

How to Identify the Research Abilities That Instructors Anticipate Students Will Develop in a Biochemistry Course-Based Undergraduate Research Experience (CURE)

Stefan Mark Irby,[†] Nancy J. Pelaez,[‡] and Trevor R. Anderson^{†*}

[†]Department of Chemistry and [‡]Department of Biological Sciences, Purdue University, West Lafayette, IN 47906

ABSTRACT

Course-based undergraduate research experiences (CUREs) have been described in a range of educational contexts. Although various anticipated learning outcomes (ALOs) have been proposed, processes for identifying them may not be rigorous or well documented, which can lead to inappropriate assessment and speculation about what students actually learn from CUREs. In this essay, we offer a user-friendly and rigorous approach based on evidence and an easy process to identify ALOs, namely, a five-step Process for Identifying Course-Based Undergraduate Research Abilities (PICURA), consisting of a content analysis, an open-ended survey, an interview, an alignment check, and a two-tiered Likert survey. The development of PICURA was guided by four criteria: 1) the process is iterative, 2) the overall process gives more insight than individual data sources, 3) the steps of the process allow for consensus across the data sources, and 4) the process allows for prioritization of the identified abilities. To address these criteria, we collected data from 10 participants in a multi-institutional biochemistry CURE. In this essay, we use two selected research abilities to illustrate how PICURA was used to identify and prioritize such abilities. PICURA could be applied to other CUREs in other contexts.

INTRODUCTION

In recent years, there has been a concerted effort to include more authentic research practices within science, technology, engineering, and mathematics courses across various science disciplines and educational levels (e.g., National Research Council [NRC], 2003, 2012; American Association for the Advancement of Science, 2011; President's Council of Advisors on Science and Technology, 2012; National Academies of Sciences, Engineering, and Medicine, 2015, 2017). In postsecondary education, the incorporation of authentic research practices in the classroom are often referred to as course-based undergraduate research experiences (CUREs). CUREs are aimed at developing student knowledge and competence to perform more authentic research rather than the ability to follow traditional recipe-like protocols (Auchincloss *et al.*, 2014). There have been various attempts to classify the meaning of a CURE (e.g., Weaver *et al.*, 2008; Lopatto and Tobias, 2010; Brownell *et al.*, 2012). Commonly, CUREs incorporate five components identified by Auchincloss *et al.* (2014): 1) they involve the use of scientific research practices; 2) they have elements of discovery; 3) their products have broader relevance or importance; 4) they afford opportunities for collaboration; and finally, 5) they emphasize the purpose of and allow for iteration of experiments. For the purpose of this essay, we adopt a broad definition of a CURE as a course wherein students engage in activities resembling those done by scientists in a particular field to conduct novel investigations about relevant phenomena that are currently unknown.

Ross Nehm, *Monitoring Editor*

Submitted December 11, 2017; Revised January 16, 2018; Accepted January 30, 2018

CBE Life Sci Educ June 1, 2018 17:es4

DOI:10.1187/cbe.17-12-0250

*Address correspondence to: Trevor R. Anderson (ander333@purdue.edu).

© 2018 S. M. Irby *et al.* 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.

In response to the abovementioned calls for more integration of authentic scientific practices into undergraduate curricula, there has been an increase in the number of CUREs being implemented and studied across a range of disciplines, formats, and academic levels (see CUREnet [n.d.] for examples of different CURE projects). Most notably, these efforts have been primarily documented within biology and related subdisciplines (e.g., Jordan *et al.*, 2014; Mordacq *et al.*, 2017) in which the motivation has been to better prepare students to keep pace with ongoing research advances in the life sciences (Pelaez *et al.*, 2014). Whereas several examples of CUREs have also been cited in the chemistry education literature (e.g., Weaver *et al.*, 2006), fewer have been published for other scientific disciplines, including biochemistry (e.g., Gray *et al.*, 2015; Craig, 2017), the focus of this essay. Typically, when CUREs are implemented, they are either adapted from a pre-established CURE (e.g., Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science; Jordan *et al.*, 2014; Hatfull, 2015); or they can be backward-designed to create an experience that targets a specific learning, departmental, or university goal (e.g., Shapiro *et al.*, 2015); or, as in the present project (Craig, 2017), they can be developed and molded from a faculty member's own research.

There are few examples of published CURE studies with explicitly defined learning outcomes (LOs; e.g., Gray *et al.*, 2015; Makarevitch *et al.*, 2015; Olimpo *et al.*, 2016; Staub *et al.*, 2016). In general, when it comes to LOs in CUREs, they fall into three categories: either they are not explicitly stated, or they are adapted from the abovementioned five components of a CURE (Auchincloss *et al.*, 2014) or another similar published source (e.g., Brownell *et al.*, 2015; Olimpo *et al.*, 2016), or they are actually based on the CURE's content or technical activities (e.g., Gray *et al.*, 2015; Makarevitch *et al.*, 2015; Kowalski *et al.*, 2016; Staub *et al.*, 2016). Also, the focus tends to be on broadly applicable skills across CUREs in general. These might be good for assessing whether a course is CURE-like or not and for establishing broad goals for the students or faculty members to strive toward. But in order to more fully establish the development of research competence in students, it is essential to also identify and assess the unique and specific course-based undergraduate research abilities (or CURAs) that students might acquire during the course. In this regard, a literature review by Corwin *et al.* (2015) outlined various proposed benefits and outcomes of CUREs, mostly pertaining to the five components of CUREs proposed by Auchincloss *et al.* (2014), and commented on how well they were documented. But again, Corwin and coworkers did not focus on the range of CURAs students would be expected to develop in particular CUREs.

When CURE projects are uploaded to the CUREnet database (CUREnet, n.d.), the LOs are usually included, but not all such projects are accompanied by explanations as to what process was used to identify them. Usually, the LOs are either contrived by instructors or, at the most, identified by a group of instructors in a consensus-building session. It is often not clear whether the claimed learning outcomes are merely anticipated learning outcomes (ALOs) or objectives suggested by instructors or verified learning outcomes (VLOs) confirmed from analysis of student responses to assessments. Indeed, it seems

that ALOs and LOs are often used interchangeably. Furthermore, in some cases, scientists who incorporate their own research in a CURE classroom may have difficulty identifying their ALOs, let alone their VLOs. As stated by Anderson (2007), the establishment of ALOs or objectives for a course is an essential step in the educational process. ALOs will inform what and how students will be assessed, what and how instructors will teach, and what and how students will learn if they are to achieve the desired LOs. In line with this thinking, Shortlidge and Brownell (2016), Anderson (2007), and many others have pointed out that the assessment method and instrument must be selected with the specific outcomes of the CURE under study in mind. But particular problems arise when trying to evaluate a CURE with an assessment instrument that was developed for a different CURE or context. To address the above concerns, a framework has been proposed to help robustly evaluate CUREs and the learning that takes place in them (Brownell and Kloser, 2015). A key component of the framework is to first identify the outcomes of the course, but once again, these authors do not provide a clear way to do so (Brownell and Kloser, 2015; Shortlidge and Brownell, 2016). Thus, there is a need to develop a rigorous process for identifying the CURAs that compose the ALOs associated with CUREs. In our view, this is key to more fully understanding how CUREs can benefit undergraduate education and the future of novice researchers. To our knowledge, no rigorous, data-driven process is currently available to identify such CURAs unique to a specific CURE that instructors anticipate students will develop (i.e., ALOs). In our view, this is a crucial first step before assessment design and confirmation of the actual VLOs of a CURE can be achieved.

Thus, the goal of this essay is to report on our development of a data-driven process for identifying the CURAs that instructors anticipate students will develop while experiencing the biochemistry laboratory CURE designed by Craig (2017) and coworkers. For reader convenience and clarity, an overview of our developed process is presented upfront in Figure 1, while the details of the process are included later in the *How to Apply the Five Steps of PICURA* section. To be in line with the terms CURE and CURA, we use the acronym PICURA to stand for Process for Identifying Course-Based Undergraduate Research Abilities. A list of all the acronyms used in this essay and their meanings is included in Table 1.

