CBE—Life Sciences Education
Vol. 9, 266–276, Fall 2010

Article

Evaluating a Modeling Curriculum by Using Heuristics for Productive Disciplinary Engagement

Julia Svoboda and Cynthia Passmore

School of Education, University of California–Davis, Davis, CA 95616

Submitted March 16, 2010; Revised June 22, 2010; Accepted June 23, 2010
Monitoring Editor: John Jungck

The *BIO2010* report provided a compelling argument for the need to create learning experiences for undergraduate biology students that are more authentic to modern science. The report acknowledged the need for research that could help practitioners successfully create and reform biology curricula with this goal in mind. Our objective in this article was to explore how a set of six design heuristics could be used to evaluate the potential of curricula to support productive learning experiences for science students. We drew on data collected during a long-term study of an undergraduate traineeship that introduced students to mathematical modeling in the context of modern biological problems. We present illustrative examples from this curriculum that highlight the ways in which three heuristics—instructor role-modeling, holding students to scientific norms, and providing students with opportunities to practice these norms—consistently supported learning across the curriculum. We present a more detailed comparison of two different curricular modules and explain how differences in student authority, problem structure, and access to resources contributed to differences in productive engagement by students in these modules. We hope that our analysis will help practitioners think in more concrete terms about how to achieve the goals set forth by *BIO2010*.

INTRODUCTION

Seven years ago, the National Research Council (NRC) published the report *BIO2010*: *Transforming Undergraduate Education for Future Research Biologists*, written by a specially commissioned Committee on Undergraduate Education (NRC, 2003). As the title suggests, this report emphasized the need to shift undergraduate education in a direction that would prepare biology students to succeed in a rapidly evolving discipline. In making their recommendations, the committee argued that to keep pace with the changing face of biological research, biology education would need to undergo substantial changes. Students would need to be prepared to deal with the exponential growth of biological data sets and the increasing complexity of biological problems. *BIO2010* emphasized the importance of developing and supporting curricula that are interdisciplinary, with a particular emphasis on integrating biology with more quantitative disciplines. The committee further argued that the design and implementation of reformed curricula should attend to research on how people learn. The report advocated for inquiry-based instructional techniques that actively engage students in conceptually deep problems, rather than asking them to memorize factual content. By combining these perspectives, *BIO2010* provided a compelling argument for the need to create learning experiences for undergraduate biology students that are more authentic to the modern fields of the biological sciences.

The committee acknowledged, however, that designing and implementing such curricula would pose a formidable challenge for practitioners. In support of this challenge, the report featured a number of successful case examples that exemplified many of the committee's recommendations. These examples provide a good starting point for educators and curriculum designers. However, we recognize that the guiding principles outlined in *BIO2010* may not be sufficient to address the practical details of curriculum design and reform. As the report itself acknowledged, research is needed to evaluate the success of such reformed curricula. Even more importantly, research is needed to understand which aspects of such programs are effective and exactly how particular aspects contribute to student learning. Without a better understanding of the link between curriculum design and learning goals, we run the risk of implementing reforms that do not further the intent of the report.

The *BIO2010* report suggests that an increased emphasis on the role of mathematical modeling in biology can support several different learning goals. These include helping stu-

DOI: 10.1187/cbe.10–03–0037
Address correspondence to: Julia Svoboda (jmsvoboda@ucdavis.edu).

dents develop an increased conceptual understanding of biology, improved quantitative skills, an appreciation of the role of mathematics in biology, improved reasoning skills, research skills, laboratory techniques, problem-solving skills, increased interest in pursuing scientific careers, and the ability and desire to collaborate (Steitz, 2003).

We would agree that these are all important learning outcomes for students. We would add that there is, perhaps, a more useful way to think about these outcomes that goes beyond a mere list. To that end, in this research we report on the ability of model-based curricula to support what Engle and Conant (2002) refer to as productive disciplinary engagement. This term encompasses the idea that students are actively and substantively participating in their learning; that their actions and words reflect the rules and norms of the discipline; and that by engaging in such discipline-specific activity, students are making some kind of intellectual progress. Thus, productive as we use it here, essentially means that students are acting, speaking, or thinking in ways that resemble a group of scientists in the midst of making some intellectual progress on a problem. Research suggests that a focus on mathematical modeling has the potential to support just this kind of progress (Gilbert, 2004; Lehrer and Schauble, 2006; Halloun, 2007). In this article, we explore how the productivity of model-based learning environments can be realized.

Utility of Cognitive Apprenticeship

Part of the difficulty inherent in designing curricula with the goal of introducing students to the practices of real scientists is that scientists engage in a wide range of activities, not all of which are particularly cognitively productive. In the context of science, the goal of a cognitive apprenticeship is to structure students' activities around authentic scientific practice (Brown *et al.*, 1989). This makes them different from more typical research apprenticeships in which students conduct research projects under the guidance of a faculty mentor. Although laboratory apprenticeships are generally positive experiences for students and faculty mentors (Seymour *et al.*, 2004), there is limited evidence that students are engaged in deep reasoning or reflection (Kardash, 2000; Bell *et al.*, 2003; Hunter *et al.*, 2007).

One criticism of laboratory apprenticeships is that although students are in some sense embedded in a context of authentic science, this position does not guarantee that they will be challenged intellectually (Barab and Hay, 2001). Students who spend hours processing samples, entering data into databases, or counting organisms are engaged in a kind of scientific activity, but we would argue that it is not a very productive activity. This is why the notion of a cognitive apprenticeship is so powerful in science education; it makes students apprentices to the cognitive practices of the scientific community rather than apprentice to a single mentor.

We use the idea of cognitive apprenticeships here to focus attention on the tasks students are asked to engage with, and the kinds of activity that potentially emerge from these tasks. We ask how tasks can be designed with the idea of productive disciplinary engagement in mind. To explore this question, we propose using a set of heuristics first proposed by Engel and Conant (2002) and later adapted and expanded by Windschitl and colleagues, which we call Heuristics for Productive Disciplinary Engagement (HPDE) after Engle and Conant's original terminology (Box 1).

