

## RESEARCH ARTICLE

# Biology exams rarely use visual models to engage higher-order cognitive skills

Crystal Uminski<sup>1</sup>, Christian Cammarota<sup>1</sup>, Brian A. Couch<sup>2</sup>, L. Kate Wright<sup>1\*</sup>, Dina L. Newman<sup>1</sup>

**1** Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, New York, United States of America, **2** School of Biological Sciences, University of Nebraska – Lincoln, Lincoln, Nebraska, United States of America

\* [lkwsbi@rit.edu](mailto:lkwsbi@rit.edu)



## OPEN ACCESS

**Citation:** Uminski C, Cammarota C, Couch BA, Wright LK, Newman DL (2025) Biology exams rarely use visual models to engage higher-order cognitive skills. PLoS One 20(7): e0317077.

<https://doi.org/10.1371/journal.pone.0317077>

**Editor:** David R Wessner, Davidson College, UNITED STATES OF AMERICA

**Received:** January 13, 2025

**Accepted:** June 12, 2025

**Published:** July 2, 2025

**Copyright:** © 2025 Uminski et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** This research involved collecting information that could be used to easily identify research participants. To protect the confidentiality of our research participants, our protocol (University of Nebraska – Lincoln protocol 21082) required that identifiable data be shared in de-identified and aggregated form and that certain data,

## Abstract

Visual models are a necessary part of molecular biology education because submicroscopic compounds and processes cannot be directly observed. Accurately interpreting the biological information conveyed by the shapes and symbols in these visual models requires engaging visual literacy skills. For students to develop expertise in molecular biology visual literacy, they need to have structured experiences using and creating visual models, but there is little evidence to gauge how often undergraduate biology students are provided such opportunities. To investigate students' visual literacy experiences, we surveyed 66 instructors who taught lower division undergraduate biology courses with a focus on molecular biology concepts. We collected self-reported data about the frequency with which the instructors teach with visual models and we analyzed course exams to determine how instructors incorporated visual models into their assessments. We found that most instructors reported teaching with models in their courses, yet only 16% of exam items in the sample contained a visual model. There was not a statistically significant relationship between instructors' self-reported frequency of teaching with models and extent to which their exams contained models, signaling a potential mismatch between teaching and assessment practices. Although exam items containing models have the potential to elicit higher-order cognitive skills through model-based reasoning, we found that when instructors included visual models in their exams the majority of the items only targeted the lower-order cognitive skills of Bloom's Taxonomy. Together, our findings highlight that despite the importance of visual models in molecular biology, students may not often have opportunities to demonstrate their understanding of these models on assessments.

## Introduction

The content covered in molecular biology courses exists at a scale so small that biologists rely on visual models, such as drawings and diagrams, to portray the nature of

such as the exam artifacts we collected in this study, would not be publicly shared at all. We have provided data about the instructors who participated in this research in de-identified and aggregated form in [S1 Table](#) and [S2 Table](#). We have provided the de-identified data necessary to reproduce our summary statistics and linear models in [S1 Data](#). The data in [S1 Data](#) is coded from exams artifacts, but per our protocol, we cannot share the exam artifacts themselves in a public format. Should you have any additional questions about this research protocol, you can contact the University of Nebraska – Lincoln Institutional Review Board at 402-472-6965 or at [irb@unl.edu](mailto:irb@unl.edu).

**Funding:** This material is based upon work supported by the National Science Foundation (NSF) Graduate Research Fellowship (DGE-1610400 awarded to CU), NSF Improving Undergraduate STEM Education Grant (DUE-2044243 awarded to BAC), and the NSF Research Traineeship (DGE-2222337 awarded to LKW and DLW). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

and relationships between invisible molecules. Given the complexity of some of the molecules and molecular processes, biologists often simplify the subject matter by using abstract shapes and symbols to convey meaning in models [1,2]. For example, letters, boxes, and lines can be used to represent concepts such as nucleotides, genes, and chromosomes. Correctly interpreting the abstract shapes and symbols depends on both content knowledge and contextual clues in the representation [3]. Consider that a single line in different contexts could represent the sugar-phosphate backbone of a few nucleotides in a ladder representation of DNA structure, thousands of nucleotides as the line of a box-and-line representation of a gene, or millions of nucleotides in a representation of an unreplicated chromosome. As molecular biologists use abstraction as a form of shorthand, it is critical for students to develop the visual literacy skills necessary to “read” the correct interpretation of the shapes and symbols commonly used to communicate in molecular biology.

Visual literacy is defined as the ability to interpret, create, and communicate with visual models [4], and it is a set of skills that are so important for biology students that they are encapsulated within multiple frameworks for science education. Visual literacy skills are encompassed within the scientific practice “Developing and Using Models” that is part of the three-dimensional framework for science education [5] and are reflected in several core competencies from the *Vision and Change* framework for undergraduate biology education, such as “Ability to use modeling and simulation” and “Ability to communicate and collaborate with other disciplines” [6]. Evidence suggests, however, that communicating science through visual methods is far from standard in biology curriculum [1]. Even though most biology students see visual models in textbooks and lecture slides, visual literacy skills are rarely explicitly taught in biology courses [7].

While it can be difficult to measure the frequency that visual literacy skills are taught without directly observing instructors during class sessions, self-reported data about teaching practices and summative assessments (e.g., tests and exams) can provide insights into how instructors incorporate visual literacy into their courses. The Measurement Instrument for Scientific Teaching (MIST) [8] is one instrument that can be used to gather self-reported data about teaching practices. The MIST specifically includes an item in which instructors can report the frequency that they ask their students to make or interpret models to summarize scientific processes. Models, as broadly defined by the MIST, include drawings, diagrams, schematics, concept maps, flow charts, illustrative models, and box-and-arrow diagrams [8]. Making and interpreting models are key visual literacy skills, thus instructor responses to this item can indicate the frequency in which students encounter instruction about visual literacy in their biology courses. The MIST and other similar measures of teaching practices, such as the Three-Dimensional Learning Observation Protocol [9], can provide information about how often students are asked to make and interpret models, but these measures do not provide context about what types of models students are encountering and how specifically students are being asked to engage with the models.

To get a broader picture of visual literacy practices in introductory-level molecular biology courses, we can analyze the content on instructors’ exams. What is included on

exams can indicate the type of content instructors have taught in their courses and can signal what content instructors think was the most important for their students to learn [10,11]. Thus, we would expect to see visual literacy skills, such as interpreting figures and creating diagrams, incorporated into summative assessments if these visual literacy skills are prioritized learning outcomes in a course [7]. As such, if interpreting or creating visual models were prioritized learning objectives in a course, we might expect the instructor to have exam questions which ask students to interpret or create models. Instructor priorities may also be reflected in the point value of exam items [12]. Instructors may signal to students the importance of a concept through the percentage of points certain concepts comprise on an exam, either with higher point value associated with a single item or a large number of points across multiple questions about the same topic. Therefore, if models are emphasized on an exam, we can infer that models were likely to also be an emphasized part of the course content.

