Authentic Inquiry through Modeling in Biology (AIM-Bio): An Introductory Laboratory Curriculum That Increases Undergraduates' Scientific Agency and Skills

Susan D. Hester,[†] Michele Nadler,[†] Jennifer Katcher,[‡] Lisa K. Elfring,[†] Emily Dykstra,[†] Lisa F. Rezende,[†] and Molly S. Bolger[†]*

[†]Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721;
[‡]Department of Mathematics and Sciences, Pima Community College, Tucson, AZ 85709

ABSTRACT

Providing opportunities for science, technology, engineering, and mathematics undergraduates to engage in authentic scientific practices is likely to influence their view of science and may impact their decision to persist through graduation. Laboratory courses provide a natural place to introduce students to scientific practices, but existing curricula often miss this opportunity by focusing on confirming science content rather than exploring authentic questions. Integrating authentic science within laboratory courses is particularly challenging at high-enrollment institutions and community colleges, where access to research-active faculty may be limiting. The Authentic Inquiry through Modeling in Biology (AIM-Bio) curriculum presented here engages students in authentic scientific practices through iterative cycles of model generation, testing, and revision. AIM-Bio university and community college students demonstrated their ability to propose diverse models for biological phenomena, formulate and address hypotheses by designing and conducting experiments, and collaborate with classmates to revise models based on experimental data. Assessments demonstrated that AIM-Bio students had an enhanced sense of project ownership and greater identification as scientists compared with students in existing laboratory courses. AIM-Bio students also experienced measurable gains in their nature of science understanding and skills for doing science. Our results suggest AIM-Bio as a potential alternative to more resource-intensive curricula with similar outcomes.

INTRODUCTION

Recent calls for science educators to place greater emphasis on scientific skills (American Association for the Advancement of Science, 2011; National Research Council [NRC], 2012) reflect the fact that traditional forms of classroom science learning are often disconnected from the authentic practices of science. Although science skills are often included in learning objectives for laboratory courses, current research in science education points to the need to *integrate* learning goals for science skills and content throughout science learning. Reasons for this are twofold. First, instructional models that focus separately on teaching science concepts and skills often emphasize science facts without context and may impede integration of knowledge and contribute to persistent conceptual difficulties among students at the undergraduate level (Songer and Mintzes, 1994; Southard *et al.*, 2016). Approaches that teach science facts separately from science skills could also contribute to the inaccurate views about the nature of science that have been documented among undergraduate students (McComas *et al.*, 1998). Second, separation of science concepts from the practice of doing science is an artificial construct that does not mirror authentic science. Authentic research

Learning Explanations
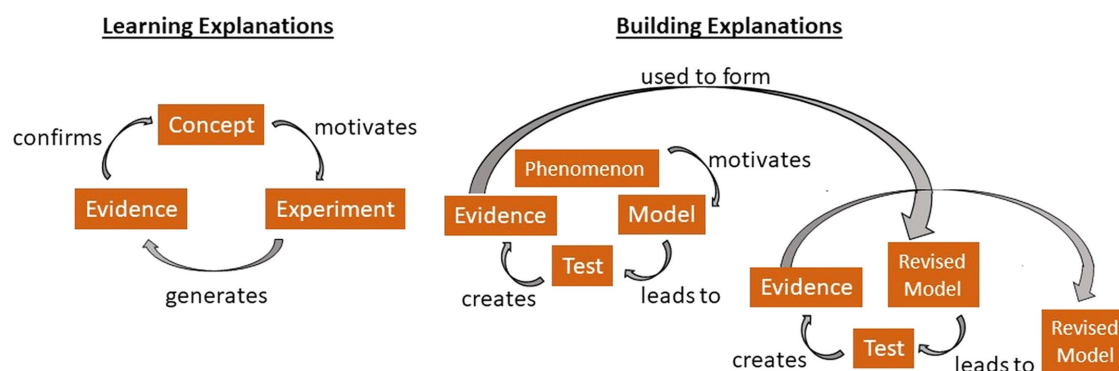
Building Explanations

FIGURE 1. Diagram to illustrate difference between a traditional laboratory curricular approach and an Authentic Inquiry through Modeling in Biology (AIM-Bio) approach. The primary goal of the traditional approach is to use experiments to confirm instructor-provided, conceptual knowledge. The AIM-Bio approach aims to engage students in the process of science, using cycles of modeling to allow students to build their own explanations for scientific phenomena.

experiences within a course or mentored lab experience can increase students' confidence in studying biology, their interest in pursuing biological research, and their persistence to graduation in biology (Jones *et al.*, 2010; Brownell *et al.*, 2012). Traditional laboratory courses, which lack authenticity, do not appear to have these same positive effects on students (Brownell *et al.*, 2012). Given current challenges in undergraduate persistence in science, technology, engineering, and mathematics (STEM) majors, these differences present issues of equity, because many students will never have an opportunity to work in a research laboratory (Bangera and Brownell, 2014). Despite widespread arguments to increase integration of authentic science practices within science courses, many questions remain about the best ways to facilitate this integration. Because so few undergraduate courses have fully integrated science-learning environments, relatively little is known about student learning that occurs in these contexts (Corwin *et al.*, 2015). Additionally, many practical barriers prevent broader application of an integrated approach, including relatively few exemplary curricula (especially in undergraduate biology).

In this paper, we advocate for an approach to undergraduate science teaching that engages students in authentic disciplinary practices throughout their course work. These forms of engagement can be integrated in different ways into diverse courses. Examples of undergraduate curricula that integrate scientific practices with content exist in physics, biology, and chemistry (Brewe, 2008; Walker *et al.*, 2011; Zagallo *et al.*, 2016). In addition, course-based undergraduate research experiences (CUREs) have been demonstrated as a successful approach to integrating aspects of an authentic research experience with learning in a laboratory classroom setting (Auchincloss *et al.*, 2014). We suggest that consistent use of an integrated approach throughout an academic program could reframe the process of undergraduate student learning. With this new frame, from the start of their undergraduate careers, students could view themselves as solving authentic and relevant problems while developing expertise as scientists. Current research suggests that this frame could increase student learning, motivation in science and perhaps their persistence in a STEM field (Capon and Kuhn, 2004; Lopatto, 2007; Harrison *et al.*, 2011; Pease and Kuhn, 2011; Graham *et al.*, 2013; Simon *et al.*, 2015; Rodenbusch *et al.*, 2016). Meeting this goal will require development

of new integrated curricula that address the learning objectives and challenges of diverse contexts. There is a particular need for integrated curricula in introductory courses, which can greatly impact students' decisions to continue within a STEM field. In response to this need, we present a novel curriculum for an introductory biology laboratory course: Authentic Inquiry through Modeling in Biology (AIM-Bio).

AIM-Bio is a one-semester curriculum that integrates the disciplinary practices of modeling, experimental design, and data interpretation with key concepts in molecular and cellular biology. To design this curriculum, we drew from epistemological research on how scientists use models as well as educational research on use of models in the classroom. Figure 1 provides an overview of the shift in instructional focus that was the driving force behind our design. Our goal was to move away from helping students to understand scientists' existing explanations and to move toward helping students build their own plausible explanations for scientific phenomena. The cycle on the left represents our view of a traditional laboratory curriculum in which students are presented with a science concept that motivates performance of an experiment. In this type of course, the purpose of experiments is often to generate evidence that confirms the scientific concept or explanation at hand. Although such a course may help students learn technical laboratory skills, scientific practices are not often integrated in the learning process. The spiral on the right represents our view of a curricular design based on the authentic practice of science. Here, students are presented with a phenomenon that motivates formation of an explanatory model, which leads to development of experimental tests that may provide evidence to support that model. Evidence often warrants revision of the model, and the cycle of experimentation is repeated until, if possible, a satisfactory model is reached. The second approach to instruction, which was the goal for our curricular design, places a greater emphasis on how scientific explanations must flexibly respond to scientific evidence. Furthermore, this type of instruction provides much greater motivation for students to generate and develop their own ideas. In our *Results* section, we illustrate how students in our curriculum generated and tested diverse plausible hypotheses, resulting in learning gains in both the affective and cognitive domains. In our view, the success of our curriculum depended in part upon leveraging students' ability

to engage in model-based reasoning. This process was facilitated by our reliance on previously established principles for design of model based–inquiry instruction. Thus, we provide background on each of these aspects from the literature.

BACKGROUND
Model-Based Reasoning
Model-based reasoning (MBR) is useful for scientists, as it allows for the creativity that is necessary for forming new ideas in science. MBR differs from some other forms of scientific reasoning in that it "cannot be reduced to an algorithm in application," and even if it is used well, an incorrect solution can be produced (Nersessian, 2002). Evidence from historical studies of scientists' work, think-aloud interviews with scientists, and "in vivo" observation of scientists doing their work all point to the use of models as thinking tools for development of new ideas in science (Dunbar, 1999; Nersessian, 1999; Odenbaugh, 2005). MBR is a complex form of reasoning that incorporates use of analogy, imagery, conceptual change, and decision making (Nersessian, 2002).

