Supplemental Material

CBE—Life Sciences Education

Uminski *et al*.

Undergraduate biology lecture courses predominantly test facts about science rather than scientific practices

Crystal Uminski, Sara M. Burbach, Brian A. Couch

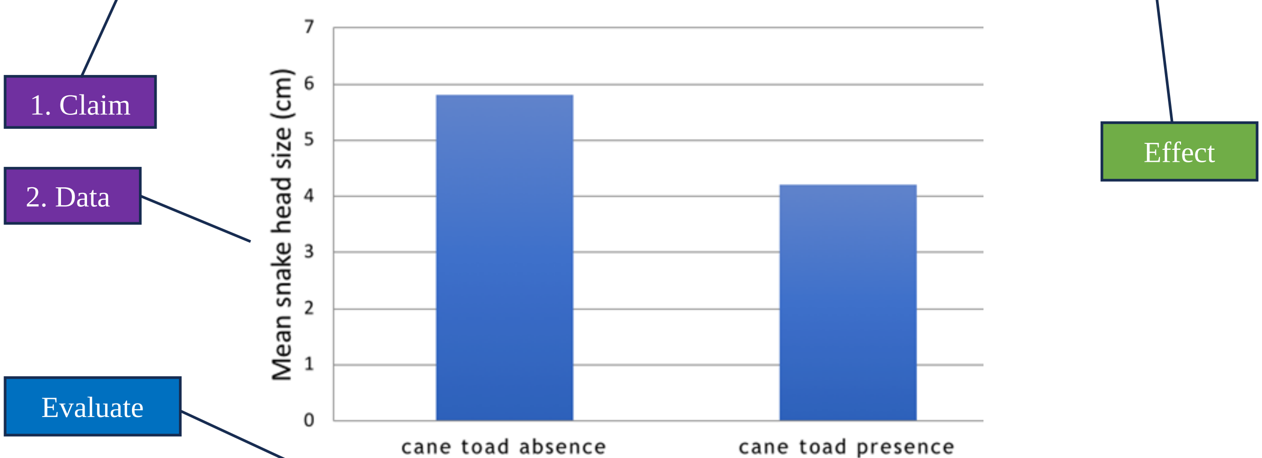
Supplemental Materials – Table of contents

Supplemental Figure 1.

Sample Item #1
Scientific Practice: Analyzing and Interpreting Data
Crosscutting Concept: Cause and Effect
Core Idea: Evolution
Bloom's Taxonomy Level: Evaluate

Cause

The invasive cane toad was introduced to Australia in the 1930s. The toads are extremely poisonous to red-bellied black snakes, and there was some concern that the snake population might be decimated. However, the snakes and toads seem to be co-existing. Imagine a biologist began investigating the situation by surveying red-bellied black snakes in areas with and without cane toads and found one major difference between the two populations. The results are depicted in the hypothetical graph below.



Which of the following scenarios is most likely given the data? — 3. Interpretation

A. Where cane toads are present, red-bellied black snakes evolved a tolerance to cane toad toxins that allowed them to survive and reproduce.
B. Where cane toads are present, red-bellied black snakes with large heads died while those with small heads survived and reproduced. — Mechanism; Evolution
C. The presence of cane toads caused mutations that resulted in smaller heads.
D. The red-bellied black snakes mutated in response to the cane toads so that they could tolerate cane toad toxins.

Applied coding protocol for Sample Item #1		
Code	**Criteria**	**Rationale**
Scientific Practice: Analyzing and Interpreting Data	1. Question gives a scientific question, claim, or a hypothesis to be investigated. 2. Question gives a representation of data (table, graph, list of observations, etc.) provided to answer the question or test the claim or hypothesis. 3. Question asks student to select an interpretation of the results or an assessment of the validity of the conclusions in the context of the scientific question, claim, or hypothesis.	1. Question gives a scientific claim about differences in populations of red-bellied black snakes in the presence of cane toads. 2. Question gives a representation of data about the mean head size of red-bellied black snakes in the presence and absence of cane toads. 3. Question asks student to select an appropriate interpretation and draw conclusions about the red-bellied black snake head size data.
Crosscutting Concept: Cause and Effect	The question provides at most two of the following: 1) a cause, 2) an effect, and 3) the mechanism that links the cause and effect, and the student is asked to provide the other(s).	Cause: Presence of cane toads Effect: Change in mean snake head size Mechanism: Differential survival of snakes based on head size
Core Idea: Evolution	The characteristics of populations change over time due to changes in allele frequencies. Changes in allele frequencies are caused by random and nonrandom processes – specifically mutation, natural selection, gene flow, and genetic drift. Not all of these changes are adaptive.	The characteristics of the red-bellied black snake populations change over time because of non-random processes – specifically natural selection.
Bloom's Taxonomy Level: Evaluate	Students have to interpret data (graph, table, figure, story, problem, etc.) then determine whether the data are consistent with a given scenario or whether conclusions are consistent with the data, critique validity, quality, or experimental data/methods, or make a judgment and/or justify their answer.	Students have to interpret data about mean snake head size then determine whether the data are consistent with a given scenario.

Sample three-dimensional item with alignment to the scientific practice, crosscutting concept, core idea, and Bloom's Taxonomy level indicated within the item and justified in a table.

Supplemental Figure 2.

Sample Item #2

Scientific Practice: Constructing Explanations and Engaging in Argument from Evidence
Crosscutting Concept: Cause and Effect
Core Idea: Evolution, Information Flow
Bloom's Taxonomy Level: Analyze

1. Phenomenon

Cause

Recently, two chiropractors in Queensland, Australia published a research article stating that the external occipital protuberance (AKA: EOP or horn-like bone spurs) are growing on human skulls due to "sustained aberrant postures associated with the emergence and extensive use of hand-held contemporary technologies, such as smartphones and tablets."[1,2]

Effect

To put it simply, the "horns" are said to be growing at the bottom of the skull, where the head bends to look down at a cell phone, and the authors say that human bodies are physically adapting to use modern technology.

Is this an example of natural selection or evolution, and why?

Information flow

Mechanism

Analyze

2, 3. Make a claim based on scientific principles

Evolution

4. Reasoning

Notes:

[1]Shahar, D. & Sayers, M. G. L. (2018). Prominent exostosis projecting from the occipital squama more substantial and prevalent in young adult than older age groups. *Scientific Reports*, 8, ar:3354.