In addition to determining how frequently students encounter visual models on their exams, it is also important to consider what students are asked to do with those models. Visual literacy skills span from rote memorization of the parts of a familiar diagram to the creation of new visual representations to communicate a new concept or idea, and these visual literacy skills can be mapped onto the levels of the cognitive domain of Bloom's Taxonomy [4,7,13]. Bloom's Taxonomy [14,15] is a widely-used framework for categorizing tasks into ones that can elicit lower-order cognitive skills (remembering, understanding, and applying) and higher-order cognitive skills (analyzing, evaluating, and creating). These higher-order cognitive skills are often equated with critical thinking and tend to be the types of skills instructors report wanting their students to know by the time they finish a course or a degree program [16–19]. By examining the cognitive skills of items with models, we can determine if students are mainly being asked to use models to demonstrate rote content knowledge or if they are being asked to use models to engage higher-order cognitive skills in ways that are more aligned to scientific practices and *Vision and Change* core competencies.

Along with determining the frequency with which instructors are using models on their exams and the cognitive skills of such items, we can also categorize the specific types of models that instructors are using. For a subset of this research, we narrow the focus from visual models in a broad sense to more specifically study only models of DNA. We focus here on DNA because it is a central topic in molecular biology courses and there is an existing framework to categorize visual models of DNA. The DNA Landscape framework [2], is a 3-by-3 matrix that shows variation in DNA representations along axes of scale (nucleotide, gene, and chromosome) and abstraction (literal shape, elements of shape and abstraction, very abstract). To develop molecular biology visual literacy skills, it is important for students to see and interpret representations of DNA across levels of scale and abstraction [1,3,20]; however, an analysis of textbook figures suggested that students may be encountering certain representations in the DNA Landscape matrix much more frequently than others [2]. For example, textbooks were more frequently using abstract "X" shaped representations of chromosomes compared to less-abstract representations of chromosomes as maps or as string-like chromatin [2]. As with textbook figures, we anticipate that certain representations of DNA on biology exams may be more common than others, particularly when considering both the typical topics covered in introductory-level molecular biology courses as well as the ease or the challenge of reproducing certain representations in an exam format.

Visual literacy is a critical skill for molecular biology, but there remains little evidence as to how often these skills are taught or assessed in undergraduate courses [7]. Our work here aims to provide a snapshot of visual literacy in molecular biology courses by addressing the following research questions:

- 1) How often do students encounter visual models on exams in introductory-level molecular biology courses?
- 2) What cognitive skills are assessed when items from introductory-level molecular biology exams do or do not include a visual model?
- 3) How do instructors assess student understanding of visual representations of DNA on introductory-level molecular biology exams?

## Methods

### Survey design and administration

This current study expands on the methods and data reported in our previous studies about lower-division biology exams [12,21]. Briefly, we developed a survey for instructors who taught lower-division biology courses (100- and 200-level and their equivalents). We administered the survey through the online platform Qualtrics to collect course artifacts (i.e., a course syllabus, a summative course exam, the exam answer key) as well as relevant information about the biology course and the context in which it was taught. We collected responses from 111 biology instructors at 100 unique undergraduate institutions in the United States. For this research, we focused our analysis on a subset of instructors who taught courses with focus on cell and molecular biology. This subset contained 66 instructors who taught courses in introductory-level cell and/or molecular biology (n=32), introductory-level general biology (n=26), genetics (n=3), microbiology (n=2), and cell/molecular biology (n=2). Our sample included broad representation across institution type as defined by Carnegie classifications (S1 Table) and from instructors across career stages (S2 Table). Class sizes in the sample ranged from 4 to 600 students, with an average class size of  $83.8 \pm 111.4$  SD.

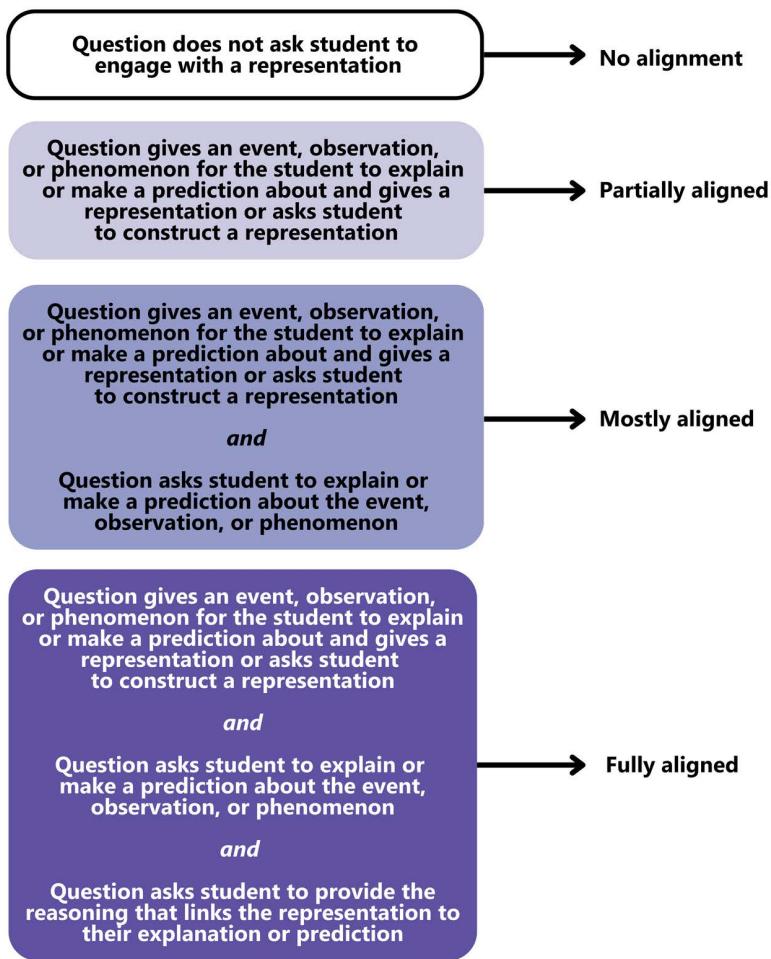
In our survey, we asked instructors to self-report on a series of factors potentially related to the structure and design of their exams (for full description of the survey design, see our previous publication [21]). One self-report measure in the survey was an abbreviated version of the Measurement Instrument for Scientific Teaching (MIST) [8,22]. For this study, we focus on responses to MIST Item #39 which asked instructors to self-report on the approximate frequency that their students were asked to make or interpret models to summarize scientific processes. To ensure that instructors were interpreting the term “models” similarly, the MIST item defined models as tools to summarize a scientific process that include but were not limited to pathways, diagrams, schematics, concept maps, flow charts, illustrative models, and box-and-arrow diagrams [8]. We binned responses to the MIST into three categories for statistical analysis. Instructors who selected that they taught using models “Zero times” were binned into the “Zero times” category. Instructors who selected that they taught using models “1-2 times per semester,” “1 time per month,” or “2-3 times per month” were binned into the “Low frequency” category. Instructors who selected “1-2 times per week,” “3-4 times per week,” or “>4 times per week” were binned into the “High frequency” category.

This research was classified as exempt from human-subjects review by the University of Nebraska–Lincoln (protocol 21082). Informed consent was collected electronically from participants at the start of the survey.