MBR is thought to rely on humans' ability to form mental models or internal representations that allow us to carry out thought experiments (Craik, 1967; Johnson-Laird, 1983). Such models act as analogies to the real or imagined world, capturing the spatial, temporal, and causal aspects of the world, as well as the constraints of that world (Johnson-Laird, 1983). Representations of models can be *propositional* (primarily based on language and rule-based logic) or *iconic* (primarily based on symbols or images that demonstrate properties of the real world through similarity; Nersessian, 2008). The nature of the representation is likely to impact the ways that it can be used to reason about a situation. For example, one might use an iconic mental representation of "changing a diaper" to decide what to throw in your diaper bag without ever creating a list of "materials needed to change a diaper" or "checklist for how to change a diaper."

Nancy Nersessian (1999) proposed three MBR practices that are useful for creating new ideas in science: 1) creating analogies, 2) employing visual representations, and 3) thought experimenting (or mental simulation). First, when there is no currently available model to explain a phenomenon, scientists may "borrow" a model from elsewhere (creating analogies). Most often, the borrowed model is taken from a nearby situation; for example, in biology, one might create a hypothesis to explain a process in one organism from a mechanism demonstrated in another organism (Dunbar, 1999). Second, mental models may be made explicit through visual representations (employing visual representations). Drawings and gesture may serve more than an explanatory role: in many cases they are considered to serve as an integral part of the "cognitive system," helping to support understanding of the causal and structural constraints and organize cognitive activity (Nersessian, 2002). Significant literature from the cognitive and learning sciences supports the role of drawing and external representations in scientific reasoning (Van Meter and Garner, 2005; Quillin and Thomas, 2015). Third, the practice of conducting thought experiments allows scientists to engage a powerful set of cognitive tools when reasoning with models. Scientists can use the tacit or explicit constraints of a model to mentally simulate the interactions within a mechanism (Nersessian, 2002). The results of this simulation can uncover new constraints or

affordances of the system, forming the basis for new hypotheses. Significant evidence from cognitive science suggests that mental simulation or "animation" can tap into the parts of the brain used to process sensory and motor inputs (Hegarty, 2004; Barsalou, 2008). In short, the frequent use of MBR among scientists can be attributed to its utility, as it takes advantage of a variety of mental processes for creative problem solving.

Application of MBR in Molecular and Cellular Biology
For molecular biologists, mental models and their external representations are often mechanistic in form. Research on mechanistic reasoning (Machamer *et al.*, 2000; Darden, 2002) describes the features that are common among the models biologists often use. Namely, these models include the spatial, temporal, and causal information that is needed to reason about the entities and activities in a model. For example, in a model of protein synthesis, a ribosome is an entity that performs the activity of adding amino acids to the growing polypeptide chain. Thus, mechanistic models in biology have several of the same features as "mental models" described by Johnson-Laird (1983). In addition, mechanistic reasoning includes "forward and backward chaining" (Darden, 2002), which is analogous to "mental simulation" (Hegarty, 2004). In other words, the mechanistic models often used by biologists allow for simulative reasoning using constraints specified by the discipline.

The application of MBR to gain knowledge in biology is a complex interaction between data and model (Passmore *et al.*, 2009). A model is useful if it allows a scientist to conceptualize or explain a natural phenomenon or make predictions about that phenomenon (Odenbaugh, 2005). Thus, biologists constantly evaluate a model based on how well it fits with observation and data and by its ability to predict experimental outcomes. Experiments are designed to test hypotheses formed through MBR, and data are examined for patterns that may support or disprove an aspect of the model or require its revision. In some cases, a model may be completely discarded and replaced. Thus, models serve an essential role in the ways scientists interact with data to construct explanations of natural phenomena.

Model-Based Instruction
Extensive use of models and modeling in K–12 classrooms has taken place in the form of "model-based inquiry" (MBI; Lehrer and Romberg, 1996; Penner *et al.*, 1997; Stratford *et al.*, 1998; Windschitl *et al.*, 2008; Passmore *et al.*, 2009; Schwarz *et al.*, 2009), through which students engage in the scientific practice of building, testing, and refining models based on inquiries conducted in the natural world. At the undergraduate level, MBI has been described in chemistry (Khan, 2007; Tien *et al.*, 2007; Walker *et al.*, 2011) and physics (Brewe, 2008; Zwickl *et al.*, 2014). In undergraduate biology, curricula have been developed for content-based courses that engage students in developing conceptual models (Speth *et al.*, 2014; Bierema *et al.*, 2017; Baze and Gray, 2018) or using models to aid in interpretation of data (Zagallo *et al.*, 2016). Svoboda and Passmore (2010) also described a small-enrollment course for biology students to engage in authentic mathematical modeling. However, as yet, we are unaware of published examples using an MBI approach of the type described here to develop a curriculum for an undergraduate biology laboratory course. This setting provides a unique opportunity for students to engage in

the authentic scientific practices that surround building, testing, and revising models based on evidence from experimental data.

Just as modeling has been shown to support reasoning among scientists, previous work suggests the potential for model-based instruction to support students' cognition. For example, Clement (2008) demonstrated that asking a student to develop models supported conceptual change. Other studies have shown how a modeling approach could encourage students' sense-making in a classroom setting (Stewart *et al.*, 2005; Bierema *et al.*, 2017). Based on what we understand about MBR among experts, it has been suggested that the ability to conduct productive mental modeling develops with learning in a domain (Ippolito and Tweney, 1995; Nersessian, 2002), further emphasizing the importance of allowing students to practice this skill through schooling.

EDUCATIONAL SETTING

This study took place at a large, public university and a local community college, both recognized as Hispanic-serving institutions (having at least 25% Hispanic students). Both institutions offer Introduction to Molecular and Cellular Biology (IMCB), a one-semester course requirement for a range of life science majors. The course topics include the chemistry of life, biomolecules, cellular respiration and photosynthesis, the central dogma, control of gene expression, DNA replication, meiosis and mitosis, cell structure and signaling, and Mendelian genetics. More than 1800 students enroll in this course each year; many of these students are from groups traditionally underrepresented in science.

At the university, the IMCB lecture and the IMCB laboratory are offered as separately graded courses. The unrevised laboratory course differs from a traditional biology laboratory course in that many of the activities involve students working through tutorials accompanied by online simulations with the aim of reinforcing course content. Wet-lab components of the laboratory involve students making predictions and designing experiments to confirm principles covered in the course. Although attempts are made to coordinate the content covered and the order of topics, the two courses are not tightly aligned. Students may choose to take the two courses concurrently or during different semesters. As a result, students come to the laboratory with a diversity of backgrounds, ranging from freshmen with relatively little background knowledge to seniors in their last semester of a life science major (e.g., physiology). Students in a single laboratory section do not necessarily take the same section of the lecture course. Each laboratory section enrolls approximately 25 students and meets for 3 hours once per week for 14 weeks. Instructors are typically graduate or undergraduate teaching assistants (TAs) working under the supervision of the laboratory director. Resources for materials and equipment in the IMCB laboratory are limited relative to many other institutions, as the department wishes to keep course fees low for students.

At the community college, IMCB is taught as a combined lecture and laboratory course. Instructors are free to use the combined course time flexibly to achieve lecture and laboratory goals. Sections of the course typically enroll 25 students. Both portions of the course are typically taught in the same room. Before the experimental semester of this project, the laboratory portion of the course was recognizable as a traditional laboratory course in which students primarily reinforced course material by conducting confirmatory experiments. Instructors are typically full-time (or sometimes adjunct) faculty members. As at the university, students come with a wide range of preparedness, with the additional challenge that many students are taking classes while balancing full- or part-time employment.

INSTRUCTIONAL DESIGN

The research described here takes place within the context of an instructional design project. This project aims to create and test a MBI curriculum for an introductory molecular and cellular biology laboratory course. We refer to our curriculum as Authentic Inquiry through Modeling in Biology (AIM-Bio).

Our goals for the project include that the AIM-Bio curriculum will

1. follow integrated learning objectives that address both key concepts and scientific practices;
2. center on models as thinking tools;
3. equip students with skills and confidence to build and refine different types of biological models;
4. focus on the relationship between models and experimental data; and
5. connect students to authentic research by allowing them to conduct inquiries related to ongoing projects by local scientists.

To meet these instructional goals, we developed our curriculum using design principles derived from the literature on MBR and MBI. The first principle was building units around *cycles of modeling*. We aimed for these cycles to provide students with multiple opportunities to create and modify models based on evidence gathered and/or comparison with other possible models (Lehrer and Schauble, 2006). The second design principle was *scientific inquiry*. Effective MBI instruction is highly student centered. Students are expected to form their own hypotheses and explanations and to adapt these ideas through experimental tests. This approach depends on students' willingness to generate and share their own ideas. Thus, the AIM-Bio approach depends on students' agency in science and also provides the potential for growth of student agency through the curriculum. The third principle was *model diversity*. We included this principle based on prior research suggesting that tasks that elicit variation in students' hypotheses and models allow for greater learning about scientific practice (Lehrer and Schauble, 2006). We also reasoned that such diversity would provide a place to support students' agency in the science classroom. The fourth design principle was *model testing*. Students were asked to use their models to make predictions about what they might observe through experimentation and to use their results as evidence for or against a particular model (Windschitl *et al.*, 2008). This principle ensured that students would gain experience in navigating between the theoretical and empirical aspects of scientific practice. The fifth principle was inclusion of *public discussion of models*. It was necessary to develop a classroom culture to support MBI. Previous work highlights the importance of asking students to share their models and evidence with classmates and to develop arguments to defend their models against peer critiques (Stewart *et al.*, 2005). The final principle was ensuring the appropriate *task difficulty level*. An adequate level of challenge in the laboratory activity is important for supporting

MBI. If one simple model obviously describes a data set, conversations about competing models and model testing will be hard to support (Stewart *et al.*, 2005). At the same time, our student population is very diverse in their levels of preparation. Ensuring appropriate difficulty level for different students required building a curriculum that was open enough to allow different students to reach different endpoints. In addition, task difficulty is related to how well a student may expect to perform that task, which is heavily tied to his or her motivation for classroom engagement (Wigfield and Eccles, 2000). Opportunities to overcome obstacles and succeed in a domain are important for development of students' self-efficacy and willingness to persist in that domain. Thus, the difficulty level of tasks is likely to be particularly important in MBI curricula.

