International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Generative mechanistic explanation building in undergraduate molecular and cellular biology

Katelyn M. Southard, Melissa R. Espindola, Samantha D. Zaepfel & Molly S. Bolger

To cite this article: Katelyn M. Southard, Melissa R. Espindola, Samantha D. Zaepfel & Molly S. Bolger (2017): Generative mechanistic explanation building in undergraduate molecular and cellular biology, International Journal of Science Education, DOI: 10.1080/09500693.2017.1353713

To link to this article: http://dx.doi.org/10.1080/09500693.2017.1353713

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2017
https://doi.org/10.1080/09500693.2017.1353713

Generative mechanistic explanation building in undergraduate molecular and cellular biology*

Katelyn M. Southard, Melissa R. Espindola, Samantha D. Zaepfel and Molly S. Bolger

Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ, USA

ABSTRACT

When conducting scientific research, experts in molecular and cellular biology (MCB) use specific reasoning strategies to construct mechanistic explanations for the underlying causal features of molecular phenomena. We explored how undergraduate students applied this scientific practice in MCB. Drawing from studies of explanation building among scientists, we created and applied a theoretical framework to explore the strategies students use to construct explanations for 'novel' biological phenomena. Specifically, we explored how students navigated the multi-level nature of complex biological systems using generative mechanistic reasoning. Interviews were conducted with introductory and upper-division biology students at a large public university in the United States. Results of qualitative coding revealed key features of students' explanation building. Students used modular thinking to consider the functional subdivisions of the system, which they 'filled in' to varying degrees with mechanistic elements. They also hypothesised the involvement of mechanistic entities and instantiated abstract schema to adapt their explanations to unfamiliar biological contexts. Finally, we explored the flexible thinking that students used to hypothesise the impact of mutations on multi-leveled biological systems. Results revealed a number of ways that students drew mechanistic connections between molecules, functional modules (sets of molecules with an emergent function), cells, tissues, organisms and populations.

ARTICLE HISTORY
Received 8 July 2016
Accepted 6 July 2017

KEYWORDS
Biology education; science practices; explanation; multi-level mechanistic reasoning

Introduction

Many national reform movements have called for educators to make science education more authentic to the practices of scientific researchers (American Association for the Advancement of Science [AAAS], 2011; Auchincloss et al., 2014; Next Generation Science Standards [NGSS], 2013). This includes providing students opportunities to construct hypotheses and explanations as scientists do. In particular, *A framework for K-12 science education* (National Research Council, 2011) draws attention to the importance of 'constructing explanations' in scientific practice. This document explains that asking students to develop their own explanations for natural phenomena can not only allow

CONTACT Molly S. Bolger mbolger@email.arizona.edu Department of Molecular and Cellular Biology, University of Arizona, 1007 E. Lowell Street, PO BOX 210106, Tucson, AZ 85721, USA
*All research was conducted at the University of Arizona.

them to engage in science but also contribute to their conceptual change. Ideas from the philosophy of science have proved useful tools for rethinking how science classrooms may align with scientific practices (Matthews, 2014). For example, research on experts' mechanistic reasoning (Machamer, Darden, & Craver, 2000) was adapted to evaluate young students' reasoning patterns (Bolger, Kobiela, Weinberg, & Lehrer, 2012; Russ, Scherr, Hammer, & Mikeska, 2008). Other examples include encouraging Model-Based-Reasoning (Nersessian, 2008) in K-12 classrooms (Lehrer & Schauble, 1998; Passmore, Stewart, & Cartier, 2009). In this study, we demonstrate how a framework for scientists' investigation of biological phenomena in an authentic molecular and cellular biology (MCB) research setting (based on Craver, 2001, 2002a; Darden, 2002; Machamer et al., 2000; Van Mil, Boerwinkel, & Waarlo, 2013) can be a useful tool for considering the forms of reasoning utilised by undergraduate students as they construct explanations. By understanding the types of reasoning students use to build these explanations of biological phenomena, we as educators can be better equipped to design and implement instructional tasks that encourage and challenge students to engage in scientific reasoning in our classrooms.

When providing students with opportunities to construct explanations, it is important to realise that some instructional tasks are more 'generative' than others, meaning that some tasks will more readily encourage students to formulate their own ideas (Perkins, 1993). We define 'generative reasoning' as the thought processes used when constructing meaning or solving a problem in a situation that is novel or unfamiliar. Duncan, 2007 provided a similar definition for this type of reasoning and pointed out that in science, generative reasoning leads to explanations or solutions that are plausible in a particular field but are not necessarily accurate. We use the term 'generative mechanistic explanations' to refer to those explanations created by students that provide a plausible, mechanistic account of a physical phenomenon. This form of explanation is of particular importance in MCB.

In order to build explanations of phenomena explored in the field of MCB, scientists must coordinate mechanistic reasoning with multi-levelled reasoning (Van Mil et al., 2013). For example, molecular mechanisms may be used to explain cellular events, which often interact to give rise to events at the tissue or even organismal level. The relationship between the components at these varied biophysical levels is frequently not straightforward due to properties that emerge only at the level of the system (Dupre, 2009; Kaiser, 2015). Furthermore, being able to consider phenomena at the system level is an integral part of the 'molecular vision' that a biologist in MCB must use in order to understand how cellular phenomena emerge from the functional interactions of macromolecules (Morange, 2008). In addition to the complexity of scale, one must often consider time as a variable in biological explanations. Explanations in the short term may rely on proximal causes of change, whereas explanations that take into account evolutionary time scales often require a different frame. Some have suggested that scientists explore the former by asking 'how' questions and the later by asking 'why' questions (Abrams & Southerland, 2001; Bock, 2017; Mayr, 1961).

Previous work has established a number of difficulties that student encounter as they begin to learn about MCB phenomena. When asked to make sense of complex phenomena, for example, the genetic basis of evolution, students may find it confusing to sort through the various levels of organisation involved in the system (Duncan & Reiser,

2007; Ferrari & Chi, 1998; Wilensky & Resnick, 1999). Researchers have also demonstrated that students frequently lack appropriate molecular explanations to describe MCB phenomena, such as patterns of inheritance or gene expression. Various studies have postulated explanations for this educational challenge. Stewart, Hafner, and Dale (1990) focused on students' lack of a mechanistic understanding of meiosis. Marbach-Ad and Stavy (2000) described students' difficulty with connecting the 'micro' and 'macro' levels. Van Mil et al. (2013) proposed that students may lack understanding of how interactions at the molecular level can give rise to phenomena that emerge at a higher level of organisation. In addition to the challenges of navigating the multi-levelled nature of complex phenomena in MCB, biology students may also struggle to use a mechanistic frame when building explanations. That is, when asked a 'how' question in which the expectation would be to provide a mechanistic explanation, students may instead provide a non-causal response. This confusion about answering 'how' questions has been demonstrated to persist in student thinking even among students at the Grade 12 level (Abrams & Southerland, 2001). Given the significant challenges in developing mechanistic explanations for MCB phenomena, which have been particularly well documented among secondary students, we found it important to further investigate explanation construction among students at a higher level of education. Understanding a student's ability to construct explanations is important for supporting their development of this scientific practice. Currently, little is known about this skill among undergraduate biology students. In particular, we wanted to know if it would be feasible for students to construct generative, mechanistic explanations related to the detailed biological phenomena they are typically asked to consider as students in MCB undergraduate courses. Prior work by Duncan (2007) suggests that such explanations are possible among this population. Our work seeks to extend Duncan's findings by providing a scientific practice lens with which to understand student reasoning in this context.

Previously, Van Mil et al. (2013) compiled a useful framework for characterising the types of reasoning patterns scientists use when constructing explanations of phenomena in MCB. Using this framework, they proposed instructional design principles for developing students' molecular mechanistic reasoning about cellular behaviour. Recently, some of these principles have been applied to design and test curricular materials to support secondary students in *interpreting* visual mechanistic models (Van Mil, Postma, Boerwinkel, Klaassen, & Waarlo, 2016). In our study, we use a similar framework as a basis for understanding how undergraduate students can begin to *build* mechanistic explanations of phenomena in MCB through characterising student responses to instructional probes that ask students to explore 'novel' phenomena in MCB. To this end, we both developed and selected previously published probes which, we hypothesised, would prompt students to engage in mechanistic explanation building that spanned multiple levels of organisation, scale and time. Because we wanted to investigate generation of explanations, rather than recitation of learned ideas, we chose probes that included biological contexts unfamiliar to our students ('novel' contexts). However, we also chose phenomena for which the relevant underlying biological mechanisms should be accessible to students from their prior coursework. In particular, we selected probes that provided students with the opportunity to generate multi-level mechanistic explanations.

Through clinical think-aloud interviews with both introductory and upper division students in biology courses, we sought to investigate the following question: What characterises undergraduates' explanations of 'novel' phenomena in MCB? Specifically, we asked:

(1) Is there evidence of generative mechanistic reasoning in students' explanations?
(2) What strategies seem to guide students' explanation building?
(3) How do students make connections across the multiple biophysical levels of complex biological systems in the context of mutations?

Theoretical framework

Generative reasoning in science

Authentic scientific inquiry includes generative reasoning, i.e. hypothesising, creative problem-solving and building of explanatory models. Generative reasoning is defined as the creation of domain-appropriate explanations that, while not necessarily accurate explanations for phenomena, must plausibly align with current understanding in the domain without violating any assumptions or constraints (Duncan, 2007), or the 'leap' of reasoning that creates a new explanation in a novel context (Clement, 2013). When approaching novel or untested systems in science, experts likely employ prior knowledge schemas, but must extend those schemas and integrate them into the context of the novel system (Clement, 2006; Darden, 2002). Despite existing research on how scientists utilise generative reasoning, we have relatively little understanding of the forms of reasoning that undergraduate biology students might use to generate explanations for novel phenomena.

Mechanistic reasoning

Scientific research often involves explanation through mechanistic description. Mechanisms comprise entities, the physical doers of a system, and activities that the entities perform. These entities and activities are then organised temporally and spatially in such a way as to give rise to the overarching behaviour of the mechanism (Craver, 2002b; Machamer et al., 2000). For example, when considering the mechanism of DNA replication, entities like helicase and DNA polymerase undergo activities such as 'binding' and 'elongation'. According to Machamer et al. (2000), 'thinking about mechanisms as composed of entities and activities provides a resource for thinking about strategies for scientific change' and can provide a starting point for building hypotheses about mechanistic schema that can then be explored and modified through experimentation. In order to build mechanistic explanations, scientists can use a number of mechanistic reasoning strategies. By 'forward and backward chaining', one can reason about one component of the mechanism based on known or hypothesised activities or properties of another component (Darden, 2002; Machamer et al., 2000). One can also abstract a general framework from an adjacent context, apply it to the mechanism under study using 'analogical reasoning' and instantiate the schema by filling in the framework with context-relevant components (Clement, 2013; Darden, 2002).

A full mechanistic description is not always required in science; in some scenarios, a truncated, abstract description of the mechanism, known as a 'mechanistic schema', will

suffice as explanation (Darden, 2002; Machamer et al., 2000). These schemas are often proposed by considering the observable factors in the phenomenon and hypothesising the involvement of key entities. Because these schemas are created without knowing exactly the entities or activities involved in each component of the phenomenon, experts often draw on mechanistic components from adjacent and relevant contexts. This often happens by hypothesising about the involvement of a key entity based on an observable activity from within the system. Darden calls this 'schema instantiation' and considers it a truncated abstract description of a mechanism that can be filled with more specific descriptions of component entities and activities (Darden, 2002; Kitcher, 1989).

Mechanistic reasoning in MCB

Most phenomena in biology span multiple biophysical levels (such as biochemical, molecular, cellular, organismal and population levels), and as a result development of mechanistic explanations requires multi-level reasoning. Mechanisms in MCB are often nested in hierarchies (Machamer et al., 2000). Within these nested hierarchies, entities at lower levels interact to create a higher level event, and those entities and activities interact to give rise to an even higher level event (Craver, 2002b; Machamer et al., 2000). Relationships among these levels work as elements of a whole, with each lower level acting as an interacting component to higher levels (Craver, 2002a; Van Mil et al., 2013; Wilensky & Resnick, 1999).

