# Probing the Relevance of Chemical Identity Thinking in Biochemical Contexts 

Courtney Ngai ${ }^{\text {t** }}$ and Hannah Sevian ${ }^{\dagger}$<br>${ }^{\dagger}$ Department of Chemistry, University of Massachusetts Boston, Boston, MA 02125; ${ }^{\ddagger}$ The Institute<br>for Learning and Teaching, Colorado State University, Fort Collins, CO 80523-1052


#### Abstract

The solving of problems in biochemistry often uses concepts from multiple disciplines such as chemistry and biology. Chemical identity (CI) is a foundational concept in the field of chemistry, and the knowledge, thinking, and practices associated with Cl are used to answer the following questions: "What is this substance?" and "How is it different from other substances?" In this study, we examined the relevance of Cl in biochemical contexts and first explored the ways in which practicing biochemists consider Cl relevant in their work. These responses informed the development of creative exercises (CEs) given to secondsemester biochemistry students. Analysis of the student responses to these CEs revealed that students incorporated precursors to Cl thinking in more than half of their responses, which were categorized by seven previously identified themes of Cl relevant to the presented biochemical contexts. The prevalence of these precursors in student responses to the CEs, coupled with the examples provided by practicing biochemists of contexts in which Cl is relevant, indicate that Cl thinking is relevant for both students training to be biochemists and practicing biochemists.


## INTRODUCTION

## The Use of Chemistry Knowledge in Biochemistry Work

Biochemistry has long been recognized as a discipline that integrates concepts from both biology and chemistry (National Research Council, 2009; American Association for the Advancement of Science, 2011; Wright et al., 2013). To perform well in biochemistry courses, students must have a fundamental understanding of the concepts and practices comprising biology and chemistry. Many chemical concepts drive the phenomena and patterns observed in biochemistry; for example, the concept of interand intramolecular forces is presented in many general chemistry textbooks and is responsible for the behavior of protein folding, an essential phenomenon studied in biochemistry. The union of chemical approaches and biochemical contexts has even given rise to a new discipline-chemical biology (Colón et al., 2008).

Given the obvious links between chemistry and biology in biochemistry, researchers have argued it is important to investigate both how biochemists use chemistry in their work and how students apply their chemistry knowledge as they learn biochemistry. These researchers have examined the transfer of certain chemistry concepts to biochemical contexts (Wolfson et al., 2014; Warfa and Odowa, 2015), designed methods to assess student understanding of relevant foundational chemical concepts for biochemistry (Villafañe et al., 2011; Haudek et al., 2012; Xu et al., 2017), and investigated the effect of the disciplinary context of students' understanding and explanation of chemistry concepts (Kohn et al., 2018). Instruments have also been developed to assess student understanding of biochemical concepts, which can include concepts students previously learned in chemistry (Wright and Hamilton, 2008; Shi et al., 2010).

Moving toward the end of understanding how biochemists-in-training use their knowledge, assessments have been developed for measuring the content

Julia Gouvea, Monitoring Editor
Submitted Dec 18, 2017; Revised Sep 6, 2018;
Accepted Sep 10, 2018
CBE Life Sci Educ December 1, 2018 17:ar58
DOI:10.1187/cbe.17-12-0271
*Address correspondence to: Courtney Ngai (court.ngai@gmail.com).
© 2018 C. Ngai and H. Sevian. CBE-Life Sciences Education © 2018 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution-Noncommercial-Share Alike 3.0 Unported Creative Commons License (http:// creativecommons.org/licenses/by-nc-sa/3.0).
"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.
knowledge of students before they enter upper-level biochemistry and biology courses (Shi et al., 2010; Villafañe et al., 2011). Researchers have also investigated students' understanding of specific concepts such as enzyme-substrate interactions within biochemistry (Linenberger and Bretz, 2015). Other studies have characterized overall conceptual difficulties for biochemistry and biology students (Loertscher et al., 2014).

## Studying Chemical Identity in Biochemistry

The interdisciplinary nature of biochemistry implies that many chemical concepts are relevant to the contexts and phenomena explored by biochemists. We have previously argued that chemical identity (CI) is a foundational concept for other concepts in the field of chemistry (Ngai et al., 2014), so it stands to reason that CI thinking informs knowledge and understanding in the field of biochemistry. Like other researchers, we recognize the importance of investigating the link between chemical concepts and their application in the field of biochemistry. Although the application of some foundational chemistry concepts (e.g., free energy, acid-base behavior) to biochemistry have been explored by other researchers (Villafañe et al., 2011; Haudek et al., 2012; Wolfson et al., 2014; Warfa and Odowa, 2015; Xu et al., 2017), how students transfer their knowledge of CI has not been explored.

CI thinking rests on the premise that all substances have at least one characteristic that allows them to be differentiated from other substances. CI thinking answers two core questions around this premise: 1) "What is this substance?" (a question of classification), and 2) "How is this substance different from other substances?" (a question of differentiation). CI thinking is composed of three facets: domain-general reasoning; CI-specific knowledge; and the practice of CI, which integrates both knowledge and reasoning (Ngai, 2017).

The concept of CI is tightly linked to the notion of substance, which has been debated by chemists and philosophers for centuries. In the early history of chemistry, substances were believed to be ratios of classical elements (earth, air, fire, water), and thus a substance's CI was defined by its proportions of these elements (Ball, 2004). The concept of substance has since evolved into two lenses to view substances: macroscopic, or the manifest view of substances (van Brakel, 2000; VandeWall, 2007), and submicroscopic, or a reduction of substances to the scale of their atoms (Ellis, 2002).

How a substance is defined in the discipline of chemistry, and subsequently classified and differentiated, is crucial for answering many questions using chemistry. The characteristics used to classify and differentiate substances can be submicroscopic or macroscopic, and there can be many chemical characteristics appropriate for establishing CI. The knowledge, reasoning, and practice of using a chemical feature to classify and differentiate substances comprises CI thinking (Ngai, 2017).

We developed the Chemical Substance Identification (CSI) Survey (Ngai and Sevian, 2017) to characterize how students use CI thinking across a broad range of educational levels, including grade 8 through undergraduate completion. The instrument asks students to classify and differentiate a variety of substances that include solids, liquids, and gases, substances typically studied in biology and chemistry, and substances that are familiar and unfamiliar. Our analysis revealed eight themes in how students classify and differentiate substances, which we called the "themes" of CI thinking (Ngai, 2017), because they describe the ways in which students use chemistry to answer: What is this substance? The themes are outlined in Table 1.

Because CI thinking is foundational in the practice of chemistry, we argue it also has major relevance in the practice of biochemistry. As in chemistry, many challenges in biochemistry involve the identification, separation, and transformation of substances. How biochemistry students classify and differentiate substances can impact how these students solve more complex biochemical problems. The application of CI to biochemical contexts by biochemists at any level of expertise has not been explored. Thus, in this study, we explore the ways and extent to which CI thinking is considered consequential in biochemistry.