Following these principles, we designed a multi-unit, one-semester curriculum, outlined in Table 1. In the following sections, we describe ways in which our design principles drove creation of our AIM-Bio curriculum.

Engaging Students in Cycles of Modeling
The curriculum was deliberately structured to engage students in cycles of modeling. Each unit focused on iteratively building models of a target phenomenon. Students first observed the phenomenon, then worked together in groups of two or three to draw an explanatory model of the phenomenon. Next, they used their models to formulate testable hypotheses and predictions, which they tested by designing and carrying out their own experiments. On the basis of the resulting data, they revisited and revised their models. Finally, they reflected on the process in individual end-of-unit reports in which, among other things, they were asked how their models had changed, how their models were supported by the data, and how they might continue to test and refine their models if given the time and resources to do so.

Supporting Scientific Inquiry
We aimed to design a curriculum that gives students a leading role in guiding the specific direction of inquiry. Whenever possible, students observed and noticed the focus phenomenon themselves rather than being told what to look for. This was the case in the black box, membrane transport, bacterial growth, and *Chlamydomonas* phototaxis units. To achieve this, we chose phenomena that were readily observable by our students and that reliably pique students' interest enough to beg explanation. Although the questions posed by the phenomenon in each unit were constrained by design (we as the designers intended for the students to ask questions aimed at explaining a particular phenomenon upon making their initial observations), allowing

TABLE 1. Outline of units, phenomena, concepts, and scientific tasks in Fall 2017 AIM-Bio curriculum

Week	Unit	Phenomenon	Biology concepts	Pose a question	Create a model	Design experiments	Interpret data	Revise a model
1	Black Box	Different-sized balls produce different outcomes when put through a "black box" maze	• What is a model and how do scientists use models • How to use a pipet		•	•	•	•
2 3	Membrane Transport	Different cell types react differently in hypotonic and hypertonic conditions	• Plant and animal cell structure • Osmosis • Protein function		•		•	•
4 5 6	Bacterial Growth	Bacterial species thrive differentially in different environments (including the presence of other bacterial species)	• Prokaryotes • Gram staining • Cellular response to environment • Enzymes • Metabolism		•	•	•	•
7 8	Computational Cancer	Tumors metastasize despite likely fatality of migration from tumor for any given cell	• Cell cycle/cellular proliferation • Mutations • Metastasis • Hallmarks of cancer • Types of chemotherapy				•	•
9 10 11	*Chlamydomonas reinhardtii* Phototaxis	*Chlamydomonas* cells swim toward a light source	• Cellular motility • Protein function • Molecular systems thinking • Photosynthesis • Mutant screens	•	•	•	•	•
12 13	Pathway Thinking in Yeast	Cells express different genes in different environments	• Inheritance • Genotype/phenotype relationships • Cell signaling		•		•	•
14	Flex week—final exam or Pathway Thinking in Yeast continued							

students to observe and notice the phenomenon themselves gave them greater ownership and motivated them to engage in the modeling cycle as they attempted to explain the phenomenon. Moreover, although the broad question—"How/Why does _____ happen?"—was determined by the designers' choice of phenomenon, more specific questions of "What is the role of _____ in making this happen?" were left up to the students as they created their models. Whereas they were all seeking to explain the same phenomenon, different students proposed and investigated different potential mechanisms.

Such open-endedness was made possible by two decisions. First, we chose noncanonical phenomena as the foci for the units. Second, we chose to not provide or drive students toward "correct" explanations for the phenomena at any point. By choosing noncanonical phenomena, we shifted the goal from familiarizing students with a phenomenon and its canonical explanation to supporting them in building, testing, and revising biologically plausible—although potentially wrong—explanations using domain knowledge and evidence. In other words, we shifted the expectation from *learning* explanations to *building* explanations in an authentically scientific way through cycles of modeling. We reinforced this shift by choosing not to "lift the curtain" at the end of the unit. Avoiding an instructor-sanctioned answer maintained the focus on students' explanations, arguments, and evidence. In our view, providing an answer at the end of the unit would likely remove students' motivation to invest the energy to strive for the most defensible explanation. The decision to use noncanonical target phenomena made this possible: students were not disadvantaged by not arriving at a correct or complete explanation.

Supporting Model Diversity

The choice of phenomenon is also important for supporting model diversity. As explored in the *Results* section, different phenomena lent themselves to diverse models in different ways. Some target phenomena (employed in the black box, membrane transport, bacterial growth, and pathway thinking in yeast units) are relatively simple phenomena that nonetheless have multiple biologically plausible explanations that are accessible to students. Others (employed in the *Chlamydomonas* phototaxis and computational cancer labs) have sufficient complexity that students can choose to explore different aspects of the same phenomenon.

Engaging Students in Model Testing

Model testing is at the heart of practicing science. In each modeling cycle, students were required to use their models to predict experimental outcomes and relate experimental data back to their models to determine whether or not elements of their models were supported. Although hypothesis-testing and experimentation are present in some form in most biology lab courses, the agency afforded students in how they chose to test their models was a key element of our AIM-Bio curriculum design. A diversity of student models called for a diversity of potential experimental tests. The black box, bacterial growth, computational cancer, and *Chlamydomonas* phototaxis units each involved different student groups designing, justifying, and carrying out different experiments to test components of their own particular models. Realizing this design element required anticipating a broad range of student ideas and making appropriate materials available for

testing those ideas. We also constrained students somewhat by providing them with a list of available materials and tasking them with formulating a hypothesis that they would be able to test. Our experience was that students' hypotheses were not overly limited by the materials available—all groups were able to test central hypotheses of their models with the given materials.

Supporting Public Discussion of Models

In each unit, we built in multiple opportunities for students to share and discuss their models. For instance, after students drew their initial models, we typically chose a few different models—representative of different types of ideas in the classroom—and asked those groups to share and describe their models. This gave students insight into the diversity of hypotheses and representations being created by their peers. We also posted student models for each unit in a shared online lab notebook so that students could reference different class models at will. To foster more in-depth discussion and comparison of models, we sometimes combined two or three groups and directed them to compare their models and ask any questions necessary to understand one another's models and the decisions that went into making them. When students shared data with one another, we instructed them to do so in relation to their models: Why did they perform the experiment that they did? What were they expecting to see? How did the resulting data relate to their model? This provided a concrete context in which students could consider the mechanisms within and evidentiary support for different models. Finally, we continually supported informal sharing of models and ideas through instruction. We frequently referred student groups to other groups when we felt that they had complementary ideas and/or data.

Ensuring Appropriate Task Difficulty Level

Because the entire modeling cycle is driven by student-generated ideas, there is a gradient of potential beginning and endpoints. Students' explanations can vary in level of sophistication, detail, strength of evidence, and quality of argument from evidence. Once again, choice of phenomena is important. We aimed for focus phenomena that were straightforward enough to be described by poorly prepared students. At the same time, we chose phenomena that could potentially be explained with multiple mechanisms that would be plausible to students with a more sophisticated background and understanding.

Identifying and Developing Target Phenomena for the AIM-Bio Approach

The choice of phenomena is central to the AIM-Bio approach. The key design criteria have been discussed: phenomena should be noncanonical, they should ideally be observable by students, they should have multiple plausible hypotheses that are accessible to students, and those hypotheses should be testable given the material and logistical constraints of the lab. In practice, identifying and developing these phenomena for instruction is an iterative process, which we describe in the paragraph that follows.

We relied heavily on the primary literature and on research scientists from the university to identify and develop appropriate phenomena for instruction. Guided by course content objectives, we typically began by identifying a system to explore (e.g., we chose *Saccharomyces cerevisiae* as a system in which to develop instruction integrating pathway thinking). Our choice

of system was often informed by the research carried out at our university. We felt that this increased authenticity, as it connected the curriculum to local ongoing research. It also provided us the valuable resource of local expertise in the system. Once we identified a system, we combed primary literature and/or discussed options with researchers, applying the criteria described earlier to identify an appropriate target phenomenon. Phenomena for which multiple hypotheses had been explored in the literature were particularly attractive, as they were likely to elicit multiple hypotheses from our students. After identifying a potential target phenomenon, we informally piloted it to test its ability to elicit multiple models from individuals with a wide range of expertise. In addition to brainstorming potential explanations ourselves, we piloted the phenomenon in an interview setting by asking students to pose an explanation and how they might test it. Potential target phenomena that succeeded in eliciting multiple possible explanations were considered promising. Finally, we developed instructional materials and laboratory protocols for developing the target phenomenon into an AIM-Bio unit. Instructional materials were intended to appropriately frame the phenomenon and provide students with background information that could help them engage productively in attempting to explain the phenomenon. Laboratory protocols were found in the literature and shared by local research scientists and typically adapted to meet our constraints.