[2]See also: Shahar, D. & Sayers, M. G. L. (2019). Author Correction: Prominent exostosis projecting from the occipital squama more substantial and prevalent in young adult than older age groups. *Scientific Reports*, *9*, ar:13707.

Applied coding protocol for Sample Item #2		
Code	**Criteria**	**Rationale**
Scientific Practice: Constructing Explanations and Engaging in Argument from Evidence	1. Question gives an event, observation, or phenomenon. 2. Question gives or asks student to make a claim based on the given event, observation, or phenomenon. 3. Question asks student to provide scientific principles or evidence in the form of data or observations to support the claim. 4. Question asks student to provide reasoning about why the scientific principles or evidence support the claim.	1. Question gives the phenomenon of the external occipital protuberance (EOP). 2. Question asks student to make a claim about the EOP. 3. Question asks students to justify claim about EOP using scientific principles of natural selection and evolution. 4. Question asks students to provide reasoning to support their claim about the EOP.
Crosscutting Concept: Cause and Effect	The question provides at most two of the following: 1) a cause, 2) an effect, and 3) the mechanism that links the cause and effect, and the student is asked to provide the other(s).	Cause: Extensive use of smartphones and tablets Effect: Growth of EOP Mechanism: Non-heritable change in phenotype
Core Idea: Evolution	Species evolve over time, and new species can arise, when allele frequencies change due to mutation, natural selection, gene flow, and genetic drift.	Students distinguish that the EOP is not an example of natural selection or evolution.
Core Idea: Information Flow	A genotype influences the range of possible phenotypes in an individual; the actual phenotype results from interactions between alleles and the environment.	The EOP phenotype is the result of interactions with the environment (smartphone use).
Bloom's Taxonomy Level: Analyze	Students are asked to compare/contrast information, have to interpret data (graph, table, figure, story problem, etc.) and come to a conclusion about the data mean (they may or may not be required to explain the conclusion), and/or have to decide what data are important to solve the problem (i.e., picking out relevant from irrelevant information).	Students compare the information about EOP to their understanding of natural selection and evolution and explain their conclusion.

Sample three-dimensional item with alignment to the scientific practice, crosscutting concept, core idea, and Bloom's Taxonomy level indicated within the item and justified in a table.

Supplemental Figure 3.

Sample Item #3

Scientific Practice: Developing and Using Models
Crosscutting Concept: Energy and Matter: Flows, Cycles, and Conservation
Core Idea: Transformations of Energy and Matter
Bloom's Taxonomy Level: Evaluate

Energy and matter

1. Phenomenon

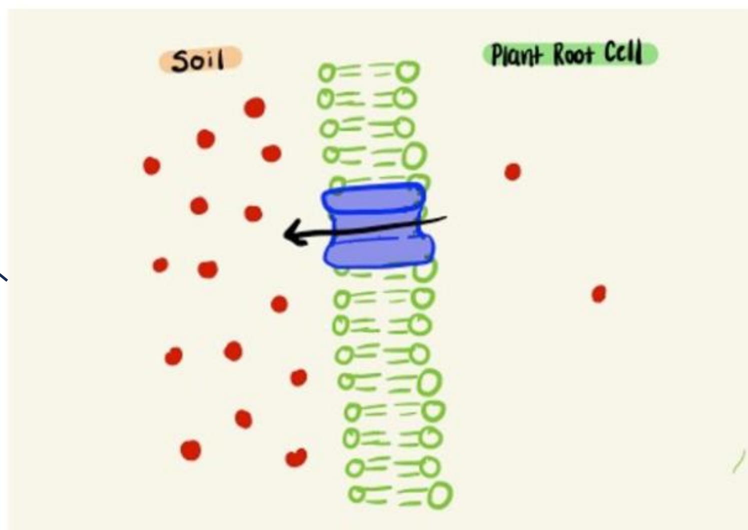
Minerals in the soil move into a plant root cell via facilitated diffusion. Image 4 is meant to show this process of a mineral shown in red moving in the direction of the black arrow across the membrane in the green through a membrane protein in blue. However, something is wrong with the image. What is wrong with the image? (select all that apply)

3. Explanation

Evaluate

Image 4

2. Representation



4. Reasoning

Transfer of matter

Energy and matter

▢ The minerals are moving in the wrong direction, they should be moving down their concentration gradient (from an area of high concentration to low concentration)

▢ Minerals should not be shown being taking in by the roots of the plant, minerals only enter the plant via the stomata

▢ ATP should be shown in the image along the membrane protein to show that ATP is used for facilitated diffusion

Applied coding protocol for Sample Item #3		
Code	**Criteria**	**Rationale**
Scientific Practice: Developing and Using Models	1. Question gives an event, observation, or phenomenon for the student to explain or make a prediction about. 2. Question gives a representation or asks student to select a representation. 3. Question asks student to select an explanation for or prediction about the event, observation, or phenomenon. 4. Question asks student to select the reasoning that links the representation to their explanation or prediction.	1. Question gives phenomenon of soil minerals moving into plant root cells. 2. Question provides Image 4 as a representation of a plant root cell. 3. Question asks student to select an explanation for how Image 4 is an incorrect representation of the phenomenon. 4. Question asks student to select an answer with appropriate reasoning.
Crosscutting Concept: Energy and Matter: Flows, Cycles, and Conservation	To code an assessment task with Energy and Matter: Flows, Cycles, and Conservation, the question asks the student to describe the transfer or transformation of energy or matter within or across systems, or between a system and its surroundings.	Question asks student to describe the transfer of minerals between a system (a plant root cell) and its surroundings (soil).
Core Idea: Transformations of Energy and Matter	Intracellular and intercellular movement of molecules occurs via 1) energy-demanding transport processes and 2) random motion. A molecule's movement is affected by its thermal energy, size, electrochemical gradient, and biochemical properties.	Question asks students to consider molecular movement in relation to the concentration gradient.
Bloom's Taxonomy Level: Evaluate	Students have to interpret data (graph, table, figure, story, problem, etc.) then determine whether the data are consistent with a given scenario or whether conclusions are consistent with the data, critique validity, quality, or experimental data/methods, or make a judgment and/or justify their answer.	Students interpret data about the concentration gradient presented in Image 4 and determine whether the data are consistent with the given situation (facilitated diffusion).