Box 1. Heuristics for Productive Disciplinary Engagement

1. **Role-Modeling by Mentors**. A good mentor is someone who can model the ways of thinking, talking, and acting of productive scientists. By productive, we mean to emphasize the kinds of activities that comprise the intellectual lives of practicing scientists, such as using evidence to convince a colleague, reasoning through a difficult problem, or asking insightful questions
2. **Opportunities for Students to Try on Roles**. For students to benefit from their interactions with role models, they must have opportunities to practice these roles themselves. Students need repeated opportunities to refine their skills in arguing, explaining, and using evidence. Ultimately, the goal is that students begin to incorporate these roles into their developing identities as scientists
3. **Holding Students Accountable to Disciplinary Norms**. Of course, opportunities to practice scientific ways of thinking and talking will not take hold unless students are given adequate feedback. Students will make the most progress if they are explicitly reminded and held to the criteria that guide authentic scientific practice
4. **Making Tasks Problematic**. Opportunities to engage in productive reasoning or discourse are most likely to emerge if students are faced with problems to think through as opposed to answers they must get to. Encouraging students to approach what is known critically, and explore the unknown without fear of getting it wrong is a necessary component of a productive curriculum
5. **Granting Authority and Ownership**. Students are much more likely to remain productively engaged with a problem if they feel connected to it in some way. Giving students authority means giving them the freedom to make important choices about what is being asked and how to go about answering it. Students will not have opportunities to practice thinking about scientific inquiry if all the important decision points are made for them ahead of time
6. **Providing Relevant Resources**. Students need basic resources such as access to computers, primary literature, and faculty mentors to make progress in science. They also need the time and space to engage with problems in depth. We want to emphasize however, that students can be given too many resources. When students have the answers at their fingertips, they lose the opportunity to have to think though problems for themselves.

Windschitl *et al.* (2008) used these heuristics to design a 10-wk secondary science methods course for preservice teachers. One of the primary instructional goals of this course was to introduce science teachers to the importance of scientific models. The authors found that even in the short span of 10 wk, the participants developed deeper ideas about the nature and function of modeling in scientific research thus showing the promise of these heuristics in the design of learning environments.

We believe that the HPDE heuristics, as we've defined them here, align with the motivations of the *BIO2010* report that concern productive, engaging experiences that are relevant to the discipline. Our main objective in this article is to explore what these heuristics might look like in practice and how they can be used to evaluate the potential for model-based learning environments to provide productive learning experiences for science students. We want to caution that these heuristics are not meant to be used as a checklist. Instead, we envision them as features of a complex and dynamic system. Each of these elements interacts with the others, and it is the synergistic effect of these elements that we believe can allow productive disciplinary learning to emerge. Our aim is to unpack the overt and subtle ways that these interacting elements can influence the learning environment for students. We do so by linking each of these heuristics to concrete examples from an undergraduate traineeship called Collaborative Learning at the Interface of Mathematics and Biology (CLIMB). We hope that our analysis of CLIMB will help practitioners think in more concrete terms about how to achieve the goals set forth by *BIO2010*.

METHODS

CLIMB Program

The CLIMB program is a yearlong, National Science Foundation (NSF)-sponsored traineeship for upper-division undergraduates majoring in mathematics and the biological sciences. Students accepted into the program form a cohort and work collaboratively throughout the year to solve biological modeling problems by using quantitative methods. During the fall and winter training periods, a series of short topic-specific modules are used to introduce students to a variety of mathematical models. Students complete a problem set for each module and collaboratively write up their findings in the form of a research paper (typically 15–20 pages). In the spring and summer, the CLIMB cohort transitions into working on an independent research project. Unlike many programs in which students work on a project closely tied to the research of their major professor, the CLIMB students are asked to develop a research project independent of the influence of the major mentor's research agenda. The CLIMB cohort spends months collectively developing a research plan. Part of this planning involves looking for a content-area expert to serve as a secondary mentor to the project.

In this research, we focused on the activities that took place in the fall quarter of this yearlong program. We chose to focus on the fall for several reasons. First, the fall quarter was the most like a traditional undergraduate course. Thus, the findings we present here can potentially be applied to curricula that fit within the constraints of the traditional university structure. Yet, because the fall was designed with the goal of preparing students to engage in independent research, the curriculum aligned strongly with the recommendations of *BIO2010*. In many ways, the fall quarter can be conceived of as a cognitive apprenticeship that was meant to prepare students for independent research by introducing them to a diverse set of interesting problems that interface mathematics and biology along with associated methodologies and patterns of reasoning. Finally, because five different guest instructors led the five fall modules, we had the opportunity to study the variation in task structure and the impact these structures had on opportunities for student learning.

Participants

Students were selected for the CLIMB program based on both their individual achievements (minimum GPA of 3.0) and potential to contribute as productive members of a group. An effort was made to include students from a range of different backgrounds with a range of academic interests within mathematics and biology. All participants were in either their junior or senior year and participated in CLIMB for 1 yr only. The cohort we followed in this study had seven students: five biology majors, one mathematics major, and one physics major. All of these students expressed at least some interest in pursuing an advanced degree in either mathematics or biology. There were five women and two men, and ethnicities represented include African American, Caucasian, Chinese, Korean, and South Asian. One of the students was recruited from an enrichment program for underrepresented undergraduates. Overall, we would characterize the CLIMB students as high-achieving and highly motivated students.

Data Collection and Analysis

One of us (J.S.) spent a full year deeply immersed in all aspects of the CLIMB program. She attended all class meetings and observed the majority of students' problem-solving sessions. She collected a wide range of data, including detailed field notes; video of all problem-solving sessions; interviews with CLIMB guest instructors before each module; interviews with students at three times during the year (October 2008, March 2009, and September 2009); and written documents including drafts and final versions of group problem-set write ups, and individual statements of contribution written by each student for each module.