## Methodology

To obtain a more comprehensive understanding of the ways and extent CI is applied to biochemical contexts, we chose to investigate the application of CI by both practicing biochemists and biochemistry students. We began with the practicing biochemists and aimed to uncover the use of CI thinking in biochemistry work. By determining the ways and contexts in which practicing biochemists use CI, we were able to design authentic biochemical contexts in which CI could be deemed relevant (Duit et al., 2012). The second part of our study identifies how biochemistry students view and use CI within biochemical

TABLE 1. Themes of CI thinking

| Theme | Description |
| :--- | :--- |
| Change | Students focus on the ways a substance did or did not change, and how this behavior indicates CI of the <br> substance. <br> Placing substances in a class or category allows students to use the general behavior or properties associated <br> with substances in that class to determine CI. |
| Class | Students use the macroscopic or microscopic composition and structure of a substance to infer CI. <br> The function or purpose of a substance is used to determine its CI and whether it has changed. <br> Composition and structure <br> Function |
| Organism effect | Information about a substance that can be obtained using the senses is used to establish the CI of a substance. <br> The source of a substance or the environment in which the substance is typically encountered is used to figure <br> out the CI of a substance. |
| Source | Students recommend the use of certain experiments to determine the CI of a substance, or they suggest testing <br> specific values they expect a substance to have (e.g., boiling point). |
| Tests and experimental values |  |

contexts. By providing biochemical contexts that are relevant to the work of practicing biochemists, we exposed biochemistry students to science content in a manner that prepared them to enter the workforce (Duit et al., 2012).

Thus, our study is guided by the overarching research question: "In what ways is CI used in the practice of biochemistry?" This is further divided into two questions:

1. In what ways do practicing biochemists deem CI relevant in their work?
2. How do students who are training to be biochemists use CI in problems that biochemistry addresses?

## RESEARCH DESIGN AND RESULTS <br> Research Question 1-The Use of Cl by Practicing Biochemists

Little research exists on what knowledge practicing biochemists rely upon in their work. Recent surveys of practicing biochemists have identified concepts central to biochemistry and the skills and knowledge found to be deficit in recent graduates (Rowland et al., 2011; Talgar and Goodey, 2015). Although the results of these studies indicate that many chemical concepts are important for biochemistry, no study has investigated the relevance of CI for practicing biochemists. Thus, it was necessary to first identify how practicing biochemists think about and apply CI in their work.

Data Collection. An online survey was designed and distributed to people who self-identified as practicing biochemists. Initially, the survey was distributed to people in industry and academia in the United States and several other countries (e.g., Germany, Israel) who the authors knew had received an education in biochemistry or a related field and were currently working on biochemical problems. Because the participants received degrees (in many cases, advanced degrees) related to biochemistry and acknowledged that all or part of their current work was biochemical in nature, the assumption was made that these participants could be considered to have expertise in biochemistry greater than that of undergraduate students who have not yet completed a degree in biochemistry. Participants could choose to participate in the survey if they identified as biochemists or viewed their work as contributing to the field of biochemistry. Participants were encouraged to advertise the survey to other known biochemists (e.g., via personal communications, posting on LinkedIn or research website), and in this manner the survey was distributed nationally and internationally. Following institutional review board (IRB) guidelines, the sur-
vey was kept anonymous to remove the need for formal consent forms (although our intentions with the survey responses were articulated at the beginning of the survey to inform participants). Only basic demographic information was collected to ensure diversity within the participant pool.

The survey began by asking participants to give a brief description of their own research. This served to establish that the participants conduct work in a wide range of biochemical areas and not just focused on one area of biochemical research. Next, the survey provided participants with a brief definition of CI (see the Supplemental Material for the complete survey) and asked whether they considered CI thinking to be relevant or useful for their own work. Following this question, participants were asked to provide an example of a problem in biochemistry they considered to require CI thinking. This provided a perspective on CI through the lens of practicing biochemists: How did they interpret CI thinking, and in what biochemical contexts did they perceive CI as useful?

Results from the Practicing Biochemist Survey. Thirty-four biochemists participated in this survey, and almost all participants ( $n=33$ ) provided a summary of their own research. At the time of survey completion, the participants were pursuing a range of research, clinical, and industrial interests. These included protein identification and purification, enzyme characterization, examination of biochemical pathways, and gene regulation, to name a few broad topics mentioned by participants. Although some focuses overlapped, for the most part, the participants pursued unique biochemistry work. Table 2 provides the general demographic information of the participant pool, showing that the participants came from different backgrounds. Although biochemistry can stand alone as a field of research, many types of chemical and biological work can fall under the umbrella of biochemical research. As noted by a philosopher of chemistry, research relevant to advancing the field of biochemistry may also be classified as relevant for molecular biology, molecular genetics, and clinical chemistry, to name a few fields (Strand, 1999). Consequently, the practicing biochemists who participated in this survey have a wide range of expertise, as indicated by the diversity of their terminal degrees.

A majority of the participants ( $n=27$ ) provided examples of biochemical problems in which they considered CI to be relevant. In several cases, the examples involved the separation of components in a mixture and then identification of the molecule or protein of interest. Another common response was the classification of substances, either for the purpose of finding similar substances or for determining the properties of a specific

TABLE 2. Practicing biochemist participant demographics ( $N=34$ )

| In the United States | In academia |
| :--- | :---: |
| 26 | 32 |
| Outside the United States | In industry |
| 8 | 2 |
| Most common terminal degree(s): PhD biochemistry, PhD chemistry, PhD molecular biophysics, PhD molecular cell biology |  |
| Most common professional societies in which currently a member: American Society for Biochemistry and Molecular Biology, American Chemical |  |
| $\quad$ Society, American Association for the Advancement of Science, RNA Society, American Society for Microbiology |  |
| Journals frequently read for research purposes: Science, Nature, Journal of Biological Chemistry, Nature Chemical Biology, Nature Cell Biology |  |

substance of interest. In most responses, multiple ways of CI thinking were considered, such as using the composition and structure of substances in a mixture to separate them via a biochemical test. Some participants identified their research as clinical, and they stressed the importance of knowing the precise CI of the substances they work with when preparing drugs or medicines. For example, Participant 4 commented,

Impurity identification is crucial in my job. It's important because to put a drug into people you need to know what's exactly in the vial. The way we do this is by making large amounts of our target molecule and then through chromatography separate all the impurities and try to identify them through mass spec, sequencing, and bio-assay.

This type of problem seeks to answer the first core question of CI: What is this substance? In this case, a major practice belonging to CI underlies the goal of this participant's research, which is creating a pure product. Participant 21 provided another example of a biochemical problem that seeks to answer the first core question of CI, in this case focusing on the identity of an intermediate:

> In discovering and elucidating new biochemical pathways we have on numerous occasions been faced with the task of identifying intermediates in the pathway. Knowing the identity of these intermediates is crucial to understanding the pathway as a whole and how it fits in to the overall metabolic network.