### Coding for item content

Our sample of 66 lower-division cell and molecular biology exams consisted of 2687 items. We used the point values and numbering schemes specified by the instructor to determine the boundaries of individual items. In line with prior recommendations [23], we coded items that shared a common stem and/or used a sub-part numbering scheme (e.g., 3a, 3b, 3c) as a single clustered item. To normalize item point values across courses, we calculated the percentage of total points by dividing individual item point values by the total number of points on the exam and multiplying by 100.

As previously reported for this dataset [12], we coded individual exam item content using existing coding protocols. We used the Three-Dimensional Learning Assessment Protocol (3D-LAP) [23] to code for alignment to each of the scientific practices outlined in the *Framework for K-12 Science Education* [5]. The 3D-LAP uses a set of 2–4 criteria statements for each scientific practice, and we previously operationalized these criteria statements into a scale to indicate degrees of alignment to a scientific practice [21]. For this current study, we narrow our analysis to item alignment to the scientific practice “Developing and Using Models,” and provide a schematic of our operationalized coding scheme in Fig 1. For some analyses, we further collapse this coding scheme to categorize items based on whether or not they contain a visual model. Items that are partially aligned, mostly aligned, or fully aligned to “Developing and Using



**Fig 1. Operationalized 3D-LAP coding scheme for the scientific practice “Developing and Using Models”.** We coded each exam item that asked students to engage with a visual model as being “partially aligned,” “mostly aligned,” or “fully aligned” to the 3D-LAP criteria [23] for “Developing and Using Models.” We coded items that did not ask students to engage with a visual model as “No alignment” to “Developing and Using Models.” The 3D-LAP criteria for selected-response and constructed-response items differed in how they describe student interactions with the item (e.g., the criteria indicates that students are asked to construct an explanation rather than select an explanation). The constructed-response criteria statements were operationalized in an identical manner to the selected-response criteria statements here.

<https://doi.org/10.1371/journal.pone.0317077.g001>

Models” contain a visual model. Items that are not aligned to any of the “Developing and Using Models” criteria do not contain a visual model.

The 3D-LAP was designed so that it could be used across science disciplines, and as such, the 3D-LAP criteria for “Developing and Using Models” provides a broad definition of the term “representation” which encompasses mathematical, graphical, computational, symbolic, and pictorial representations [23]. We broadly view the 3D-LAP characterization of “representations” as being equivalent to “visual models,” and note that visual models may be used and applied differently across science disciplines. In the context of biology, we interpreted mathematical representations to include equations when the focus of the model was to identify the symbols or the relationships between variables in the model. If students were provided an equation and asked to use it to perform a calculation, this was coded as a separate scientific practice (“Using Mathematics and Computational Thinking”). We interpreted graphical representations to include models

that illustrate generalized trends (e.g., a logistic growth curve). When students were asked to interpret graphs containing actual or simulated data, we coded those items as aligned to a separate scientific practice (“Analyzing and Interpreting Data”). We considered symbolic representations to include letters (e.g., the letters “A,” “T,” “C,” and “G” as symbols for nucleotides), and pictorial representations to include images such as those produced by microscopy. We did not observe any instances of computational models in this sample.

In our prior work [12], we coded items for alignment to levels of Bloom’s Taxonomy [14,15] using the Bloom’s Dichotomous Key [24]. We coded the highest Bloom’s level the item was capable of eliciting and the subsequently categorized the cognitive skills “remember,” “understand,” and “apply” as lower-order cognitive skills (LOCS) and “analyze,” “evaluate,” and “create” as higher-order cognitive skills (HOCS). We treated Bloom’s Taxonomy levels as an ordinal scale for statistical analysis.

Full details on item coding procedures and interrater reliability are in the Materials and Methods of our previous publications on this dataset [12,21].

For this research on the subset of 66 cell and molecular biology exams, we used the DNA Landscape [2] to characterize the visual models of DNA that were present in exam items. Visual models were coded based on the size (nucleotide, gene, chromosome) and the level of abstraction (very abstract, elements of shape and abstraction, literal shape). In some cases, items contained multiple representations of DNA within the same figure (e.g., a gene and a plasmid were both represented in the model). When the same DNA representation was referenced across multiple items (e.g., the figure is used in both exam items 13 and 14), both items were counted as having a representation but the figure is only coded once for alignment to the DNA Landscape. Only representations of DNA were coded for alignment to the DNA Landscape—representations of RNA or microscopy images of DNA were not coded. For a small subset of items from exams administered through an online learning management system (e.g., Canvas, Blackboard) (n=13) the item stem contextually referenced an image of DNA but the image file was corrupted during data collection efforts. We considered these items to have images of DNA and included them in our statistical analysis, but we did not code these items for alignment to the DNA Landscape.

To establish interrater reliability for alignment to the DNA Landscape, one member of the coding team (CU) first identified all items in the sample that contained a representation of a nucleotide(s), gene(s), and/or chromosome(s). Items that had missing images, contained images duplicated from previous items on the same exam or that contained identical images that were present on multiple exams were removed, leaving a total of 87 unique items in the coding set. Two raters (CU and CC) independently coded the DNA representations in the 87 items for alignment to the DNA Landscape. Coding disagreements were discussed between members of the research team (CU, CC, DLW, and LKW) and consensus values for each item were retained in the final dataset.

### Coding for item format

As described in our previous publication [21], we coded for item format based on response format (constructed-response and selected-response). Briefly, we classified constructed-response items as those that required students to generate an original response and selected-response items as those that asked students to select from a provided set of responses. Constructed-response items included fill-in-the-blank, short-answer, and essay. Selected-response items included multiple-choice, true-false, multiple true-false, and matching.

### Statistical analysis and data availability

We conducted statistical analysis in R (v 4.3.3) using tidyverse [25] for data processing and visualization. We used lme4 [26] to conduct our generalized linear mixed model. As we had multiple items per instructor in the sample, we included instructor as a random effect in our generalized linear mixed model. De-identified data necessary to recreate the generalized linear mixed model and summary statistics presented in the results are in [S1 Data](#).

## Results

### Biology exams rarely ask students to engage with visual models

We analyzed exams from 66 lower-division cell and molecular biology courses and found that exams rarely ask students to engage with visual representations. Of the 2687 items in the sample, only 16% ( $n=435$  items) provided the opportunity for students to engage with a visual model (Fig 2). While there were very few items with visual models in comparison to the total number of items in the pool, most exams (91%,  $n=60$ ) contained at least one visual model.