Sample three-dimensional item with alignment to the scientific practice, crosscutting concept, core idea, and Bloom's Taxonomy level indicated within the item and justified in a table.

Supplemental Figure 4.

Sample Item #4

Scientific Practice: Analyzing and Interpreting Data
Crosscutting Concept: Patterns
Core Idea: Information Flow
Bloom's Taxonomy Level: Evaluate

1. Claim

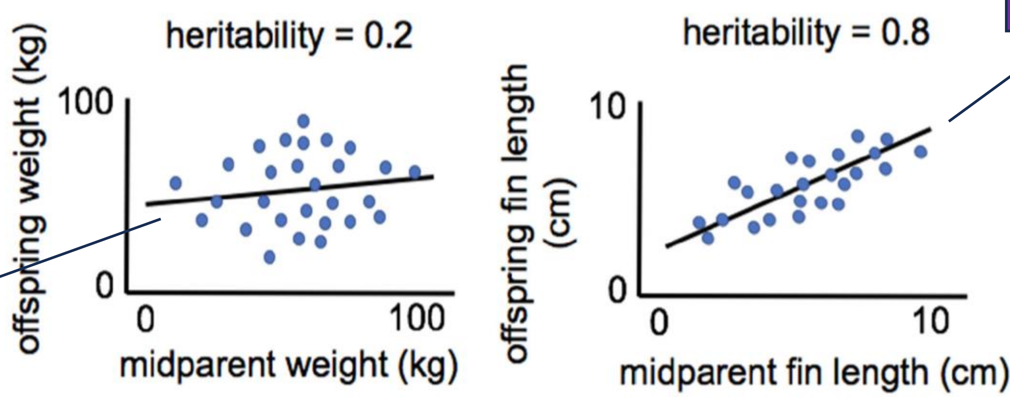
3. Interpretation

This figure shows data from a study of the heritability of weight and fin length in fish, finding a heritability of 0.2 for weight and a heritability of 0.8 for fin length. Which of the following statements is NOT supported by the figures and those findings?

2. Data

Evaluate

Pattern



A. More of the variation in weight can be explained by environment than by genotype, whereas more of the variation in fin length can be explained by genotype than by the environment.
B. A parent who weighs a lot and has long fins is more likely to have offspring with long fins than to have offspring who weighs a lot.
C. Fin length is under directional selection for larger fins.

Information flow

Applied coding protocol for Sample Item #4		
Code	**Criteria**	**Rationale**
Scientific Practice: Analyzing and Interpreting Data	1. Question gives a scientific question, claim, or a hypothesis to be investigated. 2. Question gives a representation of data (table, graph, list of observations, etc.) provided to answer the question or test the claim or hypothesis. 3. Question asks student to select an interpretation of the results or an assessment of the validity of the conclusions in the context of the scientific question, claim, or hypothesis.	1. Question gives scientific claim about the heritability of weight and fin length in fish. 2. Question provides graphs showing the relationship between parent and offspring weight and fin length. 3. Question asks student to interpret which conclusions about heritability are not supported by the data.
Crosscutting Concept: Patterns	To code an assessment task with Patterns, the question asks the student to identify patterns or trends emerging from three or more events, observations, or data.	Question asks student to use patterns in the graphs to make claims about heritability of weight and fin length.
Core Idea: Transformations of Energy and Matter	Individuals transmit genetic information to their offspring; A genotype influences the range of possible phenotypes in an individual; the actual phenotype results from interactions between alleles and the environment.	Question asks student to consider the relationship between parent and offspring phenotypes for weight and fin length and how variation in these traits may be explained by the environment.
Bloom's Taxonomy Level: Evaluate	Students have to interpret data (graph, table, figure, story, problem, etc.) then determine whether the data are consistent with a given scenario or whether conclusions are consistent with the data, critique validity, quality, or experimental data/methods, or make a judgment and/or justify their answer.	Students interpret data about heritability presented in the graphs and determine which option is not consistent with the data.

Sample three-dimensional item with alignment to the scientific practice, crosscutting concept, core idea, and Bloom's Taxonomy level indicated within the item and justified in a table.

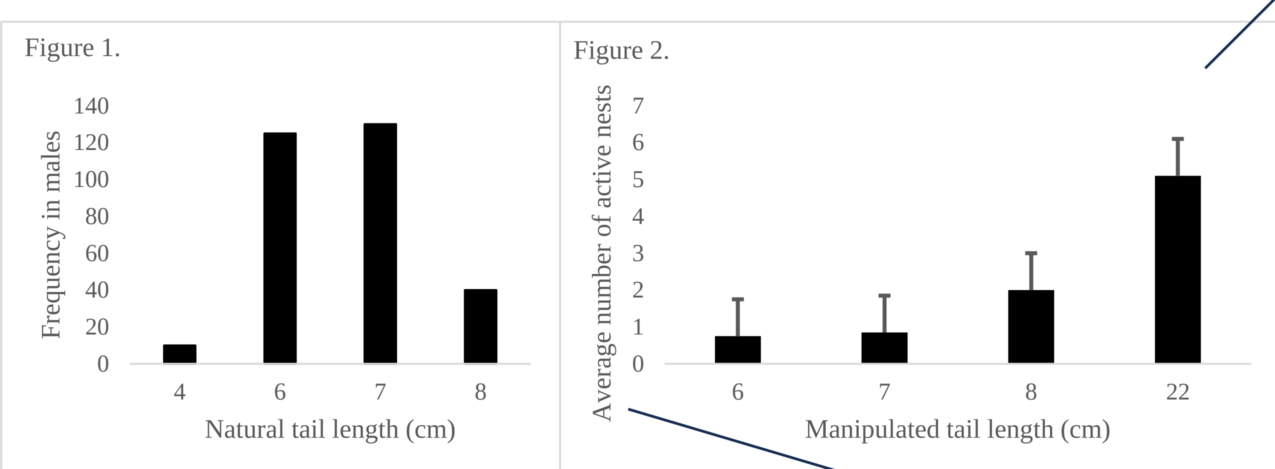
Supplemental Figure 5.