The participants also described biochemical problems addressing the second core question of CI: How is this substance different from other substances? Participant 34 explained a method used for distinguishing proteins:

> In protein chemistry, in which you are expressing a recombinant protein in an expression system such as Escherichia coli, we must consider the chemical identity of the protein being produced, so that it can be distinguished from the background proteins of the expression system. To do this, we standardly utilize PAGE [polyacrylamide gel electrophoresis] analysis to assess by size, however it is important to also confirm this with enzyme assays (if an enzyme), mass spec analysis, or western [blot] analysis to be completely sure as different proteins could have the same mass, and therefore be indistinguishable on a PAGE gel.

Other examples offered by the practicing biochemists included determining the structures of proteins, developing new biochemical analogues, and using enzymes to transform substances. The participants almost always included one or more experimental strategies for determining the CI of a substance, including mass spectrometry, gel electrophoresis, nuclear magnetic resonance spectroscopy, cell cultures, and other laboratory techniques. These survey responses indicate that there are many biochemical contexts in which CI knowledge is relevant and that biochemists have a plethora of available methods to establish CI.

The examples of biochemistry work provided by the practicing biochemists gave an overview of CI in biochemical contexts. In most cases, the interpretations of CI made by the biochemists coincided with the previously established themes of CI thinking.

TABLE 3. Coded themes of Cl thinking in responses to practicing biochemist survey

| Identified theme | Count of <br> instances | Percentage of all <br> instances |
| :--- | :---: | :---: |
| Change | 0 | 0 |
| Class | 3 | 11 |
| Composition \& structure | 6 | 21 |
| Function | 4 | 14 |
| Organism effect | 1 | 4 |
| Sensory information | 0 | 0 |
| Source | 0 | 0 |
| Tests \& experimental values | 14 | 50 |
| Total | 28 | 100 |

Out of the 30 examples provided by the participants, 21 examples (70\%) contained at least one or more references to CI. The diversity of the CI themes identified in the participants' responses is outlined in Table 3. Although not all previously outlined CI themes are represented in the data, several of the CI themes are used by participants in their responses.

In several instances, the responses included multiple themes of CI thinking. Participant 22 provided the following example where he/she considered CI to be relevant to a biochemical challenge:

Chemical identity is the core of our mass spectrometry-based
analysis of protein interactions. We use nano-LC-MS/MS
[liquid chromatography-tandem mass spectometry] (run at a
separate facility) to identify proteins that interact with our
protein of interest. Once we learn what the identities of inter-
acting proteins are, we can know more about the function of
our protein of interest.
Participant 22 references using a specific instrument (nano-LC-MS/MS) for identification of proteins, which was coded as "tests and experimental values." Participant 22 then further links knowing the CI of a protein to making inferences about the function of another protein, which was coded as the "function" theme of CI.

These responses supported the hypothesis that CI is used by practicing biochemists in their work, and participant responses to the final, closed-ended question about the relevance of CI in their work ("To what extent do you consider answering questions of chemical identity to be significant in your biochemistry work?") corroborated this conclusion. These responses are represented in Figure 1. The majority ( 26 participants, $76 \%$ ) of the biochemists responded that questions of CI are either the major part or essential to their work as biochemists.

Participants were given the opportunity to explain their choice when answering this question, and 11 participants (32\%) elaborated on their selections. Four of these participants mentioned that the context of the problem determines whether they are answering a question of CI and that, in some instances, the question of CI is resolved by others (e.g., partner chemists who identify the targeted substances) before they begin working on the problem (e.g., analyzing the targeted substance in a biological system). Thus, although these biochemists are not personally answering a question of CI, they acknowledge it is an important step in their problem-solving processes. Four


FIGURE 1. Responses of practicing biochemists to the question "To what extent do you consider answering questions of chemical identity to be significant in your work?"
new contexts and strive to organize new concepts within their existing knowledge and understanding.

CEs have been used more recently in chemistry (Lewis et al., 2010; Ye and Lewis, 2014) and biochemistry courses (Warfa and Odowa, 2015) to assess what previously learned chemistry concepts students deem relevant to the "new" context of the CE. Using CEs allowed these researchers to characterize the types of previously learned concepts students applied to new contexts. This use of CEs to capture the variety of linkages students make between content knowledge and new contexts inspired their use in this study. Previous studies have used a variety of contexts for CEs, such as gas laws and molecular shapes (Ye and Lewis, 2014), while another study used thermodynamics and enzyme kinetics (Warfa and Odowa, 2015).
participants further commented on the importance of CI in their work. For example, Participant 4 wrote, "My entire biochemistry life has been somehow related to determining the identity of compounds in order to understand the system I'm working on." The other three responses clarified the types of substances relevant to the individual's work or related his or her research to chemistry.

## Research Question 2-CI Use by Students Studying Biochemistry

Choice of Creative Exercises as an Instrument for Revealing the Use of Cl . To answer the second research question, an instrument was needed that could reveal whether and how students naturally use CI when presented with a biochemical context. In the past several decades, many instruments have been developed to elicit and evaluate student thinking, but creative exercises (CEs) best fit the requirements of this study. CEs were originally developed by Trigwell and Sleet (1990) as an alternative assessment of students' knowledge that could also benefit students' learning processes. CEs present a context to students, who are instructed to provide statements relevant to the context. Responses are graded on their correctness and uniqueness. Trigwell and Sleet compared student responses to a CE, a traditional closed-ended problem, and a concept-mapping activity, all of which focused on acid-base equilibria. They discovered that students demonstrated the necessary content knowledge to successfully solve the closed-ended exam question and to relate specific, preidentified concepts for the con-cept-mapping task, but these same students did not successfully apply these concepts to the CE. Trigwell and Sleet concluded that students are more easily able to apply previously learned content knowledge when the goals of the task are explicitly defined, and that CEs could be used to help foster students to make the connections between new contexts and previously learned content knowledge on their own. This, they hypothesized, would result in deeper learning, wherein students intentionally make connections between material and

Development of CEs. The first creative exercise (CE1) for this study mimicked a CE designed by Warfa and Odowa (2015), in which they used the structure of an amino acid (glutamic acid) as the context. No information about the amino acid was provided other than the structure, and the instructions simply asked students to write down correct, distinct, and relevant facts. Mimicking an already tested CE allowed us to focus on resolving a grading scheme and to see whether our students produced statements similar to those observed by Warfa and Odowa. Additionally, as the students in this study were in their second semester of biochemistry, they were very familiar with amino acids and dipeptides. Because the structure of CEs was unfamiliar to students, having a familiar topic was expected to make the introduction of CEs less stressful to students. Thus, CE1 served as a norming exercise for students to learn how to respond to the structure of a CE and provided practice data for designing scoring and coding schemes.

CE1 used a dipeptide (glutamyl cysteine) as the context. The instructions provided for CE1 were the same as the following CEs. CE1 is shown in Box 1, and the other CEs designed in this study are included in the Supplemental Material.