We conducted our analysis of the CLIMB curriculum as follows. We first constructed detailed summaries of each module including the nature and structure of tasks, the verbal and written instructions given by professors, and the nature of student activity. We then analyzed each module according to the HPDE framework (see Box 1) so that we had concrete examples of how each module corresponded to each of the six heuristics. We then compared the five modules, looking for patterns of similarity and difference. We divided the heuristics into those that were consistent across modules and those that varied, and then we went back to the individual examples to identify how specific aspects of the curriculum were linked to specific heuristics.

FINDINGS

We have organized our findings into three sections. In the first section, we describe the heuristics that were shared among all five modules. In the second section, we describe the heuristics that varied among the five modules and provide an overview of how these differences were expressed in each module. In the third section, we present a more in-depth comparison of two modules and use concrete examples from each to provide a more detailed description of how differences in HPDE heuristics contributed to different learning

experiences for CLIMB students. We want to make it clear from the outset, however, that the goal of this article was not to make direct comparisons among the five different modules. This would have been difficult for many reasons, two of which we highlight here. First, the modules are sequential, which meant that students were arguably more experienced by the end of Module 5 than at the beginning of Module 1. Second, the subject matter was different in each module, which given the range of background and interests of the students, could have influenced both their motivation and success.

It would be more appropriate to view the examples we present here as chosen to illustrate the potential utility of considering the HPDE heuristics when constructing or evaluating science curricula. The examples should not be considered representative of all of the experiences students had in a particular module, because there were instances of more and less productive activity in all modules. We intentionally chose examples that could demonstrate, in a concrete way, the potential consequences variation in curricular structure could have for students learning and engagement.

Heuristics Consistent across All Five Modules

1. Role Modeling by Faculty Mentors. Throughout the fall, and across all five modules, students were exposed to a range of faculty role models with different interests and philosophies. At the start of each module, the guest instructor began with a brief overview of some interesting problems in his or her field and the ways in which models could be used to address these problems. The guest instructors often presented detailed examples from their own research that illustrated how mathematical models had been successful in helping them make progress. They also spoke frequently about the process of modeling itself, sharing strategies for model construction and criteria for model evaluation. The introductory sessions were consistently rich in examples of both specific uses of models set in the biological context and more general strategies for thinking about modeling. The following excerpt in an example of the kind of statements that were common as guest instructors described their own work (field notes 10/30/2008).

> Professor 3: If you are doing a mathematical model, what is it that you have to include? The point is that you want to explain something in nature. And there are all kinds of complicated stuff. What do you put in and what do you ignore? I decided I was going to ignore age-structure in my model. I did it because it was simple, and I only discovered last year that I hadn't embarrassed the family.

Although the instructors' general motivations were quite similar, their specific motivations and approaches varied across modules. Some experts favored empirical approaches, whereas others were more theoretically driven. Some began by presenting interesting models, whereas others began by presenting interesting data patterns. Some instructors highlighted the importance of biological intuition, whereas others extolled the importance of mathematical clarity or computational power. In the span of 10 wk, the CLIMB students were exposed to a variety of ways to approach modeling in biology. Perhaps more significantly, in all joint faculty–student sessions, more than one expert was present in the room; at least one, but up to three CLIMB mentors, a graduate student teaching assistant (TA), and the guest instructor were often present simultaneously. As guest faculty discussed their research, opportunities arose for the other experts in the room to ask questions or make comments. Disagreements were common, and the resulting justifications and explanations modeled for students, for the most part unintentionally, how to construct logical arguments and how to use evidence to support claims. Many of these disagreements achieved resolution, but the experts also left room in their discussions for open disagreements and differences of opinion. Below we provide one short example of this from Module 1. In this excerpt, the Module 1 instructor, Professor 1, had just presented a biomechanical model that he had used to understand suction feeding in fish. Two of the CLIMB mentors, Professor R and Professor S probe for more explanation (field notes 10/2/2008):

> Professor 1: In this model, it essentially works out that the area of the muscle is the force.
>
> Professor R: What I want to know is, why does it work?
>
> Professor 1: Well, imagine mouth cavity expanding with equal force. A lot of these forces will cancel each other out, so it works out mathematically that it is force.
>
> Professor S: Ah, but that only works for a cylinder.
>
> Professor 1: Yes, we made a bunch of simplifications. You have to assume that levers accurately reflect the linkage system. So, one would be justified in being a little skeptical.

In the example above, Professor R pressed Professor 1 for a further explanation of how this model captures the biology. Professor 1 elaborated on his reasoning, which prompted Professor S to point out that the simplifications are not generally valid, but only apply when the fish mouth is assumed to be a cylinder. Professor 1 conceded that this skepticism was warranted, and revisited these concerns later in the conversation when he provided empirical support for the simple cylinder model, as well as evidence that using a more complex shape did not yield significantly different results.

We argue that opportunities for students to witness exchanges like these functioned in two ways. First, the instructors were modeling the kinds of discourse and reasoning that they expected the students to emulate; they were modeling how to argue and disagree in a scientifically appropriate way. Second, disagreements among professors undermined the idea that there was a single infallible authority figure. Instead, we believe it impressed upon students the importance of multiple perspectives for making progress in science.

2. Opportunities for Students to Take on Various Roles. The collaborative nature of the CLIMB program provided students with repeated opportunities throughout the year to take on various roles and stances. In the fall quarter, most of these opportunities happened in the context of group problem-solving sessions during which faculty were absent. During these sessions students could practice proposing ideas, explaining and arguing with one another. Some division of labor did occur initially: math students tended to gravitate toward calculations, and biology students felt more comfort-

able offering interpretations. In response, the CLIMB mentors suggested that students make a conscious effort to rotate roles for each module. In addition, they asked each student to write an individual reflection after each module. These essays helped students consider their roles and made them more likely to try on new roles when they found themselves in the same role for >1 wk. This structure helped ensure that biology students did not always take on roles as background experts and more mathematically inclined students did not always specialize in methods and results. Within the first few weeks, students began encouraging their peers to step out of their comfort zones. Math majors invited biology majors to the chalkboard, and biology majors challenged math majors to make connections to the biology. The frequency of cohort meetings allowed the group to develop the trust needed to try on new roles. For example, in Module 4, Romy[1], a mathematics major, and Nina, a biology major, paired up spontaneously and spent time together at the blackboard. Nina asked questions about the objective of the module and Romy explained why they needed to solve for a particular variable. Romy then asked Nina if she wanted to try to get the solution, and the two of them worked through the math together (field notes, 11/17/2008).