Sample Item #5

Scientific Practice: Asking Questions
Crosscutting Concept: Proportion and Quantity
Core Idea: Evolution
Bloom's Taxonomy Level: Create

In certain species of birds, the males with longer tail feathers tend to have a greater number of active nests, but these males with long tails face additional costs associated with movement and resource acquisition and may experience higher rates of predation. Based on this observation, researchers conducted a tail manipulation experiment in a species of bird that do not naturally produce long tails. The researchers first measured the natural tail length of males in the population (Figure 1). Then researchers randomly selected some male birds and elongated the tail feathers by cutting and gluing the long tail feathers from another similar species of bird with naturally long tails. To control for the effects of tail manipulation, all male birds had their tail feathers cut and glued, even if their tail was not elongated. After the tail manipulation, the researchers measured the number of active nests each male had in its territory (Figure 2)[1].

1. Data



Proportion and Quantity

Based on this passage:

1) Identify the hypothesis of the researchers' study.
2) Describe the results in Figures 1 and 2.
3) Explain whether the results in Figures 1 and 2 support or refute the hypothesis.
4) Identify the type of selection occurring in this population.
5) Generate a new research question that could be tested with a similar tail manipulation experimental design and use your understanding of selection to hypothesize what results you might expect to observe.

2. Testable question

Evolution

Create

Notes:
[1]Data based on Pryke, S. R., & Andersson, S. (2002). *Proc. R. Soc. Lond. B.* 269, 2141-2146.

Applied coding protocol for Sample Item #5		
Code	**Criteria**	**Rationale**
Scientific Practice: Asking Questions	1. Question gives an event, observation, phenomenon, data, scenario, or model. 2. Question asks student to generate an empirically testable question about the given event, observation, phenomenon, data, scenario, or model.	1. Question gives the phenomenon and associated data for the relationship between tail feather length and the number of active nesting sites. 2. Question asks student to generate a testable research question about tail length in birds.
Crosscutting Concept: Proportion and Quantity	To code an assessment task with Proportion and Quantity, the question asks the student to predict the response of one variable to changes in another or identify the relationship between two or more variables from data.	Question asks student to use data to determine the relationship between the variables of tail length and number of active nest sites.
Core Idea: Evolution	Fitness is an individual's ability to survive and reproduce. It is environment-specific and depends on both abiotic and biotic factors. Evolution of optimal fitness is constrained by existing variation, trade-offs and other factors.	Question asks student to consider fitness of male birds based the trait of tail length.
Bloom's Taxonomy Level: Create	To code for Create/Synthesize, students must be synthesizing information into a bigger picture (coherent whole) or creating something they haven't seen before (a novel hypothesis, a novel model, etc.), building up a model or novel hypothesis from data, or putting information from several areas together to create a new pattern/structure/model/etc.	Question asks student to generate a novel, testable research question about tail length in birds.

Sample three-dimensional item with alignment to the scientific practice, crosscutting concept, core idea, and Bloom's Taxonomy level indicated within the item and justified in a table.

Supplemental Table 1.

Supplemental Table 1. Self-reported demographic information of undergraduate biology instructors		
Characteristic	**n**	**%**
Gender[a]		
Female	67	60
Male	42	38
Self-described	0	0
Preferred not to disclose	2	2
Race/ethnicity[b]		
Non-underrepresented	97	87
Underrepresented	11	10
Self-described	1	1
Preferred not to disclose	2	2
Teaching experience as an instructor of record		
0-1 year	5	5
2-5 years	20	18
6-10 years	30	27
11-15 years	29	26
16-20 years	10	9
21-25 years	11	10
> 25 years	6	5

[a]This was the original terminology included in the survey, but we note that the categories here are more representative of sex rather than gender based on current guidelines (American Psychological Association, 2022).

[b]We use the term "underrepresented" here to convey our focus on racial/ethnic groups that have faced disproportionate challenges within STEM disciplines, including Black/African American, Hispanic/Latinx, American Indian/Alaska Native, and Native Hawaiian/Pacific Islander. This grouping is not intended to obscure the unique histories and identities of any group.

Supplemental Table 2.

Supplemental Table 4. Categories of lower-division biology courses included in the sample		
Course category[a]	**n**	**%**
Introductory – Cell/Molecular	32	29
Introductory – Organismal	31	28
Introductory – General Biology	26	23
Ecology/Evolution	6	5
Genetics	3	3
Microbiology	3	3
Anatomy/Physiology	3	3
Cell/Molecular Biology	2	2
Environmental Science	2	2
Plant Biology	2	2
Zoology	1	< 1
Lab courses		
Course has an associated lab component	95	86
Course does not have an associated lab component	16	14

[a]If course category was not evident based on the title of the course, we used the content in the course syllabus to designate the categories. We categorized introductory-series courses that primarily deal with molecules, cells, and genetics as "Introductory – Cell/Molecular," introductory-level courses that primarily deal with animal systems, biodiversity, ecology, and evolution topics as "Introductory – Organismal," and courses that broadly span both cell/molecular biology and ecology/evolution topics as "Introductory – General Biology."

Supplemental Table 3.

Supplemental Table 3. Codebook for scientific practices, crosscutting concepts, biology core ideas, and Bloom's Taxonomy levels	
Code name	**Code criteria**
Scientific Practice	Indicates that the item does (1) or does not (0) assess a Science Practice (as defined by the 3D-LAP protocol). To code a 1, the item must meet the highest criteria for at least one of the following, "Asking Questions," "Developing and Using Models," "Planning Investigations," "Analyzing and Interpreting Data," "Using Mathematics and Computational Thinking," "Constructing Explanations and Engaging in Argument from Evidence" or "Evaluating Information."
Science Practice: Asking Questions	This code only applies to constructed response items. Student is asked to generate a scientific question about a real-world event, observation, phenomenon, data, scenario, or model. 1. Question gives an event, observation, phenomenon, data, scenario, or model. 2. Question asks student to generate an empirically testable question about the given event, observation, phenomenon, data, scenario, or model.
Science Practice: Developing and Using Models	Constructed Response: Student is given or asked to construct a mathematical, graphical, computational, symbolic, or pictorial representation and use it to explain or predict an event, observation, or phenomenon. 1. Question gives an event, observation, or phenomenon for the student to explain or make a prediction about. 2. Question gives a representation or asks student to construct a representation. 3. Question asks student to explain or make a prediction about the event, observation, or phenomenon. 4. Question asks student to provide the reasoning that links the representation to their explanation or prediction. Selected Response: Student is given or asked to select a mathematical, graphical, computational, symbolic, or pictorial representation and select an appropriate explanation or prediction about an event, observation, or phenomenon based on the representation. 1. Question gives an event, observation, or phenomenon for the student to explain or make a prediction about. 2. Question gives a representation or asks student to select a representation. 3. Question asks student to select an explanation for or prediction about the event, observation, or phenomenon. 4. Question asks student to select the reasoning that links the representation to their explanation or prediction.
Science Practice: Planning Investigations	Constructed Response: Student is asked to design an experimental method or identify a set of observations that can be used to answer a scientific question or test a claim or hypothesis. 1. Question poses a scientific question, claim, or hypothesis to be investigated. 2. Question asks student to describe or design an investigation, or identify the observations required to answer the question or test the claim or hypothesis. 3. Question asks student to justify how their description, design, or observations can be used to answer the question or test the claim or hypothesis. Selected Response: Student is asked to select an appropriate design of an experimental method or an observation that can be used to answer a scientific question or test a claim or hypothesis. 1. Question poses a scientific question, claim, or a hypothesis to be investigated. 2. Question asks student to select a description of or a design for an investigation or select the observations that could be used to answer the question or test the hypothesis. 3. Question asks student to select a justification of how the description, design, or observations can be used to answer the question or test the claim or hypothesis.