RESEARCH QUESTIONS

In this paper, we describe how the model based–inquiry curriculum we designed supports students' development of cognitive skills for doing science. We first looked for evidence to test whether our instructional design fostered the kinds of student-driven generative reasoning that we aimed to encourage. Specifically, we examined the extent to which students developed diverse models and hypotheses, designed appropriate tests for those hypotheses, and revised their models in response to evidence (research question 1). Developing one's own hypotheses and experiments requires independent or creative thought. We asked whether the opportunity to engage in this type of activity within the AIM-Bio curriculum influenced students' agency and identification of themselves as scientists within the curriculum (research question 2). Finally, by engaging directly in the cognitive as well as technical aspects of the scientific process, students should gain more sophisticated views of the nature of science and increase their skills for doing science. We aimed to measure the extent to which this occurred in the curriculum (research question 3). Thus, the research questions that we address in this paper are

1. Can an AIM-Bio curriculum support students in generating, testing, and revising a diversity of explanatory hypotheses?
2. To what extend do students identify as scientists and take agency of their own ideas within an AIM-Bio curriculum?
3. Does an AIM-Bio curriculum impact students' views of the nature of science and/or scientific skills?

METHODS
Implementation

The AIM-Bio curriculum was phased in over two semesters. A pilot semester (Spring 2017) for the AIM-Bio curriculum included two coinstructors (M.S.B. and S.D.H.) in a single IMCB section at the university. The following semester (Fall 2017), we implemented a revised version of the entire curriculum in three sections of IMCB at the university. M.S.B. and S.D.H. cotaught one section, MB and an experienced undergraduate TA cotaught one section, and the same TA taught a third section. At this institution, well-supervised undergraduate TAs often teach lab sections. At the community college, two authors (J.K. and S.D.H.) cotaught one section of IMCB that incorporated three units of the AIM-Bio curriculum (black box, membrane transport, and bacterial growth), adapted for the community college setting and taught over eight class periods. The curriculum for the community college AIM-Bio section also concluded with a CURE unit conducted in collaboration with a research group from the university. In the CURE unit, students prepared insect samples for genetic bar-coding using a given protocol, then compared their sequences with databased species using a BLAST search. At the end of the semester, the students' sequences were databased by the research group for use as part of an ongoing project. Students at both the university and community college enrolled in the course without being notified that they would be experiencing a new curriculum. This allowed us to draw comparisons between students in the AIM-Bio and unrevised curricula without accounting for possible self-selection effects.

In all AIM-Bio sections, students self-selected into groups that were kept stable for the duration of the course, except for a few particular instances in which the instructor adjusted groups as deemed necessary to ensure productive interaction between students. Anecdotally, students perform well on complex tasks when allowed to maintain stable groups, as it allows them to develop relationships with those with whom they are working. This choice also decreases the logistical demands on the instructor.

At the university, each lab meeting was preceded by a pre-class assignment (typically consisting of a short reading and an online multiple-choice quiz) and began with a brief quiz typically consisting of five to seven multiple-choice and short-answer questions focusing on the prelab assignment and previously discussed concepts. In-lab assignments included worksheets guiding students through lab tasks (e.g., initial observations of the phenomena, brainstorming experiments) and in-class formulation of group models. At the end of each unit, we assigned individual reports. These reports gave students the opportunity to describe and interpret their data, propose models (either the group models from class or their own individual models, if they preferred), support their models with the data, and describe necessary revisions to their original models based on the data. We gave extensive formative feedback on reports. In particular, we emphasized the importance of including explanatory mechanisms in their models (as opposed to mere descriptions of the phenomena), interpreting experimental data, and supporting their models by directly relating specific results to components of their models. At the end of the semester, we administered a final exam that included both content and skill-based items.

At the community college, the AIM-Bio units were structured to roughly adhere to the format of other lab units for the course. We assigned preclass reading before each unit. In-lab assignments included worksheets guiding students through lab tasks (similar to those used at the university) and in-class formulation of group models. At the end of each AIM-Bio unit, we gave students a take-home assignment (formatted as a worksheet) in which they described their models, supported them with data from the class, and described necessary revisions to their

original models based on the data. We provided extensive formative feedback on these assignments. As at the university, we focused particularly on having explanatory models, interpreting data, and supporting their models by relating model components directly to the data. The final exam for the course included a section assessing the skills practiced in the AIM-Bio units.

At both the community college and the university, course points and instructor feedback for the AIM-Bio units were structured with the goal of emphasizing scientific process skills of modeling, data interpretation, and relating data to models.

Data Collection

Data from a total of 781 students are included in this study. We collected data across four semesters (Fall 2016–Spring 2018) in four different contexts: unrevised and AIM-Bio sections of the IMCB course at the community college and unrevised and AIM-Bio sections of the IMCB laboratory course at the university. Table 2 shows a breakdown of student participation in the parts of the study reported on here. All research activities were approved by the University of Arizona Internal Review Board.

Data collection included the administration of the Project Ownership Survey (POS; Hanauer and Dolan, 2014) and the adapted Classroom Test of Scientific Reasoning (Benford and Lawson, 2001, adapted from Lawson, 1978). Both instruments were administered at the university via an online format. The POS was administered in a pencil-and-paper format at the community college. We also conducted pre/post student interviews outside class time using protocols adapted from Russell and Weaver (2011) and Southard *et al.* (2017) at both the university and community college. Students were compensated for these interviews with small gift cards. On the basis of analysis of these interviews, we created and administered a written survey to assess nature of science understanding and identity as a scientist (interview analysis and development of the written survey are described in more detail later). The written survey was administered as part of an in-class quiz in the AIM-Bio section at the university, as part of a written take-home final in the AIM-Bio section at the community college, and as an in-class survey in unrevised sections of the IMCB laboratory at the university. Finally, we collected extensive documentation of the instructional intervention (audio recordings of instructors and consenting student groups, copies of consenting students' work, and researcher field notes).

This paper focuses on a subset of collected materials. To address our first research question, we used written artifacts collected from students. For the second research question, we drew from the results of the POS, interviews with students, and the written survey. For the third research question, we drew

from the results of the adapted Classroom Test of Scientific Reasoning and the written survey.

Qualitative Data Analysis

To address our first research question, we analyzed student models, experimental designs, in-class written reflections, and end-of-unit lab reports. To better understand how students generated, tested, and revised models, two researchers reviewed and discussed student artifacts from two units (the bacterial growth unit and the *Chlamydomonas* phototaxis unit). Examples of artifacts were chosen to highlight apparent themes. Finally, we developed a coding scheme to categorize the different explanations students proposed for the same biological phenomenon in the bacterial growth unit. This coding scheme is presented in the Supplemental Material. Two coders independently coded each model drawing with one or more codes, with an interrater agreement of 74%. The coders then came to agreement on all codes presented in the *Results* section.

To address research questions 2 and 3, we conducted pre/post student interviews focusing on the nature of science and students' identities as scientists. We adapted the interview protocol from Russell and Weaver (2011), who themselves adapted the questions from the Views of Nature of Science Form (Lederman *et al.*, 2002). In adapting the interview protocol, we added questions intended to gauge students' identities as scientists. We conducted pilot interviews with students from sections of the unrevised university laboratory and community college IMCB courses in Fall 2016. On the basis of a subset of transcripts from these interviews, we developed a coding scheme to capture indicators of the sophistication of students' understanding of nature of science (for nature of science questions) and recurring themes in students' descriptions of whether, when, and why they identified as scientists. We applied this coding scheme to the remainder of the transcripts, revising the coding scheme as necessary for clarification. We conducted interviews with students from the AIM-Bio university course in Spring 2017 and applied the same coding scheme to transcripts from those interviews, revising the coding scheme as necessary to capture themes that were prominent in the interviews with students from the AIM-Bio course, but not the unrevised course. On the basis of this analysis, we narrowed the interview protocol to a list of four questions that gave us the most insight into students' understanding of the nature of science and their identities as scientists:

- What is an experiment?
- Please describe an example of an experiment.
- What is a theory?
- Have you ever felt like a scientist? Please explain.

TABLE 2. Study participation data including numbers of participants in the parts of the study reported on in this paper[a]

Semester	Course type	Location	Project Ownership Survey	Adapted Nature of Science interview	Written survey	Classroom artifacts
Fall 2016	Unrevised	University	631	28	–	–
Spring 2017	AIM-Bio	University	20	7	–	23
Spring 2017	Unrevised	Community College	21	9	–	–
Fall 2017	AIM-Bio	University	64	19	43	50
Fall 2017	AIM-Bio	Community College	23	–	13	13
Spring 2018	Unrevised	University	–	–	43	–

[a]Many students chose to participate in multiple parts of the study. Data from a total of 781 students were collected for this study.

We administered these questions in two forms in Fall 2017: as the first part of a larger interview protocol with students from the university AIM-Bio course (pre/post) and as a written version given to students in university AIM-Bio (pre/post), community college AIM-Bio (post), and university unrevised (post) sections of IMCB. We verified that students gave similar answers for verbal and written forms of the questions by comparing written and verbal responses for students to whom we administered the questions in both forms. We applied the coding scheme developed for the interviews to students' responses to university AIM-Bio students' entire written surveys, and to the last question for all students' surveys. We modified the coding scheme slightly for clarification. The final coding scheme is available in the Supplemental Material. Two independent researchers (M.S.B. and S.D.H.) applied the coding scheme, with an interrater agreement of 79%. The coders then came to agreement on all codes presented in the *Results* section.