Despite encouragement, students still tended to gravitate toward familiar roles. Certainly, mathematics majors did more of the mathematics than the biology students and vice versa. The presence of peer experts may have in some ways allowed students to avoid taking on unfamiliar roles. Furthermore students realized that when deadlines were looming it was often more efficient to distribute tasks by expertise. Nevertheless, the need to collaboratively write up their work gave the group motivation to make sure everyone understood the problem enough to contribute. Rather than converting mathematics majors to biology major or vice versa, the collaboration that took place in CLIMB allowed students to extend their comfort zones enough that they could work productively with one another. One student described her experience as a collaborator as follows (final interview, 9/18/2009).

> Eve: I think … some of the things I've increased is my skills in is collaborating with other people, because I mean you don't really get the chance to do that as an undergrad - to work with a group of people for that long of a time. Sometimes you do small group projects, but I think that's really important … especially for this merging of biology and mathematics. Because I know that there are some people out there that just don't have the ability to do like really strong mathematical modeling, and then there's mathematicians who … might not care too much about the biology; they are more interested in you know, doing math. And so I think it's really important for those people to work together to get something done.

3. Holding Students to Scientific Norms. During the fall CLIMB course, students were introduced to the standards and norms of scientific practice in several ways. Three CLIMB mentors, Professor R, Professor S, and Dr. Marcia, provided students with feedback both on their written work and informally during classroom discussions. Mentors gave students detailed feedback on how to write scientific papers; they instructed students to contextualize the problem in the existing literature, articulate and justify methods, report relevant findings, and discuss the theoretical and practical implications of the results. Comments such as, "you need to justify your assumptions" or "explain how this is relevant" or "why do you think this?" were found in the margins of each paper.

In addition to written feedback, mentors sometimes used classroom conversations as opportunities to push students to think scientifically and speak clearly. Both CLIMB mentors and guest instructors sometimes asked students to consider possible explanations for problems and brainstorm about the kinds of evidence they would need to support them. These kinds of conversations did not happen very often however—perhaps once or twice per 90-min class. Because guest instructors were often concerned with covering the background material for the assignments, opportunities for discussion were fairly limited during the fall. Students also seemed much more intimidated in the presence of faculty. They spoke much less often in class than they did during problem-solving sessions. During these sessions it was often the graduate student TA who pushed students to think critically about the claims they were making and to communicate these ideas clearly both verbally and in their written reports. As the following example demonstrates, the TA was often instrumental in keeping students focused on the question they were trying to answer as opposed to treating the task like a mathematics problem (field notes, 11/7/2008).

> TA: What are you looking for? This system is nonlinear. Have you dealt with complex nonlinear systems of equations?
>
> Kevin: (shaking head no) How can I do it?
>
> TA: Why do you need an analytical solution?
>
> Kevin: Didn't you say to do that?
>
> TA: No …what is the question?

Because the CLIMB mentors and the TA were consistent figures in the program, the students had a relatively constant source of feedback across all five modules. Moreover, this feedback was cumulative, which allowed students to feel that they were making steady progress. Crucially though, most of this feedback occurred in the context of problem-solving sessions and not in the context of lectures. Without the opportunity for students to try out scientific practice for themselves, they would have missed out on opportunities to develop an appreciation for scientific norms.

Summarizing across all five modules, we saw consistent opportunities for students to observe experts, practice taking on scientific roles, and get feedback on their performance relative to a set of community norms. What is common about these three heuristics is that they were each embedded in the larger structure of the CLIMB program. This higher-level structure ensured that they were present in each of the five modules. The links between each heuristic and the particular structures of the CLIMB program are summarized in Table 1.

Differences in HPDE Heuristics across Modules

When we looked more closely at each of the five modules, we found that major differences in the learning environments stemmed primarily from differences in the nature and structure of instructor-assigned tasks. Tasks varied in the

[1] All names are pseudonyms.

Table 1. HPDE heuristics that were consistent across CLIMB modules

HPDE heuristics	Features of CLIMB program structure
Role modeling by faculty mentors	Multiple CLIMB mentors, five different guest instructors, graduate student TA
Opportunities for students to take on various roles	Cohort structure collaborative problem sets and write-ups, individual reflections, time away from faculty
Holding students to scientific norms	Written feedback on papers, classroom conversations, feedback from graduate student TA

degree to which they were problematic, the degree to which students were in positions of authority or had ownership over the tasks, and the amount and type of resources to which students had access.

4. Few Tasks Are Problematic. Across the five modules students were engaged in problem solving of some kind, but the degree to which students had to make decisions about how to solve the problem was highly variable. In the first four modules, students were, for the most part, led step by step through the articulation of the problem, the choice of methods, and the expected results. Often students were even given substantial hints about the desired interpretation of the results. The one exception to this pattern is Module 5, in which students were asked to engage with a problem with no solution. The instructor provided guidance, but the path to a solution was left open for the students. This meant that students had to struggle to make decisions about how to move toward a solution without knowing ahead of time what that solution should look like. In contrast, tasks were much less challenging or engaging for students when they were able to resolve a problem by mimicking the approach used by the instructor. The following discussion that occurred toward the end of Module 3 illustrates that the students were aware of the difference (field notes 11/9/2008):

> Rose[1]: The thing is, he [the professor] has already done this; our discussion is already published.
>
> Eve: Well, we are re-establishing it.
>
> Romy: I feel like it's more of a problem-set than anything else.
>
> Eve: Maybe we can discuss the applications of the findings.
>
> Romy and Rose: (together) But he already did that!