Science Practice: Analyzing and Interpreting Data	Constructed Response: Student is given a question, claim, or hypothesis and data collected from an experiment or observation and is asked to analyze the resulting data and interpret their meaning. 1. Question gives a scientific question, claim, or hypothesis to be investigated. 2. Question gives a representation of the data (e.g., table or graph, or list of observations) provided to answer the question or test the claim or hypothesis. 3. Question gives an analysis of the data or asks student to analyze the data. 4. Question asks student to interpret the results or assess the validity of the conclusions in the context of the scientific question, claim, or hypothesis. Selected Response: Student is given a question, claim, or hypothesis and data collected from an experiment or observation and is asked to select an interpretation of its meaning. 1. Question gives a scientific question, claim, or a hypothesis to be investigated. 2. Question gives a representation of data (table, graph, list of observations, etc.) provided to answer the question or test the claim or hypothesis. 3. Question asks student to select an interpretation of the results or an assessment of the validity of the conclusions in the context of the scientific question, claim, or hypothesis.
Science Practice: Using Mathematics and Computational Thinking	Constructed Response: Student is asked to use mathematical reasoning or a calculation and interpret the results within the context of the given event, observation, or phenomenon. 1. Question gives an event, observation, or phenomenon. 2. Question asks student to perform a calculation or statistical test, generate a mathematical representation, or demonstrate a relationship between parameters. 3. Question asks student to give a consequence or an interpretation (not a restatement) in words, diagrams, symbols, or graphs of their results in the context of the given event, observation, or phenomenon. Selected Response: Student is expected to perform a mathematical manipulation and asked to select an interpretation of the results within the context of a given event, observation, or phenomenon. 1. Question gives an event, observation, or phenomenon. 2. Question asks student to perform a calculation or statistical test, use a mathematical representation, or derive a relationship between parameters in order to obtain the correct answer. 3. Question asks student to select a consequence or an interpretation (not a restatement) in words, diagrams, symbols, or graphs of their results in the context of the given event, observation, or phenomenon.
Science Practice: Constructing Explanations and Engaging in Argument from Evidence	Constructed Response: Student is asked to provide reasoning based on evidence to support a claim. 1. Question gives an event, observation, or phenomenon. 2. Question gives or asks student to make a claim based on the given event, observation, or phenomenon. 3. Question asks student to provide scientific principles or evidence in the form of data or observations to support the claim. 4. Question asks student to provide reasoning about why the scientific principles or evidence support the claim. Selected Response: Student is asked to select reasoning and evidence to support a claim. 1. Question gives an event, observation, or phenomenon. 2. Question gives or asks student to select a claim based on the given event, observation, or phenomenon. 3. Question asks student to select scientific principles or evidence in the form of data or observations to support the claim. 4. Question asks student to select the reasoning about why the scientific principles or evidence support the claim.

Science Practice: Evaluating Information	Constructed Response: Student is asked to make sense of information or ideas presented to them. 1. Question gives an excerpt from a conversation, article, student solution, or video (or similar form of communication) that makes one or more assertions. 2. Question gives a conclusion about the validity of the assertion(s) made or asks student to make a conclusion about the validity of the assertion(s) or reconcile multiple assertions with each other. 3. Question asks student to provide reasoning to support their conclusion(s) about the validity of the assertion(s) or reconciliation with data, observations, or scientific principles. Selected Response: Student is asked to make sense of information or ideas presented to them. 1. Question gives an excerpt from a conversation, article, student solution, or video (or similar form of communication) that makes one or more assertions. 2. Question gives a conclusion about the validity of the assertion(s) or asks student to select a conclusion about the validity of the assertion(s) or reconciliation of multiple assertions. 3. Question asks student to select reasoning to support their conclusion(s) about the validity of the assertion(s) or reconciliation with data, observations, or scientific principles.
Crosscutting Concept	Indicates that the item does (1) or does not (0) assess a Crosscutting Concept (as defined by the 3D-LAP protocol). To code a 1, the item must meet the criteria for at least one of the following, "Patterns," "Cause and Effect: Mechanism and Explanation," "Scale," "Proportion and Quantity," "Systems and System Models," "Energy and Matter: Flows, Cycles, and Conservation," "Structure and Function," and "Stability and Change."
Crosscutting Concept: Patterns	To code an assessment task with Patterns, the question asks the student to identify patterns or trends emerging from three or more events, observations, or data.
Crosscutting Concept: Cause and Effect: Mechanism and Explanation	To code an assessment task with Cause and Effect: Mechanism and Explanation, the question provides at most two of the following: 1) a cause, 2) an effect, and 3) the mechanism that links the cause and effect, and the student is asked to provide the other(s).
Crosscutting Concept: Scale	To code an assessment task with Scale, the question asks the student 1) to compare objects, processes, or properties across size, time, or energy scales, or to dimensions of familiar objects, timescales, or energies or 2) to identify non-negligible/relevant interactions at various scales.
Crosscutting Concept: Proportion and Quantity	To code an assessment task with Proportion and Quantity, the question asks the student to predict the response of one variable to changes in another or identify the relationship between two or more variables from data.
Crosscutting Concept: Systems and System Models	To code an assessment task with Systems and System Models, the question asks the student to identify a system (by defining its components or boundaries), any assumptions made, and the surroundings (if necessary), and how the system and surroundings interact with each other.
Crosscutting Concept: Energy and Matter: Flows, Cycles, and Conservation	To code an assessment task with Energy and Matter: Flows, Cycles, and Conservation, the question asks the student to describe the transfer or transformation of energy or matter within or across systems, or between a system and its surroundings~~, with explicit recognition that energy and/or matter are conserved~~. *The phrase "with explicit recognition that energy and/or matter are conserved" is restrictive, and as a result, few items meet this crosscutting concept. We removed this phrase from our operational definition of the crosscutting concept "Energy and Matter: Flows, Cycles, and Conservation."*
Crosscutting Concept: Structure and Function	To code an assessment task with Structure and Function, the question asks the student to predict or explain a function or property based on a structure, or to describe what structure could lead to a given function or property. *To meet this crosscutting concept, the item needs to clearly address both structure and function. The function does not have to be immediate and may be either proximal or distal. Items that only ask to identify a structure do not meet this crosscutting concept.*
Crosscutting Concept: Stability and Change	To code an assessment task with Stability and Change, the question asks the student to determine 1) if a system is stable and provide the evidence for this, or 2) what forces, rates, or processes make a system stable (static, dynamic, or steady state), or 3) under what conditions a system remains stable, or 4) under what conditions a system is destabilized and the resulting state.

