This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.

JOURNAL OF CHEMICAL EDUCATION
Article
pubs.acs.org/jchemeduc

Adapting Assessment Tasks To Support Three-Dimensional Learning

Sonia M. Underwood,[*,†] Lynmarie A. Posey,[‡] Deborah G. Herrington,[§] Justin H. Carmel,[†] and Melanie M. Cooper[‡]

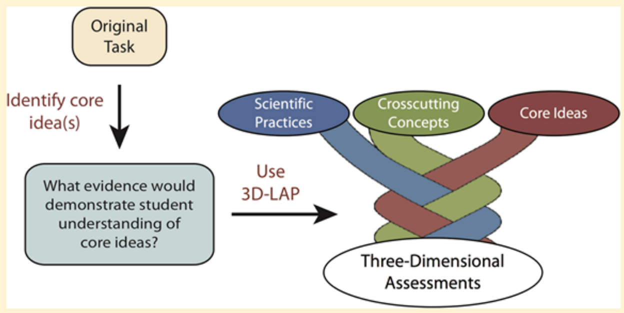
[†]Department of Chemistry & Biochemistry and STEM Transformation Institute, Florida International University, 11200 SW Eighth Street, Miami, Florida 33199, United States

[‡]Department of Chemistry, Michigan State University, 578 South Shaw Lane, East Lansing, Michigan 48824, United States

[§]Department of Chemistry, Grand Valley State University, 1 Campus Drive, 312 Padnos Hall, Allendale, Michigan 49401, United States

Supporting Information

ABSTRACT: As chemists, we understand that science is more than a set of disconnected facts. It is a way of investigating and understanding our natural world that involves things like asking questions, analyzing data, identifying patterns, constructing explanations, developing and using models, and applying core concepts to other situations. This paper uses the concept of three-dimensional (3D) learning, presented in *A Framework for K-12 Science Education*, to reconceptualize and develop assessment items that require students to integrate chemistry core ideas with scientific practices and crosscutting concepts. Developing 3D assessments from scratch is time-consuming and beyond the scope of most faculty work. Here we present an alternate approach: We provide a detailed description of ways in which instructors can take current assessment questions and modify them to align with three-dimensional learning by focusing on the evidence that is sought about what students know and can do with their knowledge.



KEYWORDS: *Chemical Education Research, First-Year Undergraduate/General, Second-Year Undergraduate, Curriculum, Student-Centered Learning, Learning Theories, Testing/Assessment*

FEATURE: Chemical Education Research

Most instructors would agree that the ultimate goal of education is to help students develop deep, transferrable knowledge; however, there is less agreement about what this would look like in practice. In our earlier paper in this project,[1] we presented an approach to curriculum development based on the current understanding of how learning occurs.[2,3] That is, knowledge should be connected into a coherent framework organized around ideas that are central to the discipline to make existing knowledge accessible and useful and to provide a foundation upon which new knowledge can be built. This approach, based on the vision offered by *A Framework for K-12 Science Education*[3] (referred to as the *Framework* in this paper), begins with recognizing that most important topics in chemistry can be understood in terms of one or more core ideas, such as the relationship between the structure of molecules and their properties.[1] As part of a larger project to transform the introductory science courses at Michigan State University (MSU), and based on a prior curriculum transformation project,[4] we proposed four chemistry core ideas (Box 1)[5] and provided examples of how various topics can be supported by and connected to these core ideas in our previous paper.[1]

Development of such a connected understanding, however, requires not only restructuring of curricula to support students in making the connections necessary for construction of a coherent framework for their knowledge; it also necessitates changes in instructional activities and in how learning is assessed, both formatively and summatively.[5] Furthermore, knowledge alone is not sufficient to meet the demands of a workplace that places a premium on how that knowledge is used, which has necessitated redefining science proficiency.[6] To help students learn to use their knowledge in ways that reflect the work of scientists and engineers, the *Framework* describes *scientific and engineering practices* (Box 1). These practices include activities such as developing and using models to predict and explain phenomena and constructing arguments from evidence, along with more recognizable elements of inquiry-based education, such as asking questions, planning and carrying out experiments, and analyzing and interpreting data. The *Framework* also describes crosscutting concepts, those ideas that transcend multiple disciplines and can help students make connections across disciplines, such as cause and effect: mechanism

Received: August 20, 2017
Revised: November 2, 2017
Published: December 12, 2017

© 2017 American Chemical Society and Division of Chemical Education, Inc.

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207−217

Box 1. Components of Three-Dimensional Learning[3]

(1) Core Ideas in Chemistry[1,3]
- (1A) Electrostatic and Bonding Interactions
- (1B) Atomic/Molecular Structure and Properties
- (1C) Energy: Macroscopic, Atomic/Molecular, and Quantum Mechanical Energy Levels and Changes
- (1D) Change and Stability in Chemical Systems

(2) Scientific and Engineering Practices[3]
- (2A) Asking Questions (for Science) and Defining Problems (for Engineering)
- (2B) Developing and Using Models
- (2C) Planning and Carrying Out Investigations
- (2D) Analyzing and Interpreting Data
- (2E) Using Mathematics and Computational Thinking
- (2F) Constructing Explanations (for Science) and Designing Solutions (for Engineering)
- (2G) Engaging in Argument from Evidence
- (2H) Obtaining, Evaluating, and Communicating Information

(3) Crosscutting Concepts[3]
- (3A) Patterns
- (3B) Cause and Effect: Mechanism and Explanation
- (3C) Scale, Proportion, and Quantity
- (3D) Systems and System Models
- (3E) Energy and Matter: Flows, Cycles, and Conservation
- (3F) Structure and Function
- (3G) Stability and Change

and explanation, patterns, and structure and function (Box 1). Blending these three strands (core ideas, scientific and engineering practices, and crosscutting concepts) in both instruction and assessment produces what is known as three-dimensional (3D) learning. In this paper, we address the development of 3D assessment items and discuss how such items might be used in instruction, and why they provide improved evidence about student understanding.