In the phenomena investigated in MCB, discrete biological functions are rarely a product of individual entities in a system, but instead arise through the interactions of many entities. Therefore, it is often necessary to consider the molecular subassemblies that give rise to a particular behaviour. Hartwell, Hopfield, Leibler, and Murray (1999) were the first to introduce the 'functional module' as the level of organisation between the molecular and cellular biophysical levels. These functional modules consist of an ensemble of molecules working together to give rise to a discrete and separable function. Functional modules may also serve as mental organisers allowing compartmentalisation of knowledge about the mechanics of molecular mechanisms according to core functional features (Southard, Wince, Meddleton, & Bolger, 2016; Darden, 2002).

Van Mil et al. (2013) proposed a framework for this multi-level molecular mechanistic explanation building in MCB, drawing heavily from Darden's work on mechanism discovery (Darden, 2002) and a historical analysis of Adler's discovery of the mechanism of bacterial chemotaxis (Adler, 1966, 1975). They proposed that in order to tackle explaining a novel MCB phenomenon, scientists typically divide the overarching activity into composite parts. By '*functionally subdividing the overarching* activity into sub-activities', scientists are practising 'modular subassembly' by considering how interactions of entities at a lower biophysical level give rise to the observed activity (Darden, 2002; Van Mil et al., 2013).

In addition to functional subdivision, scientists investigating phenomena in MCB often rely on the use of mutant organisms that display a 'broken' activity. By observing mutant phenotypes, scientists may add potential entities to the list of component parts. However, isolating and confirming the existence of an entity in the system does not define the role that it plays in the mechanism. Therefore, experts must use the properties of entities to hypothesise about the possible activities in which the identified molecules might engage. Other experimentation can reveal sub-activities, which can be used to predict the involvement of an entity, based on the known composition and function of that

entity. Therefore, Van Mil et al. (2013) additionally proposed that 'predicting molecular properties from activities and vice versa' is an essential reasoning tool for explanation construction in MCB. This practice among biologists is often carried out through 'forward and backward chaining' through the entities and activities: piecing together a complete mechanism by describing the mechanistic events from starting to termination conditions (Darden, 2002; Machamer et al., 2000; Van Mil et al., 2013).

Finally, Van Mil et al. (2013) proposed that 'hypothesising and predicting organisation in the mechanism' is the culminating step of biologists' explanation construction. This final requirement is often the most difficult for molecular and cellular biologists (Van Mil et al., 2013). It requires consideration of the temporal and spatial location of all entities and activities and an explanation of how these components interact to give rise to the whole.

The central aim of this study is to describe the forms of explanation used by undergraduate students in novel biological contexts. Specifically, we seek to explain *how* students' reasoning is adapted to the particular demands of explaining phenomena in MCB. Although the scope of our analysis was not limited to forms of reasoning that have been described among practising biologists, the preceding framework was useful for explaining many patterns of reasoning that we observed in the students we interviewed.

Methods

Interview probes: rationale and description

Bacteria Sensing Probe

The Bacteria Sensing Probe (Figure 1) was designed and implemented by Ravit Duncan and colleagues, and has been used to elicit student ideas about molecular biology and genetics in studies exploring student ideas ranging from middle school to undergraduate biology students (Duncan, 2007; Duncan & Reiser, 2007; Duncan & Tseng, 2011). This probe was selected for the present study because: (1) the phenomenon of bacterial

Bacteria Sensing Probe (Duncan, 2007)

Most bacterial cells can sense and respond to substances in their surroundings. For example if you put bacteria near food substances (sugar) they will sense the food and move towards the food; if you put them near poisonous substances they will sense the poison and move farther away from it. Some bacteria have mutations and can no longer sense substances in their environment. These bacteria do not move towards food or away from poisons and therefore are more likely to die from starvation or poisoning.

1) *How do you think "normal" bacteria sense stuff in their surroundings and move accordingly? What is going on inside the bacteria cell?*
2) *What does "mutation" mean in this context? How do you think a mutation can cause a bacterium to lose its ability to sense substances in the surroundings?*
3) *Do you think there could possibly be a mutation that made bacteria more sensitive to substances? How would that work?*

Figure 1. Bacteria Sensing Probe (Duncan, 2007; Duncan & Reiser, 2007; Duncan & Tseng, 2011).

chemotaxis is novel to our student population, but students were expected to have the background knowledge to generatively hypothesise about underlying mechanisms; (2) the probe has been successful previously in eliciting student reasoning and demonstrating a range of conceptual understandings; (3) the probe asks students to explore a mechanism that spans from the molecular to cellular level, which has been described as a particularly difficult range for students (Van Mil et al., 2013); and (4) the probe is directly aligned with the topic of bacterial chemotaxis used by Van Mil et al. (2013) as a historical example of expert heuristics in generating mechanistic explanations for novel phenomena.

Poisonous Peruvian Plant Probe

The Poisonous Peruvian Plant Probe (Figure 2) was created by our research group to enable comparison of student explanations in an additional context. Like the Bacteria Sensing Probe, students have been exposed through prior coursework to explanatory mechanisms relevant to this scenario, such as cell differentiation, differential gene expression, mutations, natural selection and evolution. We therefore expected that students would have the knowledge needed to approach explanation building in this context. However, most of the students did not have previous course experience with plant development or plant evolution. While the probe first asks students to consider a linear causal mechanism in plant development (analogous to the reasoning that might be expected in the Bacteria Sensing Probe), it additionally asks students to consider a complex system involving natural selection and evolution that spans many biophysical levels.

Study population

Participants were student volunteers from one of three courses at a large public university: Introductory Biology ($N = 17$), Molecular Genetics (upper division, $N = 20$) and Cell & Developmental Biology (upper division, $N = 7$). All three courses are large enrolment

Poisonous Peruvian Plant Probe

Ecologists studying vegetation in the rainforest of Peru have found a rare shrub that is poisonous to the foraging animals in the region. The leaves and fruit of the plant are poisonous to the animals, but the branches and roots are non-poisonous. The plant grows from a seed

1) *How does a single seed give rise to both the poisonous and non-poisonous parts of the plant?*

In a Peruvian mountain range, the foraging animal population is smaller than in the rainforest. Here the plant life is less diverse and more sparse. One of the plant species found in this mountain range is quite similar to the one found in the rainforest, but is not poisonous.

2) *Describe any possibilities you can think of for why this similar plant is not poisonous.*

Figure 2. Poisonous Peruvian Plant Probe, developed by authors.

courses taught through the MCB Department: Introductory Biology (typical enrolment = 1800 students per year), Molecular Genetics (typical enrolment = 200 students) and Cell & Developmental Biology (typical enrolment = 200 students). Additionally, each course is required for graduation with an undergraduate MCB degree, and they cover (to varying extents) the prior knowledge needed for explanation building of the two probes (e.g. cell signalling, signal transduction, cell motility, cell differentiation, differential gene expression and mutations). We recruited students on a volunteer basis with specific language expressing our desire to include students who have a range of comfort with course materials, and $5 gift cards to a coffee shop were given as an incentive for participation. We interviewed participants (referred to by pseudonym) at the end of each respective course, the majority within the 1–3-week period before their final exam. Transcripts of interviews were made from audio recordings.

Clinical think-aloud interviews

Student interviews

We conducted semi-structured, clinical think-aloud interviews with student participants. Each interview lasted approximately one hour and was conducted by one of six trained interviewers in the research group using a standardised set of questions and probes. One upper division student interview was excluded from analysis due to an inaudible audio recording. During the interview, we asked all students to read and respond to the Bacteria Sensing Probe ($N = 44$, 17 introductory, 27 upper division) and a subset of these interviews also included the Poisonous Peruvian Plant Probe ($N = 24$, 17 introductory, 7 upper division). The interview also included describing mechanisms for DNA replication, transcription and translation and making concept maps for these ideas. These additional activities were the focus of a previous study and will not be discussed here. Three students in the current study did not receive these additional interview questions due to time constraints during the interview.

Additional interviews with MCB experts

While functional modules for 'bacteria sensing' have been reported to include the 'sensing', 'signalling' and 'motor' modules through a historical analysis of Adler's work on bacteria chemotaxis (Adler, 1966, 1975; Van Mil et al., 2013), we interviewed two experts in the field of MCB to identify the relevant functional modules involved in the Poisonous Peruvian Plant Probe. Expert 1 is a plant geneticist and Expert 2 is a developmental biologist; both are tenured professors in the MCB Department at a large research (R1) university. Excerpts from these interviews are provided in Figure 3. Through iterative transcript read-throughs, we identified two functional modules within the experts' answers to the question, 'How does a single seed give rise to both the poisonous and non-poisonous parts of the plant?'. These modules were defined as 'cell differentiation' and 'tissue-specific poison generation' modules.

Analysis

Iterative transcript read-throughs

The investigation of student reasoning began with a systematic read-through of a portion of the interview transcripts and an examination of any related interview artefacts. This

Excerpt from Interview with Expert 1:

Expert 1: [Reads question 1] "So the answer in my mind would be that a seed, a seed actually has a lot of differentiated cells in it already and in some of those differentiated cells you are going to generate the, well in the meristems that are in the seed you are going to generate the leaves and the fruit, but you just have... and those cells are not yet differentiated, but once they differentiate, then you are going to have... differentiation will lead to specific cell types and in those cell types you will have transcription factors that bind to the enhancers of whatever proteins are needed to make this particular protein. So that's one possible mechanism. Another possible mechanism might be that there's some kind of environmental component, maybe herbivory, of the plant that actually leads to induction of the transcription of the poison. Ok. That's less likely because..."

Interviewer: "Can you tell me more [about] what that would be?"

Expert 1: "Well that would mean you have a chewing response, wound response in the leaves that signals to the cells in the leaf that are wounded and to the surrounding cells, and then you generate the toxin."

Interviewer: In other words as-needed

Expert 1: 'As-needed' [inaudible additional comment]

Interviewer: "Gene expression?"

Expert 1: "Yah, exactly. So but this says that the leaves and the fruit are poisonous and it doesn't really say that they are necessarily always poisonous or not always poisonous or only poisonous after feeding, so that seems less obvious."

Interviewer: "Ok so the first, just to recap then, would be that you start of as a seed with the, the different cells will differentiate differently, and then that will…"

Expert 1: "And the seeds, the cells that give rise to the leaves and the fruit are definitely in seed they're stems cells, they're complete stem cells there's no differentiation at all at that point. So…"

Interviewer: "Ok and then as they then develop into these tissues there is going to be a change?"

Expert 1: "Yah there'll be some, a series of changes in cell type and in the suite of transcription factors that are activated in those particular cells, is how I would describe it."

Excerpt from Interview with Expert 2:

Expert 2: [Reads question 1] "So the first one to me, that you have some protein is likely to be responsible for the poison. And it's likely under different temporal and spatial control via its promoter region. So that the promoter has regions specific for, what's the poisonous part?"

Interviewer: "The leaves and the fruit"

Expert 2: "The leaves and the fruit. So that's how, that's how I would look at that problem. And then the second problem, why is the mountain one... So we don't know if it's the same species, it could just be a look-alike, and so the way to approach finding that out would be to get in and look at the DNA and see how similar the DNA is, alternatively if we discover that it is the same species but you just have a variant population that may have lost, it would have a small deletion it is promoter, or even just a SNP that could give rise to some difference in the expression of the poison protein."

Figure 3. Excerpts from interviews with two MCB experts in response to the Poisonous Peruvian Plant Probe, Question 1.

initial phase of analysis allowed us to identify themes that became the focus of further analysis, including development of schemes for qualitative coding. We also conducted case studies of a few students to better understand their reasoning. The results of case-study analysis, however, are not included here due to space restrictions, but were used to inform the qualitative analysis presented. Finally, in Part 2 of the Results section, we provide some examples of transcript analysis that were the result of a systematic read-through of all transcripts, guided by an initial round of qualitative coding.

Qualitative coding

Guided by our theoretical framework and our initial observations of student reasoning in this context, we created two coding schemes to capture features of students' generative

mechanistic reasoning, as follows. The Bacteria Sensing Probe and Poisonous Peruvian Plant Probe were analysed separately.