The five steps of PICURA shown in Figure 1 are 1) a content analysis of the lab protocols, 2) an open-ended survey about how scientists conduct similar research to what the students do when performing the lab protocols, 3) a follow-up semistructured interview, 4) an alignment check of the generated ability statements across the previous steps in the process, and 5) a Likert survey to prioritize the identified CURAs. For Steps 1–3 and Step 5, there are inputs (materials or participants used as a data source) and outputs (the resulting products from completing each step of the process). More specifically, the data sources include the CURE protocols (Step 1), nine course instructors (Steps 2 and 5), and the lead designer (Steps 2, 3, and 5). Step 4 does not have an input or output like the others. Rather, it is an alignment check to assure consensus between Steps 1–3 before proceeding to Step 5.

We identified the following criteria that we consider important to ensure that the structure and design of PICURA meets

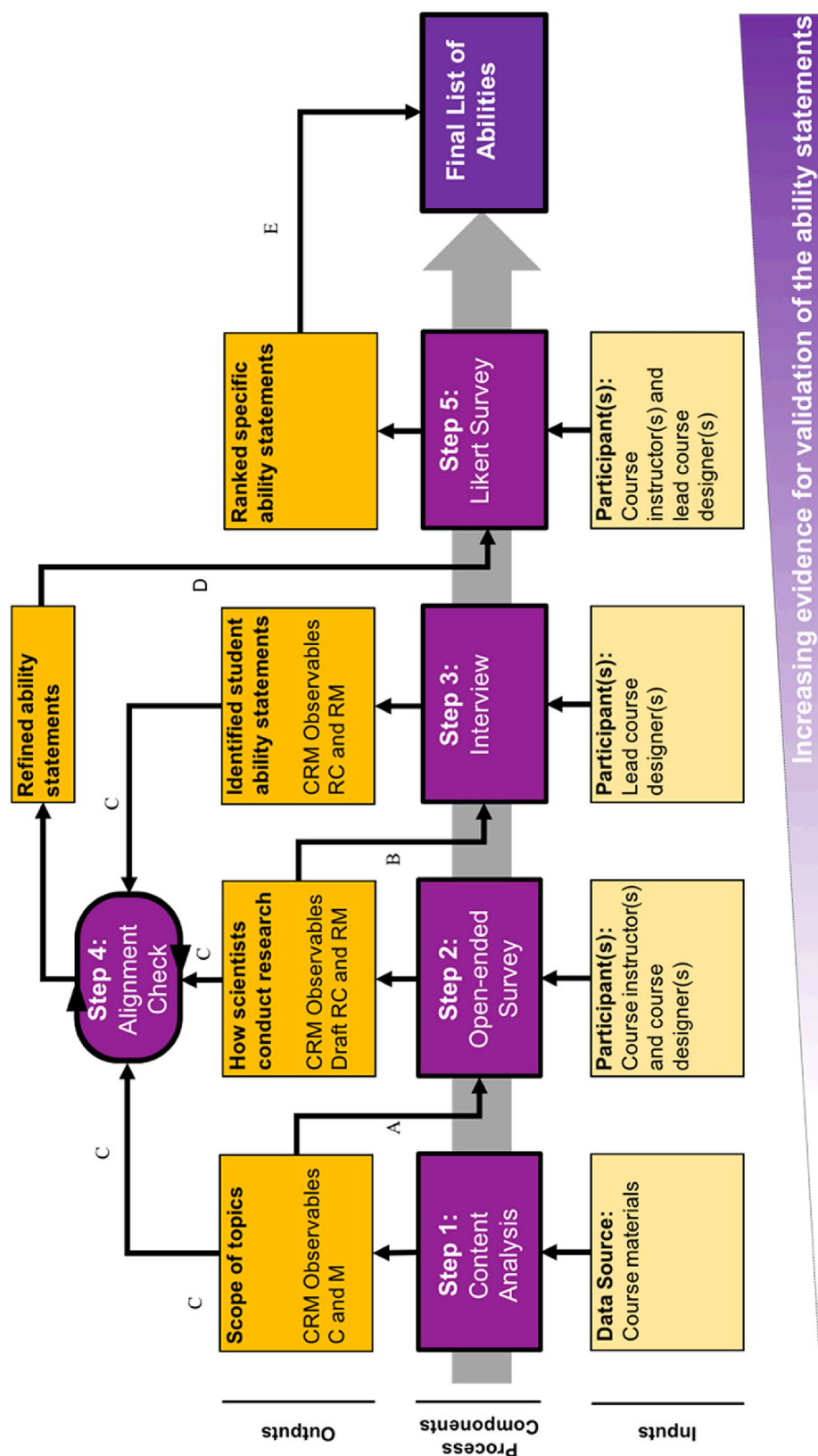


FIGURE 1. Diagram detailing the inputs and outputs for each of the components of the five-step PICURA. Arrows A, B, and D = informing the process; arrow C = alignment steps showing consensus; arrow E = prioritization. The information gained from each component is additive, meaning that each component is informed by the previous process component. Because of the nature of this process as it moves through the steps (going from a content analysis, to an open-ended survey about the course activities, to an interview, to a Likert survey), data to identify and support the abilities increased with each subsequent step. There is one feedback loop, which acts as an alignment check between the generated abilities from the interview, the scope of the topics covered identified from the content analysis, and how scientists conduct research from the open-ended survey. The final process component provided additional evidence and filtering of the abilities via a Likert survey in which participants rate their importance and uniqueness to the course.

TABLE 1. Acronyms used in this essay and their meanings

Acronym	Meaning
ALO(s)	Anticipated learning outcome(s)
CURA(s)	Course-based undergraduate research ability(ies)
CURE(s)	Course-based undergraduate research experience(s)
LO(s)	Learning outcome(s)
LR	<u>An example of a “lower-rated” CURA statement:</u> <i>Recognize how proteins that are closely related by evolution can have dramatically different functions</i>
PICURA	<u>Process for Identifying Course-Based Undergraduate Research Abilities</u>
TR	<u>An example of a “top-rated” CURA statement:</u> <i>Determine using computational software whether, and where a ligand may be binding to a protein</i>
VLO(s)	Verified learning outcome(s)
WR	Weighted-relevance

our abovementioned goal of providing a rigorous process for identifying CURAs that would be well supported by the data collected at each step. For each criterion, we indicate how it relates to the various arrows and links in Figure 1 and refer to various supporting data, presented later in the *Illustration of the Five-Step Process* section, that in our view address the following four criteria:

1. Is the process iterative? That is, do the data generated by each step inform the development and design of the instrument used in the subsequent step(s)? Arrows A, B, and D in Figure 1 and Tables 2–5 show how data inform the process.
2. Does each part of the process give additional data so that together they give more clarity on the expected learning, in terms of CURA statements, than would be known from any one part of the process on its own? Arrows C, E, and the unlabeled arrows in Figure 1 and Tables 5–10 show how the final list of CURA statements are derived.
3. Do the different types of data from each step of the process combine well to achieve consensus and internal alignment about the CURAs? Arrow E in Figure 1 and Tables 8 and 9 (Step 4) lead to the final list of CURA statements.
4. Does the process prioritize the CURAs for instructors, researchers, or other stakeholders? That is, what is the sequence of CURA importance in terms of being relevant to the course of interest? Arrow E in Figure 1 and Table 10 demonstrate the prioritization of the final list of CURA statements.

HOW TO APPLY THE FIVE STEPS OF PICURA

Context: A Biochemistry Lab CURE

We used the biochemistry CURE described by Craig (2017) to develop PICURA. This CURE has only recently been developed by instructors and course designers and is currently being implemented at seven different institutions. The lab is based on modern biochemistry research techniques (e.g., McKay *et al.*, 2015) and aims to provide students with a scientific research experience through a unique combination of computational and wet-lab experiences for which a range of independent protocols have been developed as activities for the CURE (Craig, 2017). The

students uncover new knowledge about the function of some of the many proteins listed on the Protein Data Bank (PDB; www.rcsb.org) whose structure is known but whose function has not yet been established. In addition to developing technical skills, students get the opportunity to develop their thinking, reasoning, and visualization abilities through the application of their biochemistry knowledge to solving problems in novel situations, generating and testing hypotheses through rigorous experimentation, and processing and evaluating results.

Each lab protocol is formatted to contain a background section that introduces either the computational programs and databases or the biochemical techniques being used as part of that particular activity. The background sections also include discussion of the method used, as well as some rationale for why a method was chosen and information about how it works. The protocols also provide either simulations to run or instructional steps that need to be completed when performing the computational modules and biochemical assays (for a list and description of all protocols, see Craig, 2017).