The biochemical contexts (e.g., dipeptide) for the other three CEs designed in this study were derived from the problems or scenarios identified by the practicing biochemists as instances in which they considered CI to be relevant. Using the biochemists' responses to inspire the contexts ensured the CEs elicited CI knowledge relevant to biochemistry and gave students the opportunity to practice linking CI to biochemical contexts in a manner similar to practicing biochemists. The contexts were intended to be broad enough so there were many acceptable responses; thus, any evidence of CI in students' responses was present because the students thought it relevant to the problem. Furthermore, basing the contexts on problems that practicing biochemists encounter in their own research provided authenticity to the CEs (Eilks and Hofstein, 2015).

The second creative exercise (CE2) presented the structures of two molecules: molecule A is arachidonic acid and molecule B

## Box 1. Creative Exercise 1

CE1 - Dipeptide (this heading was not included in student version)
Write down as many correct, distinct, and relevant facts as you can about:


Ten (10) statements will get you full credit for the problem, which is worth a total of 5 points. The information you use should be information you learned in a chemistry course, including general chemistry, organic chemistry, and any biochemistry courses. All other outside information, combined, will only count as one distinct fact towards the correct responses.
is prostaglandin E1. Neither molecule was labeled, other than "molecule A" and "molecule B." This context was inspired by Participant 8, who wrote about classification tasks in biochemistry:

We try to categorize enzymes into different reactive classes using a library of enzyme substrates where we vary the chemical reactive group on the substrates and then screen them against different enzymes to classify them into different chemical groupings.

It was anticipated that students would classify molecules A and B in addition to pointing out compositional features in the provided chemical structures. At the point when CE2 was implemented, students had already learned about fatty-acid synthesis, metabolism, and hormones.

The third creative exercise (CE3) was derived from Participant 27's description of the challenge of differentiating proteins and RNA molecules. This participant stated,

If a protein or RNA is the product of the reaction I'm studying then I need to prove that it was indeed synthesised. I routinely differentiate between protein and RNA molecules based on their chemical composition or physical properties, i.e. length, charge etc. Modern molecular biology techniques allow us to specifically label proteins and RNA with fluorescent markers so many times we prove chemical identity by following fluorescent signals.

DNA and RNA are commonly studied in introductory biochemistry; thus, students were expected to be familiar with these substances. The context for CE3 therefore presented two solutions of DNA: one healthy and one damaged from UV radiation. No pictures or structures were provided.

The context for the fourth creative exercise (CE4) was inspired by biochemists' responses about application of CI thinking for clinical research questions. Participant 21's comments about biochemical pathways and intermediates, along with other participants' responses about the transformation of biochemical substances, prompted the use of acetaminophen excretion pathways as the context for CE4. Metabolism is a central topic in biochemistry, and students typically encounter examples of metabolic pathways early in the curriculum. CE4 presents three different excretion pathways for acetaminophen, each adding a different substituent to the original structure.

## PARTICIPANT SAMPLE AND IMPLEMENTATION OF THE CES

The CEs were implemented in a second-semester undergraduate biochemistry course that is typically taken by students in their third or fourth year of the biochemistry and chemistry BS programs at this university. The class was taught in a traditional lecture style and covered concepts such as protein synthesis and structure and metabolic processes. When consulted, the professor of the course agreed that general concepts related to each of the biochemical contexts for the CEs were covered either in the first semester of the biochemistry course or were covered in the second semester of the biochemistry course by the time the CEs were given. CI was not explicitly mentioned in the course. In the class of 27 students, all students completed at least three out of the four CEs. The students were primarily biochemistry and chemistry majors, which was representative of the class in other semesters. The students in the course were, in turn, representative of the larger enrollment at the university, where the racial and ethnic diversity of the student body in science courses at the time of the study was $32 \%$ Caucasian (nonHispanic), 20\% Asian, 15\% Black/African American, 13\% Hispanic/Latino, $4 \%$ from other races/ethnicities, and $15 \%$ not specified.

Although the CEs were timed so that they generally followed the presentation of biochemical topics relevant to the CE (e.g., CE2 followed a unit on fatty acids), it was expected that students would also respond with other previously learned chemistry knowledge. The study was designed so that the CEs would be course work required of all students in the course. All CEs were designed with input from the instructor and were approved by the instructor before being used in the course. The students' responses were blinded by an external researcher before data were analyzed, and then scores on the CEs were unblinded by the same external researcher before the graded work was returned to the instructor of the course. Upon consultation, an administrator with the university's IRB determined that this study did not require IRB review.

CE1 served to familiarize students with the format and expectations of CEs. It was implemented as an online homework assignment, and students received feedback on the quality of their statements in addition to their grades. Because this was the first encounter students had with CEs, and because it was given as a homework assignment and not in class, the responses to CE1 were not used for data analysis, as students had access to outside resources to complete the assignment. When validating their chemistry CEs, Lewis et al. (2011) established that the CEs administered outside of class are not
as reflective of students' chemistry knowledge, because the environment is less controlled. Thus, the other three CEs for this study were implemented during class time as announced quizzes. The context of the CE was not released before implementation, and students were given 5 minutes to respond to the CE . This implementation format corresponds to other published research using CEs (Lewis et al., 2011; Ye and Lewis, 2014; Warfa and Odowa, 2015).

For each CE, the first author (C.N.) brainstormed a list of correct, unique, and relevant statements for which the students might receive credit. Correctness of a statement implies it is factually accurate, uniqueness implies the statements do not overlap in content, and relevant implies the statement can be linked to the context of the CE. This list was checked by the instructor of the biochemistry course, and immediately after the CEs were due, this initial rubric was made available to students so that they might have immediate feedback on their responses to the CE. The rubric was revised as the student responses were graded, and it was expanded to accommodate statements made by students that were not initially part of the rubric but that fit the grading criteria. The final rubric was made available to students with their grades for each CE. For CE1, students were asked to provide 10 correct, unique, and relevant statements, and they received 0.5 points for each statement that satisfied these criteria. For CE2-CE4, students were only asked to provide five such statements to receive full credit, and thus were awarded 1 point per correct, unique, and relevant statement. Students were encouraged to provide more than the requested number of statements, so even if students provided irrelevant or factually incorrect statements, they could still receive a maximum score if they provided enough correct, unique, and relevant statements. The average scores for the graded CEs ranged from 4.3 to 4.5 ( 86 to $89 \%$ ). These averages are based only on the scores of those students who completed the CEs.

Analysis of CE Data. For analysis, the student responses that were deemed relevant to the CE were categorized by the generic statements they fell under on the rubric (regardless of whether the statement was correct or incorrect). For example, in response to CE3 (healthy DNA and UV-irradiated DNA), Liam (students have been assigned code names) wrote "UV radiation causes dimer formation in adjacent thymine base of the same DNA," while Sophia noted "UV radiation will lead to thymine dimer, intrastrand." Both of these responses were categorized as belonging to the rubric statement: "during one type of UV damage, pyrimidine (thymine \& cytosine) base pairs that are stacked on top of each other dimerize into pyrimidine dimers." After this categorization, the prevalence of student responses that fell into each statement on the rubric was examined. The edited rubrics were thus a comprehensive collection of the ideas that students provided in response to the CEs.