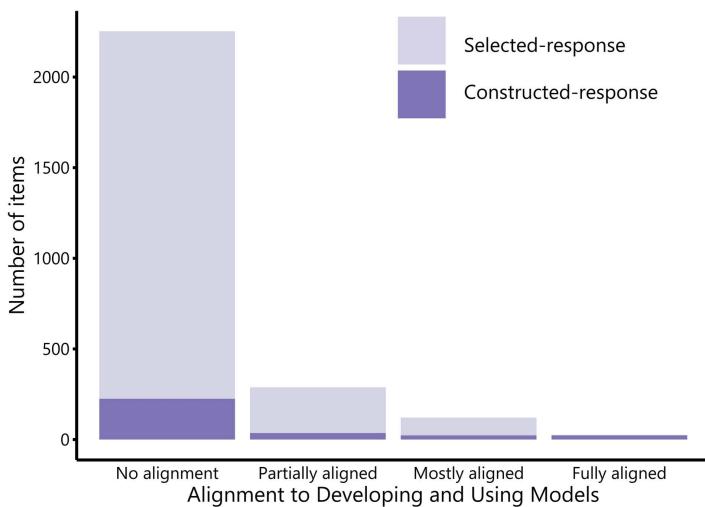
Of the 435 items that contained a visual model, the majority of those (66%,  $n=288$ ) were only *partially aligned* to the scientific practice “Developing and Using Models.” At this level of partial alignment, items include a representation, but students are only asked to engage with the model at a superficial level, such as labeling parts of a familiar diagram (e.g., identifying organelles in a model of a cell). About a quarter of the items containing a visual model (28%,  $n=121$ ) were *mostly aligned* to the scientific practice “Developing and Using Models.” These items asked students to engage with representations to explain a phenomenon beyond memorizable facts but did not explicitly ask students to select or provide the specific reasoning to justify their explanations. Only 6% of items with visual model ( $n=26$ ) asked students to explain a phenomenon using appropriate reasoning in ways that *fully aligned* with the scientific practice “Developing and Using Models.”

Items that contained visual models that were *partially*, *mostly*, or *fully aligned* to “Developing and Using Models” were much more often written in a constructed response format than items without models (Fisher’s exact two-tailed test,  $p<.001$ ). Only 10% of items without models were constructed-response (225 of 2252), yet nearly 90% of items that fully aligned to “Developing and Using Models” were constructed-response items (23 of 26).

When an instructor included a visual model in an item, the items tended to be worth a significantly greater percentage of exam points compared to items without models ( $\beta=0.41$ ,  $t=4.4$ ,  $p<.001$ ; S3 Table). Across instructors, items with models were worth on average 3.0% of exam points ( $SE=0.17$ ) compared 2.4% ( $SE=0.06$ ) for items without models.

### Self-reports on frequency of teaching with models does not predict use of visual models on exams

On average, instructors assigned 19.6% of exam points ( $SE=2.09$ ) to items containing visual models. We compared the percentage of exam points associated with visual models to instructor responses to an item from the Measurement



**Fig 2. Alignment of cell/molecular exam items to the scientific practice “Developing and Using Models”.** Exam items are color coded to distinguish between selected-response and constructed-response format.

<https://doi.org/10.1371/journal.pone.0317077.g002>

Instrument for Scientific Teaching [8] in which instructors self-reported the frequency with which their students were asked to make or interpret visual models (Fig 3). We found that self-reported frequency of teaching with models is not significantly related to the percentage of exam points with visual representations (Type II ANOVA:  $F(2,63) = 1.03$ ,  $p=0.36$ ).

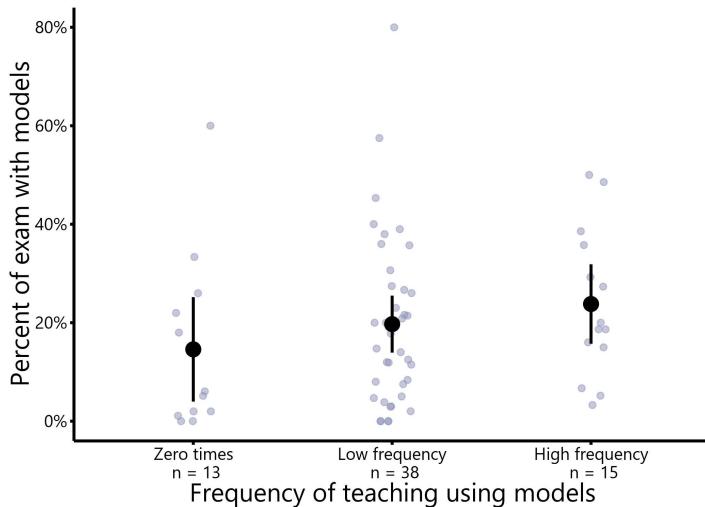
### Items with models are slightly more likely to assess higher-order cognitive skills

As we previously described [12], the majority of items in this dataset only targeted the lower-order cognitive skills “remember,” “understand,” and “apply.” Here, we wanted to see if the distribution of lower-order and higher-order cognitive skills differed between items that did and did not have visual models. We compared the proportion of items with and without visual models at each cognitive skill level (Fig 4). The majority of items with models (87%; 378 of 435) and without models (96%; 2157 of 2252) assessed lower-order cognitive skills; however, a chi-square test indicates a statistically significant association between containing a model and testing higher-order cognitive skills, suggesting that these two item features are not independent [ $\chi^2(1) = 53.9$ ,  $p < .001$ ].

Although there were very few items targeting the higher levels of Bloom’s Taxonomy ( $n=131$  items at the levels of “analyze,” “evaluate,” or “create”), these items tended to have a greater degree of alignment to the criteria for the scientific practice “Developing and Using Models” (Fig 5). There is no significant difference between the Bloom’s level of items that were not aligned or were only partially aligned to “Developing and Using Models” (Table 1). In contrast, items mostly aligned or fully aligned to the scientific practice had significantly higher Bloom’s levels compared to items without models.

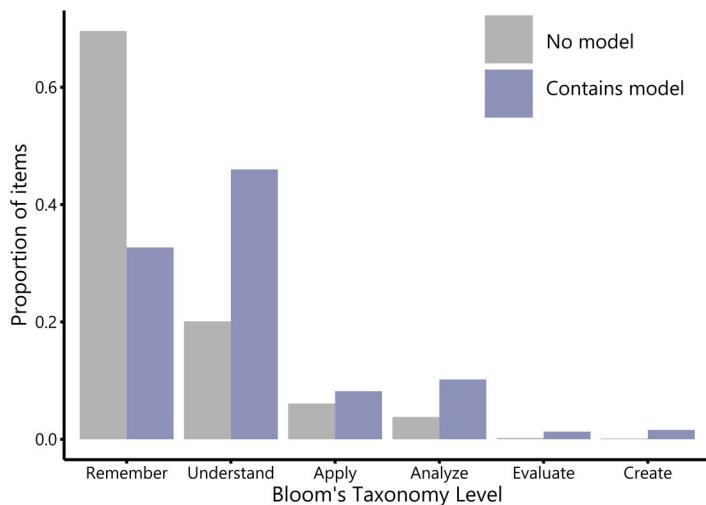
### Most representations of DNA on biology exams are very abstract

We examined a subset of exam items with visual models of DNA. Across the 66 cell and molecular biology exams in the sample, 55% of exams ( $n=36$ ) had at least one representation of DNA. In total, there were 146 exam items with DNA representations which comprised only 5% of the total number of items in the sample. Of these, 87 contained a unique representation and 46 duplicated images from prior items on the same exam. There were an additional 13 items from