Participants

A total of 10 project members volunteered to participate in different aspects of the development of PICURA, including acting as informants and responding to our data-gathering instruments. Of the participants, nine were instructors of our biochemistry CURE at seven different institutions. Of these volunteers, all had played a role in developing the CURE. The other participant was the “lead designer,” who was also the lead principal investigator and creator of the CURE project. The authors (S.M.I., N.J.P., and T.R.A.) developed the PICURA described in this paper and are also responsible for the educational evaluation of the whole project. The work reported in this essay was officially reviewed, and the data-collection protocols were approved by the Purdue University Institutional Review Board (IRB #1604017549).

The Guiding Framework for the Development of PICURA

Because the goal was to design a process for the identification of the CURAs that instructors anticipated students will develop in our biochemistry CURE, a guiding framework was needed to take inventory of the important components of the course and turn them into ability statements. Toward this end, the conceptual-reasoning-mode (CRM) model of Schönborn and Anderson (2009) was chosen to guide this work. This model has previously been successfully applied to the identification of visual competencies in biochemistry learning (Schönborn and Anderson, 2009, 2010), the structuring of verb–noun reasoning statements, and the design of assessments of student reasoning and problem-solving abilities (Anderson *et al.*, 2013). In addition, the CRM model has been used to guide research looking at how scientists use evolutionary trees (Kong *et al.*, 2017) and how they explain molecular and cellular mechanisms (Trujillo *et al.*, 2015) and, together with the assessment triangle (NRC, 2001), to guide the development of assessment instruments (Dasgupta *et al.*, 2016). The CRM model offers a useful way to account for all the concepts (C), types or modes (M) of representations, and ways of reasoning (R) with those concepts (RC) and representations (RM; Schönborn and Anderson, 2009, 2010). During the development of PICURA, the CRM model was used to inform the construction of ability statements composed of verbs (reasoning

skills) paired with nouns (important concepts or representations; Schönborn and Anderson, 2009). In addition, it was used to guide the data-collection and data-analysis processes at each step of PICURA and to inform the identification of the nature of the reasoning composing each CURA that is used when scientists engage with concepts and representations of phenomena.

Details of the Five Steps of PICURA: Instruments and Data Analysis

This section describes PICURA, the purpose of each step, and how data were collected and analyzed at each step to progressively identify the CURAs. We decided to use various qualitative research methodologies, which in our view and that of Creswell (2012), would afford an effective approach for thoroughly understanding the learning opportunities presented in our biochemistry CURE learning environment. Whereas the five steps of PICURA in Figure 1 were outlined in the *Introduction*, in this section, we provide readers with sufficient details to repeat the process at their own institutions. As will become apparent, PICURA was designed to be both reflective and iterative (see Criterion 1), meaning that each data-collection step is informed by the results of the previous step (Figure 1, arrows A–D). The reflective nature of PICURA, in conjunction with prioritizing the course-specific ability statements in Step 5, allows for progressively gaining more insight into the nature of the CURAs.

Step 1: Content Analysis. Content analysis is a process used to interpret and better understand what is being conveyed in text or other forms of communications, resulting in exact data about the information contained within a document (Cohen *et al.*, 2000; Hsieh and Shannon, 2005). In education research, it is common to perform a content analysis on a variety of course documents (lecture slides, student lab reports, notes, etc.) to identify what students should be learning and to evaluate the content covered in a course. The purpose of the content analysis in PICURA was to identify the concepts (C) and representations (M) of relevance to the various biochemical and computational techniques described in the lab protocols and to gain insight into how information is provided to students and how this CURE is structured.

To analyze the lab protocol data, we used an inductive approach (Thomas, 2006), guided by the CRM model (Schönborn and Anderson, 2009, 2010). In practice, this means there were no specific concepts or representations that were predetermined; rather, the identity of the concepts and representations emerged from analysis and inductive coding of the narratives in the protocols. This allowed for the development of a list of common and recurring concepts and representations that students are exposed to during the course and are accountable for knowing (see Table 5 for an example of this). Additionally, the content analysis gave insight into how information was pre-

sented in the protocols to the students and provided a better context of how a given course using this curriculum would be conducted. The outputs of the content analysis (Step 1) informed the development of the open-ended survey used in Step 2 (see Figure 1, arrow A).

Step 2: Open-Ended Survey. The next step was to survey, with open-ended questions, all instructors and designers about the important concepts and representations that scientists would use and reason with when they do research using the types of methods included in the biochemistry CURE. The open-ended survey was distributed using Qualtrics (www.qualtrics.com), an Internet-based survey and data-collection platform. The open-ended questions provided participants with an opportunity to elaborate on the observed concepts and representations identified in the content analysis by additionally probing participants for examples of how they as scientists would use (reason with) such concepts (RC) and representations (RM) while conducting research similar to research in this biochemistry CURE (see Criterion 1). The questions were intentionally worded to have participants reflect upon how they or other scientists would approach the situation, rather than eliciting any personal beliefs about the learning objectives or ALOs of their CURE, which might vary across institutions. Two examples of the questions are provided in Table 2; the entire set of open-ended survey questions is provided as Supplemental Material. Also, as part of the open-ended survey, participants were asked to list representations that were important to the course (chemical equations, visualizations of molecules, outputs from experiments, graphs, etc.). The open-ended survey results, or outputs from Step 2, were analyzed for common themes and for how the participants reported they were using or reasoning with the different concepts and representations. This led to the emergence of basal-level RC and RM statements, which were noted for further consideration in Step 3 (in line with Criteria 1 and 2). In addition, responses needing more elaboration to generate CURA statements were noted for discussion during the interview, Step 3 (see Criteria 1 and 2). All the analyzed data from Step 2 informed the development of an interview protocol for Step 3 (Figure 1, arrow B).

Step 3: Interview. An interview was conducted with the lead designer to elaborate on the instructors' open-ended survey responses and to identify a clearer and more accurate range of RC and RM statements. Additionally, the interview provided the opportunity to capture the reasoning that scientists would apply to those representations mentioned during Step 2 (Table 2) and provided for discussion during the interview (Table 3). It also afforded the opportunity to discuss how students would use the representations and the concepts related to them as part of the course as well as other difficulties the students might face

TABLE 2. Selected prompts from the course survey given to the CURE instructors and development team members

Open-ended survey prompts about computational techniques and representations
Explain how you, or other scientists, use computational work and protein structural data to investigate protein function.
Please list and describe the types of representations you use, and how you use them, when thinking about or explaining protein function. Representations include but are not limited to items such as the following: Coomassie-stained gels, graphs, computer models, activity assays, protein structures, sketches, diagrams, Bradford assays.

TABLE 3. Example of a portion of the semistructured interview protocol informed by the open-ended survey (Figure 1, arrow B) question about important representations used when thinking about protein function (Table 2)

Interview prompts pertaining to representations
1. In the survey, there was a question asking you to list and describe the types of representations you use, and how you use them, when thinking about or explaining protein function. You provided the following representations: A. <i>Enzyme assays: chemical reaction drawings help us to understand how the parts of a protein catalyze a reaction.</i> B. <i>Molecular visualization: ligand binding demonstrates if a substrate binds to an active site.</i> i. Enzyme assays ii. Chemical reaction drawings 1. How the parts of a protein catalyze a reaction iii. Molecular visualization 1. Ligand binding Could you please talk me through: a) How you would use each representation to reason about protein function? b) What types of biochemistry representations are useful for students to be familiar with to help them in this course? c) How you would like students to use them? d) What you observed students doing with each representation? e) Whether these representations are new to the students, or did they have some previous experiences with them? If so, describe the experiences they had. f) How would you know if students were having difficulties and whether they were improving? g) What type of things did students do to practice and overcome these difficulties?

(the full interview protocol is provided as Supplemental Material). The interviewee was allowed to explain their reasoning and interpretation of these representations, the concepts related to them, and how they would expect students to use them as part of this CURE. The interviewee was also asked to discuss what students would be accountable for knowing when it comes to various representations or techniques, as well as any difficulties students have encountered with any of the CURE activities. However, the interviewer never directly asked for any opinions about the expected CURAs of this CURE. Instead, the development of the interview protocol was directly informed by the open-ended survey, Criteria 1 and 2 (Figure 1, arrow B).