Core Idea	Indicates the that the item does (1) or does not (0) assess a Core Idea (as defined by the BioCore Guide). To code a 1, the item must meet the criteria for at least one of the following, “Evolution,” “Information Flow,” “Structure Function,” “Transformations of Energy and Matter,” and “Systems.”
Core Idea: Evolution	To meet the Core Idea, the exam item must align with at least one of the following criteria: • Overarching Principle: All living organisms share a common ancestor. • Overarching Principle: Species evolve over time, and new species can arise, when allele frequencies change due to mutation, natural selection, gene flow, and genetic drift. • Molecular: Multiple molecular mechanisms, including DNA damage and errors in replication, lead to the generation of random mutations. These mutations create new alleles that can be inherited via mitosis, meiosis, or cell division. • Molecular: Mutations and epigenetic modifications can impact the regulation of gene expression and/or the structure and function of the gene product. If mutations affect phenotype and lead to increased reproductive success, the frequency of those alleles will tend to increase in the population. • Physiology: Mutations that change protein structure and/or regulation can impact anatomy and physiological function at all levels of organization. • Physiology: Most organisms have anatomical and physiological traits that tend to increase their fitness for a particular environment. • Physiology: Physiological systems are constrained by ancestral structures, physical limits, and the requirements of other physiological systems, leading to trade-offs that affect fitness. • Ecology/Evolutionary Biology: The characteristics of populations change over time due to changes in allele frequencies. Changes in allele frequencies are caused by random and nonrandom processes – specifically mutation, natural selection, gene flow, and genetic drift. Not all of these changes are adaptive. • Ecology/Evolutionary Biology: All species alive today are derived from the same common ancestor. New species arise when populations become genetically isolated and diverge due to mutation, natural selection, and genetic drift. Phylogenetic trees depict relationships among ancestral and descendant species, and are estimated based on data. • Ecology/Evolutionary Biology: Fitness is an individual’s ability to survive and reproduce. It is environment-specific and depends on both abiotic and biotic factors. Evolution of optimal fitness is constrained by existing variation, trade-offs and other factors.
Core Idea: Information Flow	To meet the Core Idea, the exam item must align with at least one of the following criteria: • Overarching Principle: Organisms inherit genetic and epigenetic information that influences the location, timing, and intensity of gene expression. • Overarching Principle: Cells/organs/organisms have multiple mechanisms to perceive and respond to changing environmental conditions. • Molecular: In most cases, genetic information flows from DNA to mRNA to protein, but there are important exceptions. • Molecular: Gene expression and protein activity are regulated by intracellular and extracellular signaling molecules. Signal transduction pathways are crucial in relaying these signals. • Molecular: The signals that a cell receives depend on its location, and may change through time. As a result, different types of cells express different genes, even though they contain the same DNA. • Physiology: Information stored in DNA is expressed as RNA and proteins. These gene products impact anatomical structures and physiological function. • Physiology: Organisms have sophisticated mechanisms for sensing changes in the internal or external environment. They use chemical, electrical, or other forms of signaling to coordinate responses at the cellular, tissue, organ, and/or system level. • Ecology/Evolutionary Biology: Individuals transmit genetic information to their offspring; some alleles confer higher fitness than others in a particular environment. • Ecology/Evolutionary Biology: A genotype influences the range of possible phenotypes in an individual; the actual phenotype results from interactions between alleles and the environment.
Core Idea: Structure Function	To meet the Core Idea, the exam item must align with at least one of the following criteria: • Overarching Principle: Biological structures exist at all levels of organization, from molecules to ecosystems. A structure’s physical and chemical characteristics influence its interactions with other structures, and therefore its function. • Overarching Principle: Natural selection leads to evolution of structures that tend to increase fitness within the context of evolutionary, developmental, and environmental constraints.