■ ASSESSMENT AS AN EVIDENTIARY ARGUMENT

Changes in curricula to support 3D learning must be accompanied by designing appropriate assessments to align with the new instructional goals. As has been noted on numerous occasions,[7−12] the nature of the assessments in a course sends a strong signal to students about what is important, meaning that the ways in which students approach learning are often motivated by what will be on the test. Therefore, using assessment items that focus on single ideas can inadvertently lead to fragmentation of knowledge, despite instructors' best efforts to connect course content to core ideas as discussed in a preceding paper.[1] Furthermore, correct responses to such assessment items provide an incomplete picture of student understanding and can be mistakenly interpreted as evidence that students both possess the underlying knowledge and can apply it appropriately to support the correct response, when in reality they have learned a procedure by rote, or are employing a heuristic,[13,14] that may have no basis in actual scientific principles, to reach an answer. In order for students to construct a robust knowledge framework, both formative and summative assessment tasks must require students to link multiple ideas.

Knowing What Students Know[15] presents a model for assessment design, the assessment triangle (Figure 1), that is applicable

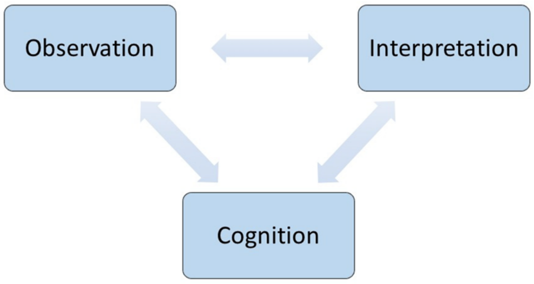


Figure 1. Assessment triangle.[15]

to all assessments irrespective of the purpose because assessment is a process of reasoning from evidence to make claims about student learning.[16,17] The assessment triangle consists of three interacting elements: cognition, observation, and interpretation. Cognition refers to the underlying cognitive model of learning being used to design the assessments. It describes how students learn in a domain and thereby specifies the knowledge and skills that must be measured in order to assess student competence/achievement. In our case the cognitive model is 3D learning. Observations are the assessment tasks designed to provide data on what students know and can do with their knowledge, which become evidence about student knowledge when interpreted in the context of a model for learning.

In order to apply the assessment triangle to our target of 3D learning, we need to consider the underlying premises. 3D learning rests on a foundation of what is known about (1) the ways in which experts in science organize and use their knowledge and (2) learning as a developmental process in which students can build increasingly sophisticated understanding over time by connecting ideas when provided with proper support. Consequently, assessment tasks designed to measure students' progress toward the goal of 3D learning must elicit evidence of the connections that students make between ideas as they build knowledge structures that are more expert in character, and the ways in which students use this knowledge that reflect the work of science. There are a number of approaches that have been used to design assessments in order to provide explicit evidence of student understanding, including evidence-centered design (ECD)[18] and construct modeling.[19] Both offer approaches to addressing the components of the assessment triangle and are being used to develop formative assessments to support 3D learning aligned with the *Next-Generation Science Standards* (NGSS).[20,21] However, both ECD and construct modeling are time-consuming and iterative processes to develop assessments that are not practical for most instructors. In the next section, we will take a more pragmatic approach showing how the principles of ECD can be adapted to develop existing assessments to support 3D learning.

■ DEVELOPING ASSESSMENTS THAT ELICIT EVIDENCE OF 3D LEARNING

As noted, there are a number of approaches that can be used to design 3D assessment tasks, the most daunting of which is to design such tasks from "scratch" to target a desired learning outcome or performance expectation. An excellent review of this process is provided by Harris et al. (2016). It involves a complex process of domain analysis requiring "unpacking" of each of the three dimensions, followed by the creation of new integrated dimension maps. Learning performances (smaller descriptions of 3D tasks) that emerge from these maps are

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207−217

Original Question:

Consider the acid ionization reaction of carbonic acid (H_2CO_3), which has a pK_a of 6.4:

$H_2CO_3(aq) + H_2O(l) \rightleftarrows HCO_3^-(aq) + H_3O^+(aq)$

Which of the following will decrease the pH of the resulting solution?

I. Addition of HCO_3^-
II. Addition of H_2CO_3
III. Removal of HCO_3^-

A. I B. II C. III D. I and II E. II and III*

Limitations of the Question

1) It does not relate directly to the core idea of Change and Stability.
2) It does not address a Scientific Practice.
3) It allows students to use a heuristic to predict the answer without understanding why.

Revised Question – Constructed Response:

Consider the acid ionization reaction of carbonic acid (H_2CO_3), which has a pK_a of 6.4:

$H_2CO_3(aq) + H_2O(l) \rightleftarrows HCO_3^-(aq) + H_3O^+(aq)$

1. What is the condition for chemical equilibrium at the molecular level?
2. Using what you know about chemical equilibria and **reaction rates**, explain why adding sodium bicarbonate ($NaHCO_3$) shifts equilibrium to the left favoring reactants. Your answer should include discussion of what is happening at the molecular level to control the relevant reaction rates. (Note: Restating LeChatelier's principle does not answer this question.)

Revised Question – Selected Response:

The acid ionization reaction has a pK_a of 6.4:

$H_2CO_3(aq) + H_2O(l) \rightleftarrows HCO_3^-(aq) + H_3O^+(aq)$

1. How would the composition of the system change as it returns to equilibrium following the addition of sodium bicarbonate ($NaHCO_3$)?

I. The concentration of H_2CO_3 would increase and the concentrations of HCO_3^- and H_3O^+ would decrease.
II. The concentration of H_2CO_3 would decrease and the concentrations of HCO_3^- and H_3O^+ would increase.
III. There would be no change.

Why?

III. The reaction shifts to the left.
IV. The reaction shifts to the right.
V. More collisions between the reactants will decrease the rate of the forward reaction.
VI. More collisions between the products will increase rate of the reverse reaction.

A. I and IV B. I and VI* C. II and III D. II and V E. I and V

Figure 2. Transformation of the original traditional question on Le Chatelier's principle to a three-dimensional cluster of constructed-response questions and a three-dimensional selected-response question.

developed and then deconstructed to identify the prior knowledge that would be required to meet the performance and evidence that should be elicited in the assessment tasks. Then, the assessment tasks themselves and grading rubrics can be developed. However, this process can be time-consuming and, for most faculty, is unrealistic given the constraints of faculty life and their expectations.