The interview was semistructured with a single participant, the lead designer (performed by S.M.I.). This allowed for guidance concerning what the interviewer wanted to learn from the interview, but allowed for flexibility to discover new, relevant, ideas being shared by the interviewee (Cohen *et al.*, 2000). The lead designer was selected for the interview because the lead designer was the most intimately familiar with the CURE and so was representative of other knowledgeable faculty members designing and implementing it. This provided an opportunity to gain greater insight into the responses to the open-ended survey and into the CURE of interest.

The interview transcripts acted as the primary data source for generating and fine-tuning the relevant CURA statements. The transcript was analyzed and coded, per the CRM model (Schönborn and Anderson, 2009; Anderson *et al.*, 2013), for reasoning with concepts (RC) and representations (RM). CURA statements were generated from the interview transcript by constructing CRM verb–noun pairs from the RC and RM codes. The initial ability statements were generated by looking at RC- or RM-coded segments and doing thought experiments about not only the direct verb–noun pair articulated by the interviewee but any others that would apply to the coded segment. In other words, taking account of the skills that the interviewee discussed led to additional possible CURA statements that could apply to the discussion of the CURE.

Step 4: Alignment Check. After the initial list of CURA statements had been generated from the interview (Step 3), they were first related back and matched to the data obtained from the content analysis (Step 1) and the open-ended survey (Step 2) to ensure that there was alignment across all data sets (Criterion 3; see Tables 8 and 9 for an example). If there was a case in which a CURA statement was not aligned with all data sets (e.g., a CURA statement generated from a portion of the interview was not directly related to the CURE), it would be considered erroneous and ignored. This alignment process ensured that the identified CURA statements were supported by the data gathered about the CURE and that all the observed prominent concepts and representations had been accounted for by the CURA statements (Criterion 3). As part of this step, CURA statements were also refined and optimized by comparing statements to reduce overlap and by conducting a member check with the interview participant, which reduced the total number of CURA statements. After this step, the CURA statements were ready to be prioritized by participants using a two-tier Likert survey (Step 5; Figure 1, arrow D).

Step 5: Likert Survey. After the specific CURA statements had been generated, a Likert-scale survey was used to prioritize them according to what the participants perceived as important to their courses, thereby addressing Criterion 4. Though the proposed utility of PICURA is that the CURAs authentically arise from data, before they can be used to assess student learning and to evaluate this CURE, they must be agreed upon in terms of their appropriateness to this CURE. To achieve this, the CURA statements were evaluated by all instructors and designers with a three-option, two-tiered, Likert survey using Qualtrics. Each CURA was rated on whether the participants expected it to be acquired or further developed by students in the course and whether it was important to the course or specific goals of each participant (Table 4).

For discerning which of the CURAs had the most consensus for this CURE, both Likert questions were used in tandem and a

TABLE 4. Example of the three-option, two-tiered, Likert-scale questions for ranking the CURA statements

Tier	Likert question	Option 1	Option 2	Option 3
First	This ability should have been acquired:	NOT acquired in this lab course	In BOTH this lab and some other course	ONLY in this lab course
Second	How important is this to your students' functioning as scientists?	Unimportant	Undecided	Important

novel weighted-relevance (WR) value was calculated. WR assigns a weight to the options from the Likert scale, such that the abilities could be ranked by how relevant they were to their course. WR is a sum of the participants' scores for their responses to the Likert-scale questions times a multiple between -1 and $+1$ (see Eq. 1 for an example).

$$\begin{aligned} \text{WR} = & (\# \text{Important}) + (\# \text{ONLY in this lab course}) \\ & + 0.5(\# \text{Undecided}) + 0.5(\# \text{In BOTH this lab and some other}) \\ & - (\# \text{Unimportant}) - (\# \text{NOT acquired in this lab course}) \quad (1) \end{aligned}$$

A weight of $+1$ was given for preferred responses indicating that the CURA was unique to an individual's course and/or important (option 3 in Table 4). For responses that were good, but not the preferred responses, such as in cases in which the CURA was acquired in both the individual's course and another course and/or the interviewee was undecided on its importance (option 2 in Table 4), a multiple of $+0.5$ was assigned. For the negative cases, in which a participant thought that the CURA was not acquired in his or her course or was unimportant (option 1 in Table 4), a multiple of -1 was assigned to remove contributions to the WR scores from those abilities with these responses. Thus, the WR scale can range from -2 times the number of participants (the CURA was rated as "NOT acquired in this lab course" and "unimportant" by all participants) to positive $+2$ times the number of participants (the CURA was rated as "ONLY in this lab course" and "important" by all participants). Because there were 10 participants for the Likert survey, the WR score could range from -20 to $+20$. This score indicates the most relevant CURAs to the curriculum, which is important for prioritizing feasible outcomes for our CURE. This step put an emphasis on unique CURAs that students might not have experienced or developed fully if they had not taken a course, which was highlighted by

Shortlidge and Brownell (2016) as a crucial step for the evaluation of CUREs.

ILLUSTRATION OF THE FIVE-STEP PROCESS

In this section, we use two selected research abilities to illustrate how PICURA (Figure 1) was used to identify and prioritize the CURAs of relevance to the biochemistry CURE of Craig (2017). Additionally, the data will be used to demonstrate how each of our four criteria were met while designing the process.

Examples of How Steps 1–3 of PICURA Were Applied to Our Biochemistry CURE

Step 1: Content Analysis. The CURE protocols (see Craig, 2017) were designed in a modular manner so that they could be stand-alone protocols (i.e., they did not directly reference other protocols and could be done independently from one another) and were organized with a background section providing details on the computational program, technique, or biochemical assay that the students would be doing. This was followed by, in general, either a tutorial in the case of computational protocols or an outlined procedure in the case of biochemical protocols. The content analysis of all the lab protocols revealed many different concepts (C) and related representations (M; see Figure 1). An example of how the protocols were used in Step 1 is provided in Table 5, with a small excerpt from the ProMOL protocol (Craig, 2017).

In Table 5, the excerpt provided shows key words bolded to indicate concepts, while phrases that mention skills that students would be using pertaining to each concept are underlined. Within this example, the concepts of protein homology and protein motifs were portrayed and detected. The underlined phrases describe how to compare motifs computationally and what to look for. This protocol also contained an image of a protein of unknown function mapped onto a known protein's active site, using a protein stick representation (Figure 2A). The protocols outlined the key concepts and gave some

TABLE 5. Example of conducting Step 1 of PICURA: Content analysis of the protocols—ProMOL module

Excerpt from protocol ^a	Analysis →	Output ^a
The first step in our function-prediction process is to compare a protein of <u>unknown function against a library of enzyme active sites from the Catalytic Site Atlas</u> that constitute the motif template library of ProMOL. Each catalytic site motif template typically consists of 2–5 amino acid residues that have a <u>fixed spatial and distance relationship</u> . The example shown [Figure 2A] is an alignment for a serine protease.	Read through all lab protocols highlighting passages and coding them for whether they are pertaining to concepts or representations. Then take note of the underlying concepts or representations being portrayed. Additionally, take note of how protocols are organized, how information is presented to the reader, and how protocols connect to one another.	Concepts detected - Protein homology - Protein motifs Representations detected - Protein visualization: software screen shots and stick representations of active site

^aThis portion of a single protocol is intended to showcase some of the breadth of concepts and representations covered by the protocols.