Using the codes assigned to AIM-Bio students' pre/post survey responses for the first three questions, we calculated pre/post Nature of Science scores for each student. Following an algorithm (included in the Supplemental Material), we gave a score between 0 and 3 for each of the three questions, with higher numbers indicating greater degrees of sophistication evident in the response, as reflected in the codes assigned. We then summed the scores for the three questions to generate a Nature of Science score with a value between 0 and 9.

Statistical Data Analysis

To further address our second research question, we administered the POS to students in unrevised and AIM-Bio sections at both the university and community college. Initially, we compared means on the entire instrument between unrevised and AIM-Bio sections at each institution, using Welch's two-sample *t* test. To gain a finer-grained picture of how the populations differed, we also compared means for each item, as originally reported in Hanauer and Dolan (2014). Because we noticed differences in certain groups of items between sections, we performed an exploratory factor analysis using maximum-likelihood extraction and oblimim with Kaiser normalization rotation. We identified four latent variables that together explained 67% of the variance in the data. We assigned items to four categories according to how they loaded onto the four latent variables (shown in Supplemental Tables S1 and S2). On the basis of the items in these categories, we characterized them as Personal Investment, Real-World

Contributions, Positive Emotions, and Self-Efficacy/Unexpected Experiences. We compared the means within these categories between unrevised and AIM-Bio sections at each institution using Welch's two-sample *t* test.

In addressing our third research question, we asked whether there was a positive shift in university AIM-Bio students' Nature of Science scores (described earlier) from pre to post. We compared the pre and post score distributions using the Wilcoxon signed-rank test, which is appropriate for determining significant differences between the distributions of paired sets of ordinal data. We also asked whether students' skills increased over the course of the semester. Before analyzing the adapted Classroom Test of Scientific Reasoning, we removed questions 14, 16, 19, and 20, due to potentially confusing wording in the published instrument. We compared distributions of university AIM-Bio students' pre and post scores on the remaining questions using Welch's two-tailed paired *t* test.

RESULTS

Can an AIM-Bio Curriculum Support Students in Generating, Testing, and Revising a Diversity of Explanatory Hypotheses?

Our first instructional design principle was to engage students in cycles of modeling that would allow students to produce and refine models to explain observed biological phenomena. During this process, students engaged in model testing by designing and carrying out experiments and analyzing data from those experiments. Importantly, we aimed to foster model diversity across the classroom. To accomplish these goals, we needed to design activities in which more than one model was plausible and testable with the materials we had available. Here, we provide examples from two different AIM-Bio units to highlight the kinds of models students produced and the experiments they designed to test their models.

Example 1: Bacterial Growth. In this unit, students were presented with six different culture conditions (summarized in Figure 2). Two bacterial species, *A* and *E*, were grown separately and together in two different growth media: a rich culture medium (ATCC#3) and a sparer medium containing colominic acid (CA; medium #2). Upon performing a Gram stain of the cultures, students observed a puzzling phenomenon: although both bacterial species thrive in the ATCC#3 medium and species *A* thrives in medium #2, species *E* only thrives in medium #2 when cocultured with *A*. We then asked students to

Phenomenon: Bacterial growth unit

Medium:	ATCC #3	ATCC #3	ATCC #3	Medium #2	Medium #2	Medium #2
Inoculated with:	*Species A*	*Species E*	*Species A and Species E*	*Species A*	*Species E*	*Species A and Species E*
Observed with Gram stain:						

FIGURE 2. Students were presented with cultures of bacteria grown in two different media. ATCC #3 is a rich medium containing meat extract and peptone. Medium #2 is a minimal medium containing colominic acid (CA). CA is a polymer of N-acetylneuraminic acid (NA). NA is also a component of complex sugars found on the surface of mammalian cell membranes. Two species of bacteria were cultured alone or together in the different growth media. Students performed Gram stains and saw that one of the species (species *A*) grew readily in both media, but the other species (species *E*) thrived in ATCC #3 alone but did not thrive in medium #2 unless it was cocultured with species *A*.
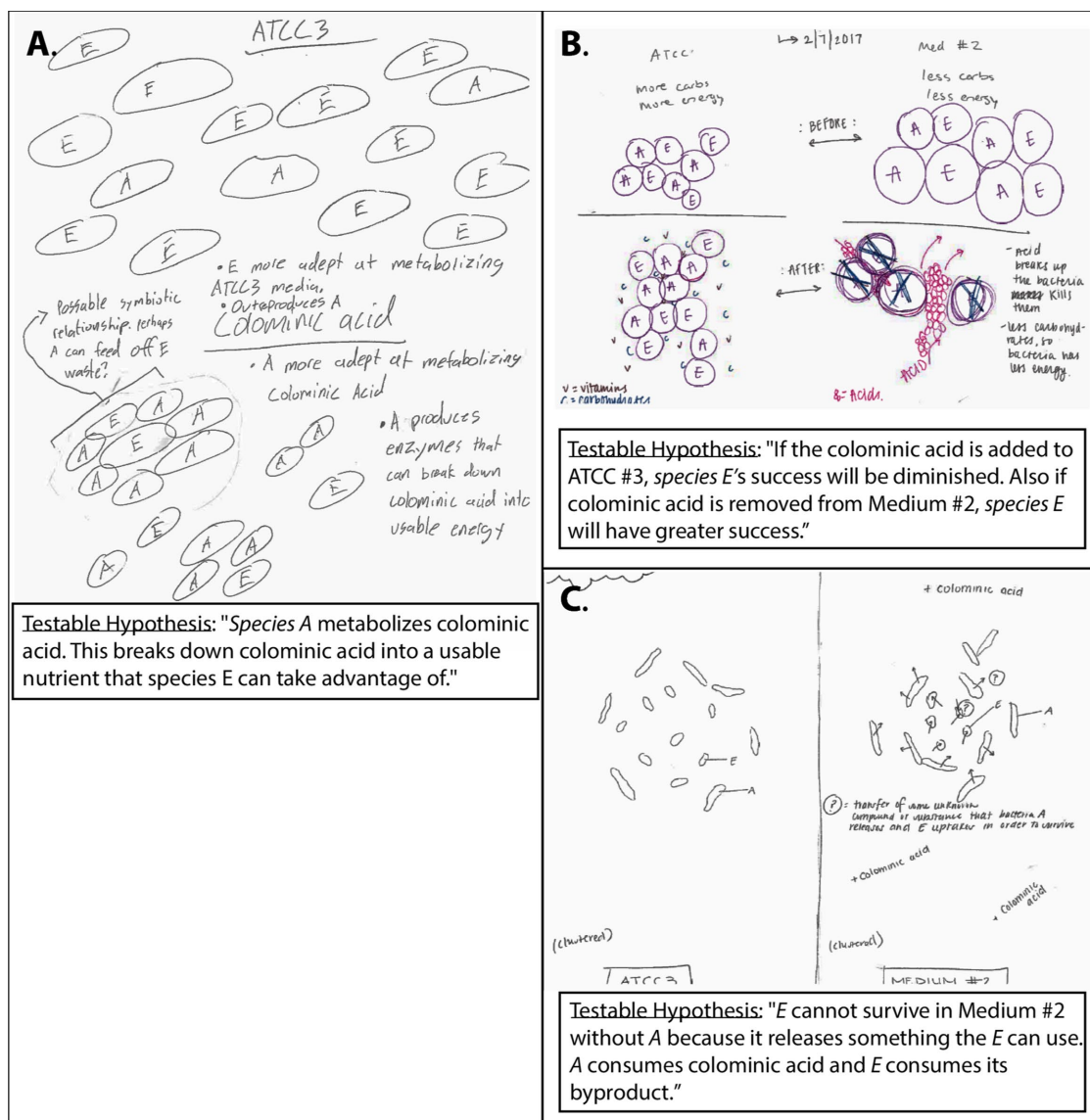
FIGURE 3. **Students worked in groups of two or three students to generate models to explain the bacterial growth phenomenon. (A–C) Three different models are shown to illustrate the diverse hypotheses that students generated in this instructional context. Students worked in combined groups of four to six students to compare their models and devise a testable hypothesis based on their model and a list of available laboratory materials.**

work in small groups to draw models that might explain this phenomenon.

Students generated and tested diverse models to explain the bacterial growth phenomenon (Figure 3). For example, one group proposed the hypothesis that species *A* metabolizes CA, which provides nutrients for species *E* (Figure 3A). Another group hypothesized that species *A* protects species *E* from harm by CA (Figure 3B). A third group hypothesized that species *A* transfers something to species *E* that it needs to survive in medium #2 (Figure 3C). When we examined models collected from students over four different AIM-Bio sections of IMCB, we saw that there was overall diversity in the models students drew (Figure 4). The hypotheses that students formulated fell into six categories, with most fitting into one of three categories. Two student groups simply drew the biological phenomenon but did

not attempt to provide a mechanism to explain it. Model drawings for each of the four sections of the course fell into at least three different categories, suggesting reliable diversity of ideas within the classroom. The three most common categories were present in three of four sections of the course. This suggests that, when teaching this unit, an instructor can predict that these more common hypotheses will usually be present somewhere in the classroom.