Recognition of functional subdivisions and mechanism creation. The three functional subdivisions historically created by Adler to describe bacterial chemotaxis were incorporated into the coding scheme for student responses to the Bacteria Sensing Probe: 'sensing', 'signalling' and 'motor' functional modules (Adler, 1966, 1975; Van Mil et al., 2013). For the Poisonous Peruvian Plant Probe, we identified two functional modules, 'cell and/or tissue differentiation' and 'localised poison generation', based on our think-aloud interviews with two MCB experts. The coding scheme was used to indicate whether each functional module was 'Not Mentioned' or 'Identified'. Additionally, the scheme indicated whether a student's explanation for each module included an 'Isolated Molecular Entity', 'Molecular Mechanism' or 'Non-Molecular Mechanism' (for the Bacteria Sensing Probe, see Table 1), or included a 'Mechanism' (for the Poisonous Peruvian Plant Probe, see Table 2). For the Poisonous Peruvian Plant Probe, coders also indicated if the stated mechanism was 'Molecular', 'Cellular', 'Tissue' or 'Population/Evolutionary-based'. Because the aim was to capture generative hypothesis building rather than simply normative versus non-normative answers, we did not exclude students' explanations from any category on the basis of their inclusion of non-normative ideas. If a student stated a biologically incorrect and yet plausible and domain-appropriate idea, it did not affect how the student's explanation was coded. For example, if a student hypothesised that a bacterium moves via a pseudopod, which is a mode of motility found in eukaryotes but not bacteria, this technically non-

Table 1. Bacteria Sensing Probe coding scheme with examples.

Code[ab]	Description	Example
'Molecular Mechanisms' Code:	Students provided (minimum) one entity and its associated activity	'I would imagine that the bacterium has receptors in its cell membrane that are made up of integral [membrane] proteins. When the correct signalling molecule, either sugar or some kind of poison that will be toxic to the bacteria, interact with those proteins, then it will trigger some kind of cell response.' ('Evan')
'Isolated Entity' Code:	Students provided a single entity (or entities) that do not perform any molecular activities or contain causal connectors[c]	'Well because bacteria are unicellular, there is receptors on the outside of the cell.' ('Alexis')
'Non-Molecular Mechanism' Code:	Student provides a mechanism that is vague or general, and/or not primarily at the molecular biophysical level (i.e. not containing key molecular entities)	'Um so I guess things like smell have to do with it. Not necessarily smell but like um, like what comes off the food or the poison has to do with the senses of the bacteria.' ('Victor')
'Identify Only' Code:	Student simply identified that the bacterium: senses its surroundings ('sensing module'), induces some type of intracellular signalling event ('signalling module'), and/or undergoes some type of motion ('motor module')	'So, um, I am just thinking that sugar and poison are stimulus somehow and they probably affect the environment and therefore affect how the bacteria respond to it.' ('Bianca')
'Not Mentioned' Code:	Student's response did not mention any entities, activities, or general function of the target functional module	N/A

[a]Each functional module ('sensing', 'signalling' and 'motor') was coded individually.
[b]Boundaries between modules: (1) the 'sensing module' began with stimuli recognition and ended with the entity and actions of a receptor, (2) any intracellular signalling was considered part of the 'signalling module', and (3) the 'motor' module begins with proteins that bind/activate 'motor' structures and ends with movement-related actions.
[c]These single entities were often receptors for the 'sensing module', a single kinase or signalling protein for the 'signalling module', and flagella or cilia for the 'motor module'.

Table 2. Poisonous Peruvian Plant coding scheme with examples.

Code[ab]	Description	Example
Mechanism Code:	Students described a mechanistic process, including at least one entity at its associated activity. Notes were made for each module as to whether the categorised mechanisms were at the molecular, cellular, tissue or population/evolutionary level	'So, in any specific cell in the body of a multicellular organism, not all of the genes are active because then you wouldn't have any differentiation between the cells. So, there's probably specific genes that are involved in making leaves and fruit poisonous and branches and roots non-poisonous, and they're probably – whatever's coding for the poison is turned on in the leaves and the fruit, but that gene would probably not be turned on in the branches and roots […]: So, it's probably more of a regulation thing on the DNA. So, maybe in the branches and roots that DNA sequence that codes for the poison is wrapped up in histones, or maybe there's a repressor on that particular gene so RNA polymerase can't bind to it. Or there's a lack of the enhancers that RNA polymerase needs to bind to that piece of the DNA. So the mRNA transcript is never made in those particular cells.' ('Sally')[c]
Module Identified Only Code:	Students simply mentioned plant growth and/or the idea that cells/tissues must specialise for the 'cell/tissue differentiation module', and when students simply mentioned that the poison must end up in the leaf and fruit structures, but no in the root or branch structures for the 'localised poison generation module' without describing a mechanism that explains how these events might occur	'And so, some cells, for whatever reason, are destined to become these poisonous ones, and other ones in different parts of the plant are non-poisonous, the same way that an ear is supposed to become an ear and a liver is supposed to become a liver.' ('Michelle')
Module Not Mentioned Code:	Students did not mention growth, cell differentiation, or even that the seed gives rise to different parts of the plant for the 'cell/tissue differentiation module'. This code was also given to responses that did not mention the plant as containing poisonous versus non-poisonous parts for the 'localised poison generation module'	N/A

[a]Each functional module ('cell/tissue differentiation' and 'localised poison generation') was coded individually.
[b]The 'cell/tissue differentiation module' included descriptions about the process of cell differentiation during development, with different plant tissues or parts developing from a single seed. The 'localised poison generation module' included descriptions of the mechanistic events by which the leaves and fruit, but not the branches or roots, gain this poisonous feature.
[c]Mechanism noted by coders as 'molecular.

normative idea would not affect coding of the students' explanation. When students created explanations that encompassed multiple categories, their responses were categorised by the code representing the most expert-like reasoning observed.

Flexible thinking about mutations. For the Bacteria Sensing Probe, the coding scheme was used to indicate whether students: (1) understood that mutations could impact a protein molecule in some way; (2) connected the idea of mutations to the context of bacterial chemotaxis (i.e. whether they hypothesised the mutation in a single specific protein in the pathway that would give rise to the altered phenotype); or, (3) hypothesised that the mutation could arise in more than one protein along the pathway to give rise to the altered phenotype. For the Poisonous Peruvian Plant Probe, the coding scheme was used to indicate whether students: (1) proposed a connection between DNA or genes

and the phenomena of the two different plant phenotypes in different environments; (2) included discussion of evolution, natural selection, or selective pressures in their explanation; and (3) included ideas of differential gene expression in their explanation.

Coding implementation. Coding analysis was performed by two independent coders. Coding analysis was performed independently for the different schemes. Thus, agreement between raters ((total instances−disagreements)/total instances × 100%) is reported as follows: for 'Recognition of Functional Subdivisions and Mechanism Creation', an average of 77% agreement for both probes; for 'Flexible Thinking about Mutations', an average of 91% agreement for both probes. Due to the complexity of the first coding scheme, the first 25% of the transcripts were used for coding scheme development. Reported results represent consensus for all transcripts.

Counting students' use of key terms

In order to count students' use of the key sensing entities ('receptor' for the sensing module, and/or the key motor entities 'cilia', 'flagella' or 'lamellipodia' for the motor module), we used Nvivo text search query as a starting point to find instances of students using these entities in their responses to the Bacteria Sensing Probe. Each instance flagged by the program was verified by returning to the transcript to verify the student's meaning. For those transcripts that were not flagged by the program, we read each and marked transcripts as either including the key sensing entity 'receptor', or not. For example, this could include a description of the protein in context without use of normative vocabulary, such as 'membrane protein on the cell wall'. Additionally, each student's transcript was coded in a similar manner for the key entities 'cilia', 'flagella' or 'lamellipodia' that are involved in the motor module.

Statistical analysis

For the mechanistic coding sets of both the Bacteria Sensing and Poisonous Peruvian Plant Probes, we compared introductory and upper division student populations. Computations were performed using IBM SPSS for Windows, Version 22. Specific details of tests that were performed are provided in the Results section.

Results

When exploring novel phenomena in the field of MCB, experts often observe a behaviour or set of behaviours and explore 'how' questions through creating and testing hypotheses via experimentation. In an educational setting, it is often difficult for instructors to parallel this authentic scientific pursuit of inquiry and experimentation, particularly in non-laboratory-based classes. However, when students in our sample population were presented with 'how' questions through the Bacteria Sensing and Poisonous Peruvian Plant Probes in an interview setting, we discovered that they were frequently using various elements of multi-level molecular mechanistic reasoning to create generative explanations for the phenomena presented. In Results Part 1, we will explore features of students' explanations for these two probes and some of the strategies that they used to build their explanations. Specifically, we will present our findings for Research Questions 1 and 2. In Results Part 2, through exploration of our Research Question 3, we will describe the forms of reasoning

students used to explain how mutations could perturb the biological systems they explored in Part 1. This section will explore student reasoning across multiple physical and ontological levels, including their ideas about evolution in response to a 'why' question.

Part 1: students use modular mechanistic thinking to construct generative explanations for biological phenomena

When students were asked to respond to both the Bacteria Sensing and Poisonous Peruvian Plant Probes, we discovered that several features of students' explanations aligned with scientists' strategies for explanation building in the field of MCB. These features include using modular thinking to identify relevant functional modules within the system, 'filling in' these sub-modules in the system with mechanistic elements and piecing together the mechanistic elements to create a productively continuous mechanistic explanation of the phenomenon. Students appear to be using strategies such as hypothesising the involvement of specific entities within the system through considering their activities, and instantiating relevant schema into the context at hand.

Students commonly identify the relevant functional modules of a system

During iterative transcript read-throughs, we discovered that many students were identifying the same functional subdivisions as experts in the field when creating an explanation for the biological phenomena presented in both probes. For the phenomenon of bacterial chemotaxis, Alder historically identified three functional modules underlying the overarching activity of chemotaxis in bacteria: the sensing, signalling and motor modules (Adler, 1966, 1975; Van Mil et al., 2013). Coding analysis was performed on all student responses to Question 1 of the Bacteria Sensing Probe (see Methods) in order to observe students' identification of the functional modules underlying bacteria chemotaxis. Across the student population, the sensing module was identified by all students, with a large percentage also identifying the signalling and motor modules (Figure 4(A)). At an individual level, 69% of students incorporated all three functional modules when creating their explanation, with no student unable to identify at least one module (Figure 4(B)). Very little difference was observed between introductory and upper division students in identifying these functional subdivisions of the chemotaxis behaviour (Figure 4(B)). These results suggest that most students used the strategy of subdividing the phenomenon into modules to aid their production of an explanation.

Similarly, during iterative transcript read-throughs of the Poisonous Peruvian Plant Probe, we observed students primarily identifying two functional modules underlying the system: the cell differentiation module and the tissue-specific poison generation module. These two functional modules were the same as those identified in the expert interviews (see Methods). Coding analysis of student responses to Question 1 of the Poisonous Peruvian Plant Probe revealed that all students identified the tissue-specific poison generation module, and just over half included the cell differentiation module (Figure 5(A)). Little difference was observed between introductory and upper division students in their ability to identify the relevant functional modules (Figure 5(B)). Taken together, coding results from both probes suggest that the strategy of functionally subdividing the mechanism into modular elements was relatively common among students as they generated explanations for 'novel' MCB phenomena.

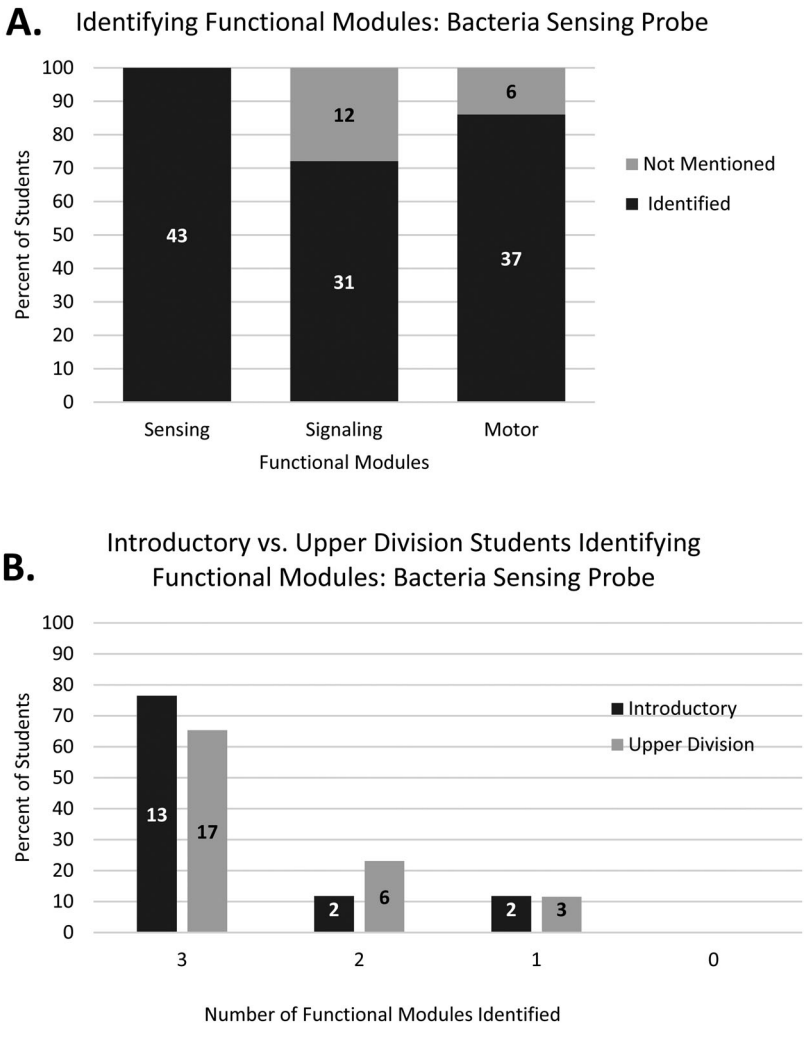

Figure 4. Coding results for functional module identification in the Bacteria Sensing Probe, Question 1, by (A) identification of functional modules by the combined student population and (B) identification of multiple functional modules by individual students at the introductory and upper division levels. Bar graphs display percentage of students while specific numbers of students per category are indicated by the data labels within each bar. $N = 43$.