	• Molecular: The structure of a cell – its shape, membrane, organelles, cytoskeleton, and polarity – impacts its function. • Molecular: The three-dimensional structure of a molecule and its subcellular localization impact its function, including the ability to catalyze reactions or interact with other molecules. Function can be regulated through reversible alterations of structure e.g. phosphorylation. • Molecular: The structure of molecules or organisms may be similar due to common ancestry or selection for similar function. • Physiology: Physiological functions are often compartmentalized into different cells, tissues, organs, and systems, which have structures that support specialized activities. • Physiology: The size, shape, and physical properties of organs and organisms all affect function. The ratio of surface area to volume is particularly critical for structures that function in transport or exchange of materials and heat. • Physiology: Structure constrains function in physiology; specialization for one function may limit a structure's ability to perform another function. • Ecology/Evolutionary Biology: Natural selection has favored structures whose shape and composition contribute to their ecological function. • Ecology/Evolutionary Biology: Competition, mutualism, and other interactions are mediated by each species' morphological, physiological, and behavioral traits.
Core Idea: Transformations of Energy and Matter	To meet the Core Idea, the exam item must align with at least one of the following criteria: • Overarching Principle: Energy and matter cannot be created or destroyed, but can be changed from one form to another. • Overarching Principle: Energy captured by primary producers is necessary to support the maintenance, growth and reproduction of all organisms. • Overarching Principle: Natural selection leads to the evolution of efficient use of resources within constraints. • Molecular: Energy captured by primary producers is stored as chemical energy. This stored energy can be converted through a series of biochemical reactions into ATP for immediate use in the cell. • Molecular: In cells, the synthesis and breakdown of molecules is highly regulated. Biochemical pathways usually involve multiple reactions catalyzed by enzymes that lower activation energies. Energetically unfavorable reactions are driven by coupling to energetically favorable reactions such as ATP hydrolysis. • Molecular: Intracellular and intercellular movement of molecules occurs via 1) energy-demanding transport processes and 2) random motion. A molecule's movement is affected by its thermal energy, size, electrochemical gradient, and biochemical properties. • Physiology: Energy captured by primary producers is stored as chemical energy. This stored energy can be converted into ATP, which is required for energetically demanding activities necessary for life, including synthesis, transport, and movement. • Physiology: Due to the inefficiency of biochemical reactions and other constraints, physiological processes are never 100% efficient. • Physiology: Organisms have limited energetic and material resources which must be distributed across competing functional demands. These include movement of material across gradients, growth, maintenance, and reproduction, inevitably leading to trade-offs. • Ecology/Evolutionary Biology: Energy captured by primary producers is stored as chemical energy. At each trophic level, most of this energy is used for maintenance, with a relatively small fraction available for growth and reproduction. As a consequence, each trophic level in an ecosystem has less energy available than the preceding level. • Ecology/Evolutionary Biology: Chemical elements are transferred among the abiotic and biotic components of an ecosystem; changes in the amount and distribution of chemical elements can impact the ecosystem.
Core Idea: Systems	To meet the Core Idea, the exam item must align with at least one of the following criteria: • Overarching Principle: Biological molecules, genes, cells, tissues, organs, individuals, and ecosystems interact to form complex networks. A change in one component of the network can affect many other components. • Overarching Principle: Organisms have complex systems that integrate internal and external information, incorporate feedback control, and allow them to respond to changes in the environment. • Molecular: Cells receive a complex array of chemical and physical signals that vary in time, location, and intensity over the lifespan of the organism; a cell's response depends on integration and coordination of these various signals.

<table>
<tr><td></td><td>
Molecular: During development the signals a cell receives depend on its spatial orientation within the embryo and its intercellular interactions. As a consequence, cells adopt different cell fates depending on their local environment and/or cell lineage.
Molecular: Alteration of a single gene or molecule in a signaling network may have complex impacts at the cell, tissue or whole-organism level.
Physiology: Organ systems are not isolated but interact with each other through chemical and physical signals at the level of cells, tissues, and organs.
Physiology: An individual’s physiological traits affect its interactions with other organisms and with its physical environment.
Physiology: In the face of environmental changes, organisms may maintain homeostasis through control mechanisms that often use negative feedback; others have adaptations that allow them to acclimate to environmental variation.
Ecology/Evolutionary Biology: The size and structure of a population is dynamic. A species’ abundance and distribution are limited by available resources and by interactions between biotic and abiotic factors.
Ecology/Evolutionary Biology: Ecosystems are not isolated and static – they respond to change, both as a result of intrinsic changes to networks of species and as a result of extrinsic environmental drivers. Within an ecosystem, interactions among individuals form networks; changes in one node of a network can cause changes in other nodes – directly or indirectly.
Ecology/Evolutionary Biology: Biodiversity impacts many aspects of ecosystems.
</td></tr>
<tr><td>Bloom’s Taxonomy</td><td>Only apply the code for the highest level of Bloom’s Taxonomy that the item is capable of assessing.

Question 1. Could students memorize the answer to this specific question?
Yes: go to question 2.
No: go to question 4.

Question 2. To answer the question, are students repeating nearly exactly what they have heard or seen in class materials (including lecture, textbook, laboratory, homework, clicker, etc.)?
Yes: See Remember
No: go to question 3.

Question 3. Are students demonstrating a conceptual understanding by putting the answer in their own words, matching examples to concepts, representing a concept in a new form (words to graph, etc.), etc.?
Yes: See Comprehend
No: Go back to question 1. If you are sure the answer to question 1is yes, the question should fit into “remember” or “comprehend.”

Question 4. Is there potentially more than one valid solution* (even if a “better” one exists or if there is a limit to what solutions can be chosen)?
Yes: go to question 5.
No: go to question 8.

Question 5. Are students making a judgment and/or justifying their answer?
Yes: See Evaluate
No: go to question 6.

Question 6. Are students synthesizing information into a bigger picture (coherent whole) or creating something they haven’t seen before (a novel hypothesis, novel model, etc.)?
Yes: See Synthesize/create
No: go to question 7.

Question 7. Are students being asked to compare/contrast information?
Yes: See Analyze
No: go to question 16

Question 8. To answer the question, do students have to interpret data (graph, table, figure, story problem, etc.)?
Yes: go to question 9.
No: go to question 14.</td></tr>
</table>

Question 9. Are students determining whether the data are consistent with a given scenario or whether conclusions are consistent with the data? Are students critiquing validity, quality, or experimental data/methods?
Yes: See
Evaluate
No: go to question 10.

Question 10. Are students building up a model or novel hypothesis from the data?
Yes: See Synthesize/create
No: go to question 11.

Question 11. Are students coming to a conclusion about what the data mean (they may or may not be required to explain the conclusion) and/or having to decide what data are important to solve the problem (i.e., picking out relevant from irrelevant information)?
Yes: See Analyze
No: go to question 12.

Question 12. Are students using the data to calculate the value of a variable?
Yes: See Apply
No: go to question 13.

Question 13. Are students redescribing the data to demonstrate they understand what the data represent?
Yes: See Comprehend
No: go back to questions 4and 8.

Question 14. Are students putting information from several areas together to create a new pattern/structure/model/etc.?
Yes: See Synthesize/create
No: go to question 15.

Question 15. Are students predicting the outcome or trend of a fairly simple change to a scenario?
Yes: See Apply
No: go to question 16.

Question 16. Are students demonstrating that they understand a concept by putting it into a different form (new example, analogy, comparison, etc.) than they have seen in class?
Yes: See Comprehend
No: go back through each category or refer to category descriptions to see which fits the best

For shorthand coding purposes, we have condensed and operationalized the Bloom's Dichotomous Key to reflect the "Yes" responses for each Bloom's Taxonomy level and assigned a numerical value to each.