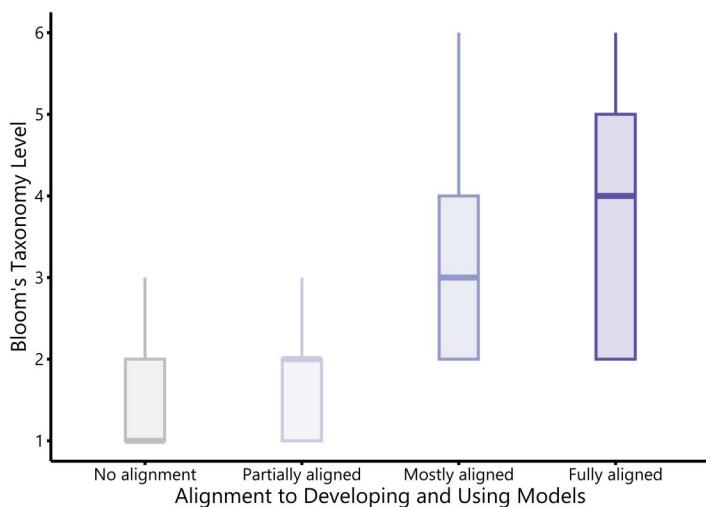
**Fig 3. Relationship between self-reported frequency of teaching using models and the percent of exam points that contain a visual model.** Teaching practices were self-reported on item #39 from the Measurement Instrument of Scientific Teaching (MIST) [8]. Instructors who reported that they taught using models zero times during a semester were binned into “Zero times”. Instructors who reported teaching using models less than or equal to 2-3 times per month were binned into “Low frequency.” Instructors who reported teaching using models greater than or equal to 1-2 times per week were binned into “High frequency.” Individual instructors are represented with the smaller purple dots. Larger black dots represent the mean and lines represent the 95% confidence interval.

<https://doi.org/10.1371/journal.pone.0317077.g003>



**Fig 4. Proportion of cognitive skills assessed by items that did not and did contain visual models.** Proportions calculated based on the total number of items with no model (n=2252) and the total number of items that contained a model (n=435).

<https://doi.org/10.1371/journal.pone.0317077.g004>



**Fig 5. Box-and-whisker plots of the Bloom's Taxonomy levels of items at each degree of alignment to the scientific practice “Developing and Using Models”.** Solid bars within each box represent the median value, boxes represent the interquartile range, and whiskers represent 1.5 times the interquartile range. Bloom's Taxonomy levels 1–6 correspond to the cognitive skills “remember,” “understand,” “apply,” “analyze,” “evaluate,” and “create,” respectively.

<https://doi.org/10.1371/journal.pone.0317077.g005>

exams administered through an online learning management system (e.g., Canvas, Blackboard) for which the item stem indicated an image but the image files were corrupted during data collection efforts.

We categorized the 87 items with unique representations of DNA for their alignment to the DNA Landscape. Most of the representations of DNA aligned to a single location on the DNA Landscape, but there were 12 items that contained two representations, and as such there were a total of 98 DNA representations in the sample. We found that there were examples of DNA representations at each scale of nucleotide, genes, and chromosomes, but across these scales, most of the representations were very abstract (Fig 6). The majority of the images (59%, 58 of 98) were very abstract representations

**Table 1. Linear regression model predicting the Bloom's Taxonomy level of items based on alignment to “Developing and Using Models”.**

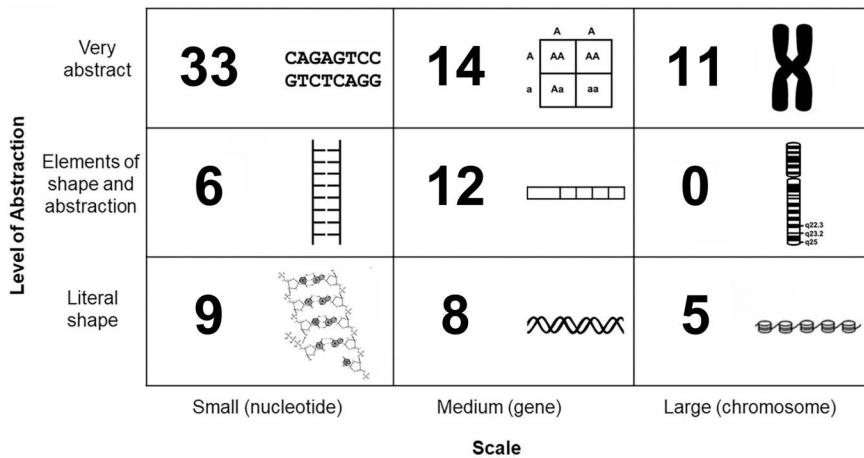
Fixed effects	$\beta$	Standard Error	df	t	p
Intercept	1.57	0.06	44	26.3	<.0001
Partially aligned	0.09	0.05	2638	1.9	0.052
Mostly aligned	1.10	0.07	2625	15.5	<.0001
Fully aligned	2.04	0.15	2667	13.2	<.0001

Model: Bloom's level ~ Alignment to “Developing and Using Models” + (1|instructor)

Reference group: No alignment

$R^2 = 0.36$

<https://doi.org/10.1371/journal.pone.0317077.t001>



**Fig 6. Categorization of the representations of DNA in molecular biology exams based on their alignment to the scale and abstraction axes of the DNA Landscape.** There were unique DNA representations on 87 exam items, and 12 of those items had representations that were coded in multiple places on the DNA Landscape.

<https://doi.org/10.1371/journal.pone.0317077.g006>

in which nucleotides, genes, and chromosomes were represented with letters and symbols. Notably, there were not any representations of gene maps in which the representation included the location of an actual gene or the distance between multiple genes on a chromosome.

The majority of items with representations of DNA tended to assess lower-order cognitive skills (93%, n = 136). In comparison, while the majority of items with non-DNA representations also assessed lower-order cognitive skills (83%, n = 239), there was a statistically significant difference between the proportion of lower-order and higher-order skills assessed between the items with DNA and non-DNA representations (Fisher's exact two-tailed test, p = .004). Items with representations of DNA often asked students to complete rote tasks, such as selecting a sequence of complementary nucleotide base pairs, matching representations of mitotic stages, or identifying nucleotides within a series of monomer chemical structures. When comparing the cognitive skills assessed, the items with representations of DNA were statistically different from the rest of the larger pool of items with representations.

## Discussion

Using models is a key scientific practice and competency for biology students [5,27] and visual literacy skills appear in nearly a third of nationally-endorsed learning objectives for introductory biology [28], yet we found that these skills are rarely assessed on undergraduate biology exams. Thus, our findings underscore the disconnect between the goals and

priorities of national calls and what is practiced in undergraduate biology courses. While most introductory-level biology exams predominately assess lower-order cognitive skills, exam items containing models were slightly more likely to ask students to engage in higher-order thinking. Here, we highlight the utility of models as a means to engage students' higher-order cognitive skills, call for more robust assessment of students' molecular biology visual literacy skills, and discuss ways to address challenges to using models to teach and assess visual literacy.