In this paper, we provide an alternative, more pragmatic approach that focuses on adapting existing questions typically used in general chemistry courses. Our approach to designing assessments aligned with 3D learning also focuses on the need for evidence of student engagement with 3D learning. That is, assessments should be constructed to elicit information on the nature of students' proficiency in using core ideas blended with scientific practices and crosscutting concepts. It is further informed by the *Three-Dimensional Learning Assessment Protocol* (3D-LAP),[5] designed to characterize assessment items across the disciplines of biology, chemistry, and physics, to determine the potential that such items have to elicit evidence of student engagement with core ideas, scientific practices, and crosscutting concepts. For scientific practices, the 3D-LAP provides detailed criteria that an assessment task must include in order to have the potential to elicit evidence of student engagement with a particular scientific practice. Separate criteria were developed for both constructed (open-ended) and selected-response (multiple choice) items: recognizing the limitations of the selected-response format but also acknowledging the reality that large-enrollment classes often necessitate the use of such items. The 3D-LAP also includes four core ideas that underlie the more traditional chemistry topics[1] (Box 1) and descriptions of what an item must include to elicit evidence of student engagement with a particular crosscutting concept. As noted earlier these core ideas were those we developed as part of a course transformation,[1,4] but there are a number of other authors who have also proposed core ideas that could be used in conjunction with the 3D-LAP.

■ ADAPTING ASSESSMENT TASKS

In the following sections, we first show how both constructed- and selected-response questions can be adapted from existing questions to meet the criteria for 3D learning as outlined in the 3D-LAP. We provide several examples of common assessment items which, although designed to address specific content learning outcomes, do not address the other dimensions of 3D learning, and we will describe the process by which they can be transformed. This process typically begins by restating the learning goals as learning performances. That is, we stipulate what students should know and how they should be able to use that knowledge in terms of core ideas and scientific practices, and then specify the evidence that the task is intended

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207−217

Table 1. Alignment of Revised Assessment Items in Figure 2 with 3D Learning

Criteria	How Criteria Are Met by Revised CR Question	How Criteria Are Met by Revised SR Question
Core Idea: Change and Stability in Chemical Systems		
Energy and entropy changes, the rates of competing processes, and the balance between opposing forces govern the fate of chemical systems. Change: Change in chemical systems results from the natural evolution of the system or occurs in response to a perturbation to the system.	Students are asked to apply their understanding of equilibrium as a balance between rates of forward and reverse reactions and the relationship between reactant concentrations, collisions, and reaction rates to explain why addition of $NaHCO_3$ will result in an increase in amount of H_2CO_3 when the system returns to equilibrium following perturbation.	Students are asked to predict the effect of adding $NaHCO_3$ on the composition of the equilibrium established in the acid ionization reaction of H_2CO_3 and select the underlying cause of the change in composition.
Scientific Practice: Constructing Explanations and Engaging in Argument from Evidence		
(1) Question gives an event, observation, or phenomenon.	(1) $NaHCO_3$ is added to the aqueous H_2CO_3 system.	(1) $NaHCO_3$ is added to the aqueous H_2CO_3 system.
(2) Question gives or asks student to make a claim based on the given event, observation, or phenomenon.	(2) Question 2 provides the claim that addition of $NaHCO_3$ shifts equilibrium to the left increasing the concentration of H_2CO_3.	(2) The first part of question 1 asks students to select the correct description of the effect of adding $NaHCO_3$ on the composition of the equilibrium reaction system (option I).
(3) Question asks student to provide scientific principles or evidence in the form of data or observations to support the claim.	(3) Question 1 asks students to identify the condition for equilibrium at the molecular level, namely, that the rates for the forward and reverse reactions are equal.	(3) In the second part of the question, which asks "Why?" students should associate the increased amount of $NaHCO_3$ with more collisions between products (option VI).
(4) Question asks student to provide reasoning about why the scientific principles or evidence support the claim.	(4) In order to correctly answer question 2, students must state that increasing the concentration of HCO_3^- will increase the rate of collisions between HCO_3^- and H_3O^+ leading to an increase in the rate of the reverse reaction which results in increased production of H_2CO_3 until equilibrium is re-established.	(4) In selecting a response to the second part of the question, students should connect the increase in the concentration of H_2CO_3 and decrease in the concentrations of HCO_3^- and H_3O^+ to an increased rate of the reverse reaction resulting from more collisions between the products HCO_3^- and H_3O^+ (answer B).
Crosscutting Concept: Cause and Effect: Mechanism and Explanation		
For an assessment task to be associated 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).	The question cluster provides the cause (addition of $NaHCO_3$) and effect (equilibrium is shifted to the left favoring products). In question 2 students are asked to provide the mechanism that considers the effect of adding $NaHCO_3$ on the relative rates of the forward and reverse reactions.	The question cluster provides the cause (addition of $NaHCO_3$) and asks students to select the effect (option I) and the mechanism (option VI).

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207−217

to elicit as discussed below. These "evidence statements" can also be used as the beginning of a grading rubric for a constructed-response task or answer choices for selected-response tasks. Student responses elicited from constructed-response questions can also provide a fruitful source of distractors (incorrect responses) when developing selected-response items. We will also address the importance of designing assessment prompts to effectively elicit the desired evidence of 3D learning by illustrating how some changes to the prompts in a 3D question can potentially elicit stronger evidence of student understanding.

Example 1: Constructing Explanations and Engaging in Argument from Evidence

A typical learning goal associated with the study of chemical equilibria requires students to predict the effect of various perturbations on the equilibrium (using Le Chatelier's principle). However, this goal alone does not explicitly connect to a core idea nor does it require use of a scientific practice. Figure 2 shows an example of a problem in which students are asked to apply Le Chatelier's principle to determine which perturbations would increase the hydronium ion concentration. Students should deduce that the correct answer choice is E, because both the addition of H_2CO_3 (option II) and removal of HCO_3^- (option III) would result in generation of more products and an increase in $[H_3O^+]$. Students may use a learned pattern to predict the effect of disturbing a system at equilibrium without an understanding of the underlying cause and mechanism of the system's response. As written, the question does not require students to engage in a scientific practice, nor does it provide an opportunity for students to show that they are using the core idea *Change and Stability in Chemical Systems*.

A 3D performance expectation would require students to go beyond predicting the effect of adding or removing a reactant or product; it would also require students to provide the underlying mechanism for the shift in the equilibrium properties of the reaction system. For example, the performance expectation might state the following: Explain why adding or removing components of a chemical system at equilibrium perturbs the position of equilibrium, and predict the effect of that perturbation. Modifying an assessment task to elicit this information would align with the scientific practice of *Constructing Explanations and Engaging in Argument from Evidence* (note that for assessment purposes the 3D-LAP criteria for constructing explanations and arguments are the same[5]) and crosscutting concept *Cause and Effect: Mechanism and Explanation*. Several ideas are required to support a full mechanistic understanding of what happens when a chemical reaction at equilibrium is disturbed: here are evidence statements that can support both how the question is structured, and provide a scoring guide for assessment tasks.