1 = Remember

- To code for Remember, students could memorize the answer to the question and students are repeating nearly exactly what they have heard or seen in class materials (including lecture, textbook, laboratory, homework, clicker, etc.).

2 = Understand/Comprehend

- To code for Understand/Comprehend, students demonstrate a conceptual understanding by putting the answer in their own words, matching examples to concepts, representing a concept in a new form (words to graph, etc.), etc., or demonstrate that they understand a concept by putting it into a different form (new example, analogy, comparison, etc.) than they have seen in class.

3 = Apply

- To code for Apply, students are using data to calculate the value of a variable or are predicting the outcome of a trend of a fairly simple change to a scenario.

4 = Analyze

- To code for Analyze, students are asked to compare/contrast information, have to interpret data (graph, table, figure, story problem, etc.) and come to a conclusion about the data mean

	(they may or may not be required to explain the conclusion), and/or have to decide what data are important to solve the problem (i.e., picking out relevant from irrelevant information). 5 = Evaluate • To code for Evaluate, students have to interpret data (graph, table, figure, story, problem, etc.) then determine whether the data are consistent with a given scenario or whether conclusions are consistent with the data, critique validity, quality, or experimental data/methods, or make a judgment and/or justify their answer. 6 = Create/Synthesize • To code for Create/Synthesize, students must be synthesizing information into a bigger picture (coherent whole) or creating something they have not seen before (a novel hypothesis, a novel model, etc.), building up a model or novel hypothesis from data, or putting information from several areas together to create a new pattern/structure/model/etc.

The coding protocol for scientific practices and crosscutting is from the Three-Dimensional Learning Assessment Protocol (Laverty et al., 2016), which is an adaptation of *A Framework for K-12 Science Education* (National Research Council, 2012). The coding protocol for biology core ideas is from the BioCore Guide (Brownell et al. 2014), which is an articulation of the biology core concepts of *Vision and Change* (American Association for the Advancement of Science, 2011). The coding protocol for Bloom's Taxonomy Levels is the Bloom's Dichotomous Key (Semsar & Casagrand, 2017), which is an articulation of Bloom's Taxonomy (Bloom et al., 1956; Anderson et al., 2001).

Supplemental Table 4.

Supplemental Table 4. Percent agreement across the training set and percent agreement in a set of items to assess rater drift

Code name	Training[a] between Rater 1 and Rater 2	Training[a] between Rater 1 and Rater 3	Overall agreement[b] between Rater 1 and Rater 3	Phase 1[c] between Rater 1 and Rater 3	Phase 2[d] between Rater 1 and Rater 3	Phase 3[e] between Rater 1 and Rater 3
Scientific Practice	97.9	95.8	94.7	92.9	94.0	96.9
Scientific Practice: Asking Questions	100	100	100	100	100	100
Scientific Practice: Developing and Using Models	97.9	97.9	99.0	98.1	99.4	99.4
Scientific Practice: Planning Investigations	100	100	99.0	98.7	98.2	100
Scientific Practice: Analyzing and Interpreting Data	95.8	100	97.5	94.2	98.2	100
Scientific Practice: Using Mathematics and Computational Thinking	100	100	99.8	100	99.4	99.4
Scientific Practice: Constructing Explanations and Engaging in Argument from Evidence	97.9	95.8	97.3	97.4	97.0	97.5
Scientific Practice: Evaluating Information	100	100	99.8	100	99.4	100
Crosscutting Concept	75.0	81.2	82.3	84.0	83.9	79.1
Crosscutting Concept: Patterns	97.9	97.9	94.9	91.7	94.0	98.8
Crosscutting Concept: Cause and Effect	91.7	93.8	89.1	91.0	89.9	86.5
Crosscutting Concept: Scale	100	100	99.2	97.4	100	100
Crosscutting Concept: Proportion	100	97.9	94.9	94.9	91.1	98.8
Crosscutting Concept: Systems and System Models	85.4	87.5	94.9	96.2	91.1	97.5
Crosscutting Concept: Energy and Matter	91.7	95.8	97.3	94.2	98.2	99.4
Crosscutting Concept: Structure and Function	87.5	83.3	90.1	91.7	89.9	89.0
Crosscutting Concept: Stability and Change	100	100	99.0	98.1	98.8	100
Biology Core Idea	79.2	89.6	82.3	87.2	76.8	83.4
Biology Core Idea: Evolution	97.9	100	93.2	93.6	95.2	90.8
Biology Core Idea: Information Flow	89.6	91.7	89.9	96.2	83.9	90.2
Biology Core Idea: Structure Function	87.5	83.3	85.8	88.5	82.7	86.5
Biology Core Idea: Transformations of Energy and Matter	93.8	95.8	97.1	94.2	98.2	98.8
Biology Core Idea: Systems	83.3	81.2	89.9	90.4	88.1	91.4
Bloom's Level (LOCS/HOCS)[f]	95.8	95.8	94.9	92.3	97.0	95.1

[a]The set of training items consisted of 48 randomly selected items. The items were coded iteratively in sets of 12, and a single averaged value is presented here.
[b]Overall percent agreement was calculated based on the 487 items from Phases 1, 2, and 3.
[c]Phase 1 consisted of 156 items across 4 exams that were randomly selected from the first third of coding completed by Rater 1.
[d]Phase 2 consisted of 168 items across 4 exams that were randomly selected from the second third of coding completed by Rater 1.
[e]Phase 3 consisted of 163 items across 4 exams that were randomly selected from the last third of coding completed by Rater 1.
[f]Lower-order cognitive skills (LOCS) consisted of "remember," "understand," and "apply." Higher-order cognitive skills consisted of "analyze," "evaluate," and "create."

Supplemental Table 5.

Supplemental Table 5. Computations of Fisher's *z*-tests concerning differences between correlations of the percent of higher-order cognitive skills and the percent of exam points in each dimension

	3D	Scientific practices	Crosscutting concepts	Core ideas
3D	-	2.92*	-3.99*	-4.46*
Scientific practices	2.92*	-	6.91*	7.38*
Crosscutting concepts	-3.99*	6.91*	-	0.47
Core ideas	-4.46*	7.38*	0.47	-

Note: An asterisk (*) indicates significant differences ($p < 0.05$) between correlation coefficients.