### What might be constraining assessment of visual literacy skills in introductory biology?

Our analysis revealed that introductory-level biology exams rarely asked students to use visual literacy skills to engage with models. Overall, there were very few items that included a visual model. Even though our findings indicate that visual models can be a means to engage higher-order cognitive skills, many items containing a model only targeted memorization and recall. We found that when instructors included visual models on their exams, many of those items were only asking students to label familiar structures. While labeling can help students learn about molecular structures [29], labeling alone does not address the suite of visual literacy skills that is needed for expertise in molecular biology. To gauge students' visual literacy and ability to use higher-order thinking, instructors should create assessments that ask students to evaluate the power and limitations of models, construct models to explain a concept, translate between multiple representations of the same concept, or to compare representations at various scales [20]. For example, an instructor could present students with an item that contains a written description of a particular point mutation and ask them to explain why an abstract visual model of a chromosome is not an effective model to represent the mutation and to draw a new model that would better represent the mutation. Such an assessment item would ask students to evaluate the limitations of a chromosome model, construct a more appropriate model to explain a mutation, ask students to translate between written and graphical representations of a mutation, and compare models of DNA across different scales. Despite the importance of these visual literacy skills in biology, we found that very few exams specifically targeted these skills.

One potential explanation for these results is that instructors have very few examples of what a visual literacy assessment "looks like." While the Drawing-to-Learn framework provides a robust set of recommendations for facilitating assessment of visual literacy skills [7], there are few examples of visual literacy assessments that instructors can readily use or adapt in their biology courses. Instructors who rely on assessments generated by curriculum publishers may encounter the challenge that most figures in "test bank" items are often only asking students to add labels to a diagram. Instructors may turn to published biology concept assessments or concept inventories (e.g., [30–34]), but models are rarely the focus of such assessments and students may be interpreting the models on concept assessments in different ways than the assessment developers intend [35]. To date, there are no valid and reliable assessment instruments that are specifically intended to measure a suite of visual literacy skills in molecular biology. We emphasize the need for biology education researchers and assessment developers to create assessment instruments that can be used to measure biology visual literacy skills. Such visual literacy assessment instruments may be directly used by biology instructors in their classes or may serve as inspiration for how instructors can create their own visual literacy assessments.

We also consider our results in light of the limitations that often constrain how instructors are able to design and administer their exams. Instructors often have limited time to write and grade exams [36], and it can take a substantial amount of time to write and grade three-dimensional exam items that address scientific practices (such as the practice "Developing and Using Models") [37]. Grading exams can be especially time consuming when instructors use constructed-response (i.e., open-ended) items where students can create models and articulate their reasoning in a written format. When asking students to create their own models on exams is not feasible, instructors may assess student understanding of models if they integrate pre-fabricated models as multiple-choice options.

Another potential explanation of our findings is that instructors may be limited in their ability to locate, create, or reproduce models for an exam context. To ensure that students are engaging visual literacy skills and not simply memorizing a familiar diagram for an exam, instructors may need to locate figures that are similar but different from the exact model

they have used in class, which can be challenging. There may be material and financial constraints involved with creating physical copies of the figures for in-person exams, especially if the figures include color. While there are certainly time and resource constraints in grading constructed-response questions, having students create their own models to demonstrate their understanding of visual literacy conventions circumvents the need for instructors to invest time and resources into locating, creating, or reproducing models themselves.

### **Backward designing biology curriculum to emphasize visual literacy in both teaching and assessment**

We found that there was a mismatch between the frequency with which instructors reported teaching using models and the percent of their exams that contained models. In some cases, instructors who reported that they frequently taught using models ( $\geq 1\text{--}2$  times per week) rarely incorporated models into their exams, while conversely, some instructors who reported never directly teaching models had used models on upwards of a third of their exam. This mismatch may reflect how instructors are thinking differently about teaching and assessing using models, and we anticipate that this mismatch can lead to confusion and frustration for students. If exams are misaligned to instruction, students may be left wondering why they spent time on content they were not tested on and they may question if it is fair that they were tested on content they were never directly taught. We recommend taking a Backward Design approach to curriculum development to address this mismatch between instruction and assessment of modeling. Following principles of Backward Design, course learning goals should be reflected in what is assessed and what is taught in a course [38]. Accordingly, if using models and visual literacy skills are prioritized learning outcomes in a biology course, modeling and visual literacy should be incorporated into exams and integrated into instruction. Instructors may also use point value of exam items to signal to students that modeling and visual literacy are priorities. Higher point values on items with models may help to strengthen students' association that modeling is a valuable and important skill for them to learn.

### **Using models is a method to engage students in higher-order thinking**

Even though visual models have the potential to be powerful tools to engage student reasoning with higher-order cognitive skills, when instructors incorporated visual models into their exam items, these items were often only targeting lower-order cognitive skills, such as labeling structures on a familiar diagram. While items with models can certainly be used to assess higher-order cognitive skills, our findings highlight that instructors may not be capitalizing upon using exams to engage students in model-based reasoning. We encourage instructors to consider ways to ask their students to critically analyze and evaluate the models they are presented. For example, instructors can ask students to explain how and why a biologist might use abstraction to simplify a model and why such simplifications may be useful for conveying certain information. Instructors can also ask students to consider when such simplifications obscure some relevant information and when alternate models may be more useful in a particular context.

We also encourage instructors to consult the 3D-LAP criteria statements for "Developing and Using Models" (reproduced in Fig 1) as methods for designing exam items that target critical thinking. A key feature of the 3D-LAP criteria is the emphasis on asking students to productively use models to engage in reasoning. Following the 3D-LAP guidelines, it is not only important for students to be able to use a model to explain a phenomena—students need to be able to explain the reason *why* that model is useful for explaining a phenomenon. This reasoning component provides important information for instructors to gauge how their students are thinking about models, yet this reasoning component is often absent from exam items [21,39,40]. As full alignment to the 3D-LAP criteria for "Developing and Using Models" was significantly associated with higher levels of Bloom's Taxonomy, we suggest instructors consider ways to ask students to demonstrate their reasoning. Instructors who have the time and resources to grade open-ended questions can incorporate explicit directions in their items to elicit reasoning in a written format. Students may be more likely to demonstrate reasoning about models if they are prompted to describe how they arrived at a conclusion or to support their argument using evidence. If

instructors are writing closed-ended questions, we recommend formats that enable multiple points of student input, such as multiple-true-false questions [31,41,42] or two-tiered questions that allow students to first use a model to explain a phenomenon and then complete a follow-up question in which they can demonstrate their reasoning.

### Emphasizing the value of abstract models in biology

We were not surprised to see that most representations of DNA in our sample were abstract models. Biologists frequently use abstract models because the shapes and symbols in abstract models highlight the most salient structures and features in simple-to-draw ways. These abstract models are often easy for instructors to recreate for an exam context and tend to be well-suited for grayscale printing. Instructors can use abstract models to create meaningful opportunities for model-based reasoning [7], yet we saw few examples of this in how abstract models were used on exams. Exam items with abstract representations mostly focused on rote tasks, such as matching and labeling. In both teaching and assessment, we encourage instructors to consider how to leverage the value of abstract models and ask students to think deeply about the purpose and efficacy of such representations. For example, in addition to asking students to use letters to complete a complementary strand of nucleotides, instructors can ask students to explain the reason why letters are a useful model for that context (as compared to the ladder model or the nucleotide chemical structures). Instructors can alternatively ask students to explain a situation when letters are not a particularly useful model. Students often encounter abstract models, so we encourage instructors to consider ways to ask students to engage with *why* and *how* those abstract models convey meaning in biology.