Evidence statements:

1. At equilibrium the rates for the forward and reverse reactions are equal.
2. Adding a substance increases the frequency of its collisions leading to an increased rate for its reaction.
3. Removing a substance decreases its collision rate and thereby slows its reaction.
4. When the rates of the forward and reverse reactions are different, the product of the faster reaction builds up because it is generated more rapidly than the product of the slower reaction.
5. The rate of the faster reaction decreases and the rate of the slower reaction increases until the rates are equal and equilibrium is reestablished.

Ideally, we might want to redesign the question to elicit these ideas, but in moving from more traditional items to 3D assessment tasks, a single question may not be sufficient to provide the evidence that we seek about student understanding. It is often necessary to ask students a series of scaffolded questions to fully engage them in a scientific practice. We will refer to a series of related questions that taken together comprise a larger task as a "cluster".

The cluster of revised constructed-response questions provided in Figure 2 recasts the original question to address not only what happens when a substance is added to a reaction system at equilibrium, but also why it happens, addressing the scientific practice of *Constructing Explanations and Engaging in Argument from Evidence* as shown in Table 1. By comparison, the original task only satisfied criteria 1 and 2 for the practice, while the evidence (criterion 3) and reasoning (criterion 4) components were absent. For the revised constructed-response cluster, the claim was provided in question 2 so that the evidence and reasoning components of the explanation become the focus of the question. This also avoids the problem of students getting the claim wrong and then constructing the rest of the explanation around an incorrect claim.

The task prompt should be carefully crafted to obtain evidence of deeper student thinking. A prompt that does not provide enough structure to cue the students about what kinds of thinking are necessary to support the answer will often result in vague or incomplete responses. A task that is too structured may provide too much information in the way that multiple choice tasks do, and therefore may overestimate what students understand.[22,23] For example, in our work on student mechanistic understanding of acid−base reactions we found that asking both questions about what is happening on the molecular level during a given reaction and why the given reaction occurs provide enough structure to signal to students that we were asking for more than a description, and are trying to elicit their understanding of the causal mechanism for the reaction.[24] Therefore, a balance must be struck between providing enough scaffolding to potentially engage students in the scientific practice of *Constructing Explanations and Engaging in Argument from Evidence* and effectively elicit their knowledge of the reasoning without giving the correct answer away. As detailed in Table 1, the constructed-response question cluster now addresses the core idea *Change and Stability in Chemical Systems*, scientific practice of *Constructing Explanations and Engaging in Argument from Evidence*, and crosscutting concept *Cause and Effect: Mechanism and Change*.

Though it is often easier to write constructed-response questions that are 3D, situations such as large class sizes often necessitate the use of selected-response questions. Accordingly, once the prompts for the constructed-response question have been developed to elicit the targeted information from students, it is possible to use those prompts, and associated student responses, to generate selected-response questions. A corresponding revised selected-response question is provided in Figure 2. Students are first asked to predict (make a claim about) what happens when sodium bicarbonate ($NaHCO_3$) is added to an aqueous solution containing carbonic acid (H_2CO_3) and bicarbonate (HCO_3^-) at equilibrium. Students' understanding of why equilibrium shifts toward reactants (option I) is targeted in the second part of the question, which asks students to select the evidence and reasoning that best support their claim (option VI). Taking the two parts of the question together (correct answer B) satisfies the criteria for *Constructing*

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207−217

Original Question:

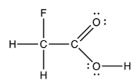
Calculate the percent ionization of a 0.030 M solution of fluoroacetic acid (CH_2FCOOH), which has a pH of 2.12. Fluoroacetic acid has a pK_a of 2.59.

Limitations of the Question

1) It does not relate directly to the core idea of Change and Stability.
2) It does not address a Scientific Practice.
3) It only asks the student to do a calculation without an understanding of what that number means.

Revised Question – Constructed Response:

Fluoroacetic acid (CH_2FCOOH) ionizes in water and has a pK_a of 2.59.



1. Draw out the reaction.
2. A 0.030 M solution of fluoroacetic acid has a pH of 2.12. What is the percent ionization? Show your work.
3. Using your result from question 2, would you classify fluoroacetic acid as a strong or weak acid? Explain your reasoning.
4. Draw a particulate (molecular-level) representation of the equilibrium reaction mixture. (Hint: Consider your answers from questions 1-3 when making your representation.)
5. What would happen to the pH if you diluted the solution by adding water? (no need for a calculation). Explain your reasoning.

Figure 3. Adaptation of an acid–base equilibrium calculation to include the three dimensions.

Explanations and Engaging in Argument from Evidence (Table 1). One would ideally like to obtain evidence about student understanding of ideas that further support a mechanistic understanding of the effect of disturbing a reaction system at equilibrium; however, the selected-response question format imposes limits on the amount of detail that can be provided for evidence and reasoning in explanations and arguments. While selecting a correct response is not the same as constructing a response, some selected-response items are a reality of large-enrollment classes. Like the revised constructed-response question cluster, the revised selected-response question also incorporates the core idea *Change and Stability in Chemical Systems* and crosscutting concept *Cause and Effect: Mechanism and Change* as shown in Table 1.

Example 2: Using Mathematics and Computational Thinking

In addition to constructing explanations, many general chemistry courses require students to perform calculations of various types. The practice of *Using Mathematics and Computational Thinking*, however, requires students to go beyond the calculation to provide a consequence or interpretation of their results. Here we consider performing a calculation associated with acid–base equilibria. Learning objectives for such calculations often take the form "calculate the pH, percent ionization, or K_a" when given appropriate data. A typical question might ask students to "Calculate the percent ionization of 0.030 M solution of fluoroacetic acid, which has a pH of 2.12." either in a constructed- or selected-response format (Figure 3). However, production of the correct response does not (of course) mean that the student understands the physical meaning of the calculation. To transform this question into one that could elicit evidence of mathematical thinking, we have developed (and used) the cluster of scaffolded constructed-response questions as shown in Figure 3. Here students are required to calculate the percent ionization as before (question 2) and then use the result to determine (with reasoning) whether the acid is strong or weak, an interpretation of the result of the calculation (question 3). Furthermore, students are also asked to rerepresent the answer from their calculation as a molecular-level diagram which shows the relative proportion of ionized and un-ionized acid (question 4). Both questions 3 and 4 fulfill the 3D-LAP criterion for *Using Mathematics and Computational Thinking* in that students should provide a consequence or an interpretation of their calculated number (Table 2). This is often the portion of the scientific practice that is omitted within traditional question prompts.