In this excerpt, the students expressed their confusion and dissatisfaction with the task. It was clear to them that the professor owned the task; he had already published a paper in which he solved the problem he assigned to the students. Eve tried to find a way for them to contribute, but Romy and Rose were clearly frustrated with having to redo someone else's work.

5. Limited Opportunities for Authority or Ownership. The overall goal of the CLIMB program was to move the students toward independent inquiry. However, in the fall, the guest instructors maintained the traditional role as the authority figures in the classroom. They did so implicitly, perhaps unintentionally, by assigning tasks in which they were the experts. As we mentioned above, in the first four modules, the tasks largely involved redoing or resolving a problem that had already been solved, in most cases by the instructor himself. In three of the modules, the instructor had already published the results of the problems. Not only were such "problems" not problematic at the level of content, they also kept students from developing their independence as scientists, because ultimately, the instructor was the authority on the subject. We argue that this inherent power differential kept students focused on getting an answer that matched the one that already existed as opposed to following their own interests. The instructors did not explicitly undermine the students' authority; they did not shut students down or intentionally intimidate students. However, the tasks they asked students to complete implicitly positioned the instructors in the seat of authority.

In the CLIMB setting, authority and ownership were closely linked to how open-ended, or problematic, the tasks were. However, we can envision other ways in which instructors can undermine student authority that are independent of task structure, by creating a classroom culture in which students contributions are not valued, for example. It is for this reason that we keep these two heuristics separate despite the large degree of overlap we saw in the CLIMB curriculum.

6. Access to Too Many Resources. Both the challenge level of the tasks and the degree of authority were linked in some way to resources. The CLIMB students had access to many different resources including expert faculty and primary literature. As we already mentioned, during the first four modules it was common for the instructor to provide significant guidance, sometimes going through the calculations first in class or discussing the implications that students were supposed to be discovering. In contrast, the instructor for the fifth module worked through examples in class that were analogous to the problem he presented to students, but that did not lead them directly to the answer.

In the first four modules, students were often assigned reading from the primary literature that addressed the very problem that they were asked to solve. In these cases, the methods, findings, and interpretations could be read out of the papers. This was not the case in the fifth module for which no published result existed. Our impressions from faculty interviews suggested to us that it was not the intention of the instructors to have students rely too heavily on published work. However, they wanted students to have access to relevant background information. In the face of this tension, most instructors chose to give students access to the papers. We argue that this tendency contributed both to students' sense that they were meant to emulate the expert and also to their confusion and frustration over not adding anything "new" to these problems.

In summary, we found that in four of the five modules, tasks were relatively unproblematic, authority lay primarily with the instructors, and students had access to resources that, in many cases, solved the problem for them. However, the structure of Module 5 was different in several important ways. To further explore this difference, we compare Module 1, which typified the first four modules, with Module 5. We hope that the examples we present from each of these modules serve to illus-

trate how different degrees of adherence to the HPDE heuristics might look in practice.

Comparing Module 1 and Module 5

1. Comparing the Problems. *Module 1: The Instructor Specifies the Problem.* The instructor outlined the steps for the four-bar linkage task on the first day of class (see Box 2). Along with these instructions, the professor provided the students with a substantial amount of guidance as to how to proceed. They were told which angles they needed to use to calculate the transmission coefficient (Box 2, step 2). They also were provided with a general strategy for constructing the morphospace plot. The instructor told them, "you could, for example, make a morphospace plot using the L_O (output lever) and the L_I (input lever); but you will need to figure out how to make it 3-D." The final step, in which the instructor asked students to make sense of the patterns they generated, was a series of leading questions. The main pattern they were meant to discover was literally highlighted in the task instructions. The only open aspect of this task was proposing possible implications for fish evolution. However, as we will see later, even this task became unproblematic when we consider the resources that students had available to them.

Box 2. Four-Bar Linkage Task

The following[1] steps were articulated verbally at the end of the first day of class (field notes 9/30/2008).

Step 1. Build a physical model of the four-bar linkage model. Use stiff cardboard and fasteners.

Step 2. Figure out how to calculate the maxillary transmission coefficient (MaxKT) from the four-bar linkage model. You should try it independently, but you may need to get together.

Step 3. Start collaborating on coding it up. You will use the inputs the lengths I give you and calculate the MaxKT. Take the data I have g[1]iven you and calculate the MaxKTs for all these species.

Step 4. Contemplate the results. Figure out how to visualize the shapes of the four-bar. Make a graph of the morphospace and map onto that space the MaxKTs of the 30 linkage shapes I have given you. Somehow express the world of four-bar linkage in three-dimensional (3-D) space. And for each shape put a point and then color-code it.

The final step of the project was outlined in a handout:

Step 5. The write-up. Consider and answer these questions: What patterns do you see in the data concerning the relationsh[1]ip between the four-bar shape and MaxKT? **Does every four-bar shape have a unique MaxKT?** [emphasis original] Given your answer to the last question, what do you think are some implications of this for the evolution of jaw mechanics in fishes? Try to come up with three ways in which this relationship between the four-bar shape and MaxKT could impact evolution of fish jaws.

Module 5: Setting Up the Struggle. In contrast, the instructor for Module 5 began by telling students that they should be prepared to struggle through a difficult problem:

> The goal of this module is to work through some of the earliest steps of a research project. As a student, I found that the earliest steps are the most confusing. And typically, when we are taught science, the confusions and mistakes are left out. I want to give you an inside look at the struggle part of research. We are going to struggle to generate a mathematical representation of a biological problem. I wanted the example to be simple enough that we won't get bogged down in the math. But we are going to have to think about how to represent biological assumptions. And optimistically you will be able to see how the theory can help us think about the empirical problems (field notes 11/24/2008).

At first the task was simply to "apply the logic of the resources model that we went over in class to the new problem of top-down versus bottom-up control." The resources model explored resource allocation strategies that would maximize fitness for organisms that had to decide how to allocate energy to two necessary resources. The top-down, bottom-up problem was also an allocation problem except that it asked, how should an organism allocate its energy given the need to both obtain resources and avoid predation? The problems are related, but the latter problem had not been explored by scientists as an allocation problem. The instructor was careful to tell students that there "is no answer to this problem." He allowed them to struggle with the problem for a few days in between class meetings, and then he introduced some more structure to help them get a hold on the problem (see Box 3). When he outlined the problem, the instructor specified some of the parameters of the model that he wanted them to explore. However, he left several important choice points open to the students, such as how to define the relationship between defense and survival and the details of the model structure itself.