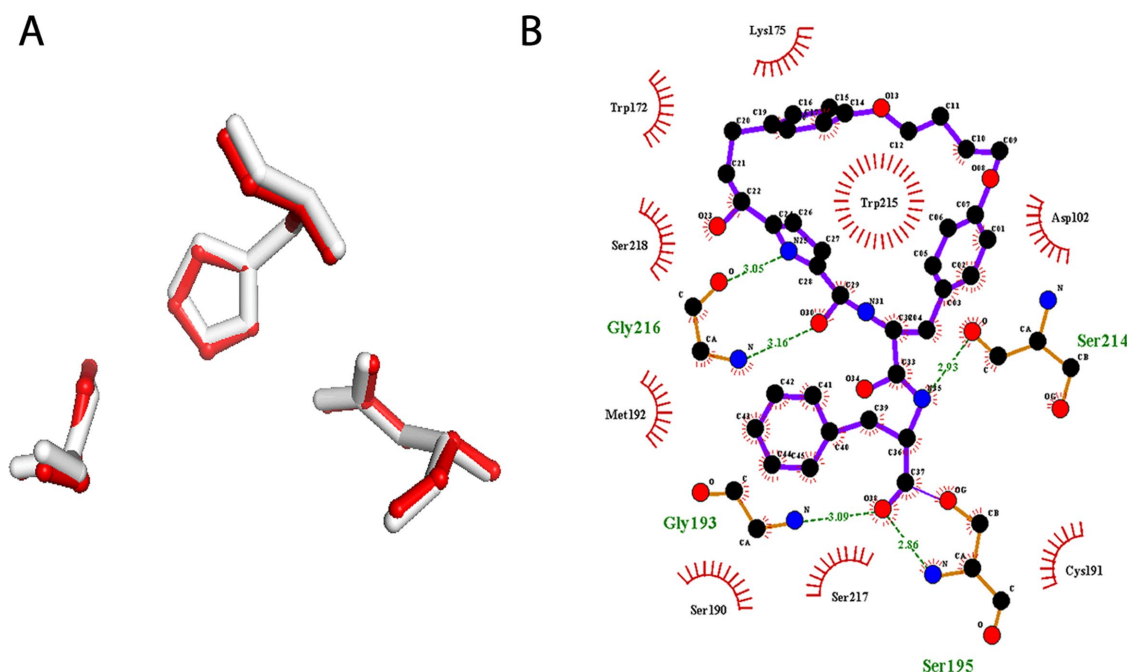


FIGURE 2. Two representations identified by PICURA. (A) Example of an alignment with a serine protease in the ProMOL module lab protocol, as part of the content analysis. (B) LigPlot+ graph mentioned in the open-ended survey and provided for the interview by the lead designer.

examples of representations students may encounter. However, the nature of the protocols did not allow direct analysis of how the students would be connecting all the computational and biochemical techniques together or what types of explorations students might do after learning about the computational techniques or procedures to represent and link together all their collected data. Additionally, we asked the question, “Which representations other than what was presented in the protocols will the students produce or encounter, that scientists would likely use when conducting this type of research?” In other words, we did not assume that representations presented as part of the background information in the protocols provided a complete list of all the representations students would encounter while working through the protocols. For example, students are encouraged to use additional computational simulations, while some of the biochemical protocols give an example of the results (the SDS–PAGE protocol gives an example of a stained gel), but others do not (the protein purification protocol does not provide any representations of the chromatographic results). Thus, to begin to gain greater insight into the nature of the reasoning skills necessary for performing the protocols and to document a more exhaustive list of representations, we conducted an open-ended survey as Step 2 of PICURA (see Figure 1, Step 2, and Table 6). This in turn would allow us to meet Criterion 2 of our process.

Step 2: Open-Ended Survey. The open-ended survey yielded data about how scientists (instructors and designers of the CURE) would perform the research described in the protocols. This, in turn, generated preliminary information about the types of reasoning the participants would expect students to use

during the lab. Selected example responses from the lead designer are provided in Table 6. These responses were typical of the type of responses received from the other participants (unpublished data). When analyzing the lead designer’s responses (Table 6), some of the reasoning (shown in *italics*) with various concepts (RC) and representations (RM; both **bolded**) became much clearer to the authors compared with those identified in Step 1, but the data were still not comprehensive enough to generate clear CURA statements. The nature of the responses was listed stepwise and aligned with the progression of the CURE activities and how they were presented in the protocols (Table 6). Though these response examples (Table 6) were relatively brief, some insight into exactly what scientists use each program or technique for and how they use them started to emerge. Thus, Step 2 suggested that we were starting to meet Criterion 2, in that it was yielding additional information to build on the findings from Step 1. The question about the types of representations provided an opportunity to gain greater insight regarding what other representations may be encountered or generated by students during this CURE. Some of the representations mentioned during this data-collection step had not been previously identified in the protocols (Figure 2B).

Though the open-ended survey provided more details about the possible set of CURAs students may be expected to develop, responses were not detailed enough to determine the extent of student proficiency that would be required for each CURA or to provide a more comprehensive range of possible CURAs for the CURE (Criteria 1 and 2; see Table 6). For example, the lead designer stated the need to “identify ligands that bind [substrate binding] to members of these protein families” (Table 6), but there are a range of CURAs that a student may need to employ

TABLE 6. Example of data yielded by Step 2 of PICURA: Open-ended survey

Selected questions from the survey ^a	Analysis →	Output: coded responses from the lead designer ^a
Q: Explain how you, or other scientists, use computational work and protein structural data to investigate protein function.	Once all the participants (instructors and designers) had completed the open-ended survey, the responses were analyzed for the concepts and representations (bolded), as well as, the reasoning (italicized) applied. ^b Additionally, participant responses and the level of detail they provided were recorded.	A: 1. Look for <u>structural alignments of the full protein backbone to identify folds or families^{RM}</u> . 2. Look for <u>template based alignments of small motifs to identify active sites or ligand binding sites^{RM}</u> . 3. Look at <u>sequence alignments (BLAST) to identify protein families^{RM}</u> . 4. Explore Pfam and UniProt to study the families. 5. <u>Identify ligands that bind [substrate binding] to members of these protein families^{RC}</u> . 6. <u>Dock the ligands to the proteins and see [<i>determine</i>] if there are favorable binding energies^{RM}</u> .
Q: Please list and describe the types of representations you use, and how you use them, when thinking about or explaining protein function.		A: 1. Enzyme assays: <u>chemical reaction drawings help us to understand how the parts of a protein catalyze a reaction^{RM}</u> . 2. Molecular visualization: <u>ligand binding demonstrates if a substrate binds to an active site^{RC}</u> [e.g., a LigPlot+ figure was provided and is shown in Figure 2B].
		Output: general observations • Participants tended to discuss items in sequential order, often similar to how they are presented to the students. • Though reasoning elements were identified, they were at a basal level.

^aFull set of survey questions and responses for the lead designer is provided in the Supplemental Material.

^bUnderlining indicates either an RM or RC (superscript) coded segment with verbs (italics) showing reasoning associated with the noun (bold), which is either a concept or representation.

to identify substrate binding. It was therefore clear to us that the responses to the open-ended surveys tended to reveal only less-than-optimal participant knowledge and that interviews would be necessary to yield a more thorough articulation of their thinking. Thus, a follow-up interview was conducted with the lead designer (whose survey responses are highlighted in Tables 6 and 7) to more deeply understand the open-ended survey responses, to discuss the role of representations mentioned in the survey, and to further probe how scientists perform this type of research, including the reasoning skills that they use (Figure 1, arrow B).

Step 3: Findings from the Interview. To meet Criterion 1, we used the data for all participants from the open-ended survey to inform the design of an interview protocol (see the Supplemental Material). During the interview, to gain greater clarity regarding the open-ended responses, we probed deeper into the nature of the concepts, representations, and related reasoning and how such knowledge is applied by students during the course (Criterion 2). In addition, the lead designer was asked how the representations were generated, what students should be able to extrapolate from them, and what meaning was attached to the different symbolism (Table 7). As shown by the interview quote (Table 7), by discussing the provided representations with the lead designer, we gained information about how scientists and students use such representations to understand the functions of proteins.

To generate the initial ability statements, we analyzed the interview transcript for direct mention of RC and RM abilities. For example, the lead designer first mentioned how to use a “LigPlot+ graph” to “count the hydrogen bonds” between a ligand and a protein (RM segment, Table 7), and

how difference in binding energy can come from these interactions, because “the more hydrogen bonds the more negative the free energy is when binding” (RC segment, Table 7). These interview quotations (Table 7) provided richer detail of the type of reasoning skills involved in the CURE compared with the responses from the open-ended survey (Table 6), but together they provided great insight into some potential CURAs (Criterion 2). For example, in Step 2, there was a brief mention of identifying substrate binding (Table 6), whereas during the interview (Step 3), the lead designer gave a much more detailed explanation of how hydrogen bonds impact the favorability of protein-ligand binding interactions (Table 7). Thus, the interview data acted as a primary data source for the fine-tuning of specific CURA statements, but such statements were also progressively informed through the previous steps of the process in an iterative manner (Criteria 1 and 2).

After the interview, quotes were coded for RC and RM components. A thought experiment was also conducted to construct the possible CURA statements that would apply to a segment (Table 7, outputs). An example of this is the interview quote in Table 7, in which one RM and two RC segments were noted. These three segments could account for at least five separate initial CURA statements. Because the interview acted as the primary data source for the generation of these initial CURA statements, and because it was the first time that there were in-depth reasoning responses, there was a need to check for alignment with the previous two steps to make sure that the CURA statements generated were also supported by the content analysis and the open-ended survey. Thus, in combination, we were seeking consensus between the different data sources about each CURA (Criterion 3).