For students to complete the CEs successfully, they were required to provide correct, relevant, and unique statements in response to the prompt. The CEs did not specifically ask students about the chemical identities of the substances presented in the prompts, as this would have violated the open-ended nature of the CEs. Only relevant statements were evaluated for their correctness. Relevance implies usage of the facts for a purpose, and because this was a biochemistry course and the CEs counted as quiz grades, the students most likely consid-
ered that "relevance" meant solving a biochemical problem. Characterization of biological molecules, contrasting an altered biological molecule with its normal counterpart, and thinking about biochemical pathways are typical biochemistry problems. In any of these types of problems, as well as other biochemistry problems, identifying and differentiating what substance is involved could be relevant. However, because students were given instructions to provide facts and were not directly asked questions related to CI, it was unlikely that students would display full arguments that would contain CI thinking. Instead, students were more likely to state their knowledge but not their reasoning as to why this knowledge was relevant to the provided context. Thus, the statements students provided gave evidence of knowledge that could potentially be used to solve questions of CI but did not exhibit the full characteristics of CI thinking. This knowledge that has the potential to lead to CI thinking will be referred to as "precursors to CI thinking." Our challenge was 1) to infer whether the students' responses contained precursors to CI thinking in students' statements, and if so, 2) to characterize them using the eight themes of CI, summarized in Table 1.

The rubrics for CE2, CE3, and CE4 were coded by the two authors for the presence of CI themes. Nearly all of the statements not coded as related to CI were identified as belonging to structure-property relationships (SPR) thinking, which has been previously defined as a type of thinking in chemistry that answers a different question than CI thinking does (Sevian and Talanquer, 2014). SPR thinking answers the question "How do we predict the properties of chemical substances?" Differentiating CI and SPR is based on the purpose of the knowledge that is being applied; for example, in response to CE2 (fatty-acid comparison) Jackson responded "Both molecules have unsaturation, but molecule A has four double bonds instead of molecule B's one." This was coded as CI thinking related to the "composition and structure" theme, because a student would likely have considered this knowledge to be relevant to solving problems that would require analysis based on composition and structure. Jackson also wrote, "Molecule A has a lower melting point than molecule B." Because this was related to specific properties of each molecule being relevant to answering some question about the substances, this statement was coded as belonging to "structure-property relationships" thinking.

The authors individually coded the rubric statements using the previously established themes of CI. For most rubric items, the coders concurred on whether the ideas expressed were precursors of CI thinking or not, and if so, which CI theme was relevant. In cases in which the authors were not in agreement, they discussed which category the rubric statement belonged to and came to a consensus for the coding of that rubric statement until agreement was reached for all statements. Because the CE rubrics are not designed to be used for research outside of this study, interrater reliability was not calculated as a form of validation for the rubrics.

As an illustration of the difference between student statements versus rubric statements, a portion of one student's responses (code-named Riley) to CE2 has been reproduced in Box 2. Each numbered statement represents a statement made by Riley in response to CE2, as written by Riley. Each statement made by Riley was assigned to a rubric statement, and the corresponding rubric statement has been reproduced beneath each

## Box 2．A Portion of Riley＇s Responses to CE2

1．Molecule A is an unsaturated fatty acid
a．Mention of type of molecule（Molecule A is a fatty acid and is arachidonic acid，Molecule B is a type of eicosanoid and is prostaglandin E1）—class
2．Molecule A has 20 carbon，so it＇s very hydrophobic in x area［student labeled the hydrocarbon tail of molecule A as＂$x$＂］
a．The chemical formula for each molecule $\left(\mathrm{C}_{20} \mathrm{H}_{32} \mathrm{O}_{2}\right.$ ， $\mathrm{C}_{20} \mathrm{H}_{34} \mathrm{O}_{5}$ ）or referencing the types and number of atoms－composition and structure
3．The reactive portion of molecule $A$ is the $y$ area［student labeled the carboxyl group as＂$y$＂］
a．Mention of the reactivity or stability of parts of the molecules（functional groups，double bonds，nonre－ ducing ends）—structure－property relationships
4．Molecule A has a carboxylic group，which can be used in H －bonding．
a．Functional groups：Molecule A has a carboxylic acid group，Molecule B has a ketone，also has a terminal carboxylic acid group and two alcohol groups－com－ position and structure
5．This molecule can be added or combined with two other fatty acids to make TAG or triacyl glycerol
a．Can be utilized to make diacylglycerol or triacylglyc－ erol molecules－function
of Riley＇s statements．The rubric statements were coded for the presence of precursors to CI thinking，structure－property relationships，or coded as not related to either，and the coding for each rubric statement has been provided in italics．

For better understanding of the different ways in which stu－ dents responded to the CEs，two student profiles that character－ ize how students used CI in different ways in their CE responses have been created．These two students received markedly dif－ ferent scores overall（ 100 vs． $80 \%$ ）on their CEs．These profiles can be found in the Supplemental Material and can be used to understand the different CI themes students used to respond to the CEs．Creating profiles for a class could help instructors to see whether there are subsets of students who consistently use
the same CI themes in their responses，and whether these pat－ terns are correlated to performance．

Results from CE Data：Observed Cl Themes in CE Responses．The CI theme corresponding to each rubric state－ ment that contained precursors for CI thinking was identified． The prevalence and distribution of the CI themes identified in student responses to each CE are outlined in Table 4．These numbers represent how frequently students responded with statements related to a CI theme for each CE．They are counts of the statements students made and are not reflective of whether students received credit for each statement．

Although seven unique themes of CI were identified as related to the student responses（＂change，＂＂class，＂＂composi－ tion and structure，＂＂function，＂＂organism effect，＂and＂tests and experimental values＂），＂composition and structure＂rep－ resents the majority of the statements coded as related to CI．In response to all three CEs，students provided 450 statements （examples of one student＇s statements can be found in Box 2）． Out of the students＇responses， 185 statements belonged to rubric statements coded as＂composition and structure．＂This represents $\mathbf{~ 4 1 \%}$ of the students＇responses．
＂Class＂and＂function＂were the only other CI themes for which CI－thinking precursors were observed in all three of the CEs．Student statements coded as＂class＂typically involved cat－ egorizing the molecule in the CE（e．g．，fatty acid，eicosanoid）， while statements related to＂function＂contained information about how a molecule might be used（e．g．，is used in signaling pathways）．Both themes appeared most frequently in CE2，rep－ resenting 14 and $8 \%$ of student responses，respectively．In CE3 and CE4，the representation of both themes was low，with $6 \%$ or fewer of the student statements coded as related to＂class＂or ＂function＂for each CE．

The CI themes＂change＂and＂tests and experimental values＂ were present as precursors of CI thinking in student responses to CE3．For＂change，＂students commented on what happened when the UV radiation interacted with the DNA．Their focus on characterizing the process（e．g．，the DNA is chemically altered） prompted these statements to be coded as related to the ＂change＂theme of CI，and represented $4 \%$ of statements made by students．The statements coded as related to＂tests and experimental values＂were the second most frequent CI theme present in CE3，at $9 \%$ of the students＇statements，and all