Box 3. Top-Down, Bottom-Up Task

The following written instructions were handed to students after the second day of class (12/5/2008):

1. What might be some reasonable choices for a function relating investment in defense to the probability of surviving an encounter with a predator? Describe three possibilities, and explain the underlying biological rationale for each.
2. Build a model that relates an individual's fitness to its investment in a) securing food resources from the environment and b) defense against predators. Explain what your model assumes about the relative timing of predation risk and reproduction in the organism's life cycle.
    a. Explore the model to obtain the optimal investment in defens[1]e in a deterministic world, where resource availability and predator density do not change.
    b. Incorporate environmental stochasticity (variation in either resource availability, predator density, or both). Does the optimal allocation change?

2. A Shift in Authority. *Module 1: Getting It Right:* The students began the module by finding the equation for MaxKT and using given data to calculate values for different species. In the excerpt that follows, the students had attempted to plot the values they calculated (see Box 2, step 3) when they were visited by Dr. Marcia, a CLIMB mentor, who found a problem with their results (field notes 10/6/2008):

> Dr. Marcia: There is something wrong—what you get should not be that …. That is not correct.
>
> Henry: What is the graph?
>
> Lillian: The graph of MaxKT and how many of each value we have.
>
> Dr. Marcia: The graph is appropriate for your calculations, but the calculations are wrong.
>
> Henry: Everyone's?
>
> Rose: Probably everyone, because we all calculated it the same way.
>
> Dr. Marcia: The data should all be lumped.
>
> Kevin: I want to look at the way Rose did things. I want to see that we all did the same thing. I am assuming that our method is right.
>
> Eve: Did we go over exactly how to calculate E-prime?
>
> Sean: The way I did it, I used different geometry than everyone else, and I got the same answer even though we calculated it differently. So we must be calculating the wrong thing.
>
> Rose: Basically all of us are wrong.
>
> Dr. Marcia: Dr. Robert (the professor) showed me that the distribution should really be split into three, as opposed to the one histogram.
>
> Eve: Are we sure we are supposed to calculating psi (the angle)?
>
> Henry: I don't see how our calculations could all be off.
>
> Kevin: I think it must have been put into the computer wrong.

When the students were unable to "get it right," they had no other option than to appeal to a higher authority. The students did not understand how the model was constructed; therefore, aside from looking for technical errors, they had no ability to reason through the problem. Eventually, the students brought their problem to the instructor, who chose to give them the correct data to interpret because he did not want calculations to be the sole focus of the module.

Module 5: The Students Take On the Problem: Although the students had an analogous problem to structure their thinking, they were being asked to think about a novel problem. In the excerpt below, students struggled to make sense of the task among themselves. They began by trying to translate the concepts of "resources" and "pumps" (the energy allocated to each resource) into the trophic level problem (video transcript 12/3/2008):

> Lillian: I think we need to figure out what the pumps are.
>
> Eve: Could one be defense?
>
> Lillian: Or foraging skills?
>
> Sean: What about time spent eating versus hiding?
>
> Eve: Energy spent on search versus hiding? This is weird for producers though; they have phenotypic plasticity.
>
> Sean: They [producers] have secondary compounds for defense.
>
> Lillian: The paper implies that a resource for producers is space.
>
> Sean: Resources that you put into growing tall will make you easier to eat.
>
> Eve: What about interspecific and intraspecific competition?
>
> Romy: What happens at the top?
>
> Sean: By definition top-down does not exist [at the top].
>
> Nina: Does fitness depend on both [top-down and bottom-up]? I say it does. If you can't eat, then you can't produce babies; if you don't survive you don't have a chance.
>
> Eve: We are assuming that the only way you die is from a predator?
>
> Lillian: It is not which way you die that matters—it is which R (resource) runs out first. The whole point is reproduction.
>
> Eve: Should we write down some equations?

The important thing to notice about this excerpt is that it featured multiple students engaged in trying to make sense of the problem. They were posing ideas, considering assumptions, and explaining ideas to one another. In contrast to Module 1, the students were not concerned with what was "right." Instead of trying to figure out what the instructor wanted from them, they were much more focused on their own ideas. This made for a much more lively and convoluted discussion as students attempted to pursue multiple ideas at once. In reflecting on this module in a March interview, several students commented specifically on their level of engagement and ownership. For example, in response to a prompt asking her to remember what happened in Module 5, Eve responded, "okay, so for Module [5], this one I actually really enjoyed because we actually really tried to build our own model. We did a lot of thinking on our own instead of working with an already established model." This comment, and others like it, demonstrates that the students were both aware of and appreciative of the chance to take more ownership of a problem.

3. Optimal Resources. *Module 1: The Papers Hold the Answers:* As a part their assignment for Module 1, the students were asked to read several papers on the four-bar linkage. The students were also expected to cite other relevant references in the final paper. One of them, entitled "Evolutionary Consequences of Many-to-One Mapping of Jaw Morphology to Mechanics in Labrid Fishes," presented output from the four-bar linkage model and proposed a series of implications for the evolution of jaw morphology. This was exactly what students were asked to do as the final step in their

problem set (see Box 2, step 5). Although it is not clear whether all of the students read this paper, Eve explicitly mentioned reading the paper in her individual reflection. It seems likely then, that in the following discussion of the model results, her reference to "the paper" was a specific reference to this paper (field notes 10/12/2008).