TABLE 7. Example of conducting Step 3 of PICURA: Interview

Excerpt from the interview with the lead designer ^a	Analysis →	Output: examples of initial CURA statements from the provided quote
Interviewer: So if you could just kind of talk to me about how these types of representations could be used when kind of forming the hypothesizes about what these proteins functions could potentially be? Like how would you, say, look at the 3D plot or the 2D representation LigPlot+ and start to hypothesize just what these proteins are doing, or if you know you have a good substrate. Lead Designer: They do their docking studies, they get a number so let's say -8 kcal/mol for binding of a ligand to a protein and so I would like them to be able to look at a <u>LigPlot ± graph</u> [Figure 2B] like this and say ok I get a -8 for this one ligand and I get a -6 for this other one and I would like them to look at this and <i>count</i> the hydrogen bonds^{RM}, because the more hydrogen bonds the more negative the free energy is when <i>binding^{RC}</i>. but to be able to <i>compare</i> that and then the next thing that they can do is look for a "goodness of fit"[,or binding, between a protein and a substrate]^{RC}.	The interview with the lead designer was transcribed verbatim. Then the interview transcript was coded for instances where the participant discussed reasoning with concepts and/or representations (i.e., RM or RC statements). After this, these segments were used to generate initial ability statements.	RC and RM abilities detected • Count the number of H-bonds in a molecular visualization • Determine using computational software whether and where a ligand may be binding to a protein • Identify ligands that bind to members of a protein family • Demonstrate if a substrate binds to an active site • Estimate the relative ligand-binding stability based on the number of protein ligand interactions

^aUnderlining indicates either an RM or RC (superscript) coded segment with verbs (italics) showing reasoning associated with the noun (bold), which is either a concept or representation.

Reaching Consensus about the Generated Ability Statements (Steps 4 and 5)

In the first part of this section, data from Steps 1–3 of PICURA were highlighted to give an example of the important role of each step in clarifying the nature of the CURAs. In this section, the focus is on the alignment check (Step 4, Criterion 3) and the Likert survey (Step 5, Criterion 4), which together with all the data sources from each step provided a consensus among participants as to which CURAs they considered to be important. Results for two identified CURAs (TR and LR, Table 1) are provided to illustrate this alignment (Tables 8 and 9) and consensus process (Table 10). These two CURA statements were from the top-rated (TR) group, meaning more consensus about their importance, and the lower-rated (LR) group (less consensus) of statements identified using PICURA. The data presented here are aimed at reinforcing how each step led to greater insight into the nature of each CURA and how its relevance to the curriculum was determined.

Step 4: Ability Alignment Check across the PICURA Steps. As described in the preceding sections, the ability statements were initially generated as the output for the interview in Step 3 of PICURA (Figure 1 and Table 7). This is demonstrated by how the interview quotes informed the generation of ability statements (Table 7). During the interview, the participant discussed a detailed example in support of a top-rated ability statement (Table 8) by outlining how computational programs can be used to determine whether a ligand was binding and where it could be binding. This was also indicated in the open-ended survey response, in which it was mentioned that the alignment programs inform hypotheses about function and candidate substrates, whereas the docking simulation produces a ranked list of how likely the candidate substrates bind to a protein

(Table 8). This CURA statement is also supported by the content analysis (Step 1), in which the concepts of motifs, homology, and ligand binding were determined to be prominent themes throughout the protocols, along with the representations from computational software that were used to depict enzymes' active sites (Table 8). The amount of evidence gained in support of this CURA as we progressed from Step 1 to Steps 2 and 3 of PICURA supports why it was rated highly in Step 5 (Criteria 3 and 4; Tables 8 and 10).

Additionally, there was evidence for the lower-rated CURA statement (Table 1) from the interview (Step 3) in which the participant discussed the idea that just because two enzymes have a similar active-site structure does not necessarily mean that they will have the same function, which could be quite different (Table 9). However, protein similarity (or homology) was discussed in a different context in the open-ended survey (Step 2) and content analysis (Step 1; Criterion 3; Table 9). Here, homology is used for hypothesis generation for a protein's function and not to emphasize that, although homologous, related proteins can have distinctly different functions (Table 9). A low WR score was measured for this item, mainly because this CURA received very few marks for being covered in this course only and had two responses each for "not in this course" and "unimportant" (Criterion 4; Table 10).

It is also important to note that the alignment check for this CURA did not generate any CURA statements that were seriously misaligned. However, there were instances of CURA statements having their wording refined as part of Step 4. For example, the LR example shown in Table 9 was originally worded as "Realize that common protein homologies can lead to drastically different behavior" but was modified to "Recognize how proteins that are closely related by evolution can have dramatically different functions" to more deeply clarify this CURA

TABLE 8. Alignment check (Step 4) for a top-rated (TR) CURA statement showing that the ability statement was supported by Steps 1–3 and each step added greater insight into the finalized CURA statement: Determine using computational software whether and where a ligand may be binding to a protein

PICURA components	Supporting outputs
Step 1: Content analysis	Concepts detected: - Protein motifs - Protein homology - Structure–function relationships - Intermolecular forces: protein–ligand interactions Representations detected: - Various computer-generated protein structures (e.g., LigPlot+ and ProMOL representations)
Step 2: Open-ended survey	“We use protein sequence alignment to find similar proteins with known function, we use domain analysis to find proteins with similar domain composition, we use structure alignment to find similar structures with known functions, we use docking to simulate interactions between enzyme and possible substrate to try to choose more likely substrate. Each type of computational evidence does not generate one answer, but rather a list that can be ordered.”
Step 3: Interview	“[A student] could <i>look at</i> the binding of two ligands to a protein that the ligands are almost identical, they’re slightly different and see that you know, let’s say, one ligand has a benzene ring attached to it and the other one doesn’t, and the one with the benzene ring binds with 2 kilocalories per mole better than one without it, and so I would hope that they would look at that and say I need to find out where that benzene ring <i>interacts</i> to cause that much better binding. And that sort of thing, so by having them look at the results they obtained... computationally in one program but then <i>test</i> that in either another program or [run an assay in the lab].”

statement. Member checking, with the lead designer, and comparing the CURA statements for overlap also led to more concise wording for each statement before the Likert survey was administered (Step 5).

Step 5: Evaluating the Importance of the CURA Statements with the Likert Survey. The two-tier Likert survey served three main roles: 1) to confirm the relevance of the generated CURA statements to the curriculum and the ability of PICURA to detect such CURAs, 2) to check for agreement among instructors about the relative importance of the CURAs, and 3) to narrow down the number of CURAs that could be used for student assessment design. Again, two selected CURA examples

(Table 1) are used to demonstrate how Step 5 of PICURA was conducted to reach a final set of CURA statements (Figure 1).

From the Likert survey data in Table 10, a WR score of +17 suggested strong agreement between participants that a top-rated CURA was unique to this lab course (10/10) and was important for conducting this type of research (7/10). In contrast, a WR score of +9 for the lower-rated CURA suggested that there was less agreement among participants about how unique this ability was to this CURE, although it was generally considered to be important (7/10) to this field of research.

This final Likert step acted in two ways to confirm the importance of the selected CURAs. The first way was through the WR score (Table 10), which reflected the extent of agreement

TABLE 9. Alignment check (Step 4) for a low-rated (LR) CURA statement showing that the ability statement was supported by Steps 1–3 and each step added greater insight into the finalized CURA statement: Recognize how proteins that are closely related by evolution can have dramatically different functions

PICURA components	Supporting outputs
Step 1: Content analysis	Concepts detected: - Protein motifs - Protein homology - Structure–function relationships Representations detected: - Amino acid single-letter code alignment - Superimposed protein ribbon structures
Step 2: Open-ended survey	“Some favor sequence homology to suggest function. While that is useful in some cases, I prefer a strong component of structural homology to suggest function.” “I use sequence and structural data to find similar proteins with a known or hypothesized function.” <i>No specific mention of how related proteins can have different functions</i>
Step 3: Interview	“Just because you have a catalytic triad that doesn’t mean that an enzyme will cut proteins, maybe it will cut lipids, maybe it will cut something else. I’m hoping that they’ll have <i>some sort of grasp</i> on the physical nature of proteins , you know, like the molecular weight of proteins, how they behave”

TABLE 10. Example of two CURA statements from the two-tier Likert survey

Ability ^a	Likert question 1			Likert question 2			Weighted-relevance (WR)
	NOT acquired in this lab course	In BOTH this lab and some other course	ONLY in this lab course	Unimportant	Undecided	Important	
TR	0	0	10	1	2	7	+17
LR	2	5	3	2	1	7	+9

Counts represent the number of participants selecting a given response.