Even with the criteria for *Using Mathematics and Computational Thinking* satisfied, however, the question cluster does not explicitly address a core idea, and while not all questions should be 3D, there are a number of ways in which this could be rectified. Here we again choose to connect to *Change and Stability in Chemical Systems*, by asking students to predict the effect on the percent ionization and pH if water is added to the solution (question 5). This brings us back to the Le Chatelier's principle question. Because students are not specifically asked to calculate, but rather reason through using the ideas discussed earlier, we can connect these calculations to the *Change and Stability in Chemical Systems* core idea and the *Cause and Effect: Mechanism and Explanation* crosscutting concept. All of these tasks might fall under a learning performance: Calculate a range of equilibrium values for chemical systems, interpret the results, and use them to predict how perturbations will change the equilibrium system.

Evidence statements for such a question cluster:

1. The relationship between pH and $[H_3O^+]$ is $pH = -\log[H_3O^+]$.
2. The equilibrium concentration of $[H_3O^+]$ divided by the original concentration of the acid yields the percent ionization.
3. In water, weak acids only ionize a small amount, and strong acids ionize almost 100%.
4. Percent ionization depends on the sum of the concentrations of the ionized and un-ionized forms of the weak acid and its K_a.
5. Percent ionization is an indirect measure of the relative amounts of ionized and un-ionized species. Dilution of solution reduces the total concentration of the un-ionized

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207–217

Table 2. Alignment of Revised Constructed-Response Question Involving an Acid–Base Equilibrium Calculation with 3D Learning

Criteria	How Criteria Are Met by Revised CR Question
Core Idea: Change and Stability in Chemical Systems	
Energy and entropy changes, the rates of competing processes, and the balance between opposing forces govern the fate of chemical systems.	In question 5, students are asked to predict how diluting the fluoroacetic acid (CH_2FCOOH) solution will impact its pH and provide reasoning to support this prediction which should address how reducing the total concentration of the acid in its ionized and un-ionized forms will impact the composition of the system at equilibrium.
Change: Change in chemical systems results from the natural evolution of the system or occurs in response to a perturbation to the system.	
Scientific Practice: Using Mathematics and Computational Thinking	
(1) Question gives an event, observation, or phenomenon.	(1) Fluoroacetic acid (CH_2FCOOH) ionizes in water.
(2) Question asks student to perform a calculation or statistical test, generate a mathematical representation, or demonstrate a relationship between parameters.	(2) What is the percent ionization in a 0.030 M solution of an acid (CH_2FCOOH) that has a pH of 2.12? (question 2)
(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.	(3) Question 3 asks students to identify whether the acid is weak or strong and justify using the data from the calculation. Then the students must rerepresent the percent ionization in a molecular-level diagram in question 4 to show the relative numbers of ionized and un-ionized molecules.
Crosscutting Concept: Cause and Effect: Mechanism and Explanation	
For an assessment task to be associated 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).	Prompt for question 5 provides the cause, dilution of the fluoroacetic acid solution. Students are asked to predict the effect of dilution on pH and then provide reasoning to support the prediction.

and ionized forms of the acid, and, as a consequence, $[H_3O^+]$ decreases.

6. Increasing $[H_3O^+]$ decreases pH; decreasing $[H_3O^+]$ increases pH.

Alignment of the revised constructed-response question with the criteria for the three dimensions is shown in Table 2. An example of how the original question in Figure 3 could be modified to a selected-response question is shown in the Supporting Information (Figure S1).

Example 3: Refining a 3D Question To Elicit More Convincing Evidence of Student Interpretation of Models

With the first two examples, we have shown how questions can be adapted to align with 3D learning; however, the development of 3D questions is an iterative process. In this example, we show how the 3D-LAP can be used to refine an existing 3D assessment item to better elicit evidence of student use of the scientific practice *Developing and Using Models*. Before using the 3D-LAP to evaluate and develop assessment items, one of the authors used the original question in Figure 4 in alignment with the following performance expectation, "Use structural representations to determine the types of attractive forces that exist within and between particles of a substance, and use these to explain relative physical properties of substances." As shown in Table 3, this question could be viewed as addressing the core idea of *Electrostatic Bonding and Interactions*, the science practice of *Developing and Using Models*, and the crosscutting concept of *Cause and Effect: Mechanism and Explanation*. However, as noted earlier,[24] when trying to elicit students' causal mechanistic understanding, prompts need to provide adequate scaffolding. As originally written, this prompt is likely to elicit a response from students stating that the CG base pair would be more stable to heating because it has more hydrogen bonding interactions, but most students are unlikely to include anything about why more hydrogen bonding interactions between base pairs would require more heat to overcome. Moreover, as students are given a molecular-level representation, interpreting the representation is also an important part of this question that is not being explicitly assessed in the original prompt. In fact, it is possible for students to give this seemingly correct response of more hydrogen bonding interactions requiring a higher temperature to overcome, even if they identify the hydrogen bond as the covalent bonds between N and H or O and H within the individual molecules. As this aspect of understanding, the nature of hydrogen bonding interactions, is not explicitly assessed in the original question shown in Figure 4, in most cases it would not be possible to tell if students held this incorrect idea. Revising the question to include the additional prompts that require students to identify which of the two types of attractive forces would be overcome upon addition of a moderate amount of heat and how heating would allow for overcoming that attractive force, using the concepts of both force and energy, we are able to get a more complete picture of student understanding of this process.

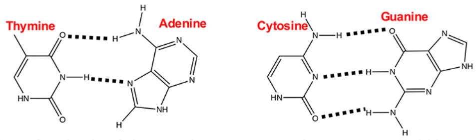
Evidence statements for such a question:

1. The attractive forces within molecules are covalent bonds.
2. The attractive forces between molecules are intermolecular forces, with the strongest types of intermolecular forces between DNA base pairs being hydrogen bonding interactions.
3. Covalent bonds are much stronger than intermolecular forces and thus require much more energy to overcome.

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207–217

Original Question:

DNA is held together by intermolecular attractions between base pairs. Adenine (A) always pairs with Thymine (T) and Guanine (G) always pairs with Cytosine (C) as shown below. One way to make the two strands of DNA separate is to heat up the DNA.