> Eve: They [the values of MaxKT] are all clustered in the same area. Multiple combinations of morphology have the same functionality.
>
> Rose: Isn't there some tradeoff?
>
> Henry: Longer length would yield slower velocities?
>
> Eve: The paper said, the more components there are to something (inaudible). Let's say that there is an optimum MaxKT. These guys clustered here have that (pointing to the cluster MaxKT values).
>
> Henry: Even though they have different morphology?
>
> Eve: Then there are outliers which have either higher or low…
>
> Henry: They could just be better for some other reason.
>
> Eve: It talked about in the paper how it would be interesting to know whether or not they are the same species.

It was the intention of the instructor to get students to think deeply about how a simple biomechanical model could be used to "get insights into what is important in the evolution of this system" (field notes 9/30/2008). However, it seems that the opportunity for the students to think about these implications was truncated by the existence of interpretations that could be found in the literature.

Module 5: The Instructor Keeps the Struggle Going: In Module 5, students ran into several different problems and became confused at several points. However, they had nowhere to go to get the answer. In fact, the instructor made an effort to keep students in the struggle and to signal to students that struggling is a part of the process (field notes 12/4/2008):

> Eve: So, we were confused about how defense could be a resource.
>
> Professor 5: The "a" (resource 1) is clear, but the "b" (resource 2) is … keep going? You may want to leave it unclear for now.
>
> Eve: seems like b is dependent on a. You need time to get energy and nurture it.
>
> Sean: b could be a function that is how many offspring you get over time?
>
> Maybe you need to live until six (inaudible). That way b is a function though, and we can't use the linear model.
>
> Professor 5: This is great! This is the same path I was struggling along.

Instead of telling the students what to do, the instructor highlighted the struggle and encouraged the students to keep struggling. Although the instructor had not yet published a paper on this problem, he had clearly thought about it more than the students. However, he did not share his model with them. Instead, he encouraged them to explore multiple possibilities. For example, the students had already decided that they needed to relate survival to defense. Instead of asking them to choose a single function, the instructor asked students to explore three different options (see Box 3). In this way, he helped structure their exploration of the problem without constraining them to a particular path.

DISCUSSION

Our findings highlight the complexity with which program elements can interact to influence the kinds of learning opportunities that are available to students. Using the CLIMB program as an example, we were able to tease apart two levels of curricular structure: the larger structure at the program level, and the structure of specific tasks. We address each of these levels below.

CLIMB Program as a Collaborative Apprenticeship

The overall structure of CLIMB embodied three important heuristics for productive disciplinary engagement: role modeling by mentors, opportunities for students to take on roles, and holding students to disciplinary norms. These heuristics are key to describing CLIMB as a collaborative apprenticeship, a term we use to emphasize the importance of collaboration, a feature that is often missing from more traditional research apprenticeships. We see this collaboration occurring in three important ways.

First, CLIMB was a collaborative effort among multiple instructors. The presence of multiple experts created the opportunity for students to interact with a variety of different mentors who could serve as role models. Because each expert had a different specialization, students got a more complete picture of scientific practice—one that began to transcend the specific idiosyncrasies of an individual mentor. Even more importantly, by observing how experts interact with one another, students had the opportunity to see scientific discourse modeled for them as well. The opportunity to observe interactions is an important way in which novices can begin to gain access to a community of experts (Rogoff, 2003). In this way, CLIMB provided a richness that is not typically present in the more traditional single student–single mentor model of apprenticeship.

Second, CLIMB featured extensive collaboration among peers. The collaborative nature of the fall problem sets gave students multiple opportunities to take on different roles. This was especially important considering the interdisciplinary nature of CLIMB because it allowed biology students to take on math problems and math students to engage with the biological interpretations. When students develop a level of comfort with peers, we believe it helps them take risks they would not take in the presence of a more advanced expert. Students in apprenticeships see peers as important resources and sources of moral support (e.g., Grindstaff and Richmond, 2008). In CLIMB, this was partially evident from the different level of participation we observed in class versus problem-solving sessions. Students also made many positive comments about peer collaboration in individual interviews.

Third, in the fall, we began to see collaboration between faculty and students in the sense that the instructors and mentors treated students like novice colleagues. Students were slowly being inducted into the scientific community. Part of this process was evident when students were held to the standards of the discipline. The experts rarely made these standards explicit, but instead communicated implicit messages about standards of scientific practice in their feedback to students. This final element is an important part of the iterative process of "bootstrapping" (Gee, 2002), which is thought to be an important mechanism through which students gradually enter the scientific community. In CLIMB, students first had the opportunity to observe expert discourse, they then had the opportunity to try it on for themselves, and finally they received feedback directly from those experts so that they could refine their behavior. Although we do not report on the remainder of the CLIMB year in this article, in future work we plan to describe how this collaboration between faculty and students deepened over the course of the spring and summer, such that by the end of the CLIMB program, students reported feeling like they were legitimate colleagues with their mentors.

As a whole, we argue that, at the program level, CLIMB offered students the chance to practice behaving like scientists. We believe that this can be explained by the program's adherence to three of the HPDE heuristics: providing students with multiple expert role models, multiple opportunities to try on these expert roles, and continual feedback on their performance relative to a set of implied disciplinary norms.

Implications of Variation at the Task Level

Despite this productive backdrop, we argue that variation at the task level seemed to dramatically influence opportunities for productive disciplinary engagement. We believe that much of this variation can be explained by the extent to which three additional heuristics were manifested at the task level: the extent to which tasks were problematic, the degree of ownership/authority, and access to resources. Each of these three heuristics relate to the degree to which tasks challenged students to think independently.

When students were asked to step through procedures they lost opportunities to reason deeply about the material. That is not to say that students were not challenged by the tasks they faced in the fall; many of these tasks were time consuming and technically challenging. Nor do we mean to say that developing quantitative skills and reading primary literature are not important tasks for undergraduates. However, we argue that there are ways to have students practice these skills while leaving them some room to grapple with some more ill-defined problems. This does not mean that students should be left on their own to solve novel problems. Rather, we would advocate for a balance of structure and opportunity to explore. In Module 5, the task was quite structured; the instructor chose the problem, the general structure of the model, and suggested possible analytical paths. However, enough choices were left open to students that they had the chance to think. We believe that the context of modeling offers many ways to achieve this balance. Students could be asked to revise an existing model in light of new data, compare two proposed models, or even construct a simple model for a new problem.