## TABLE 4．Total distribution of coded student responses ${ }^{\text {a }}$

|  |  | $\begin{aligned} & \text { ू } \\ & \text { む̃ } \end{aligned}$ |  |  |  |  |  |  | 先 | 先 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| CE2 | 0 | 25 | 78 | 15 | 0 | 0 | 0 | 0 | 51 | 10 |
| CE3 | 5 | 8 | 47 | 2 | 0 | 0 | 0 | 12 | 15 | 45 |
| CE4 | 0 | 1 | 60 | 5 | 1 | 0 | 5 | 0 | 29 | 36 |
| Total | 5 | 34 | 185 | 22 | 1 | 0 | 5 | 12 | 95 | 91 |

${ }^{a}$ Tests \＆exp．values，tests \＆experimental values；SPR，structure－property relationships；NR，not relevant to CI or SPR．
Shading has been added to indicate the frequency of statements that were coded as a particular CI theme relative to the other coded statements for that CE．Darker shading indicates a higher frequency of statements coded as that CI theme in comparison to the other CI themes．
involved using different laboratory techniques to separate and identify the normal versus damaged DNA.

Precursors of CI thinking that fell into the "source" and "organism effect" themes were only observed in student responses to CE4, and their frequency was low in comparison with precursors related to other observed themes ( 4 and $<1 \%$ of students' statements, respectively). The statement corresponding to "source" noted that acetaminophen is a component of or can be found in drugs, and the statement coded as "organism effect" contained information on the effect of acetaminophen on humans.

## DISCUSSION

Although there is evidence for the majority of the CI themes in students' responses to the CEs, students were more likely to respond using certain CI themes over others. The distribution is likely influenced by the context of the CEs, as previous work with the CI themes has indicated that they are context dependent. The distribution of CI themes that students relied on was also different compared with the CI themes elicited from the practicing biochemists. This may be dependent on the method that was used to elicit responses about CI, as the practicing biochemists were explicitly asked to consider CI in their own biochemical work, whereas students were asked to respond to CEs within biochemical contexts. The ways in which students applied CI themes and possible reasons behind the varying presence of each theme are explored in the following sections.

## Students' Usage of Cl themes in CEs

The prevalence of statements related to "composition and structure" indicates that students regarded "composition and structure" as relevant for the biochemical contexts provided in the CEs. In CE2 and CE4, molecular structures were provided as part of the context. This likely contributed to the greater prevalence of student responses in these two CEs that exhibited precursors of CI thinking related to this theme. The statements coded as "composition and structure" included general observations about the substances or molecules presented in the CE context, such as noting the degree of unsaturation and location of the double bonds for the fatty acids in CE2 or identifying the types of functional groups attached to the benzene ring in CE4.

The presence of the chemical structure in the context was not necessary to elicit "composition and structure" statements, however, as evidenced by the responses to CE3, for which the context focused on healthy DNA versus DNA damaged by light. Statements involving precursors for "composition and structure" for CE3 included, for example, noting the different units that make up DNA (base + pentose sugar + phosphate group) and descriptions of the types of bonds that form between base pairs as a result of DNA damage. In CE3, students were not provided with any images or structures related to DNA, nor did the context indicate that the UV radiation impacted the composition or structure of the DNA. Thus, the overwhelming presence of "composition and structure" precursors in students' responses to all of the CEs indicates that students have little trouble linking "composition and structure" to biochemical contexts. The composition and structure of molecules and macromolecules is discussed frequently throughout a typical biochemistry curriculum, with some textbooks (e.g., Lehninger

Principles of Biochemistry) bringing up these concepts as early as the introductory chapter (Nelson et al., 2008).

Evidence of the theme "class" appeared most frequently in students' responses to CE2. The context for CE2 presented two molecules, and although students were not explicitly directed to do so, students often included information about classifications of each molecule. CE3 and CE4 also presented multiple substances, but students' usage of "class" in their responses to these CEs was not as frequent. This may be due to the complexity of the context; in CE3 and CE4, more details were provided to students (e.g., where the DNA came from, the rates of different excretion pathways), which may have drawn students' attention away from classifying the substances in the context. These details may have cued students to rely on other ways of thinking about the substances, and suggests that, if instructors would like their students to focus on classifying substances, the context should not contain details beyond the presentation of the substances themselves.
"Function" appeared in responses to all CEs. The substances in the CEs (fatty acids, DNA, acetaminophen) are generally described at the submicroscopic or molecular level, and they were more frequently represented at this level in the contexts. The meaning of function for a substance might vary when considered at the molecular level, and the idea of function in biochemistry is likely to be tied to submicroscopic processes. For example, one of the rubric statements for CE2 coded as "function" was "Molecule A is used in phospholipids and in cell membranes, molecule B is a type of hormone used in signaling pathways." This statement can be considered a precursor to a CI argument about the function of a substance being used for classification or differentiation, as it directly links the molecules in the context to their purposes in biological systems.

The other statements coded as "function" in CE3 and CE4 related to the function of DNA as the genetic code for living organisms and the use of acetaminophen as a pain reliever. These are generally accepted purposes or uses of these two substances, but few students included these statements about function in their responses. Several reasons might explain why students did not include these statements. First, at this level in their undergraduate education, biochemistry students might be trained to look beyond (or overlook) the more obvious usage or functions of substances. Second, students may have framed the CE as a school activity (Davis and Russ, 2015) and therefore discounted common knowledge while privileging school knowledge.

The relatively low presence of "tests and experimental values" precursors in responses to the CEs suggests that, when students are not specifically asked to identify and differentiate substances, they are unlikely to talk about these types of methods. All student responses that were coded as "tests and experimental values" suggested some method for separating or distinguishing the healthy versus the damaged DNA in the context of CE3. In some cases, students included information related to "composition and structure" in the service of "tests and experimental values" (e.g., the negatively charged backbone of the DNA enables it to be separated using electrophoresis). This is one example of many in which CI themes can be used in conjunction with one another, which indicates that reasoning within a chemical or biochemical context might require the application of more than one way of classifying or differentiating a substance.

Although there was only one statement on the CE4 rubric that contained a precursor related to "source," five different students provided statements that fell within this category. The context of CE4 appears to have triggered students to consider "source" in this CE and not the others. Possible reasons could include familiarity with this compound outside of biochemistry and the fact that this was the only substance presented in the CEs that is exogenous to the body. Students are less likely to encounter arachidonic acid or DNA in commercial products, which may have prompted them to draw on their biochemistry knowledge instead.

## Explaining the Varying Presence of Cl Themes

Precursors of CI thinking that fell into seven unique CI themes were observed in students' responses to the CEs; however, the theme of "composition and structure" was more prevalent than the other themes. These trends may be linked to the context of the problem, the discipline perceived by the student to be targeted for the activity (chemistry or biochemistry), the emphasis of the biochemistry curriculum at the time the CE was administered, or a combination of these variables. Further studies would be necessary to determine whether "composition and structure" is a more dominant CI theme in biochemistry than in chemistry, and how the specific substance being examined influences the theme of CI elicited. The emphasis of certain CI themes in the biochemistry textbook and in biochemistry courses could also be examined to determine whether they might have influenced what students perceived as relevant when solving problems in biochemistry.