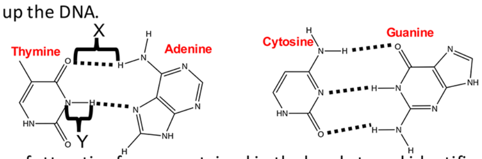
Identify which of the two base pair combinations would be more stable to heating (in other words, would need to be heated to a higher temperature before they separated), and explain why this base pair would be more stable to heating.

Limitation of the Question

Allows students to answer correctly using incorrect interpretation of the model. If students believe that "H-bonds" are the solid lines (Y in Revised Question), they could still arrive at the answer "C-G because it requires the breaking of more H-bonds."

Revised Question:

DNA is held together by intermolecular attractions between base pairs. Adenine (A) always pairs with Thymine (T) and Guanine (G) always pairs with Cytosine (C) as shown below. One way to make the two strands of DNA separate is to heat up the DNA.



1. Identify the types of attractive forces contained in the brackets and identified by the letters X and Y.

X: ______________________________

Y: ______________________________

2. Which of the two attractive forces that you identified in Question 1 would you expect to be affected if you heated a strand of DNA up to a temperature of 100 °C?

3. Explain how heating DNA would affect the attractive forces you identified in Question 2. Be sure to use the concepts of force and energy in your explanation.

4. Which of the two base pair combinations do you think is more stable to heating (in other words, would need to be heated to a higher temperature before they separated)? (circle one)

A—T or C—G

5. Explain why the base pair you chose in Question 4 would be more stable to heating. Be sure to use the concepts of force and energy in your explanation and link it to the structures shown above.

Figure 4. Revision of a 3D question to elicit more complete evidence of student understanding.

4. Upon moderate heating, the much weaker noncovalent forces are overcome rather than the covalent bonds within the molecules.
5. The more interactions formed between molecules, the stronger the attraction between them, meaning that more thermal energy (heat) would be needed to overcome the stronger attractive noncovalent forces.

The prompt for the explanation in Figure 4 was expanded to provide guidance on the depth of explanation expected, namely, to include consideration of forces and energy and connect to the structural representations provided (question 5) instead of stopping at higher temperature is required to overcome more hydrogen bonds. Without explicitly prompting students to use these ideas, they may not provide responses that reflect their full understanding. Further, by adding this model interpretation piece, we now explicitly address an important piece of the science practice. The alignment of the original and revised questions with the core idea, crosscutting concept, and criteria for the developing and using models scientific practice is summarized in Table 3.

Two additional examples showing revisions of traditional assessment tasks to include all three dimensions are found in Figures S2 and S3.

■ IMPLICATIONS OF 3D ASSESSMENTS FOR THE DESIGN OF INSTRUCTION

These and other similar assessment tasks are already being used in large-enrollment general chemistry courses for STEM majors at our institutions. We use a mixture of constructed-response items, using the evidence statements as guides for developing grading rubrics, and selected-response items. However, in order to use such assessments, curricula must be modified to explicitly connect content to core ideas to help students build the framework required to make their knowledge robust and usable. Students must also be expected to respond to low-stakes formative assessment activities that require 3D responses. For many students, this will be the first time that they are asked to connect fragments of knowledge to core ideas and to demonstrate that they can use this knowledge rather than simply recall it. In other words, students need practice with the scientific practices. Instructors can help students by scaffolding questions and providing adequate feedback. In large-enrollment classes, providing individual feedback on constructed-response tasks can be quite daunting. Indeed, this is one reason that commercial online homework systems which are designed to provide automatic grading and feedback tend to focus on facts, fragments, and numerical calculations that can be easily scored as correct or incorrect. However, it is possible to provide feedback through discussion and critiques of examples of student work. In our work, we use an online system called *beSocratic*[25,26] that allows us to administer and collect large numbers of such tasks, in which students must construct drawings, models, explanations, and arguments on a regular basis. Each student is assigned homework twice a week, completion of which counts for a small portion of the course grade. We then provide students with anonymous responses and ask them to critique a

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207−217

Table 3. Alignment of the Original and Revised Questions in Figure 4 with the Criteria for the Three Dimensions

Criteria	How Criteria Are Met by Original CR Question	How Criteria Are Met by Revised CR Question
Core Idea: Electrostatic Bonding and Interactions		
Attractive and repulsive forces govern noncovalent and bonding interactions between atoms and molecules. The strength of these forces depends on the magnitude of the charges and the distances between them.	Students must recognize that having 3 hydrogen bonding interactions between molecules is stronger than having 2 hydrogen bonding interactions and thus will require more energy (higher temperature to be overcome). However, students can answer this question without being able to correctly identify these hydrogen bonding interactions in the representation.	Students must: (1) identify types of attractive forces shown in the representation; (2) identify the hydrogen bonding interactions as weaker than the covalent bonds and thus the ones that will be overcome upon initial heating; and (3) recognize that having 3 hydrogen bonding interactions between molecules is stronger than having 2 hydrogen bonding interactions and thus will require more energy (higher temperature to be overcome).
Scientific Practice: Developing and Using Models		
(1) Question gives an event, observation, or phenomenon for the student to explain or make predictions about.	(1) The phenomenon is that DNA can be separated by heating.	(1) The phenomenon is that DNA can be separated by heating.
(2) Question gives a representation or asks student to construct a representation.	(2) A model showing the H-bonding interactions between the A−T and G−C base pairs is provided for students.	(2) A model showing the H-bonding interactions between A−T and G−C base pairs is provided for students.
(3) Question asks student to explain or make a prediction about the event, observation, or phenomenon.	(3) Students are asked to predict which base pair would be most resistant to heating and explain why that would be the case.	(3) Students are asked to explain how heating affects the intermolecular attractions between DNA base pairs as well as predict which set of base pairs would be most resistant to heating and explain why that would be the case.
(4) Question asks student to provide the reasoning that links the representation to their explanation or prediction.	(4) Students are not explicitly asked anything about their interpretation of the representation so we do not know how they are connecting their representation and their explanation.	(4) In their explanation for why one base pair would be more resistant to heating than the other, they have to link the concepts of force and energy to the model showing the H-bonding interactions that they were provided.
Crosscutting Concept: Cause and Effect: Mechanism and Explanation		
For an assessment task to be associated 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).	The cause (heating) and effect (DNA separating upon heating) are provided. Students are expected to provide a mechanism that links the two.	The cause (heating) and effect (DNA separating upon heating) are provided. Students are expected to provide a mechanism linking the two that includes the concepts of force and energy.