### Limitations

We relied on exams to make inferences about what content students were encountering in their coursework. This methodology is common in education research studies [43–46], but we acknowledge that exams do not capture the full range of assessment types that can be administered in introductory-level biology lecture courses. It is possible that students engaged with models in formative assessments or other types of summative assessments, such as projects or presentations, which were not reflected in our sampling. While our research focused on lecture courses, we anticipate that students are also likely to encounter visual models in both instruction and assessment in lower-division biology lab courses.

### Conclusion

Using models is a key skill for biologists, yet our research identified that introductory-level students often have few opportunities in their course exams to demonstrate this ability. Despite the importance of visual models in molecular biology, models are rarely incorporated into introductory-level molecular biology exams. When models are present in exam items, they can have the potential to elicit higher-order cognitive skills yet are often only used to assess students' ability to recall rote information. This finding signals a greater need for assessments to incorporate visual literacy skills and to engage students in demonstrating their understanding of why and how models are useful tools in biology.

### Supporting information

#### S1 Data. De-identified data.

(CSV)

#### S1 Table. Institutional Carnegie classifications.

(DOCX)

#### S2 Table. Self-reported teaching experience of undergraduate biology instructors.

(DOCX)

**S3 Table. Linear mixed model predicting the point value of items with visual models.**  
(DOCX)

## Acknowledgments

We thank the instructors who participated in this research. We thank Noah Courtney for contributions in early conceptualization of this project.

## Author contributions

**Conceptualization:** Crystal Uminski, L. Kate Wright, Dina L. Newman.

**Data curation:** Crystal Uminski.

**Formal analysis:** Crystal Uminski.

**Funding acquisition:** Crystal Uminski, Brian A. Couch, L. Kate Wright, Dina L. Newman.

**Investigation:** Crystal Uminski.

**Methodology:** Crystal Uminski.

**Project administration:** Brian A. Couch, L. Kate Wright, Dina L. Newman.

**Resources:** Crystal Uminski.

**Software:** Crystal Uminski.

**Supervision:** Brian A. Couch, L. Kate Wright, Dina L. Newman.

**Validation:** Crystal Uminski, Christian Cammarota, L. Kate Wright, Dina L. Newman.

**Visualization:** Crystal Uminski.

**Writing – original draft:** Crystal Uminski.

**Writing – review & editing:** Crystal Uminski, Christian Cammarota, Brian A. Couch, L. Kate Wright, Dina L. Newman.

## References

1. Schönborn KJ, Anderson TR. The importance of visual literacy in the education of biochemists. *Biochem Mol Biol Educ*. 2006;34(2):94–102. <https://doi.org/10.1002/bmb.2006.49403402094> PMID: 1638645
2. Wright LK, Wrightstone E, Trumppore L, Steele J, Abid DM, Newman DL. The DNA Landscape: development and application of a new framework for visual communication about DNA. *CBE Life Sci Educ*. 2022;21(3):ar47. <https://doi.org/10.1187/cbe.22-01-0007> PMID: 35816448
3. Offerdahl EG, Arneson JB, Byrne N. Lighten the load: scaffolding visual literacy in biochemistry and molecular biology. *CBE Life Sci Educ*. 2017;16(1):es1. <https://doi.org/10.1187/cbe.16-06-0193> PMID: 28130273
4. Trumbo J. Visual literacy and science communication. *Sci Commun*. 1999;20(4):409–25. <https://doi.org/10.1177/1075547099020004004>
5. National Research Council. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, D.C.: National Academies Press; 2012. <https://doi.org/10.17226/13165>
6. American Association for the Advancement of Science. Vision and Change in Undergraduate Biology Education: A Call to Action. Washington, DC: AAAS; 2011. Available from: <https://live-visionandchange.pantheonsite.io/wp-content/uploads/2011/03/Revised-Vision-and-Change-Final-Report.pdf>
7. Quillin K, Thomas S. Drawing-to-learn: a framework for using drawings to promote model-based reasoning in biology. *CBE Life Sci Educ*. 2015;14(1):es2. <https://doi.org/10.1187/cbe.14-08-0128> PMID: 25713094
8. Durham MF, Knight JK, Couch BA. Measurement Instrument for Scientific Teaching (MIST): a tool to measure the frequencies of research-based teaching practices in undergraduate science courses. *CBE Life Sci Educ*. 2017;16(4):ar67. <https://doi.org/10.1187/cbe.17-02-0033> PMID: 29196428
9. Bain K, Bender L, Bergeron P, Caballero MD, Carmel JH, Duffy EM, et al. Characterizing college science instruction: the three-dimensional learning observation protocol. *PLoS One*. 2020;15(6):e0234640. <https://doi.org/10.1371/journal.pone.0234640> PMID: 32544166
10. National Research Council. Assessment in Support of Instruction and Learning: Bridging the Gap Between Large-Scale and Classroom Assessment - Workshop Report. Washington, D.C.: National Academies Press; 2003. <https://doi.org/10.17226/10802>

