), *The nature of science in science education* (pp. 3–39). Dordrecht, Netherlands: Springer.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282.
- National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.
- Nehm, R. H., & Ha, M. (2011). Item feature effects in evolution assessment. *Journal of Research in Science Teaching*, 48(3), 237–256.
- Nehm, R. H., & Ridgway, J. (2011). What do experts and novices “see” in evolutionary problems? *Evolution: Education and Outreach*, 4(4), 666–679.
- 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(6), 744–777.
- Prevost, L. B., Knight, J., Smith, M. K., & Urban-Lurain, M. (2013, April 6). Student writing reveals their heterogeneous thinking about the origin of genetic variation in populations. In Annual Meeting of the National Association for Research in Science Teaching.
- Reiner, M., Slotta, J. D., Chi, M. T., & Resnick, L. B. (2000). Naive physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction*, 18(1), 1–34.
- Russ, R. S. (2018). Characterizing teacher attention to student thinking: A role for epistemological messages. *Journal of Research in Science Teaching*, 55(1), 94–120.
- Russ, R. S., Coffey, J. E., Hammer, D., & Hutchison, P. (2009). Making classroom assessment more accountable to scientific reasoning: A case for attending to mechanistic thinking. *Science Education*, 93(5), 875–891.
- Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, 92(3), 499–525.
- Sandoval, W. A. (2003). Conceptual and epistemic aspects of students’ scientific explanations. *Journal of the Learning Sciences*, 12(1), 5–51.
- Shaw, K. M., Van Horne, K., Zhang, H., & Boughman, J. (2008). Essay contest reveals misconceptions of high school students in genetics content. *Genetics*, 178(3), 1157–1168.
- Slotta, J. D. (2011). In defense of Chi’s ontological incompatibility hypothesis. *Journal of the Learning Sciences*, 20(1), 151–162.
- Slotta, J. D., & Chi, M. T. (2006). Helping students understand challenging topics in science through ontology training. *Cognition and Instruction*, 24(2), 261–289.
- Slotta, J. D., Chi, M. T., & Joram, E. (1995). Assessing students’ misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction*, 13(3), 373–400.
- Smith, M., & Knight, J. (2012). Using the Genetics Concept Assessment to document persistent conceptual difficulties in undergraduate genetics courses. *Genetics*, 191(1), 21–32.
- Talanquer, V., Bolger, M., & Tomanek, D. (2015). Exploring prospective teachers’ assessment practices: Noticing and interpreting student understanding in the assessment of written work. *Journal of Research in Science Teaching*, 52(5), 585–609.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5(4), 381–391.
- Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing* (Jossey-Bass education series). San Francisco: Jossey-Bass.
- Wiser, M., & Amin, T. (2001). “Is heat hot?”: Inducing conceptual change by integrating everyday and scientific perspectives on thermal phenomena. *Learning and Instruction*, 11(4–5), 331–355.