DOI: 10.1021/acs.jchemed.7b00645
J. Chem. Educ. 2018, 95, 207−217

range of responses to identify the components of a good response. At one of our institutions which has classes of 350−400 students, and 2,000−3,000 students per semester, smaller recitations also provide an opportunity for students to engage in 3D activities in groups with support from a graduate teaching assistant. It should be noted that employing such assessments without providing students with appropriate learning experiences and formative assessment tasks is unlikely to improve outcomes. Students who are taught in a more traditional format, as reflected in most online assessment systems, are likely to have difficulty completing these tasks.

■ IMPLICATIONS FOR THE DESIGN OF FORMATIVE AND SUMMATIVE ASSESSMENTS

This paper is intended to address the design of both formative and summative assessment tasks; we are, however, not advocating that all of the assessment items that students see should be 3D. Indeed, there are many facts, skills, and mathematical routines (such as stoichiometric calculations) that students must master, ideally to the point of automaticity, before they can address many 3D tasks. For example, for a student to understand how a permanent dipole can emerge from neutral molecules and how interactions between polar entities subsequently affect both physical and chemical properties, they must concatenate a rather long sequence. They must be able to construct 2D Lewis structure representations, convert these structures to three-dimensional structures by applying the rules of VSEPR, determine which bonds are polar (by knowing electronegativity trends), sum up individual bond polarities to determine molecular polarity (by having some understanding of vector addition), and then determine how molecules will interact. Each of these necessary skills and rule applications is typically tested in general chemistry, but none of them, in isolation, are meaningful. If students do not understand why they are learning to draw structures, or use VSEPR, it is unlikely that they will remember how to do these skills at a later date.[27−29] Indeed, studies show that students typically rely on heuristics,[13,30] or a misunderstanding of the nature of hydrogen bonding to predict properties.[14]

We believe that both formative and summative assessments need to both address simple isolated skills and facts, and also incorporate a significant proportion of 3D questions. The ideal ratio of 3D to typical questions is a researchable question, but pragmatically, we have found that if 50% of the points on a summative examination come from 3D questions, it is still quite manageable from both exam-writing and grading perspectives, and can provide students with enough time to finish the examination. The mix of questions also sends a message to students that both skills and 3D questions are important outcomes. Furthermore, working with students who have a less-developed chemistry and mathematics skill sets should not preclude the use of 3D assessment items provided proper instructional support and formative assessment opportunities are offered. One of the authors is using 3D assessment items in a chemistry bridge course that serves students with low mathematics placements.

■ LIMITATIONS OF 3D ASSESSMENT ITEMS

There are many approaches to designing assessment items, depending on the purpose of the assessment or, in other words, the evidence sought. For example, national large-scale assessment instruments (such as those provided by the ACS Examinations Institute or by ETS) must have good psychometric properties; they must produce data that is reliable (give similar results when used with similar populations) and valid (they should measure what they intend to measure). Unfortunately, in order to meet these criteria, assessment items are often centered on one specific skill or piece of knowledge rather than requiring students to connect multiple ideas to address underlying concepts and core ideas central to chemistry.

Assessment items that require students to put their knowledge to use (i.e., involve scientific practices) and integrate more than one concept are often more complex, and it will certainly be more difficult to establish the same psychometric properties as some existing assessment instruments that address isolated ideas. However, if students are never asked to construct and use their knowledge, it is highly unlikely that they will develop this kind of expertise. There exists, then, a tension between the need for validity and reliability of assessment items, and the need for those assessment items to measure something other than facts, algorithmic calculations, or pattern recognition. Indeed, in the NRC report *Developing Assessments for the Next Generation Science Standards*[31] the authors call for a suite of assessment items, that taken as a whole can be used to assess 3D learning. However, since most faculty write their own examinations and typically do not worry too much about the psychometric properties of their examinations, this should not deter the use of 3D questions.

Another drawback to using 3D items is that, at least initially, it is more time-consuming and difficult to construct such items, and there are currently no commercial online homework systems or test banks available. However, if we want students to develop deep, useful knowledge frameworks, we must provide both formative tasks and summative examinations to both help students learn and to identify when they have learned. That being said, each of the authoring team has been involved in developing 3D examinations, both collaboratively and individually. It is our experience that, after the initial period of adjustment, writing 3D questions becomes more natural and does not take much more time than writing more typical questions. Each of us prepares three examinations and a final every semester, either alone or in collaboration. We encourage instructors, when possible, to begin by writing constructed-response questions, which require less time to write and can often be used as a starting point for developing selected-response items.

■ SUMMARY

We show here how faculty can adapt their existing questions to elicit stronger evidence about what students know and can do. With incorporation of scientific practices into assessment tasks, it is possible to tie fragments of knowledge (facts, skills, calculations) to core ideas, which will help students develop more robust and transferable knowledge structures. Using the criteria in the 3D-LAP makes this process easier and less time-consuming than designing questions from scratch, because the changes required are often in the need to prompt for reasoning, interpretation, or justification, which is where the knowledge linkage is elicited.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.7b00645.

Additional examples showing revision of traditional assessment tasks to incorporate a scientific practice (Example S1) and all three dimensions (Examples S2 and S3) (PDF, DOCX)

AUTHOR INFORMATION

Corresponding Author

*E-mail: sonia.underwood@fiu.edu.

ORCID

Sonia M. Underwood: 0000-0002-4919-2758
Lynmarie A. Posey: 0000-0001-7713-0637
Deborah G. Herrington: 0000-0001-6682-8466
Justin H. Carmel: 0000-0001-9281-3751
Melanie M. Cooper: 0000-0002-7050-8649

Notes

Any opinions, findings, conclusions, or recommendations expressed here are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by the Association of American Universities' (AAU) Undergraduate STEM Education Initiative, funded by the Helmsley Charitable Trust, and by the National Science Foundation under DUE 0816692 (1359818), DUE 1043707 (1420005), and DUE 1122472 (1341987). We would also like to acknowledge the other members of the AAU project team at Michigan State University (Marcos D. Caballero, Diane Ebert-May, Cori L. Fata-Hartley, Sarah E. Jardeleza, James T. Laverty, and Rebecca L. Matz) for their contributions to the development of the Three-Dimensional Learning Assessment Protocol (3D-LAP).