Students range in how they 'fill in' functional modules with mechanistic elements

After discovering that many students identified relevant functional modules for both systems, we next sought to characterise what mechanistic elements students used to 'fill in' these sub-modules to create mechanistic explanations. Coding analysis revealed that explanations ranged from productively continuous molecular and cellular mechanisms, to those characterised by isolated entities, to non-molecular reasoning. In other cases, students identified the need for a particular functional module, but did not fill in any details to explain how that function or behaviour (such as 'sensing' or 'cell differentiation') would work. Next, we provide examples of these forms of explanation in the context of each probe.

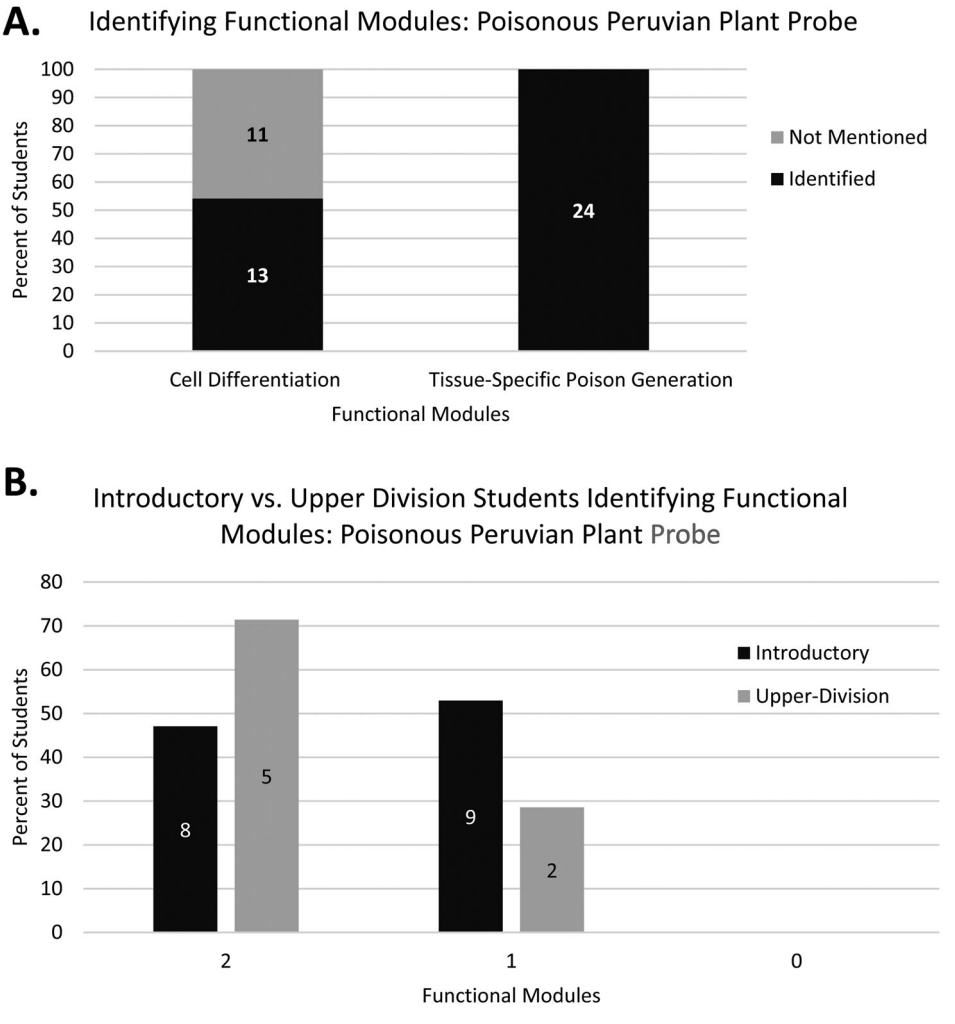

Figure 5. Coding results for functional module identification in the Poisonous Peruvian Plant Probe, Question 1, by (A) identification of functional modules by the combined student population and (B) identification of multiple functional modules by individual students at the introductory and upper division levels. Bar graphs display percentage of students while specific numbers of students per category are indicated by the data labels within each bar. $N = 24$.

For the Bacteria Sensing Probe, we discovered that more than half of students included a molecular mechanism for both the sensing and signalling module (Figure 6). For example, Sally fills in the sensing module with a basic mechanism, saying:

> Alright. So since it's a mutation that turns this off, probably there's some kind of protein, that's sensing the stuff I guess. And it's probably a membrane protein. Just because those [the membrane proteins] are probably more exposed to the outside environment and able to get information from the environment. So yah, it probably is some kind of protein that binds to poisons or foods and has a reaction inside the bacteria that makes it either go toward or away from the substance.

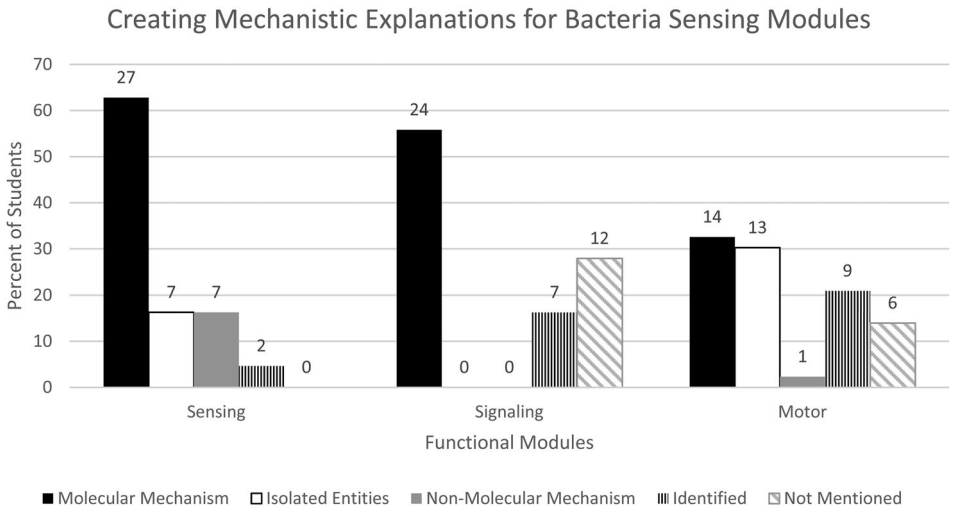

Figure 6. Coding results for mechanistic elements in the Bacteria Sensing Probe, Question 1. $N = 43$.

However, when students were not creating basic mechanistic explanations, we observed some students hypothesising the involvement of a single key entity (shown as 'isolated entities' in Figure 6) in the module, without describing any of its associated activities in the system. Many students hypothesised the involvement of a receptor protein in the sensing module, but did not continue to build a mechanistic explanation for how this entity works to create the sensing activity. For example, Alexis simply mentions that, 'well because bacteria are unicellular, there is [are] receptors on the outside of the cell'. Fewer students created a molecular mechanism for the motor module, and often instead used this method of naming a single entity, typically flagella or cilia, devoid of activities or causal connections. Coding analysis for the Bacteria Sensing Probe revealed some occasions in which students used 'non-molecular' explanations, which most often involved the bacterium's general ability to 'feel' the stimuli or described mechanisms analogous to 'smelling' in higher organisms. For example, Victor states, 'um so I guess things like smell have to do with it. Not necessarily smell but like um, like what comes off the food or the poison has to do with the senses of the bacteria'.

For the Poisonous Peruvian Plant Probe, our coding scheme captured mechanistic explanations at the molecular, cellular, tissue and population/evolutionary-based levels (Figure 7). For this probe, our scheme provided flexibility to allow us to categorise a student's explanation in more than one of these mechanistic categories. 'Molecular level mechanisms' like Sally's were most often observed in students' explanations of the 'tissue-specific poison generation' module:

> So, in any specific cell in the body of a multicellular organism, not all of the genes are active because then you wouldn't have any differentiation between the cells. So, there's probably specific genes that are involved in making leaves and fruit poisonous and branches and roots non-poisonous, and they're probably – whatever's coding for the poison is turned on in the leaves and the fruit, but that gene would probably not be turned on in the branches and roots […]: So, it's probably more of a regulation thing on the DNA. So, maybe in the branches and roots that DNA sequence that codes for the poison is wrapped up in histones, or maybe there's a repressor on that particular gene so RNA polymerase can't bind to it. Or

there's a lack of the enhancers that RNA polymerase needs to bind to that piece of the DNA. So the mRNA transcript is never made in those particular cells.

While the 'cell differentiation' module was less often identified or filled in by students, some, like Brenda, created a primarily 'cellular level description' for this particular module:

So the cells differentiate eventually. Then they end up kind of segregating themselves. Because I think, from [her course], I kind of learned that, 'well, cells that are more like each other are more likely to congregate together'. And then with like, a mosaic of cells. So that's kind of how they end up differing from each other.

Maggie, on the other hand, considers both the 'population level (evolution-based)' and the 'tissue level' when creating her explanation:

I guess the, you know maybe the leaves and fruit, maybe because … I guess this has to do with adaptation again. Maybe if their leaves and fruit have a greater probability of being eaten by animals, as opposed to branches and roots I think. I mean, I wouldn't want to be an animal in Peru eating a branch [laughs]. So maybe those leaves and fruit have been able to chan – adapt and keep that poisonous, like … they've adapted and been able to grow a certain thing that they know is toxic to the animal.

Overall, these results demonstrate a range of explanation forms that students use to provide details for 'how' these novel phenomena may take place. 'Filling in' with mechanistic ideas, i.e. entities and their activities, was relatively frequent in student explanations, especially for some functional modules. However, some students also used functional modules to support their explanation without adding any mechanistic elements. In other cases, students simply hypothesised the involvement of a single entity that could be part of a functional module but did not extend this explanation into a mechanism involving activities that that entity may perform within the system. While this type of explanation could be the starting point of a mechanistic explanation, it lacks the

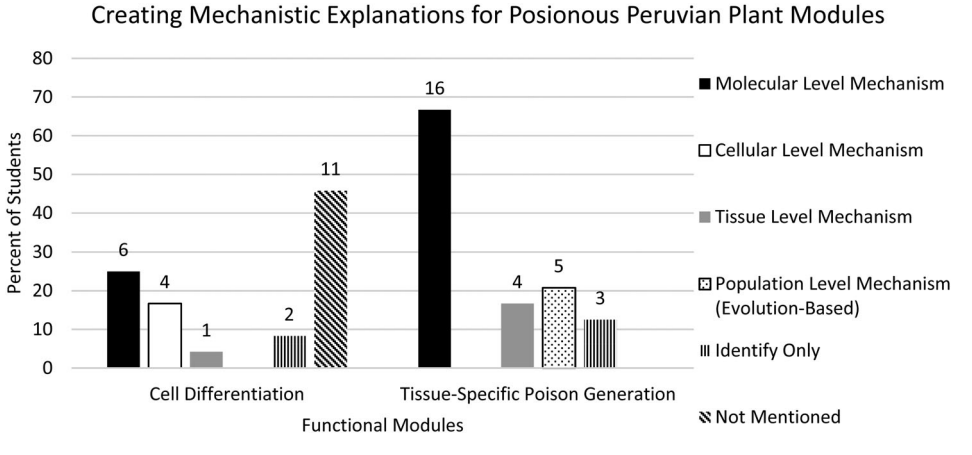

Figure 7. Coding results for mechanistic elements in the Poisonous Peruvian Plant Probe, Question 1. Questions were coded for whether they were 'mechanism', 'identify' or 'not mentioned'. The 'mechanism' codes were given additional descriptors that represented the different biophysical levels of mechanisms described. Six transcripts were given more than one of the level-representative descriptors. For this reason, values do not sum to 100%. $N = 24$.

explanatory power of a more complete mechanistic account. Overall, there were no apparent differences in explanation form between upper division and lower division students. We provide a more detailed analysis of this comparison in the final section of Part 1.