Although there were eight themes of CI thinking revealed in the responses to the CSI Survey (Ngai and Sevian, 2017), the CE responses did not elicit the same range. A variety of reasons might have contributed to this outcome. The first is that the CEs did not directly ask students to solve problems of CI. Because students were not deliberately prompted to classify and differentiate substances, the range of CI themes was likely limited.

Second, although the CSI Survey deliberately asked questions related to CI, no question elicited all eight themes of CI. CEs developed by other researchers (Warfa and Odowa, 2015; Xu et al., 2017) also captured the transfer of a limited number of chemistry concepts to the CE contexts, as the design of the CE is not meant to elicit a student's comprehensive knowledge of chemistry. Thus, from previous work and from these CEs, it can be inferred that the themes of CI relied upon are in part linked to the nature of the context. This is not surprising, as CI thinking encompasses many approaches to identifying and differentiating substances, and it is unlikely they will all be applicable to every problem in chemistry or other interdisciplinary problems. If a response based on one way of using CI is satisfactory for the problem at hand, there is no need to provide additional CI thinking. Additionally, every context will have certain features that are likely to be more salient than others, and students and experts are more likely to respond to those obvious cues. These cues are dependent on the nature of the problem.

It is also likely that certain themes of CI are more relevant for problems in biochemistry than others. The majority of the examples provided by the practicing biochemists used "composition and structure," "tests and experimental values," or both.

The prevalence of "composition and structure" in practicing biochemists' responses reinforces the claim that this theme of CI might be more relevant in biochemical contexts than in the general contexts in the CSI Survey.

The presence of "tests and experimental values" in practicing biochemists' responses might be attributed to the design of the survey. Participants were explicitly asked how they would identify and/or differentiate substances in their example biochemical problems. This format is similar to the CSI Survey, in which students were explicitly asked to identify and/or differentiate substances (and "tests and experimental values" was the most prevalent theme). The CEs, on the other hand, did not explicitly direct students to consider identification and differentiation of substances, which might explain the lower presence of precursors related to this theme in students' CE responses.

Finally, the CEs were not designed to elicit the broadest possible range of themes of CI thinking. On the other hand, this was a goal of the research using the CSI Survey, that is, the questions in the CSI Survey were specifically designed to uncover as many types of CI thinking as possible. Because the CEs asked for relevant facts, the CEs gave students opportunities to exercise CI thinking. However, students primarily responded with knowledge that they found relevant. This meant we had to infer the CI thinking that led students to consider that knowledge relevant and, from that, judge which themes these precursors of CI thinking related to. Based on the prevalence of CI themes in the data that we were able to infer from the rubric statements, it seems clear that students do consider precursors to CI thinking relevant to biochemical contexts. Thus, in answer to the overarching goal of the study, we consider that there is a link between precursors of CI thinking and biochemistry contexts.

## Broader Impacts

The CEs developed for this study can be used in other teaching and research contexts beyond the ones described in this manuscript. The rubrics obtained from the administration of these CEs with our sample will be a starting point, but may not be useful with biochemistry students at other institutions unless modified by answers collected in practice (as we did when revising from our initial rubrics to the rubrics postadministration). The rubrics that emerged from our study came out of the data collected in this study and likely depend considerably upon institution-specific variables, for example, the course sequence before students take biochemistry and/or the curriculum and instructional resources used in teaching biochemistry. In general, CEs have been used to provide evidence of and promote the behavior of linking previously learned knowledge to new contexts (Ye and Lewis, 2014; Warfa and Odowa, 2015). The CEs presented here can be used to reinforce this type of behavior, with an emphasis on linking chemical knowledge to biochemical contexts. Additionally, these CEs can deliberately be used by instructors or researchers to explore the precursors to CI thinking that students may apply to biochemical contexts. Instructors can create student profiles similar to those provided to qualitatively compare the types of CI themes that are present in student responses. These profiles can be shared with students to encourage them to explore other types of CI knowledge that may be relevant to a CE.

## Limitations and Future Work

Owing to the nature of the CEs, it is impossible to be certain why students thought their statements were relevant to the context of the exercise, or in what ways students would use the knowledge they presented in these CEs. Students were rewarded for providing unique, relevant, and correct statements and were not required to justify them; thus, we could only capture precursors to CI thinking in student responses to the CEs. To gain more certainty about the practice of CI thinking by students, open-ended questions that ask for justification or cognitive interviews could be used to probe student reasoning for linking specific precursors of CI thinking to the provided biochemical contexts.

Although practicing biochemists were surveyed to collect biochemical scenarios in which CI thinking would be used, the CE rubrics were developed based on student responses alone. This resulted in a lack of variety of biochemical ideas with regard to the themes of CI that might be applied to the CE contexts. The observation of "composition and structure" used in conjunction with "tests and experimental values" in practicing biochemists' survey responses indicates there are other relationships between CI themes that were not captured in either the CSI Survey data or CE data. To gain a better understanding of the range of responses possible for these CEs, researchers should administer them to both practicing biochemists and students.

Although the results presented in this study do include an evaluation of the quality of the precursors of CI thinking the students provided in response to the CEs, future work can build on the evidence that a link between CI and biochemical contexts exists. The CEs can be modified or new assessments developed that draw from these CEs to investigate the extent to which students link CI and biochemistry and the quality of the CI thinking they employ, similar to instruments designed by other researchers that examine the chemistry knowledge of students within biochemical contexts (Villafañe et al., 2011; Haudek et al., 2012; Wolfson et al., 2014; Warfa and Odowa, 2015). The effect of different biochemical (or other disciplinary) contexts on the CI thinking that students employ can also be investigated, similar to Kohn et al.'s (2018) study that investigated students' concepts of energy when framed within a variety of contexts.

Future studies can build from this research and more closely examine the types of CI knowledge that can be applied to biochemical contexts and the effect of expertise on what CI knowledge is applied. Following that, whether specific instructional practices have an effect on students' application of CI to biochemical contexts can be explored; for example, future studies could investigate whether deliberately linking general chemistry concepts to biochemical problems in the classroom will lead to observed behavior in students that is different from what was observed in this study. Ultimately, this could contribute to the development of supported learning environments for targeted science content such as CI in biochemistry.

## CONCLUSION

The survey of practicing biochemists and analysis of student responses to the CEs illustrate some ways in which CI is relevant to biochemical contexts. Certain CI themes, for example,
"composition and structure" and "class," occur more frequently than others. Based on these observations, students may need more assistance linking some themes of CI to biochemical contexts than others. Instructors can use the CEs to determine what themes of CI their students do not link to biochemical contexts and then facilitate students' linkages of previously learned concepts to new problems in biochemistry. CEs can also be used to reward students for linking previously learned content knowledge to new contexts.