^aExamples of top-rated (TR) and lower-rated (LR) ability statements (see Table 1 for full description of the CURA statements).

between each answer to the two-tier Likert survey. This permitted the selection of the most appropriate CURAs for the curriculum (Criterion 4). However, that is not to say that lower-rated CURAs are inappropriate or highly relevant to individual implementations of the CURE, rather, at this time, they are not as relevant across all institutions as the top-rated CURAs. Second, PICURA was used to identify CURAs for this CURE through unbiased data. For this implementation of PICURA, the WR scores could have ranged from -20 to $+20$ (see *How to Apply the Five Steps of PICURA* section for Step 5 and Eq. 1). However, the lowest-rated CURA statement actually received a WR score of $+3.5$, meaning all CURA statements had a degree of agreement about uniqueness and importance. This is consistent with the data that were generated for the lower-rated statement, in which the concepts identified in Step 1 were covered by this CURA. However, excerpts from Step 2 and Step 3 show how the concepts pertaining to homology were used in differing ways (Table 9), so that a lower WR score was not surprising (Table 10). Thus, this process made it possible to find relevant candidate ability statements and to prioritize the CURA statements for future assessment development and to guide CURE evaluation (Criteria 2 and 4).

CONCLUSION

This essay reports on the development and application of a novel five-step process (PICURA) for the rigorous identification of the ALOs, specifically the CURAs, considered by the instructors as key to student learning in our biochemistry CURE. The scientists and educators involved with this project found that, without much effort, PICURA led them to agree on ALOs, so that they are now ready to design relevant assessments to identify the actual VLOs for their CURE. In addition, we have described a novel way to interpret two-tier Likert-scale questions by using weighted relevance to rate the CURAs according to consensus of importance. We have demonstrated that PICURA (Figure 1), guided by the CRM model (Schönborn and Anderson, 2009), generates four data sources—from course materials/content analysis, an open-ended survey, an interview, and a Likert survey—that can be aligned in a consensus process with instructors to effectively identify, fine-tune, and prioritize specific CURA statements for our biochemistry CURE. Thus, we believe that PICURA meets our stated criteria. First, the process is iterative, in that the data generated by each step inform the development and design of the instrument used in the subsequent step (Criterion 1). Second, the data are additive, in that each part of the process yields more clarity on the nature of the CURAs (Criterion 2). Third, the data from each step of the process combine well to achieve consensus and internal alignment about the CURAs (Criterion 3). Finally, the process permits the

prioritization of the CURAs in terms of the level of consensus about the importance and relevance of the CURAs to the course (Criterion 4). In addition, we found that the techniques of PICURA (outlined in Figure 1) are user-friendly and that it was an efficient way to collect and process the data from participant instructors.

It should be emphasized that using PICURA led to the identification of what the participant instructors considered the most important CURAs (ALOs) for their specific context within this specific biochemistry CURE at this point in time and, therefore, it should not be assumed that these are generalizable CURAs that will necessarily be relevant to other institutions performing the same CURE now and later. Indeed, as is the case for all courses, it is likely that the emphasis on certain ALOs will continually change, necessitating a new application of PICURA as and when required. It should also be noted that PICURA is aimed at identifying the ALOs, with the longer-term goal of using such ALOs to inform the design of assessments that will yield student responses that will allow instructors to check whether what they anticipate students will learn (ALOs) equates to verified learning outcomes (VLOs). As stated by Anderson and Rogan (2011), if this crucial alignment does not exist, instructors will need to either modify the ALOs of the course and/or the nature of the assessment to ensure that all ALOs are adequately assessed and that no assessments target the wrong ALOs.

In the context in which PICURA was developed, the CURAs identified are intended to be higher-order biochemistry research abilities, rather than technical abilities like being able to pipette or do SDS-PAGE, which, although important for researchers, can be easily assessed by checking a box that says “yes” or “no” as to whether the student can do it. As is apparent from the data presented above, the ALOs we sought to identify would reflect the scientific reasoning and problem-solving abilities necessary for students to become competent and effective researchers in the area of biochemistry focused on by our CURE. Future use of PICURA will more fully establish to what extent the steps in this process may achieve this goal for other CURES.

The goals of PICURA are also in agreement with the tenets of curriculum theory that advocate for the importance of ongoing course development throughout the life of a course, including periodic review of the ranking of the expected learning statements (Anderson and Rogan, 2011). This is an important component of PICURA in the CURE context, because the CURAs may change as new instructors join, instructors gain more experience, the CURE becomes more widely disseminated, and, most importantly, the research being conducted within a CURE evolves. The nature of CURES means that, as more cohorts conduct research and novel findings are added to the collective

understanding of the research being conducted, the goals of the CURE will surely change. Although the WR scores may fluctuate, as long as the protocols, techniques, and research questions remain unchanged, the CURAs will remain relevant to that CURE.

Future work will involve using PICURA to generate an entire taxonomy of CURAs for our biochemistry CURE. A taxonomy of this kind, besides informing the relevant cognition component for assessments according to the assessment triangle (NRC, 2001), could act as a framework to study additional, and deeper, reasoning about the focus of our biochemistry CURE. In this regard, we agree with Shortlidge and Brownell (2016), who identify various assessments that are available, some of which are widely used but many of which were designed for a specific context and may not meet the unique qualities of another CURE.

In conclusion, PICURA clearly permitted the participant instructors to think more deeply about the ALOs of this CURE and their particular use of the CURE in their own context than would happen if a single instructor had brainstormed the ALOs alone. This approach, in our view, is key to situations in which a particular course is run across multiple institutions, as is the case for our biochemistry CURE.

POTENTIAL APPLICATIONS OF PICURA

Although the data presented in this essay are not intended to imply generalizability to other CUREs, we do recommend that colleagues test PICURA's usefulness in their own contexts. PICURA may be particularly useful for collaborative courses across multiple institutions and for courses in which several instructors need to come up with a set of abilities that are relevant across settings. Additionally, the nature of the process will allow a collaborative team to optimize the ALOs and come closer to what students will really learn in order to better inform course assessment development by instructors. Though this process was developed in an upper-division biochemistry CURE, this process should be transferable across education levels, disciplines, and course formats in ways that could be useful to education researchers, instructors, and administrators. There may be a need to customize some of the steps for these other purposes. However, we suggest that any modifications to the process should primarily be to the inputs, specifically the questions given to instructors, while still applying the simple structure of the process itself (Figure 1). Finally, administrators could use the results of this process to inform decisions concerning implementation of educational policy and reform and to gain greater insight into the role a particular course serves within a department or institutional curriculum. This may lead to the adoption of more CURE curricula and their dissemination.

ACKNOWLEDGMENTS

We thank the Biochemistry Authentic Scientific Inquiry Lab project team for their participation. We also thank Kathleen Jeffery for her useful ideas and feedback on the final manuscript. This work was funded by National Science Foundation (NSF) Division of Undergraduate Education grants #1503798 and #1710051. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