We next provided students with opportunities to test and revise their models. This required that materials be available to support the ideas that students wanted to test. To facilitate alignment between students' hypotheses and possible experimental tests, we provided students with a list of available laboratory materials and asked them to work with one other group to examine their models and devise a "testable hypothesis."
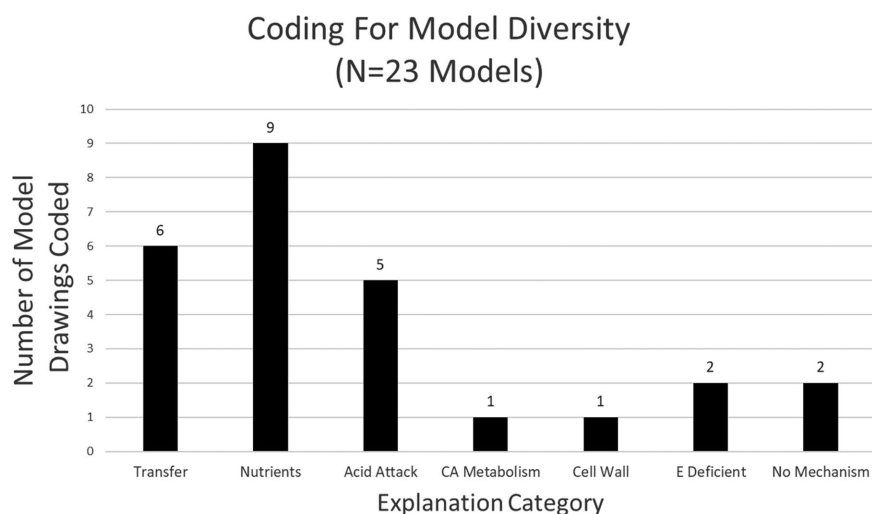
Coding For Model Diversity
(N=23 Models)

FIGURE 4. Students' initial model drawings for the bacterial growth unit were coded for types of explanations included (the coding scheme is included in the Supplemental Material).

Students generated and conducted a diverse set of experiments to test their hypotheses. Figure 5 shows experimental conditions designed by the students whose models are shown in Figure 3. For example, a central element of the model shown in Figure 3C is the speculation that there is "transfer of some unknown compound or substance that bacteria A releases and E uptakes in order to survive." Working with another group, these students stated the hypothesis, "E cannot survive in Medium #2 without A because it releases something that E can use. A consumes colominic acid and E consumes its byproduct." They reasoned that the physical presence of species A was not required for species E to thrive, as long as species A's "by-product" was present in the medium. To test this, the students designed and carried out the experimental condition described in Figure 5C. The students began with an established culture of species A in medium #2 with CA. They removed the bacterial cells from the medium by centrifuging the culture, removing the supernatant, and sterilizing it with a syringe filter. They then cultured bacteria E in the sterilized supernatant. The students predicted that species E would thrive in this culture, as it should contain the necessary "by-product" released by species A into the medium.

After conducting their experiments, students shared their findings with other groups. Then the original small groups met to draw a revised model. Figure 6 shows revised models for the groups whose initial models are shown in Figure 3, A and B. Students often included a greater level of molecular mechanism in their revised models. For example, Figure 6A shows the model that was the revision of the model shown in Figure 3A. The initial model mentions in words that "A produces enzymes that can break down colominic acid into useable energy." The revised model, however, includes a molecular mechanism for how this could occur: in the presence of CA, species A secretes an "acidase" that breaks the polymer CA into sugar monomers that can diffuse into the cells of both species and be metabolized. Moreover, although the initial model included the central idea in some form, it was encoded largely in words (the images do not do much "work" in communicating the ideas) and lacked

spatial detail. The revised model, on the other hand, depicts dynamic spatial relationships between entities in the model. The revised model also incorporates a molecular mechanism that is entirely absent from their initial model: a signal protein that detects CA and "triggers the production of acidase." Interestingly, the annotation of this mechanism in the model ends with a question mark, indicating that the students are deliberately incorporating speculative molecular mechanisms that are biologically plausible but lack current, direct evidence. By using models as a space to generate explanations that reflect current experimental knowledge as well as plausible hypotheses, students were engaging in modeling practices similar to those used by expert scientists.

Students were asked to reflect on their model revision in their lab reports. In the following excerpt from an end-of-unit lab report, one of the students involved in drawing the models shown in Figures 3A and 6A describes how their revised model was supported by the experimental evidence collected by their team:

> Species A has evolved an enzyme that can break down colominic acid into metabolizable sugars as shown in UV tests noted above. It must excrete this enzyme outside the cell due to the large size of the CA molecule and its thick cell wall (gram positive test). Species A has also developed some type of signal protein that interacts with CA so the cell only produces this enzyme when needed saving energy. This is supported by the finding that A does not produce this enzyme in ATTC3 medium. Species E is unable to metabolize CA as tests show it cannot grow alone in CA medium. However, because A must breakdown the CA in the outside environment, species E can steal and metabolize the CA monomers. This is why E can grow in CA medium with A and, why it can grow really well alone in CA medium that has been treated with A and then filtered out.

The student then goes on to reflect on how and why their original model changed from the beginning of the unit:

> Our original model depicted more of a symbiotic relationship between the two species but tests that show more limited E growth with A compared to A filtered medium caused us to revise our model.

In some cases, students' experimental data did not support their original hypotheses. For example, the group that drew the initial model in Figure 3B tested their "acid attack hypothesis" by culturing species E in ATCC#3 medium containing CA and in medium #2 without CA (Figure 5B). Counter to their prediction that species E would be unable to thrive in ATCC#3 with CA, they saw no decrease in bacterial growth compared with the condition of species E cultured in ATCC#3 without CA. Likewise, counter to their prediction that species E would thrive in medium #2 without CA, they saw no increase in bacterial growth compared with the condition of species E
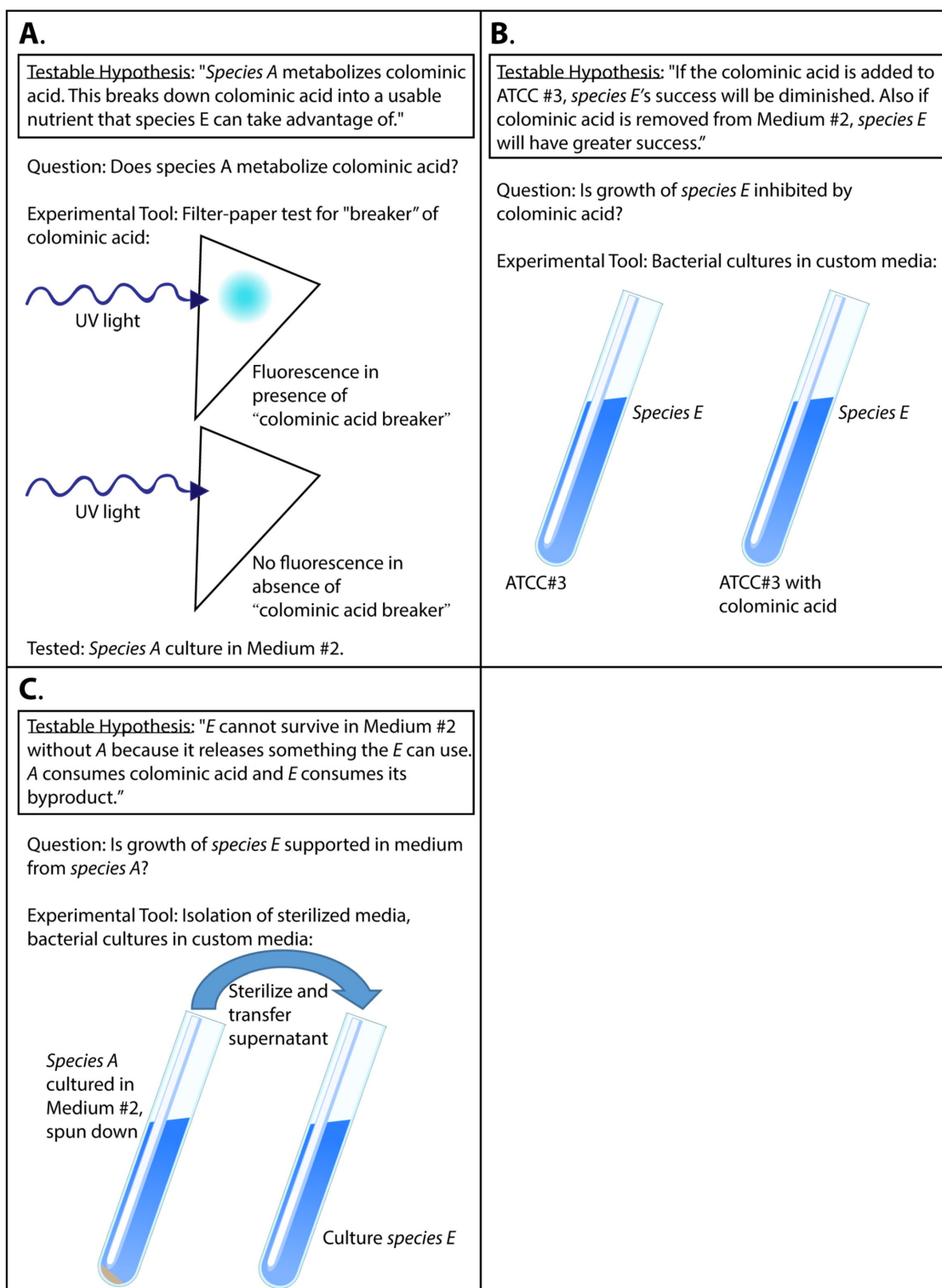
FIGURE 5. Combined groups of four to six students worked together to design and conduct experiments to test their hypotheses. Images show the overall concept of three groups' experiments. Additional conditions and controls were included by students but are not shown here. The experiments shown in A, B, and C correspond to the models drawn by students in Figure 3, A, B, and C, respectively.

cultured in medium #2 with CA. Thus, they decided to draw a model based on their understanding of the data collected by other groups—in this case, a version of the CA metabolism hypothesis (Figure 6B).