Examining the themes of CI in the context of biochemistry has added depth to the existing understanding of CI thinking. It is useful to consider these themes from another discipline, as this study demonstrates that the concept of CI is relevant outside the field of chemistry. This helps provide a more comprehensive understanding of CI and suggests that there might be other unique ways of characterizing substances in other disciplines. Future work can explore these other discipline-specific ways of investigating the identity of substances and determine what themes of CI are more relevant for the discipline of biochemistry than others.

## ACKNOWLEDGMENTS

We thank biochemist Dr. Daniel Dowling for many helpful conversations regarding this work. Additionally, this work was part of the C.N.'s doctoral research at University of Massachusetts Boston, and we acknowledge funding from the Sanofi Genzyme Doctoral Research Fellowship, the University of Massachusetts Boston Distinguished Doctoral Fellowship, and National Science Foundation awards DRL-1222624 and DUE-1348722.

## REFERENCES

American Association for the Advancement of Science. (2011). Vision and change in undergraduate biology education. Washington, DC.
Ball, P. (2004). The elements: A very short introduction (Vol. 104). Oxford University Press.
Colón, W., Chitnis, P., Collins, J. P., Hicks, J., Chan, T., \& Tornow, J. S. (2008). Chemical biology at the US National Science Foundation. Nature Chemical Biology, 4(9), 511-514.
Davis, P. R., \& Russ, R. S. (2015). Dynamic framing in the communication of scientific research: Texts and interactions. Journal of Research in Science Teaching, 52(2), 221-252.
Duit, R., Gropengießer, H., Kattmann, U., Komorek, M., \& Parchmann, I. (2012). The Model of Educational Reconstruction-A framework for improving teaching and learning science. In Jorde, D., \& Dillon, J. (Eds.), Science education research and practice in Europe (pp. 13-37). Boston, MA: Sense Publishers.
Eilks, I., \& Hofstein, A. (Eds.). (2015). Relevant chemistry education. Boston, MA: Sense Publishers.
Ellis, B. (2002). The philosophy of nature: A guide to the new essentialism. Montreal: McGill-Queen's University Press.
Haudek, K. C., Prevost, L. B., Moscarella, R. A., Merrill, J., \& Urban-Lurain, M. (2012). What are they thinking? Automated analysis of student writing about acid-base chemistry in introductory biology. CBE-Life Sciences Education, 11(3), 283-293.
Kohn, K. P., Underwood, S. M., \& Cooper, M. M. (2018). Energy connections and misconnections across chemistry and biology. CBE-Life Sciences Education, 17(1), ar3.
Lewis, S. E., Shaw, J. L., \& Freeman, K. A. (2010). Creative exercises in general chemistry: A student-centered assessment. Journal of College Science Teaching, 40(1), 48-53.
Lewis, S. E., Shaw, J. L., \& Freeman, K. A. (2011). Establishing open-ended assessments: Investigating the validity of creative exercises. Chemistry Education Research and Practice, 12, 158-166.

## C. Ngai and H. Sevian

Linenberger, K. J., \& Bretz, S. L. (2015). Biochemistry students' ideas about how an enzyme interacts with a substrate. Biochemistry and Molecular Biology Education, 43(4), 213-222.
Loertscher, J., Green, D., \& Lewis, J. (2014). Identification of threshold concepts for biochemistry. CBE-Life Sciences Education, 13, 516-528.

National Research Council. (2009). A new biology for the 21st century Washington, DC.
Nelson, D. L., Lehninger, A. L., \& Cox, M. M. (2008). Lehninger principles of biochemistry. New York: MacMillan.
Ngai, C. (2017). An investigation of chemical identity thinking. (Graduate doctoral dissertation). Retrieved from ScholarWorks Dissertations and Theses (336).
Ngai, C., \& Sevian, H. (2017). Capturing chemical identity thinking. Journal of Chemical Education, 94(2), 137-148
Ngai, C., Sevian, H., \& Talanquer, V. (2014). What is this substance? What makes it different? Mapping progression in students' assumptions about chemical identity. International Journal of Science Education, 36(14), 2438-2461.

Rowland, S. L., Smith, C. A., Gillam, E., \& Wright, T. (2011). The concept lens diagram: A new mechanism for presenting biochemistry content in terms of "big ideas." Biochemistry and Molecular Biology Education, 39(4), 267-279
Sevian, H., \& Talanquer, V. (2014). Rethinking chemistry: A learning progression on chemical thinking. Chemistry Education Research and Practice, 15(1), 10-23.
Shi, J., Wood, W., \& Martin, J. (2010). A diagnostic assessment for introductory molecular and cell biology. CBE-Life Sciences Education, 9(4), 453-461.

Strand, R. (1999). Towards a useful philosophy of biochemistry: Sketches and examples. Foundations of Chemistry, 1(3), 269-292
Talgar, C. P., \& Goodey, N. M. (2015). Views from academia and industry on skills needed for the modern research environment. Biochemistry and Molecular Biology Education, 43(5), 324-332.

Trigwell, K., \& Sleet, R. (1990). Improving the relationship between assess ment results and student understanding. Assessment and Evaluation in Higher Education, 15(3), 190-197
van Brakel, J. (2000). The nature of chemical substances. In Bhushan, N., \& Rosenfeld, S. (Eds.), Of minds and molecules: New philosophical perspectives in chemistry. New York: Oxford University Press.
VandeWall, H. (2007). Why water is not $\mathrm{H}_{2} \mathrm{O}$, and other critiques of essentialist ontology from the philosophy of chemistry. Philosophy of Science 74(5), 906-919.

Villafañe, S. M., Bailey, C. P., Loertscher, J., Minderhout, V., \& Lewis, J. E. (2011). Development and analysis of an instrument to assess student understanding of foundational concepts before biochemistry coursework. Biochemistry and Molecular Biology Education, 39(2), 102-109.
Warfa, A.-R. M., \& Odowa, N. (2015). Creative exercises (CEs) in the biochemistry domain: An analysis of students' linking of chemical and biochemical concepts. Chemistry Education Research and Practice, 16(4), 747-757.
Wolfson, A., Rowland, S., Lawrie, G., \& Wright, A. (2014). Student conceptions about energy transformations: Progression from general chemistry to biochemistry. Chemistry Education Research and Practice, 15, 168-183.
Wright, A., Provost, J., Roecklein-Canfield, J. A., \& Bell, E. (2013). Essential concepts and underlying theories from physics, chemistry, and mathematics for "biochemistry and molecular biology" majors. Biochemistry and Molecular Biology Education, 41(5), 302-308.
Wright, T., \& Hamilton, S. (2008). Assessing student understanding in the molecular life sciences using a concept inventory. ATN Assessment, 8, 216-224.
Xu, X., Lewis, J. E., Loertscher, J., Minderhout, V., \& Tienson, H. L. (2017). Small changes: Using assessment to direct instructional practices in large-enrollment biochemistry courses. CBE-Life Sciences Education, 16(1), ar7.
Ye, L., \& Lewis, S. (2014). Looking for links: Examining student responses in creative exercises for evidence of linking chemistry concepts. Chemistry Education Research and Practice, 15(4), 576-586.