Strategies that appear to guide student explanation building

Both the Bacteria Sensing and Poisonous Peruvian Plant Probes asked students to consider biological phenomena that had not been covered in mechanistic detail in their course work. Much like authentic scientific pursuit in the field of MCB, this required students to hypothesise the involvement of key entities and their activities, without knowing the components of the system, by drawing on their observations and prior knowledge in order to build appropriate or plausible mechanistic explanations. We looked deeper into some of the strategies that students appeared to use as drivers for building mechanistic explanations. Two frequently productive strategies were hypothesising the involvement of entities from their activities in the system and instantiating mechanistic schema.

For the Bacteria Sensing Probe, we observed that many students readily hypothesised the involvement of key entities in both the motor and sensing modules. Of the students who identified the motor module in their explanation, 73% also implicated a key motor protein or structure for the module. These proteins or structures were most often flagella, cilia or lamellipodia structures. Of the students who identified the sensing module, 77% implicated the role of a membrane-bound signalling protein, known as a receptor protein, to carry information or signals about the presence of food or poison from outside the cell to its interior. Students often used the salient activities of the functional modules, such as 'sensing' or 'movement', to hypothesise the involvement of these key entities. For example, Karla says, 'Sense stuff … well they've probably got, like, receptors'. Many students appeared to use their hypothesis of key entities, such as this key receptor entity, as a driver to generate a schema for the functional module at hand, like the sensing module. For example, Evan says:

> I would imagine that the bacterium has receptors in its cell membrane that are made up of integral proteins. When the correct signaling molecule either sugar or some kind of poison that will be toxic to the bacteria interact with those proteins, then it will trigger some kind of cell response.

Here, we observe Evan hypothesising the involvement of a receptor in his mechanistic explanation. By considering the receptor's location on the bacterium's membrane and its chemical make-up that includes proteins that span from the outside to the inside of the cell (via integral membrane proteins), Evan instantiates a 'receptor' schema for how the bacterium can receive an external signal through interaction with sugar or poison that is relayed inside the cell to trigger a cellular response, such as movement.

In the Poisonous Peruvian Plant Probe, we found that more than half of students hypothesised the involvement of genes as key entities for both modules. Using this hypothesised entity, they would often then instantiate a 'gene expression' schema. Students instantiated this schema in different ways, depending on the hypothesised details of the context at hand. For example, the gene expression schema may be invoked to explain both the 'cell differentiation' and the 'tissue-specific poison generation' functional modules in the Poisonous Peruvian Plant Probe. However, it is the fitting of the schema into the context that creates different mechanistic explanations for the individual modules.

For example, Allison uses the 'gene expression' schema to explain the 'tissue-specific poison generation' module:

> Ok. So … I don't know … I don't know much about seeds, but I assume it carries on the DNA of the original plant. And so, just like in bacteria and eukaryotes, it will code for RNA, which will code for proteins. And so I'm guessing that it expresses a poison protein, I guess, like a protein that controls the poison in the poisonous parts of the plant, and it just isn't expressed at all in the non-poisonous parts of the plant.

Here, we see Allison considering the origin of the poison, hypothesising that it is a protein and that proteins are expressed by the genes of a plant. By hypothesising the involvement of a key entity, a poison gene, she is able to 'forward chain' (Machamer et al., 2000) to the involvement of another key entity, a poisonous protein. This instantiates the gene expression schema where a poison gene is expressed in some plant tissues but not others.

Allison appears to use her hypothesis of the involvement of 'genes' to consider a 'gene expression' schema, namely 'turning on' or 'turning off' of certain genes. However, differences in the instantiation process seemingly generate different mechanistic explanations, tailored to the context at hand.

Through these results, we have described how students proposed explanations by utilising mechanistic schema. Despite the fact that the biological phenomena were 'novel', these students were able to relate these situations to schema that were likely learnt in another context. In some cases, they explicitly stated that they were drawing on a related idea from another context. In other cases, the schema seemed more abstract and references to other contexts were not explicitly mentioned. When generating explanations, students often formed hypotheses about the involvement of key entities and their relevant properties, which triggered a mechanistic schema upon which they elaborated to fit the current context.

Putting together the mechanistic pieces is a challenging final step for students

Our analysis suggests that, overall, many students in the target population hypothesised molecular mechanisms to explain the sub-activities involved in a particular functional module (70% of students for each probe). However, analysis at the level of the individual student revealed that few students created mechanistic explanations that spanned all three functional modules, creating a productively continuous explanation from staring to termination conditions (28% of students for the Bacteria Sensing Probe and 25% for the Poisonous Peruvian Plant Probe). Next, we describe these findings in more detail.

Results from the Bacteria Sensing Probe revealed that less than a third of the study population hypothesised molecular mechanisms for all three of the functional modules (Table 3). A similar number of students did not include a molecular mechanism for

Table 3. Individual students' mechanistic explanations for multiple functional modules.

	Mechanisms for three functional	Mechanisms for two functional modules	Mechanism for one functional module	No mechanisms coded
Bacteria Sensing Probe	27.9% (12)	25.6% (11)	16.3% (7)	30.2% (13)
Poisonous Peruvian Plant Probe	N/A	25.0% (6)	45.8% (11)	29.2% (7)

Notes: Percent of students (number of students). Bacteria Sensing Probe, $N = 43$. Poisonous Peruvian Plant Probe, $N = 24$.

any of the three modules. Most commonly, students included molecular mechanisms for some, but not all, of the functional modules. While this approach includes mechanistic reasoning, it does not allow one to fully explain the biological phenomenon as a system. To illustrate how 'filling in' functional modules with molecular mechanisms can contribute to a complete explanation of the biological system, we provide an example of Evan's explanation. Here, he creates a productively continuous chain of entities and activities that spanned all three functional modules to give rise to a cohesive, generative mechanistic explanation for the bacteria-sensing phenomenon:

> I would imagine that the bacterium has receptors in its cell membrane that are made up of integral proteins. When the correct signaling molecule, either sugar or some kind of poison that will be toxic to the bacteria interact with those proteins, then it will trigger some kind of cell response. Then that would induce some kind of process that would signal the proteins working along the intercellular structures microtubules things. That would change the shape of the cell membrane and then activate the cilia and flagella in the outside of the bacteria causing it to move through the environment. Then they would just kind of move somewhat randomly until they find which direction the ingredient is of the sugars. So they move towards the most concentrated source until they get there, then they can use it. Or vice versa for the toxins.

Complete mechanistic explanations like Evan's were rare, occurring only among those coded as including a molecular mechanism for all three functional modules.

Because we thought that the ability to construct cohesive molecular mechanism for this probe might develop during an undergraduate major, we compared introductory and upper division groups. Mechanistic codes of functional modules for the Bacteria Sensing Probe were assigned a point value (1 = 'Not Mentioned', 2 = 'Identified', 3 = 'Non-Molecular Mechanism', 4 = 'Isolated Molecular Entity' or 5 = 'Molecular Mechanism'). We summed these values for the 3 modules ('Sensing', 'Signalling' and 'Motor') to give a total score between 3 and 15, with a score of 15 indicating that the student created a molecular mechanism for all 3 modules. A two-sample, two-sided, *t*-test (pooled variances) was performed on the total scores. A small estimated mean difference between groups was observed (introductory students mean = 11.59, SE = 0.90; upper division students mean = 11.0, SE = 0.73; mean difference = 0.59, SE = 1.16), but this difference was not significant ($p = .62$). Because the data were not normally distributed, the result was verified with a two-sided Wilcoxon Rank Sum Test, which confirmed non-significance ($p = .67$).

Similarly, for the Poisonous Peruvian Plant Probe, a minority of students created mechanistic explanations for both functional modules underlying the phenomenon (Table 3). Again, no statistically significant difference was found between introductory and upper division student groups for this measure. Specifically, mechanistic codes of functional modules for the Poisonous Peruvian Plant Probe were also assigned to one of two categories for both introductory and upper division groups. The two categories indicated whether students created a 'mechanism' for both modules ('cell/tissue differentiation' and 'localised poison generation') or whether they did not create a mechanism for both modules. Fishers Exact Test (which is more appropriate than the chi-square test for the small sample size) revealed no significant difference between rates at which introductory and upper division students created mechanistic explanations for the module ($p = .67$).

With our relatively small sample size and high variability of codes between students, it is difficult to say whether or not there could be differences between introductory and upper division students with regard to their propensity to construct generative mechanistic explanations. However, our results certainly suggest that these forms of reasoning are not limited to upper division students. We find it encouraging that even students with relatively little undergraduate biology experience, likely one or two courses, are able to engage productively in aspects of this scientific practice.

Part 2: students use flexible, multi-level reasoning when hypothesising the impact of mutations

In MCB, experts use multi-level reasoning to carry ideas across biophysical and ontological levels, often considering the impact of molecular events on an organism or a population. This type of reasoning is often employed when considering the impact of genetic changes. To generate and test hypotheses for the mechanisms underlying target phenomena, scientists use mutagenesis as a tool to implicate key entities by comparing wild type (normal) phenotypes with altered mutant phenotypes. Furthermore, genetic changes are the underlying feature of evolution and biodiversity of organisms. Therefore, the ability of biology students to learn to apply generative mechanistic reasoning to contexts of genetic change is paramount. Both interview probes were designed to give students an opportunity to consider not only the mechanisms underlying biological phenomena, but to consider scenarios in which they should hypothesise the impact of mutations on a biological system. Both questions call for students to use multi-level thinking and to think flexibly across levels.

Bacteria Sensing Probe

After students were asked to hypothesise a mechanism for the chemotaxis phenomenon, they were asked to consider a situation in which the bacteria are no longer able to sense substances in their environment and move accordingly. Specifically, they were asked to describe the meaning of 'mutation' in this context and to hypothesise how a mutation could cause a bacterium to lose this ability (Figure 1). Table 4 shows the results of coding students' responses to this question (see Methods for coding scheme description). The most common response included a student hypothesis that a single specific protein would be altered by the mutation, leading to the change in phenotype. In almost every case, these students focused their hypothesis on the membrane receptor that would be responsible for sensing the food or poison molecules in the environment. For example, Crystal provides the following hypothesis:

> So, mutation is any change in DNA that causes different functions or different functional proteins. And in this case its either related to the, I don't know, it says specifically for sensing and responding to the substances in their surroundings, I say that there is a mutation that, a mutation in the receptors or anything that like respond to the surroundings.

Table 4. Student responses to bacterial mutation question.

	Single specific protein	Multiple possible proteins	Protein unspecified	Other
Percent of students (number out of 43)	60.5% (26)	16.3% (7)	13.6% (6)	9.3% (4)

A subset of students (16%) hypothesised multiple possible proteins that might be affected by the mutation. Several students (14%) acknowledged that the mutation would somehow alter a protein that was involved in the sensing phenomenon, but did not present a hypothesis for what that protein might be. Thus, over 90% of students appeared to make the connection that 'mutation' would interfere with chemotaxis by altering a protein molecule. The remaining students provided hypotheses that were not clearly linked to a protein or were unwilling to present a hypothesis. This suggests that most students in this context are able to flexibly connect the idea of 'gene' to a cellular phenomenon. Next, we looked for evidence of how students were making these connections.