In her end-of-unit lab report, one of the students involved in drawing the models shown in Figure 3B and 6B described abandoning their original model when their experimental results were not consistent with their tested hypothesis:
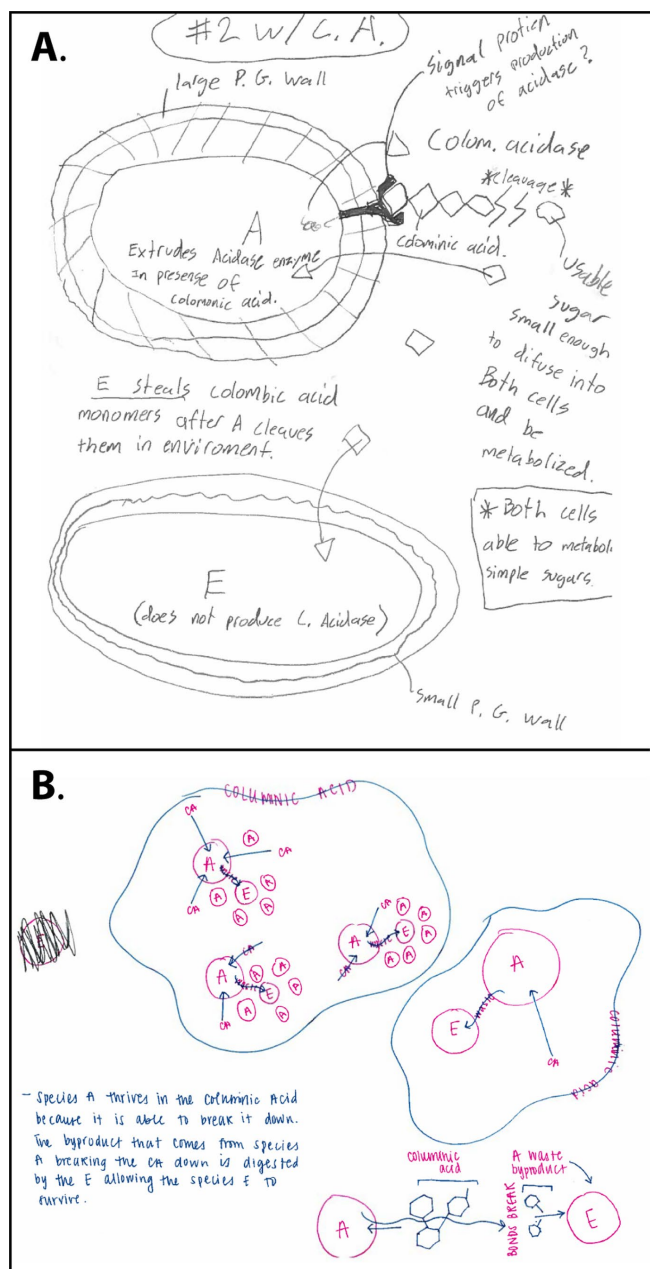
FIGURE 6. Groups of two or three students drew revised models after they completed their experiments and shared findings with other groups. The models shown in A and B were drawn by the same students who drew the models shown in Figure 3, A and B, respectively.

The refuting of our hypothesis caused us to have to drastically change our model from Week 4. Our original model included a drawing of the colominic acid breaking down the membrane of Species *E*, causing Species *E* to lose energy and carbohydrates. It also included a drawing of Species *A* somewhat enveloping Species *E* to add an additional layer of protection so that colominic acid would be unable to break down Species *E*, while Species *A* would also bring in additional carbohydrates and energy that would allow Species *E* to thrive. Our new model was completely different than our original and showed the catabolic process Species *A* performs in order to

allow Species *E* to survive in the acidic environment. This drawing shows Species *A* breaking down the colominic acid bonds, while Species *E* takes these bonds and consumes them in order to survive. Unlike the original model, our group completely dropped the idea that the colominic acid breaks down the bacteria's outer membrane and completely destroys Species *E*. Instead, we now show that Species *E* does not get completely destroyed, but is still able to survive [underlined emphasis added].

The student reports that their group's revised model was based primarily on experimental results shared by groups who conducted experiments different from their own:

After looking at other groups' data, our hypothesis was only proved more incorrect, and our findings during Week 6 were more supported….Species *A* is able to break down colominic acid and the byproduct that is created from the breakdown of the colominic acid is consumed or taken in by Species *E*, allowing *E* to survive. Essentially, the chemical structure of colominic acid is primarily held together by bonds that species *A* is able to catabolize. These broken bonds serve as energy for Species *E*, allowing Species *E* to survive. We were able to figure out about this catabolic reaction due to the findings of other groups [underlined emphasis added].

Revised models tended to have less diversity than the original models, in that students were moving toward models that were better supported by the collective data of the class. However, there remained significant diversity in the details of how the models were drawn and what details were included. More importantly, uncertainty remained among students about the details of the mechanism within the model and how they might be supported with further experimentation. As one student wrote in an end-of-unit reflection in class,

I would like to test how the enzymes break down CA. Whether it is partially digested outside and then completely digested inside or completely digested and then the monomers are broken down. This part of our model is still unclear. Also, I would like to test if *E* is metabolizing the broken down CA, in a way stealing from *A*'s food source. This is an aspect of another group's model that we have not considered [underlined emphasis added].

This remaining uncertainty helped underscore the nature of authentic science to students. Additionally, because the "correct" answer to the bacterial growth puzzle was never revealed by the instructors, students could retain agency as authentic investigators for the modeling cycles in subsequent units of the curriculum.

Example 2: Chlamydomonas Phototaxis. In the bacterial growth unit, generative reasoning was encouraged because the phenomenon presents a puzzle that can be "solved" through different hypothetical mechanisms. In another unit, students created and revised models to explain the mechanism for phototaxis (swimming toward light) for a unicellular algae *Chlamydomonas reinhardtii*. In this unit, the phenomenon is sufficiently complex and the experimental design task is relatively open, which encouraged students to explore different aspects of the phenomenon through a variety of experimental
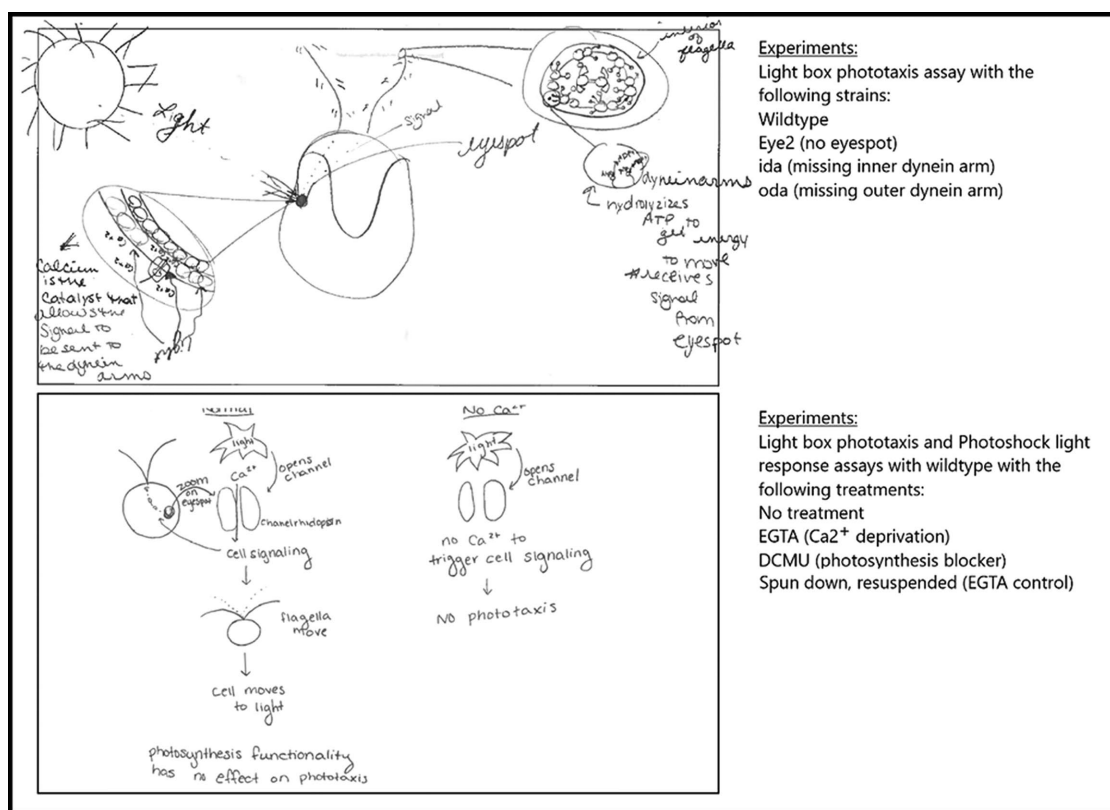
FIGURE 7. Groups of two or three students drew models to explain how *Chlamydomonas* performed phototaxis (swimming toward light). These final models reflected the results from experiments (outlined to the right) designed and performed by each group.