- American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC.
- Anderson, T. R. (2007). Bridging the educational research-teaching practice gap: The power of assessment. *Biochemistry and Molecular Biology Education*, 35(6), 471–477. <https://doi.org/10.1002/bamb.20136>
- Anderson, T. R., & Rogan, J. M. (2011). Bridging the educational research-teaching practice gap: Curriculum development, part 1: Components of the curriculum and influences on the process of curriculum design. *Biochemistry and Molecular Biology Education*, 39(1), 68–76. <https://doi.org/10.1002/bmb.20470>
- Anderson, T. R., Schönborn, K. J., du Plessis, L., Gupthar, A. S., & Hull, T. L. (2013). Multiple representations in biological education. In Treagust, D. F., & Tsui, C.-Y. (Eds.), *Multiple representations in biological education* (Models and modeling in science education, vol. 7) (pp. 19–38). Dordrecht, Netherlands: Springer. <https://doi.org/10.1007/978-94-007-4192-8>
- Auchincloss, L. C., Laursen, S. L., Branchaw, J. L., Eagan, K., Graham, M., Hanauer, D. I., ... Dolan, E. L. (2014). Assessment of course-based undergraduate research experiences: A meeting report. *CBE—Life Sciences Education*, 13(1), 29–40. <https://doi.org/10.1187/cbe.14-01-0004>
- Brownell, S. E., Hekmat-Scafe, D. S., Singla, V., Chandler Seawell, P., Conklin Imam, J. F., Eddy, S. L., ... Cyert, M. S. (2015). A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data. *CBE—Life Sciences Education*, 14(2), ar21. <https://doi.org/10.1187/cbe.14-05-0092>
- Brownell, S. E., & Kloser, M. J. (2015). Toward a conceptual framework for measuring the effectiveness of course-based undergraduate research experiences in undergraduate biology. *Studies in Higher Education*, 40(3), 525–544. <https://doi.org/10.1080/03075079.2015.1004234>
- Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. (2012). Undergraduate biology lab courses: Comparing the impact of traditionally based “cookbook” and authentic research-based courses on student lab experiences. *Journal of College Science Teaching*, 41, 36–45.
- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London: RoutledgeFalmer.
- Corwin, L. A., Graham, M. J., & Dolan, E. L. (2015). Modeling course-based undergraduate research experiences: An agenda for future research and evaluation. *CBE—Life Sciences Education*, 14(1), es1. <https://doi.org/10.1187/cbe.14-10-0167>
- Craig, P. A. (2017). A survey on faculty perspectives on the transition to a biochemistry course-based undergraduate research experience laboratory. *Biochemistry and Molecular Biology Education*, 45(5), 426–436. <https://doi.org/10.1002/bmb.21060>
- Creswell, J. W. (2012). *Analyzing and interpreting quantitative data in educational research: Planning, conducting, evaluating quantitative and qualitative research* (4th ed.). 174–203). Boston, MA: Pearson.
- CUREnet (n.d.). CUREnet: Course-Based Undergraduate Research. Retrieved September 1, 2017, from <https://curennet.cns.utexas.edu>
- Dasgupta, A. P., Anderson, T. R., & Pelaez, N. J. (2016). Development of the Neuron Assessment for measuring biology students' use of experimental design concepts and representations. *CBE—Life Sciences Education*, 15(2), ar10. <https://doi.org/10.1187/cbe.15-03-0077>
- Gray, C., Price, C. W., Lee, C. T., Dewald, A. H., Cline, M. a., McAnany, C. E., ... Mura, C. (2015). Known structure, unknown function: An inquiry-based undergraduate biochemistry laboratory course. *Biochemistry and Molecular Biology Education*, 43(4), 245–262. <https://doi.org/10.1002/bmb.20873>
- Hatfull, G. F. (2015). Innovations in undergraduate science education: Going viral. *Journal of Virology*, 89(16), 8111–8113. <https://doi.org/10.1128/JVI.03003-14>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Jordan, T. C., Burnett, S. H., Carson, S., Clase, K., DeJong, R. J., Dennehy, J. J., ... Hatfull, G. F. (2014). A broadly implementable research course for first-year undergraduate students. *mBio*, 5(1), 1–8. <https://doi.org/10.1128/mBio.01051-13>. Editor
- Kong, Y., Thawani, A., Anderson, T., & Pelaez, N. (2017). A model of the use of evolutionary trees (MUET) to inform K–14 biology education. *American Biology Teacher*, 79(2), 81–90. <https://doi.org/10.1525/abt.2017.79.2.81>

- Kowalski, J. R., Hoops, G. C., & Johnson, R. J. (2016). Implementation of a collaborative series of classroom-based undergraduate research experiences spanning chemical biology, biochemistry, and neurobiology. *CBE—Life Sciences Education*, 15(4), ar55. <https://doi.org/10.1187/cbe.16-02-0089>
- Lopatto, D., & Tobias, S. (2010). *Science in solution: The impact of undergraduate research on student learning*. Washington, DC: Council on Undergraduate Research.
- Makarevitch, I., Frechette, C., & Wiatros, N. (2015). Authentic research experience and "big data" analysis in the classroom: Maize response to abiotic stress. *CBE—Life Sciences Education*, 14(3), ar27. <https://doi.org/10.1187/cbe.15-04-0081>
- McKay, T., Hart, K., Horn, A., Kessler, H., Dodge, G., Bardhi, K., & Craig, P. A. (2015). Annotation of proteins of unknown function: Initial enzyme results. *Journal of Structural and Functional Genomics*, 16(1), 43–54. <https://doi.org/10.1007/s10969-015-9194-5>
- Mordacq, J. C., Drane, D. L., Swarat, S. L., & Lo, S. M. (2017). Development of course-based undergraduate research experiences using a design-based approach. *Journal of College Science Teaching*, 46(4), 64–75.
- National Academies of Sciences, Engineering, and Medicine. (2015). *Integrating discovery-based research into the undergraduate curriculum: Report of a convocation*. Washington, DC: National Academies Press.
- National Academies of Sciences, Engineering, and Medicine. (2017). *Undergraduate research experiences for STEM students: Successes, challenges, and opportunities*. Washington, DC: National Academies Press.
- National Research Council (NRC). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.
- NRC. (2003). *BIO2010: Transforming undergraduate education for the future research biologists*. Washington, DC: National Academies Press.
- NRC. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Olimpo, J. T., Fisher, G. R., & DeChenne-Peters, S. E. (2016). Development and evaluation of the Tigriopus course-based undergraduate research experience: Impacts on students content knowledge, attitudes, and motivation in a majors introductory biology course. *CBE—Life Sciences Education*, 15(4), ar72. <https://doi.org/10.1187/cbe.15-11-0228>
- Pelaez, N., Anderson, T. R., & Postlethwait, S. N. (2014). A vision for change in bioscience education: Building on knowledge from the past. *BioScience*, 65(1), 90–100. <https://doi.org/10.1093/biosci/biu188>
- President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: U.S. Government Office of Science and Technology.
- Schönborn, K. J., & Anderson, T. R. (2009). A model of factors determining students' ability to interpret external representations in biochemistry. *International Journal of Science Education*, 31(2), 193–232. <https://doi.org/10.1080/09500690701670535>
- Schönborn, K. J., & Anderson, T. R. (2010). Bridging the educational research-teaching practice gap: Foundations for assessing and developing biochemistry students' visual literacy. *Biochemistry and Molecular Biology Education*, 38(5), 347–354. <https://doi.org/10.1002/bmb.20436>
- Shapiro, C., Moberg-Parker, J., Toma, S., Ayon, C., Zimmerman, H., Roth-Johnson, E. A., ... Sanders, E. R. (2015). Comparing the impact of course-based and apprentice-based research experiences in a life science laboratory curriculum. *Journal of Microbiology and Biology Education*, 16(2), 186–197. <https://doi.org/10.1128/jmbe.v16i2.1045>
- Shortlidge, E. E., & Brownell, S. E. (2016). How to assess your CURE: A practical guide for instructors of course-based undergraduate research experiences. *Journal of Microbiology & Biology Education*, 17(3), 399–408. <https://doi.org/10.1128/jmbe.v17i3.1103>
- Staub, N. L., Poxleitner, M., Braley, A., Smith-Flores, H., Pribbenow, C. M., Jaworski, L., ... Anders, K. R. (2016). Scaling up: Adapting a phage-hunting course to increase participation of first-year students in research. *CBE—Life Sciences Education*, 15(2), ar13. <https://doi.org/10.1187/cbe.15-10-0211>
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237–246. <https://doi.org/10.1177/1098214005283748>
- Trujillo, C., Anderson, T. R., & Pelaez, N. (2015). A model of how different biology experts explain molecular and cellular mechanisms. *CBE—Life Sciences Education*, 14(2), ar20.
- Weaver, G., Russell, C., & Wink, D. (2008). Inquiry-based and research-based laboratory pedagogies in undergraduate science. *Nature Chemical Biology*, 4(10), 577–580.
- Weaver, G. C., Wink, D., Varma-Nelson, P., Lytle, F., Morris, R., Fornes, W., ... Boone, W. J. (2006). Developing a new model to provide first and second-year undergraduates with chemistry research experience: Early findings of the center for authentic science practice in education (CASPIE). *Chemical Educator*, 11(6), 125–129. <https://doi.org/10.1333/s00897061008a>