Further analysis of transcripts revealed several ways in which students were making mechanistic links between the genetic (or molecular) idea of 'mutation' and the cellular phenomenon of 'chemotaxis'. Examples below highlight the major modes of explanation used by students whose responses were coded as 'single specific protein' or 'multiple possible proteins'. The diagrams in Figure 8, allow for side-by-side comparison of these modes of mechanistic explanation. One of the more common ways that students explained the impact of the mutation was a somewhat general idea that the receptor protein was somehow altered or 'broken' due to a mutation and that this prevented the ability of the cell to sense (Figure 8, Explanation Form A). Another common form included the specific idea that the receptor protein was missing due to a mutation that either deleted the gene or prevented the fully formed protein from reaching the plasma membrane

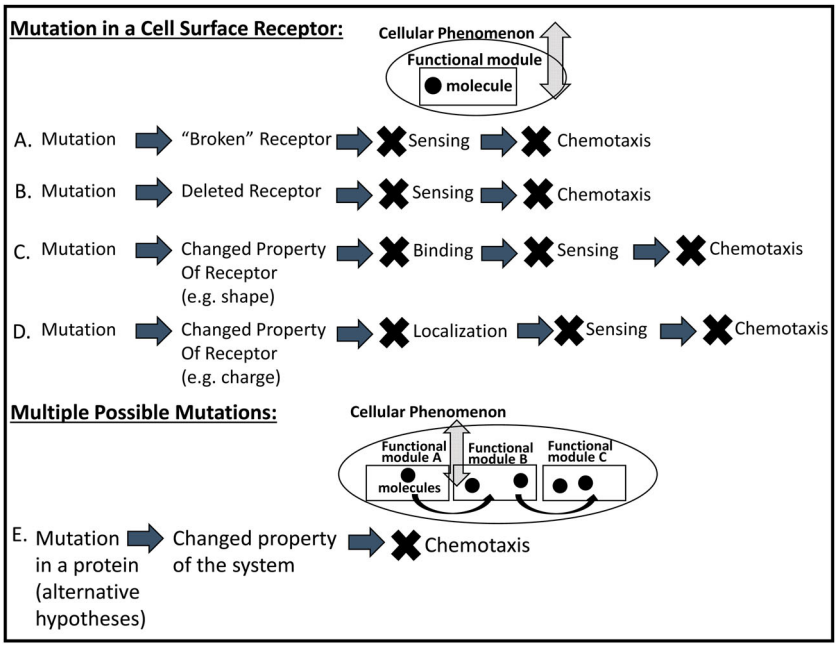

Figure 8. Diagram illustrating the varied forms of student explanations for the effect of a mutation in the Bacteria Sensing Probe. Several forms (A–D) involved students' multi-levelled, mechanistic hypotheses for how a mutation could lead to an alteration in a cell surface receptor, which would impact the 'sensing' functional module and ultimately the cell's ability to perform chemotaxis towards food. Another form of explanation (E) involved students' recognition of the potential impact of mutations at multiple points in the multi-levelled biological system.

(Figure 8, Explanation Form B). Sally begins by considering a 'broken' protein, but then alternatively considers the deletion of a functional protein due to a mutation and its impact on the bacteria's ability to sense.

> Alright so, if there's like a particular gene that encodes for this protein that senses this outside substance and for some reason this gene has the wrong nucleotide somewhere [MUTATION] and it either changes part of the sequence so you get a messed up protein ['BROKEN' PROTEIN], or if it like stops it, premature stop codon, and this protein just doesn't exist at all [DELETED PROTEIN], then if that's not there then the substance will be floating around and the cell won't be able to sense that it's there because it has … it doesn't have the correct protein to do it [LOSS OF SENSING].

Two additional forms of explanation focused on how the mutation would alter a specific property of the receptor protein and how this could affect the ability of the protein to either bind the substance being sensed or to reach the proper subcellular localisation (Figure 8, Explanation Forms C and D, respectively). The following examples from Michelle and Steven illustrate these forms of explanation:

> Michelle: 'Well, as so, if like, the sugar or whatever is binding a ligand maybe like the ligand is shaped wrong. Or what … the receptor is shaped wrong so it will no longer combine to it. Like if it is shaped like this (refers to her drawing) and then the receptor is a triangle then that wouldn't recognize it, it needs a circle. This doesn't go in there.'
>
> Steven: 'Maybe it no longer has the hydrophobic groups that it needs to in order to be planted in the cell wall.'

In each of these Explanation Forms (A–D), students seem to use mechanistic connections to causally link a molecular entity (often identified as the receptor) with the overall phenomenon (inability to perform chemotaxis in response to food) by considering the role of that entity in its functional module ('sensing') and how that is connected to behaviour at the cellular level. In this way, students are creating mechanistic connections between hypothesised molecular entities and cellular activities, creating a mechanistic explanation that spans the molecular and cellular levels through functional modules.

Finally, we examine how students reasoned within those explanations that were coded as 'Multiple Possible Proteins' (14% of students, see Figure 8, Explanation Form E). Carlos demonstrates this form of explanation:

> So, in the factors that I talked about earlier, perhaps there is a decrease in viability of the ligand to the receptor. Maybe the receptor is missing or broken. There could be a number of cascade factors that are missing or one is constitutively active or inhibited that it typically needs to be the opposite. Maybe the flagella could just … is continuously moving in one way or another. It didn't really say. So that could be broken as well. Maybe the cilia if there is a mechanical cascade of function that it can sense what's going on around it, maybe that's broken. So yah those are a number of hypotheses.

In these explanations, students seem to be taking a more expert-like view of the system: holding in mind the system as a whole, flexibly moving across entities and functional modules to propose mechanistic links between the mutation and the cellular phenomenon. In fact, students coded in this category made a diverse range of sophisticated hypotheses, including the following ideas for possible mutations: inability of receptor to

signal, inability of second messengers to transmit signals, decreased binding of ligand to receptor, missing receptor, missing or misregulated signalling cascade factors, misregulation of flagella or cilia, change in gene expression resulting from signalling and a change in protein phosphorylation.

Poisonous Peruvian Plant Probe

In the previous section, we discussed students' responses when asked to hypothesise *how* a plant from the rainforests of Peru developed both poisonous and non-poisonous parts growing from a single seed (Figure 2, Question 1). Here, we focus on their responses when asked to consider a similar plant in the sparse mountain ranges of Peru that does *not* demonstrate this poisonous feature (Figure 2, Question 2). In this second situation, students are prompted to consider *why* this similar plant species is not poisonous and to consider how this phenomenon came to be over time. Therefore, in this context, students are primed to consider the impact of possible mutations on poison production, the possible link between genes and the environment, and the potential for predation to provide selective pressure for and against traits. However, unlike the Bacteria Sensing Probe, the word 'mutation' is not explicitly mentioned in the questions. Compared to the Bacteria Sensing Probe, the Poisonous Peruvian Plant Probe is much more complex, in terms of spanning both biophysical levels and causal organisation of the system.

We applied a coding scheme to demonstrate whether students' explanations included a mechanistic connection between 'genes' and the biological system (see Methods). The majority of students (85.7%) hypothesised a specific connection between genetic change (at the molecular level) and a complex biological phenomenon (diversity of species at the population level), while a minority (14.3%) did not. Explanations with this gene-to-system connection either focused on evolution, gene expression or contained both of these features (Table 5). Next, we investigated how students made these connections.

Further analysis of transcripts revealed several different generative, mechanistic explanations that students hypothesised to account for difference between the two Peruvian plants. Schematics of these explanation forms are presented in Figure 9. Three forms of explanation proposed an evolutionary mechanism to connect genes to the observed phenomenon (Explanation Forms A, B and C). Most commonly, students noticed the differences in the environment of the two plant species and hypothesised that a spontaneous mutation in the plant genome leads to a change in protein production. In most explanations, this mutation led to the formation of a poison that was selected for in the rain forest environment (Figure 9, Explanation Form A). Jane provides an example of this form of explanation:

Table 5. Student responses to Peruvian plant mutation question: types of mechanistic connection between genes and the system.

	Evolution connection	Gene expression connection	Both evolution and gene expression connections	No connection between genes and the system
Percent of students (number out of 21)	57.1% (12)	14.3% (3)	14.3%(3)	14.3% (3)

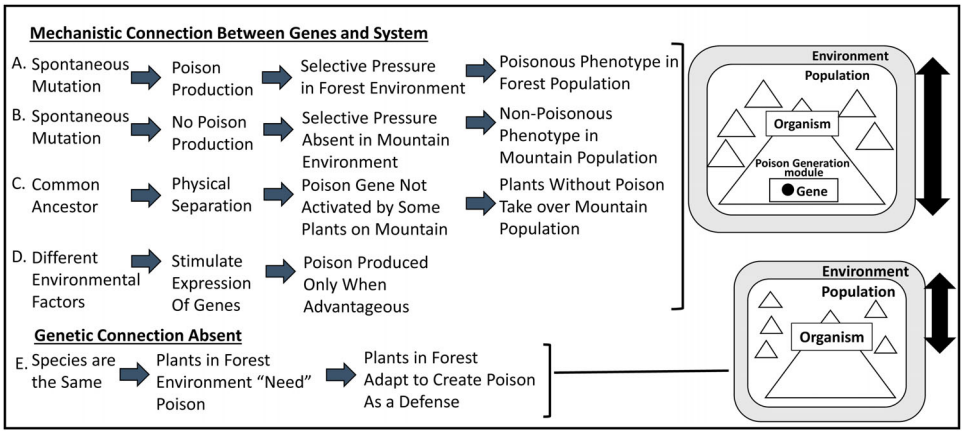

Figure 9. Diagram illustrating the varied forms of student explanation for phenotypic difference between two plants in the Poisonous Peruvian Plant Probe. Several forms of explanation (A–D) included mechanistic connections between genes and the biological system. This system spanned multiple levels including the molecular level (genes and proteins), the functional module level (the set of molecules that would generate poison), the organismal level (the plant) and the population level (the set of plants in the rainforest or mountain environment). Some forms (A–C) mechanistically connected genes to the biological system by including evolutionary ideas such as change through spontaneous mutation, selective pressure and common ancestors. One form relied on ideas of differential gene expression instead of evolution (D). One form of explanation (E) included evolutionary ideas, but did not link these to ideas of genetics at the molecular level.

Jane: 'Well in the mountain range the foraging animal population is smaller and it is probably not the case where like in the rain forest the similar plant had the need to protect itself. So, there probably was not as much selection for poisonous plants in the mountain range. Because the foraging animal population didn't drive that selection. Like if you have tuberculosis, I know that when they started using antibiotics on them the use of antibiotics selected for antibiotic resistance strains. And that would be the same case in the rainforest, if you have this animal population that is constantly targeting the leaves and fruit of the plant, then there is going to be selection for a plant that could resist those attacks from the animals, [SELECTIVE PRESSURE IN FOREST ENVIRONMENT] but in the mountain range, that probably is not the case because there is just not as many animals targeting it. So, it doesn't favor selection for those types of individuals. [LACK OF SELECTIVE PRESSURE IN MOUNTAIN ENVIRONMENT].'

Interviewer: 'Yeah, can you think of any molecular process, or specific molecules that might be different in the two plants?'

Student: 'Well it is probably that there are individuals in the mountain range that have a gene that codes for the poison [POISON PRODUCTION]. And if the poison is a mutation that the mountain range plant are the original plants and they did not have the mutation beforehand, then it is possible that in the rain forest, the poison was just some random occurrence [SPONTANEOUS MUTATION] that all of the sudden we are able to make a protein that happens to be poisonous to the animals [POISON PRODUCTION].'

In this example, Jane first focuses on the impact of selective pressure on the plant population. After being prompted to include her ideas about the molecular level, she adds that

the poison is likely produced by a 'gene that codes for the poison' and that the 'poison was just some random occurrence' that resulted from a 'mutation'. Therefore, students may enter this causal chain of events at different points, likely due the complex, non-linearity of the biological system.

In other explanations, this mutation resulted in the *inability* of the plant to make poison. This genetic change would not be detrimental in the mountain environment, and might even be beneficial if it reduced the requirement of the plant to expend energy, producing an unnecessary protein (Figure 9, Explanation Form B). Sally provided an explanation of this type:

Sally:	'That could be a difference in the DNA sequence that arose at some point in evolution whenever they had a common ancestor [SPONTANEOUS MUTATION]. And it knocked out the poisonous DNA … yeah so then the cells have the poisonous protein because it just doesn't exist in the cells because the DNA was mutated at some point. And it doesn't function [NO POISON PRODUCTION].'
Interviewer:	'Do you want to explain why that probably would've happened?'
Sally:	'Um … so, if there's a smaller foraging animal population there's less selection for a plant that can avoid being eaten by foraging animals [SELECTIVE PRESSURE ABSENT IN MOUNTAIN ENVIRONMENT]. So if there's nothing there to select for that – if the plant lost its ability to be poisonous, it's not necessarily going to die. So it'll just pass those genes on instead of being eaten and not being able to pass those genes on and continuing the poisonous plant population [NON-POISONOUS PHENOTYPE IN MOUNTAIN POPULATION].'