11. Scouller K. The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *High Educ.* 1998;35(4):453–72. <https://doi.org/10.1023/a:1003196224280>
12. Uminski C, Burbach SM, Couch BA. Undergraduate biology lecture courses predominantly test facts about science rather than scientific practices. *CBE Life Sci Educ.* 2024;23(2):ar19. <https://doi.org/10.1187/cbe.23-12-0244> PMID: 38640405
13. Arneson JB, Offerdahl EG. Visual literacy in Bloom: using Bloom's Taxonomy to support visual learning skills. *CBE Life Sci Educ.* 2018;17(1):ar7. <https://doi.org/10.1187/cbe.17-08-0178> PMID: 29351910
14. Anderson LW, Krathwohl DR, Bloom BS. A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York, NY: Longman; 2001.
15. Bloom BS, Krathwohl DR, Masia BB. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. New York, NY: McKay; 1956.
16. Allen D, Tanner K. Approaches to cell biology teaching: questions about questions. *Cell Biol Educ.* 2002;1(3):63–7. <https://doi.org/10.1187/cbe.02-07-0021> PMID: 12459794
17. Bissell AN, Lemons PP. A new method for assessing critical thinking in the classroom. *BioScience.* 2006;56(1):66. [https://doi.org/10.1641/0006-3568\(2006\)056\[0066:anmfac\]2.0.co;2](https://doi.org/10.1641/0006-3568(2006)056[0066:anmfac]2.0.co;2)
18. Moon S, Jackson MA, Doherty JH, Wenderoth MP. Evidence-based teaching practices correlate with increased exam performance in biology. *PLoS One.* 2021;16(11):e0260789. <https://doi.org/10.1371/journal.pone.0260789> PMID: 34847190
19. Zheng AY, Lawhorn JK, Lumley T, Freeman S. Assessment. Application of Bloom's Taxonomy debunks the "MCAT myth". *Science.* 2008;319(5862):414–5. <https://doi.org/10.1126/science.1147852> PMID: 18218880
20. Schönborn KJ, Anderson TR. Bridging the educational research-teaching practice gap: Foundations for assessing and developing biochemistry students' visual literacy. *Biochem Mol Biol Educ.* 2010;38(5):347–54. <https://doi.org/10.1002/bmb.20436> PMID: 21567860
21. Uminski C, Couch BA. Identifying factors associated with instructor implementation of three-dimensional assessment in undergraduate biology courses. *PLoS One.* 2024;19(10):e0312252. <https://doi.org/10.1371/journal.pone.0312252> PMID: 39436884
22. Durham MF, Knight JK, Bremers EK, DeFreece JD, Paine AR, Couch BA. Student, instructor, and observer agreement regarding frequencies of scientific teaching practices using the Measurement Instrument for Scientific Teaching-Observable (MISTO). *Int J STEM Educ.* 2018;5(1):31. <https://doi.org/10.1186/s40594-018-0128-1> PMID: 30631721
23. Laverty JT, Underwood SM, Matz RL, Posey LA, Carmel JH, Caballero MD, et al. Characterizing College Science Assessments: The Three-Dimensional Learning Assessment Protocol. *PLoS One.* 2016;11(9):e0162333. <https://doi.org/10.1371/journal.pone.0162333> PMID: 27606671
24. Semsar K, Casagrand J. Bloom's dichotomous key: a new tool for evaluating the cognitive difficulty of assessments. *Adv Physiol Educ.* 2017;41(1):170–7. <https://doi.org/10.1152/advan.00101.2016> PMID: 28235756
25. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. *J Open Source Softw.* 2019;4(43):1686. <https://doi.org/10.21105/joss.01686>
26. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models Usinglme4. *J Stat Soft.* 2015;67(1):1–48. <https://doi.org/10.18637/jss.v067.i01>
27. AAAS. Vision and Change in Undergraduate Biology Education: A Call to Action. Washington, D.C.: American Association for the Advancement of Science; 2011 pp. 100. Available from: [www.visionandchange.org](http://www.visionandchange.org)
28. Hennessey KM, Freeman S. Nationally endorsed learning objectives to improve course design in introductory biology. *PLoS One.* 2024;19(8):e0308545. <https://doi.org/10.1371/journal.pone.0308545> PMID: 39146309
29. Rotbain Y, Marbach-Ad G, Stavy R. Understanding molecular genetics through a drawing-based activity. *J Biol Educ.* 2005;39(4):174–8. <https://doi.org/10.1080/00219266.2005.9655992>
30. Couch BA, Wood WB, Knight JK. The Molecular Biology Capstone Assessment: a concept assessment for upper-division molecular biology students. *CBE Life Sci Educ.* 2015;14(1):ar10. <https://doi.org/10.1187/cbe.14-04-0071> PMID: 25713098
31. Couch BA, Wright CD, Freeman S, Knight JK, Semsar K, Smith MK, et al. GenBio-MAPS: a programmatic assessment to measure student understanding of Vision and Change core concepts across general biology programs. *CBE Life Sci Educ.* 2019;18(1):ar1. <https://doi.org/10.1187/cbe.18-07-0117> PMID: 30681904
32. Kalas P, O'Neill A, Pollock C, Birol G. Development of a meiosis concept inventory. *CBE Life Sci Educ.* 2013;12(4):655–64. <https://doi.org/10.1187/cbe.12-10-0174> PMID: 24297292
33. Newman DL, Snyder CW, Fisk JN, Wright LK. Development of the Central Dogma Concept Inventory (CDCI) Assessment Tool. *CBE Life Sci Educ.* 2016;15(2):ar9. <https://doi.org/10.1187/cbe.15-06-0124> PMID: 27055775
34. Smith MK, Wood WB, Knight JK. The Genetics Concept Assessment: a new concept inventory for gauging student understanding of genetics. *CBE Life Sci Educ.* 2008;7(4):422–30. <https://doi.org/10.1187/cbe.08-08-0045> PMID: 19047428
35. Uminski C, Newman DL, Wright LK. Probing visual literacy skills reveals unexpected student conceptions of chromosomes. *CBE Life Sci Educ.* 2025;24(1):ar17. <https://doi.org/10.1187/cbe.24-07-0176> PMID: 39982982
36. Wright CD, Huang A, Cooper K, Brownell S. Exploring differences in decisions about exams among instructors of the same introductory biology course. *Int J Scholarsh Teach Learn.* 2018;12(2). <https://doi.org/10.20429/ijstl.2018.120214>

37. Nelson PC, Matz RL, Bain K, Fata-Hartley CL, Cooper MM. Characterizing faculty motivation to implement three-dimensional learning. *Discip Interdiscip Sci Educ Res.* 2023;5(1). <https://doi.org/10.1186/s43031-023-00079-0>
38. Wiggins GP, McTighe J. *Understanding by Design*. Alexandria, VA: Association for Supervision and Curriculum Development; 2005.
39. Laverty JT, Underwood SM, Matz RL, Posey LA, Carmel JH, Caballero MD, et al. Comment on “Analyzing the Role of Science Practices in ACS Exam Items”. *J Chem Educ.* 2017;94(6):673–4. <https://doi.org/10.1021/acs.jchemed.7b00170>
40. Reed JJ, Brandriet AR, Holme TA. Analyzing the role of science practices in ACS exam items. *J Chem Educ.* 2016;94(1):3–10. <https://doi.org/10.1021/acs.jchemed.6b00659>
41. Brassil CE, Couch BA. Multiple-true-false questions reveal more thoroughly the complexity of student thinking than multiple-choice questions: a Bayesian item response model comparison. *IJ Stem Ed.* 2019;6(1). <https://doi.org/10.1186/s40594-019-0169-0>
42. Couch BA, Hubbard JK, Brassil CE. Multiple-true-false questions reveal the limits of the multiple-choice format for detecting students with incomplete understandings. *BioScience.* 2018;68(6):455–63. <https://doi.org/10.1093/biosci/biy037>
43. Matz RL, Fata-Hartley CL, Posey LA, Laverty JT, Underwood SM, Carmel JH, et al. Evaluating the extent of a large-scale transformation in gateway science courses. *Sci Adv.* 2018;4(10):eaau0554. <https://doi.org/10.1126/sciadv.aau0554> PMID: 30397646
44. Momsen JL, Long TM, Wyse SA, Ebert-May D. Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE Life Sci Educ.* 2010;9(4):435–40. <https://doi.org/10.1187/cbe.10-01-0001> PMID: 21123690
45. Stowe RL, Esselman BJ, Ralph VR, Ellison AJ, Martell JD, DeGlopper KS, et al. Impact of maintaining assessment emphasis on three-dimensional learning as organic chemistry moved online. *J Chem Educ.* 2020;97(9):2408–20. <https://doi.org/10.1021/acs.jchemed.0c00757>
46. Stowe RL, Scharlott LJ, Ralph VR, Becker NM, Cooper MM. You are what you assess: the case for emphasizing chemistry on chemistry assessments. *J Chem Educ.* 2021;98(8):2490–5. <https://doi.org/10.1021/acs.jchemed.1c00532>