When students have the chance to do things like construct, reconstruct, critique, or revise models, they become authors. That is, they are able to take at least partial ownership over the choices that they make rather than simply following a series of steps that have been predetermined for them. Giving students ownership can change the way they relate to the task. When students feel that they are being asked to contribute intellectually to an authentic problem, they are more likely to remain motivated and engaged (Lee and Songer, 2003). A sense of investment in a problem also can influence the depth and quality of students' discourse (Jimenez-Aleixandre *et al.*, 2000). As we saw in CLIMB, students were very much aware of the authenticity of the problems they were being asked to solve, and we noticed that when students perceived that they were redoing someone else's work, their level of engagement suffered.

Challenging students to think for themselves may mean withholding some of the available resources. This does not necessarily mean that students should not be given resources. Instead, it might mean considering whether and how the available resources support the learning goals (Hay and Barab, 2001). In CLIMB, we saw that in some contexts engaging students in the primary literature was an essential resource that helped the students comprehend how the problems they were working on fit into the bigger picture. However, in other contexts access to published articles undermined the learning goals by cheating students out of the opportunity to reason through the problem on their own.

Interactions between the Program and Task Levels

To bring together the program level and the task level we return to the notion of a cognitive apprenticeship. The goal of the cognitive apprenticeship is to combine the benefits of a traditional apprenticeship such as observing, imitating, and interacting with a field expert with an explicit attention to the intellectual demand of scientific tasks. We believe that attending to the six HPDE heuristics can be a useful way to think about the design and evaluation of cognitive apprenticeships. We further encourage instructors to think about how these heuristics apply at different structural levels. As can be seen in our analysis of CLIMB, some heuristics were applicable at the program level, while others were applicable at the task level. And, although this is likely to vary from program to program, ideally, these heuristics will be present and mutually reinforcing on all levels.

To end, we suggest that these heuristics provide useful inroads into the design of complex learning environments that could support the kind of transformation in biology education advocated in the *BIO2010* report. We hope that when explicit attention is paid to the notion of productive disciplinary engagement and heuristics for accomplishing it, learning will be enhanced in biology classrooms, and the promise of preparing the next generation of scientists can be met.

ACKNOWLEDGMENTS

We acknowledge CLIMB mentors Richard K. Grosberg, Carole L. Hom, and Sebastian J. Schreiber for support of this work. We are

particularly indebted to the students of the 2008–2009 CLIMB cohort for allowing us to bare witness to their learning process. This research was supported by the National Science Foundation Interdisciplinary Training for Undergraduates in Biological and Mathematical Sciences program under grant 0531935.

REFERENCES

Barab, S., and Hay, K. (2001). Doing science at the elbows of experts: issues related to the science apprenticeship camp. J. Res. Sci. Teach. *38*, 70–102.

Bell, R., Blair, L., Crawford, B., and Lederman, N. (2003). Just do it? The impact of a science apprenticeship program on high school students' understandings of the nature of science and scientific inquiry. J. Res. Sci. Teach. *40*, 487–509.

Brown, J., Collins, A., and Duguid, P. (1989). Situated cognition and the culture of learning. Educ. Res. *18*, 32–42.

Engle, R., and Conant, F. (2002). Guiding principles for fostering productive disciplinary engagement: explaining an emergent argument in a community of learners classroom. Cogn. Instr. *20*, 399–483.

Gee, J. (2002). Learning in semiotic domains: a social and situated account. Natl. Reading Conf. Yearb. *51*, 23–32.

Gilbert, J. (2004). Models and modelling: routes to more authentic science education. Int. J. Sci. Math. Educ. *2*, 115–130.

Grindstaff, K., and Richmond, G. (2008). Learners' perceptions of the role of peers in research experience: implications for the apprenticeship process, scientific inquiry, and collaborative work. J. Res. Sci. Teach. *45*, 251–271.

Halloun, I. (2007). Mediated modeling in science education. Sci. Educ. *16*, 653–697.

Hay, K., and Barab, S. (2001). Constructivism in practice: a comparison and contrast of apprenticeship and constructionist learning environments. J. Learn. Sci. *10*, 281–322.

Hunter, A., Laursen, S., and Seymour, E. (2007). Becoming a scientist: the role of undergraduate research in students' cognitive, personal, and professional development. Sci. Educ. *91*, 36–74.

Jimenez-Aleixandre, P., Rodriguez, A., and Duschl, R. (2000). "Doing the lesson" or "doing science": argument in high school genetics. Sci. Educ. *84*, 792, 757–792.

Kardash, C. M. (2000). Evaluation of an undergraduate research experience: perceptions of undergraduate interns and their faculty members. J. Educ. Psychol. *92*, 191–201.

Lee, H. S., and Songer, B. (2003). Making authentic science accessible to students. Int. J. Sci. Educ. *25*, 923–948.

Lehrer, R., and Schauble, L. (2006). Scientific thinking and science literacy: supporting development in learning contexts. In: Handbook of Child Psychology, ed. W. Damon, R. Lerner, K.A. Renninger, and I. Sigel, Hoboken, NJ: Wiley, 153–196.

National Research Council (2003). BIO2010: Transforming Undergraduate Education for Future Research Biologists, Washington, DC: National Academies Press.

Rogoff, B. (2003). Learning through guided participation. In: The Cultural Nature of Human Development, New York: Oxford University Press.

Seymour, E., Hunter, A.-B., Laursen, S., and Deantoni, T. (2004). Establishing the benefits of research experiences for undergraduates in the sciences: first findings from a three-year study. Sci. Educ. *88*, 493–534.

Steitz, J. (2003). Commentary: Bio2010—New challenges for biology educators. Cell Biol. Educ. *2*, 87–91.

Windschitl, M., Thompson, J., and Braaten, M. (2008). How novice science teachers appropriate epistemic discourses around model-based inquiry for use in classrooms. Cogn. Instr. *26*, 310–378.