A similar form of explanation included evolution as the basis for the phenomenon, but focused on physical separation of plants with a common ancestor leading to change in the localised population over time (Figure 9, Explanation Form C). For example, Karla provides the following explanation:

> How that might work in the plant … Well, they probably had a common ancestor at some point [COMMON ANCESTOR]. And then, however they came to different regions … [PHYSICAL SEPARATION] Whatever the expression gene for the poisonous part was just not activated [POISON GENE NOT ACTIVATED] and then continually reproduced, so that was the prevalent form [PLANTS WITHOUT POISON TAKE OVER].

In each of these forms of explanation, students causally attributed differences between species to a genetic change in the population, through a spontaneous mutation or a separation of species to new environments.

Finally, a subset of students causally attributed the differences between species to differences in gene expression through signals in the local environment (Figure 9, Explanation Form D). This form of explanation connected the impact of 'genes' to the biological system, but did not include evolutionary ideas. Importantly, the experts we interviewed suggested that the difference in plants could have either an evolutionary explanation or could be explained by a difference in local gene expression due to environmental conditions. Finally, a small number of students provided explanations that used the term 'evolution' but included no evidence of a mechanism requiring genetic change. In these cases, the primary cause of change seemed to be a 'need' for the plant to adapt to

environmental pressure (Figure 9, Explanation Form E). Richard's response is an example of this type of explanation.

Richard:	Well I have to say the exact same thing in terms of evolution. Um, where in this region, because the population is much more dense or whatever, then those plants develop these defenses and develop these poisonous branches, but on the mountain range it didn't do so because it didn't need it. [PLANTS IN FOREST ENVIRONMENT NEED POISON].
Interviewer:	It didn't need it?
Richard:	Yah.
Interviewer:	That makes sense.
Richard:	But I think a process takes like millions of years.
Interviewer:	Yah. How do you think that process happens, where one would have a …
Richard:	Well let's say a plant keeps on getting its branches chopped off once a week from mammals or whatever, eventually, I mean with time, years and years and decades, it develops a defense, a defense mechanism where it develops those poisonous branches. But if the plant in the mountain ranges are not getting chopped off all the time, why would it even think about … [PLANTS IN FOREST ADAPT TO CREATE POISON AS A DEFENSE].

Overall, analysis revealed several generative mechanisms that students were able to use to hypothesise the involvement of genetic changes in a complex, multi-levelled system. In many cases, students' ideas of genetic change were tightly linked to a molecular view of the functioning of an individual organism. In some cases, this molecular view even expended to students' apparent understanding of populations of organisms and their interactions with the environment. The most common explanations seemed to be rooted in the idea that a spontaneous mutation arose that caused the creation of a new poison, which from a conceptual perspective may be the most causally direct. However, it is important to note that the same probe produced a diversity of student hypotheses, including loss of poison through mutation and changes in gene expression through interaction with the environment. In addition, the probe was useful for revealing cases in which students may not have developed a genetics-driven view of evolutionary changes within a system.

Discussion

Generative reasoning, by definition, involves hypothesising plausible ideas, not simply reciting memorised mechanisms (Clement, 2013; Duncan, 2007). In this study, we asked whether undergraduate students would use generative mechanistic reasoning to construct explanations for novel phenomena (Research Question 1). We found that most students did propose biologically plausible mechanisms or partial mechanisms to explain these phenomena. This suggests that asking undergraduate students to make hypotheses about novel phenomena allows them to engage productively in the scientific practice of mechanistic explanation construction. Encouragingly, this was true even for contexts that were novel to students but still related to the material taught in undergraduate MCB courses, which most often focus on complex multi-level phenomena and detailed molecular mechanisms.

We also investigated the strategies that seemed to guide students' explanation creation (Research Question 2). We found that students often used the strategy of functionally

subdividing phenomena to aid in explaining those phenomena, similar to the strategy biologists use to explain phenomena in a research setting. In fact, when students used the functional subdivision strategy, they identified the same modules for chemotaxis and plant development as MCB experts. The most challenging aspect of constructing a mechanistic explanation appeared to be hypothesising the entities and activities that could belong in each functional module, i.e. 'filling in' the modules with mechanisms. Students often hypothesised a mechanism for only some of the modules they proposed or only hypothesised a single entity to fill a module. This 'partial' mechanistic reasoning is not surprising and seems to be common when students, at any level, are asked to construct explanations for phenomena that are novel to them (Bolger, et al., 2012). However, it is interesting to note that a subset of students did hypothesise a coherent mechanistic explanation for the entire biological system represented in each probe, despite having no experimental evidence of what entities or activities might be involved. When we examined how students were 'filling in' modules with mechanism, we found that they often used strategies such as: (1) hypothesising the involvement of entities by contemplating the necessary activities in the biological system and (2) instantiating mechanistic schema, i.e. using details of the novel setting to propagate an abstract schema. Again, these reasoning strategies are similar to those used by experts in a research setting, as described in our theoretical framework. By demonstrating the ways in which students can use these strategies to build explanations for multi-levelled biological phenomena, our findings provide empirical support for the importance of educational ideas theorised by Van Mil et al. (2013). Specifically, our work lends credence to their suggestions that instructors should pose 'how questions about cellular activities' and that instructors might ask students to explore functional modules as a 'stepping stone' to aid in multi-level reasoning from molecules to cells.

Finally, we asked how students reason about mutations within a multi-levelled phenomenon (Research Question 3). In the context of the bacterial sensing probe, we found that almost all students made a connection between 'mutations' and the concept of changing a protein molecule. In many cases, their hypotheses focused on a single mechanistic entity, a receptor, within a single functional module, 'sensing', to explain the change in behaviour at the cellular level. This approach seems to be a relatively accessible way that many students can utilise the complex idea of 'genetic change gives rise to altered phenotype' to create a basic mechanistic explanation that spans the molecular and cellular levels. Other students approached the probe by hypothesising that a mutation in multiple places within the system could influence bacterial chemotaxis. This more holistic approach seemed to indicate that these students were able to mentally manipulate their entire mechanistic explanation, spanning all levels of organisation. The flexibility of thinking in this second approach is reminiscent of the animation of mental models that has been described for expert scientists (Nersessian, 2008).

In the context of the Poisonous Peruvian Plant Probe, we found that the majority of students were able to hypothesise a specific connection between genetic change (at the molecular level) and a complex biological phenomenon (diversity of species at the population level). This was somewhat surprising given the number of physical and organisational levels that are involved in the phenomenon, as well as previous reports of student difficulty with this type of micro- to macro-level reasoning (Marbach-Ad & Stavy, 2000; Wilensky & Resnick, 1999). Similar to what we frequently observed in the bacterial sensing probe, students often created a relatively simple mechanistic link between

mutation and species diversity. Specifically, they hypothesised that a spontaneous mutation led to the production of a protein needed for poison creation and that mutation was positively selected in the rainforest environment. This form of explanation was most common, but we observed students drawing mechanistic links between mutation and species diversity in several other ways, including negative selection, physical separation of genetically diverse organisms and regulated gene expression. However, a small number of students did not make these same mechanistic links between levels. We described these students as focusing on evolution without a molecular genetic component. The forms of explanation we uncovered are similar to those previously described as 'Lamarckian' explanations for evolution (Ferrari & Chi, 1998). In both cases, students seemed to imply that traits emerged for a reason, rather than through selection. However, in our study, students appeared to understand that these changes take place over long periods of time, which is different from previous descriptions of student ideas.

Students' generative mechanistic reasoning

Other studies have considered student reasoning in MCB contexts, as well as student mechanistic reasoning more broadly. However, to our knowledge, the current work represents the first investigation of undergraduate student explanation building using a framework based on the reasoning practices biologists use to explore multi-levelled mechanisms. While our results contribute new empirical evidence of students using these scientific reasoning practices, we found many ways in which our results relate to and often support previous work on student explanation construction. Here, we consider how our results relate to several key studies.

Problematic ideas in MCB

Several studies have explored empirically or hypothesised about problematic ideas and novice reasoning patterns in molecular biology and genetics at the K-12 level and explored transitions from naïve everyday ideas to basic scientific principles (Duncan & Reiser, 2007; Duncan & Tseng, 2011; Marbach-Ad & Stavy, 2000; Southerland, Abrams, Cummins, & Anzelmo, 2001; Venville & Treagust, 1998). We observe some similar patterns of conceptual difficulty described by these studies in our undergraduate student population. For example, younger students (8–10th grade) often draw on macro-level mechanisms or anthropomorphised ideas when considering cellular activities or genetic events (Duncan & Reiser, 2007; Duncan & Tseng, 2011). When given the Bacteria Sensing Probe, eighth and ninth grade students created explanations about the bacterium's ability to 'instinctively sense' stimuli or suggested 'human-like senses and mental capabilities' to give rise to the observed cellular response (Duncan & Tseng, 2011). When students in our population did not create full or partial mechanistic explanations, a few students fell back on similar non-molecular and human-like sensing mechanisms such as 'smelling', intuitive 'feeling' and nerve-like intracellular signalling. However, our students did not appear to struggle in the same ways as K-12 students when identifying proteins as central to genetic phenomena (Duncan & Reiser, 2007; Duncan & Tseng, 2011) and understanding genes as carrying instructions for protein structure (Duncan & Reiser, 2007; Duncan & Tseng, 2011; Venville & Treagust, 1998).

The role of domain-specific knowledge in generative reasoning in MCB

In her study of student reasoning in molecular genetics, Duncan (2007) investigated the role of domain-general and domain-specific knowledge in students' generative reasoning. However, where Duncan's study utilised a conceptual change framework (diSessa, 1988; Greeno, 1983; Penner & Klahr, 1996), our study focused on students' generative reasoning using a framework derived from investigations of scientists' reasoning in the field (Craver, 2001, 2002a; Darden, 2002; Machamer et al., 2000; Van Mil et al., 2013). As a result, Duncan's work proposed a useful model for student generative reasoning, which involved interactions between domain-general solution frames and domain-specific heuristics and schemas. Our work probes more deeply into the domain-specific aspects of students' generative reasoning.

By utilising a framework that takes into account strategies for multi-level reasoning, we were able to add information about how students utilise functional subdivision to guide their reasoning about MCB phenomena, and observe specific strategies that students employ to build mechanistic elements into their explanations. Additionally, inclusion of the Poisonous Peruvian Plant Probe allowed us to expand beyond mechanisms for phenomena that affect single organisms, and allowed us to explore student reasoning across many biophysical levels (from molecular to organismal to population levels in the context of evolution). Interestingly, despite different theoretical frames, the ideas of schema instantiation emerged in both analyses. Duncan (2007) utilised the idea of schema from the cognitive sciences literature (Rumelhart, 1980); we utilised a similar idea from the philosophy of science literature (Darden, 2002; Kitcher, 1989; Van Mil et al., 2013). This finding points to the potential importance of schema instantiation as a robust reasoning mechanism for generative reasoning in MCB. Finally, Duncan (2007) identified several specific domain-specific schema that were prevalent in our investigation of students' explanations, for example, 'regulation of gene expression'. However, our study dove more deeply into how students use specific biological entities, i.e. 'propositional knowledge', to instantiate a schema. As our framework incorporated mechanistic reasoning (Machamer et al., 2000; Van Mil et al., 2013), we were able to further explore how students hypothesised causal relationships between particular entities and their activities within a biological system. We believe that our empirical analysis of students' generative explanation building through this different theoretical lens has the potential to augment Duncan's pivotal work in this area (Duncan, 2007; Duncan & Tseng, 2011).