REFERENCES

(1) Cooper, M. M.; Posey, L. A.; Underwood, S. M. Core Ideas and Topics: Building Up or Drilling Down? *J. Chem. Educ.* **2017**, *94* (5), 541−548.

(2) National Research Council. *How People Learn: Brain, Mind, Experience, and School*; National Academies Press: Washington, DC, 1999.

(3) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; National Academies Press: Washington, DC, 2012.

(4) Cooper, M. M.; Klymkowsky, M. W. Chemistry, Life, the Universe and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90*, 1116−1122.

(5) Laverty, J. T.; Underwood, S. M.; Matz, R. L.; Posey, L. A.; Carmel, J. H.; Caballero, M. D.; Fata-Hartley, C. L.; Ebert-May, D.; Jardeleza, S. E.; Cooper, M. M. Characterizing College Science Assessments: The Three-Dimensional Learning Assessment Protocol. *PLoS One* **2016**, *11*, e0162333.

(6) Pellegrino, J. W. Proficiency in Science: Assessment Challenges and Opportunities. *Science* **2013**, *340* (6130), 320−323.

(7) Momsen, J.; Offerdahl, E.; Kryjevskaia, M.; Montplaisir, L.; Anderson, E.; Grosz, N. Using Assessments to Investigate and Compare the Nature of Learning in Undergraduate Science Courses. *CBE-Life Sci. Educ.* **2013**, *12* (2), 239−249.

(8) Scouller, K. The Influence of Assessment Method on Students' Learning Approaches: Multiple Choice Question Examination versus Assignment Essay. *High. Educ.* **1998**, *35* (4), 453−472.

(9) Scouller, K.; Prosser, M. Students' Experiences in Studying for Multiple Choice Question Examiniations. *Stud. High. Educ.* **1994**, *19* (3), 267−279.

(10) Snyder, B. *The Hidden Curriculum*; The MIT Press: Cambridge, MA, 1973.

(11) Entwistle, N. J. Approaches to Learning and Perceptions of the Learning Environment. *High. Educ.* **1991**, *22* (3), 201−204.

(12) Crooks, T. J. The Impact of Classroom Evaluation Practices on Students. *Rev. Educ. Res.* **1988**, *58* (4), 438−481.

(13) Maeyer, J.; Talanquer, V. The Role of Intuitive Heuristics in Students' Thinking: Ranking Chemical Substances. *Sci. Educ.* **2010**, *94*, 963−984.

(14) Cooper, M. M.; Corley, L. M.; Underwood, S. M. An Investigation of College Chemistry Students' Understanding of Structure−property Relationships. *J. Res. Sci. Teach.* **2013**, *50*, 699−721.

(15) National Research Council. *Knowing What Students Know: The Science and Design of Educational Assessment*; Pellegrino, J. W., Chudowsky, N., Glaser, R., Eds.; National Academies Press: Washington, DC, 2001.

(16) Mislevy, R. J. Evidence and Inference in Educational Assessment. *Psychometrika* **1994**, *59*, 439−483.

(17) Mislevy, R. J. Test Theory Reconceived. *J. Educ. Meas.* **1996**, *33* (4), 379−416.

(18) Mislevy, R. J.; Haertel, G. D. Implications of Evidence-Centered Design for Educational Testing. *Educ. Meas. Issues Pract.* **2006**, *25* (4), 6−20.

(19) Wilson, M. *Constructing Measures: An Item-Response Modeling Approach*; Erlbaum: Mahwah, NJ, 2005.

(20) Harris, C. J.; Krajcik, J. S.; Pellegrino, J. W.; McElhaney, K. W.; DeBarger, A. H.; DiBello, L. V.; Gane, B.; Lee, J. *Constructing Assessment Tasks That Blend Disciplinary Core Ideas, Crosscutting Concepts, and Science Practices for Classroom Formative Applications*; SRI International: Menlo Park, CA, 2016. https://www.sri.com/sites/default/files/publications/constructing_assessment_tasks_2016.pdf (accessed Oct 2017).

(21) National Research Council. *Next Generation Science Standards: For States, By States*; National Academies Press: Washington, DC, 2013.

(22) Lee, H.-S.; Liu, O. L.; Linn, M. C. Validating Measurement of Knowledge Integration in Science Using Multiple-Choice and Explanation Items. *Appl. Meas. Educ.* **2011**, *24* (2), 115−136.

(23) Hubbard, J. K.; Potts, M. A.; Couch, B. A. How Question Types Reveal Student Thinking: An Experimental Comparison of Multiple-True-False and Free-Response Formats. *CBE-Life Sci. Educ.* **2017**, *16* (2), ar26.

(24) Cooper, M. M.; Kouyoumdjian, H.; Underwood, S. M. Investigating Students' Reasoning about Acid−Base Reactions. *J. Chem. Educ.* **2016**, *93* (10), 1703−1712.

(25) Cooper, M. M.; Underwood, S. M.; Bryfczynski, S. P.; Klymkowsky, M. W. A Short History of the Use of Technology to Model and Analyze Student Data for Teaching and Research. In *Tools of Chemistry Education Research*; Cole, R., Bunce, D., Eds.; ACS Symposium Series; American Chemical Society: Washington, DC, 2014; pp 219−239.

(26) Bryfczynski, S. P. BeSocratic: An Intelligent Tutoring System for the Recognition, Evaluation, and Analysis of Free-Form Student Input. Doctoral Dissertation, Clemson University, 2012.

(27) Cooper, M. M.; Grove, N.; Underwood, S. M.; Klymkowsky, M. W. Lost in Lewis Structures: An Investigation of Student Difficulties in Developing Representational Competence. *J. Chem. Educ.* **2010**, *87*, 869−874.

(28) Novak, J. D. *Learning, Creating, and Using Knowledge: Concept Maps As Facilitative Tools in Schools and Corporations*; Lawrence Erlbaum Associates: Mahwah, NJ, 1998.

(29) Bretz, S. L. Novak's Theory of Education: Human Constructivism and Meaningful Learning. *J. Chem. Educ.* **2001**, *78*, 1107−1117.

(30) McClary, L.; Talanquer, V. Heuristic Reasoning in Chemistry: Making Decisions about Acid Strength. *Int. J. Sci. Educ.* **2011**, *33* (10), 1433−1454.

(31) National Research Council. *Developing Assessments for the Next Generation Science Standards*; National Academies Press: Washington, DC, 2014.