Students' use of mechanistic reasoning

A broader interest in science education is providing students, at all levels of education, opportunities to practise reasoning mechanistically about phenomena they observe or learn about in class. Several studies have noted that students in various contexts, and at various levels of education, may construct explanations that do or do not address 'how things work' in a mechanistic way (Abrams & Southerland, 2001; Lehrer & Schauble, 1998; Metz, 1985). The current work joins a small set of investigations that utilise specific ideas about mechanistic reasoning among experts (Machamer et al., 2000) to understand how students construct explanations for novel phenomena (Bolger, et al., 2012; Russ et al., 2008). Similar to our previous work on young children's reasoning about simple machines (Bolger, et al., 2012), in the current study, we found that students commonly provided explanations that included pieces of mechanistic reasoning, but did not give a complete

mechanistic account to explain the phenomena. Previously, we referred to this similar phenomenon as 'elements of mechanistic reasoning' (Bolger, et al., 2012). In that context, we discovered that use of multiple elements of mechanistic reasoning was positively correlated with the ability of students to predict correctly the movement of simple machines. We found that use of multiple mechanistic elements was rare, but demonstrated how careful orchestration of multiple elements could lead to a complete mechanistic explanation. Similarly, in this study, we found that a small subset of students provided complete mechanistic explanations for a biological system. A limitation of both studies is that they were performed in an interview setting and asked students to explain phenomena with which they had little or no experience. Therefore, it is not possible to know the extent to which the 'neutral' context may have contributed to the forms of explanation students provided or whether or not partial mechanistic explanations could have developed into full mechanistic explanations with the social feedback or learning that might occur in a classroom setting.

By contrast, Russ et al. (2008) developed and piloted a framework to explore levels of mechanistic explanation in a classroom setting. They found that a group of young students in a physics classroom engaged in discourse that moved in and out of different levels of mechanistic explanation, dependent on the teacher moves and student interactions at different points in the lesson. Similar to this study, Russ's study explicitly demonstrated how a framework to describe student mechanistic reasoning could be derived from research on reasoning among scientists in the field. However, our framework takes into account additional aspects of mechanistic reasoning that are particularly important for explaining multi-levelled phenomena in MCB, for example, functional subdivision.

Implications for education

Barriers to learning that are related to domain complexity must be answered with efforts to support learning processes that encourage greater cognitive flexibility (Spiro, Feltovich, Jacobson, & Coulson, 1992). The results of this study support the use of explanatory tasks to encourage generative mechanistic reasoning and cognitive flexibility among undergraduate biology students. Our results suggest that the types of reasoning used by scientists to solve research problems can be used by students, even in a non-laboratory context. We see the potential for explanatory tasks, like the ones posed here, to augment course-based undergraduate research experiences (Auchincloss et al., 2014; Brownell et al., 2015) or laboratory research opportunities.

In order to serve as a useful educational tool in classrooms with students who have diverse conceptual understandings and educational backgrounds, tasks should be accessible to students with less expertise and yet pose little limitation on the generative reasoning of those with greater expertise. The explanatory tasks we used were accessible to students at different educational levels and different levels of content knowledge. Every introductory and upper division student in our study, with one exception, was able to hypothesise explanations for the phenomena we posed. In most cases, their explanations were plausible. These same tasks enabled a subset of students to provide detailed mechanistic explanations that included multiple plausible ideas.

While we did find evidence of some conceptual difficulties previously reported for these topics among secondary students, these difficulties were not common. Most students had a

relatively firm grasp of the role that genes and proteins played at the molecular level, and were able to connect these ideas to the cellular, organism and population levels. Similarly, the ability to reason about the role of protein molecules within biological phenomena was seen amongst undergraduate students in Duncan's (2007) study. A gain in this ability was also seen after instructional intervention amongst secondary students in studies by Van Mil et al. (2016) and Duncan and Tseng (2011), but not after typical genetics instruction (Duncan & Reiser, 2007). Together, these studies suggest that when instructors provide students with adequate information about molecular mechanisms, they can begin to use a molecular, mechanistic frame with which to view biological phenomena. Our findings suggest that this frame allows students to transfer their knowledge in order to generatively construct molecular mechanistic explanations for 'novel' phenomena.

Our results demonstrate that, despite having a grasp of these basic ideas, students formulated explanations with a wide range of form and complexity. However, we did not find any statistically significant differences between the explanations of students at the introductory versus upper division levels. This finding is consistent with previous work (Southard, et al., 2016), suggesting that significant shifts in these deeper disciplinary skills may not be detectable at the population level during the several semesters that undergraduate students spend taking MCB coursework. However, we cannot rule out that our limited sample size masked a potential difference between these groups. Furthermore, because we relied on students to volunteer for our study, it is possible that the sample we obtained was somehow skewed to mask potential differences between groups. Despite these limitations, our results do demonstrate the existence of explanations at all ranges of complexity among students at both the introductory and upper division levels.

When thinking about the challenge of asking students to generate mechanistic explanations for novel biological systems in the classroom, one might be concerned that only a minority of students in our study provided complete mechanistic accounts for the entire system. First, we suggest that this is not likely due to a *lack* of mechanistic reasoning or a problem with 'systems thinking' in general. Rather, a student's level of experience with related mechanistic entities and biological systems is likely to influence their response. Duncan (2007) posed a similar argument, suggesting that 'domain-specific dynamics are critical to reasoning about these systems'. Similarly, Van Mil et al. (2016) argue that mechanistic reasoning is intuitive, but 'domain-specific knowledge and expertise' are needed to hypothesise mechanisms. In addition, we suggest that students' comfort level with this type of flexible reasoning could be influenced by how often he or she has had the opportunity to generate explanations in this way. Second, we suggest that undergraduate biology instructors should provide students with opportunities to practise explanation construction in novel contexts. The students in this study had experienced classroom instruction that included biological concepts, but did not focus on scientific practices. Therefore, the explanations they provided as individuals in an interview setting should be considered only the beginning of what they might produce in a collaborative learning environment, which included this scientific practice.

Finally, our research points to the need for further development of such explanatory tasks in undergraduate MCB classrooms. In order to foster greater cognitive flexibility among our students, we must first work to understand how skills for explanation construction develop. While our study contributes to this current understanding, further studies are needed to fully characterise how these kinds of reasoning may be influenced

by specific biological contexts and how students' explanations may change in a classroom context. In particular, it may be important to see whether the diversity of explanations constructed by students could be a productive starting point for expanding generative reasoning through dialogue at the classroom level.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abrams, E., & Southerland, S. (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.

Adler, J. (1966). Chemotaxis in bacteria. *Science*, *153*(3737), 708–716.

Adler, J. (1975). Chemotaxis in bacteria. *Annual Review of Biochemistry*, *44*(1), 341–356.

American Association for the Advancement of Science (AAAS). (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC: Author.

Auchincloss, L. C., Laursen, S. L., Branchaw, J. L., Eagan, K., Graham, M., Hanauer, D. I., & Dolan, E. L. (2014). Assessment of course-based undergraduate research experiences: A meeting report. *CBE Life Sciences Education*, *13*(1), 29–40.

Bock, W. (2017). Dual causality and the autonomy of biology. *Acta Biotheoretica*. doi:10.1007/s10441-016-9303-2

Bolger, M. S., Kobiela, M., Weinberg, P., & Lehrer, R. (2012). Children's mechanistic reasoning. *Cognition and Instruction*, *30*(2), 170–206.

Brownell, S. E., Hekmat-Scafe, D. S., Singla, V., Seawell, P. C., Imam, J. F. C., Eddy, S. L., & Cyert, M. S. (2015). A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data. *CBE-Life Sciences Education*, *14*(2), 1–14.

Clement, J. (2006). Thought experiments and imagery in expert protocols. In L. Magnani (Ed.), *Model-based reasoning in science and engineering* (pp. 1–16). London: College Publications.

Clement, J. J. (2013). Roles for explanatory models and analogies in conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 412–446). New York: Routledge.

Craver, C. F. (2001). Role functions, mechanisms, and hierarchy. *Philosophy of Science*, *68*(1), 53–74.

Craver, C. F. (2002a). Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science*, *69*(S3), S83–S97.

Craver, C. F. (2002b). Structures of scientific theories. In P. K. Machamer & M. Siberstein (Eds.), *The Blackwell guide to the philosophy of science* (pp. 55–79). Oxford: Blackwell.

Darden, L. (2002). Strategies for discovering mechanisms: Schema instantiation, modular subassembly, forward/backward chaining. *Philosophy of Science*, *69*(S3), S354–S365.

diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. B. Pufall (Eds.), *Constructivism in the computer age* (pp. 49–70). Hillsdale: Lawrence.

Duncan, R. G. (2007). The role of domain-specific knowledge in generative reasoning about complicated multileveled phenomena. *Cognition and Instruction*, *25*(4), 271–336.

Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. *Journal of Research in Science Teaching*, *44*(7), 938–959.

Duncan, R. G., & Tseng, K. A. (2011). Designing project-based instruction to foster generative and mechanistic understandings in genetics. *Science Education*, *95*(1), 21–56.

Dupre, J. (2009). It is not possible to reduce biological explanations to explanations in chemistry and/or physics. In F. J. Ayala & R. Arp (Eds.), *Contemporary debates in philosophy of biology* (pp. 32–47). Oxford: Wiley-Blackwell.

Ferrari and Chi. (1998). The nature of explanations of natural selection. *International Journal of Science Education*, *20*(10), 1231–1256.
Greeno, J. G. (1983). Conceptual entities. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 227–252). Mahwah, NJ: Lawrence Erlbaum Press.
Hartwell, L. H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999). From molecular to modular cell biology. *Nature*, *402*, C47–C52.
Kaiser, M. I. (2015). A closer look at biological explanations. In C. T. Wolfe, P. Huneman, & T. A. C. Reydon (Eds.), *Reductive explanation in the biological sciences, history, philosophy and theory of the life sciences* (pp. 135–171). Switzerland: Springer International.
Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In Philip Kitcher & Wesley Salmon (Eds.), *Scientific explanation. Minnesota studies in the philosophy of Science* (Vol. 13, pp. 410–505). Minneapolis: University of Minnesota Press.
Lehrer, R., & Schauble, L. (1998). Reasoning about structure and function: Children's conceptions of gears. *Journal of Research in Science Teaching*, *35*(1), 3–25.
Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*(1), 1–25.
Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education*, *34*(4), 200–205.
Matthews, M. R. (2014). *International handbook of research in history, philosophy and science teaching*. Dordrecht: Springer.
Mayr, E. (1961). Cause and effect in biology: Kinds of causes, predictability, and teleology as viewed by a practicing biologist. *Science*, *134*, 1501–1506.
Metz, K. E. (1985). The development of children's problem solving in a gears task: A problem space perspective. *Cognitive Science*, *9*(4), 431–471.
Morange, M. (2008). *The death of molecular biology* (M. Cobb, Trans.). Cambridge, MA: Harvard University Press.
National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
Nersessian, N. J. (2008). *Creating scientific concepts*. Cambridge, MA: MIT Press.
Next Generation Science Standards (NGSS) Lead States. (2013). *Next generation science standards: For states, By states*. Washington, DC: National Academies Press.
Passmore, C., Stewart, J., & Cartier, J. (2009). Model-based inquiry and school science: Creating connections. *School Science and Mathematics*, *109*(7), 394–402.
Penner, D., & Klahr, D. (1996). The interaction of domain-specific knowledge and domain general discovery strategies: A study with sinking objects. *Child Development*, *67*, 2709–2727.
Perkins, D. (1993). Teaching for understanding. *The Professional Journal of the American Federation of Teachers*, *17*(3), 28–35.
Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Berwer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Lawrence Erlbaum.
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.
Southard, K., Wince, T., Meddleton, S., & Bolger, M. S. (2016). Features of knowledge building in biology: Understanding undergraduate students' ideas about molecular mechanisms. *CBE-Life Sciences Education*, *15*(1), 1–16.
Southerland, S. A., Abrams, E., Cummins, C. L., & Anzelmo, J. (2001). Understanding students' explanations of biological phenomena: Conceptual frameworks or p-prims? *Science Education*, *85*(4), 328–348.
Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in Ill-structured domains. In T. M. Duffy & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 57–76). Hillsdale, NJ: Routledge.

Stewart, J., Hafner, B., & Dale, M. (1990). Students' alternative view of meiosis. *The American Biology Teacher*, *52*(4), 228–232.

Van Mil, M. H., Boerwinkel, D. J., & Waarlo, A. J. (2013). Modelling molecular mechanisms: A framework of scientific reasoning to construct molecular-level explanations for cellular behaviour. *Science & Education*, *22*(1), 93–118.

Van Mil, M. H., Postma, P. A., Boerwinkel, D. J., Klaassen, K., & Waarlo, A. J. (2016). Molecular mechanistic reasoning: Toward bridging the gap between the molecular and cellular levels in life science education. *Science Education*, *100*, 517–585.

Venville, G. J., & Treagust, D. F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. *Journal of Research in Science Teaching*, *35*(9), 1031–1055.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3–19.