tools. For example, one group focused on testing the hypothesis that light-sensing structures (eyespots) and motility structures (flagella) were necessary for the phototaxis behavior (Figure 7, top). They chose to test their hypothesis by observing whether mutants that lacked functional eyespots and flagella were able to demonstrate phototaxis. Another group focused on the role of eyespot calcium channels in detecting light and enabling phototaxis. They hypothesized that movement of calcium ions into the cell when these channels open in response to light triggers a signal that causes flagella to move toward the detected light (Figure 7, bottom). This group also tested an additional hypothesis, that because phototaxis evolved as a way to optimize light exposure for photosynthesis, photosynthesis might be necessary for driving phototaxis. They tested these two hypotheses using pharmacological manipulation of the wild-type strain. They treated cells with either EGTA (ethylene glycol-bis(β-aminoethyl ether)-*N,N,N′,N′*-tetraacetic acid, a chelating agent that deprives cells of available calcium ions) or DCMU (3-(3,4-dichlorophenyl)-1,1-dimethylurea, an herbicide that, at low concentrations, blocks the light reactions of photosynthesis without immediately killing cells) and then assayed their ability to detect and respond to light and to demonstrate phototaxis. Although these two groups' models share similarities in the mechanisms they propose, they chose to explore different aspects of their models using very different experimental tools.

As was the case in the bacterial growth unit, groups' revised models often included more molecular mechanism than their initial models. For example, one of the students in the group that produced the model in the bottom panel of Figure 7 wrote the following in her lab report at the end of the unit:

> The most important feature the final model depicts are how the light source will excite and open channelrhodopsin ion channels in the eyespot, and how the influx of calcium results in cell signaling that guides the flagella in the direction of the light source. The model also makes the distinction that the light source may excite the channelrhodopsin channels to open, but without the calcium as a second messenger inside the cell no cell signaling can occur and the flagella will not move towards the light source and perform phototaxis. The model clearly supports our results since the EGTA Ca2+ deprivation treatment resulted in no phototaxis, but the DCMU treated cells did phototax since the necessary components were present (light and calcium).

Here, the student discusses the molecular details in her group's model. She goes on to discuss the inclusion of additional molecular detail in their revised model compared to their initial model:

> This final model is different from the original because it includes more details about the ion channel and the affect calcium has on cell signaling. The previous model simply displayed the cell moving towards light because the eyespot sensed it without any of these essential details.

Finally, she describes the limitations of the model—that it lacks a mechanism for how calcium signaling ultimately causes the flagella to move:

> The specifics of how calcium creates a cell signal that reaches the flagella are speculative, but it is known that calcium is the second messenger that begins the process.

This discussion demonstrates the student's awareness of the importance of specific explanatory molecular mechanisms to a satisfying model.

The *Chlamydomonas* phototaxis unit was similar to the bacterial growth unit in that students all began by observing and attempting to explain the same initial phenomenon. However, this later unit had a greater diversity of available tools and was more open-ended as to what questions the students might choose to investigate. This unit was near the end of the semester, and students seemed to take up the challenge of posing their own questions and hypotheses.

To What Extent Do Students Identify as Scientists and Take Agency of Their Own Ideas within an AIM-Bio Curriculum?

> So I think a large part of [feeling like a scientist] was having to draw out the models and, they say, "we're not going to tell you how, like, figure out, like, what can you do?" It needs to be something that's testable; it needs to be something that we have the resources and tools in order to actually conduct. I feel like that and having to think about it in our own way and draw conclusions and—like, we could have run an entire wrong experiment and they wouldn't have told us until, like, they would've been, like, "yeah, you probably shouldn't have done that," like, at the end. Like, that's so cool.

This quotation, and others like it, collected during end-of-semester interviews, suggest that experiencing a sense of agency or ownership over their own ideas and actions was an important force for motivating and empowering students to invest in the processes of forming and revising their models. On the basis of these data, as well as our experiences as instructors in the classroom, we predicted that AIM-Bio students would report having greater agency and ownership within their IMCB lab than students in the unrevised lab course. To test this prediction, we used the POS (Hanauer and Dolan, 2014). This survey was designed to detect differences in perceived ownership between students in traditional versus research-based laboratory experiences. We administered the survey in four contexts:

the unrevised and AIM-Bio university lab courses and the unrevised and AIM-Bio community college courses. As shown in Table 3, results for students in labs with AIM-Bio indicated higher rates of feelings of project ownership than those for students in unrevised labs at both institutions. Because the two contexts differed, we analyzed the data from the university and community college separately. Of the four categories identified in an exploratory factor analysis (described in the *Methods* section), university AIM-Bio students had more positive responses than university unrevised-section students in three categories: Personal Investment, Positive Emotions, and Self-Efficacy/ Unexpected Experiences. University students from the two curricula did not significantly differ from one another in the category of Real-World Contributions. These results are consistent with the goals and design of the AIM-Bio curriculum. Students had greater authorship of and responsibility for their experiences in the AIM-Bio sections. Neither the university AIM-Bio nor the unrevised curriculum, however, involved generating data or analyses for a publishable research project. At the community college, responses from students in the AIM-Bio and unrevised sections of the course differed in all four categories. Although it is impossible to disentangle the relative impacts of the AIM-Bio and CURE units on community college students' responses, it is plausible that the AIM-Bio units impacted students similarly in the university and community college settings and that the CURE at the community college strongly impacted students' responses in the Real-World Contributions category.

Student interview data also suggested that experiencing agency within the AIM-Bio curriculum was associated with students identifying as scientists. To further investigate this connection, we administered a written survey that included the prompt, "Have you ever felt like a scientist? (Yes/No) Please explain." We gave the survey to university students in the AIM-Bio and unrevised laboratory courses at the end of the semester. We coded each student's response as falling into one of three categories: "No," "Yes (does not specify IMCB)," and "Yes (specifies IMCB)." Figure 8 shows the comparison between responses from AIM-Bio students and unrevised-curriculum students at the university. The AIM-Bio curriculum appears to have a greater positive effect on students' identities as scientists. Ninety-five percent from the AIM-Bio sections responded "Yes," compared with 79% from the unrevised curriculum. A large percentage of students in both groups answered "Yes" without specifically stating that they felt this way in the biology lab course (IMCB). Responses from these students described in general what made them feel like a scientist or referred to an experience in another laboratory course, in an authentic research lab, or in everyday

TABLE 3. Results of the Project Ownership Survey shown for the entire instrument and four categories determined with an exploratory factor analysis using maximum-likelihood extraction and oblimim with Kaiser normalization rotation[a]

	Unrevised university	AIM-Bio university	Unrevised community college	AIM-Bio community college
Project Ownership (all items)	2.77	2.40***	3.04	2.21***
Personal Investment	2.56	2.13***	2.58	1.77**
Real-World Contributions	2.80	2.67	3.69	2.50***
Positive Emotions	3.12	2.51***	2.79	2.08*
Self-Efficacy and Unexpected Experiences	2.71	2.34***	3.04	2.40**

[a]Asterisks indicate statistically significant differences between the AIM-Bio and unrevised sections at each institution: *$p \leq 0.01$; **$p \leq 0.001$; ***$p < 0.0001$. p values were calculated with Welch's two-sample t test. Note that the responses were on a scale from 1 (strongly agree) to 5 (strongly disagree), with agreement signaling stronger feelings of positivity.
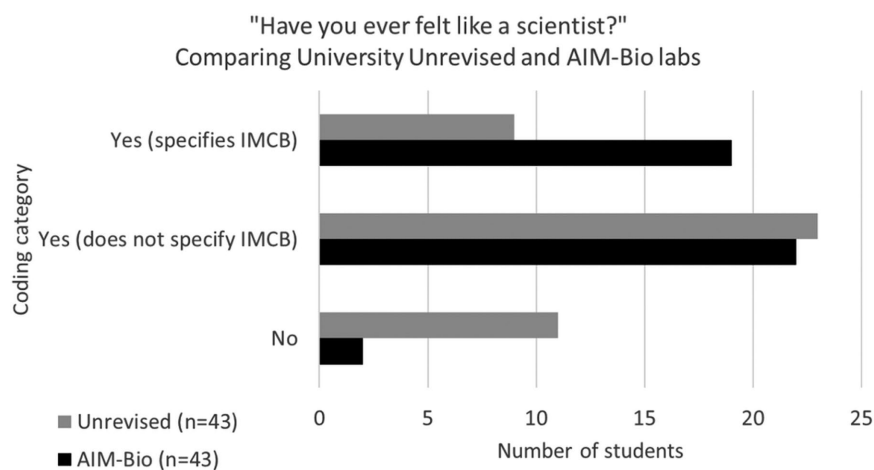
"Have you ever felt like a scientist?"
Comparing University Unrevised and AIM-Bio labs

FIGURE 8. Students responded to the prompt "Have you ever felt like a scientist? (Yes/No) Please explain" on a written survey. Students' responses were coded for whether they said "yes" or "no." "Yes" responses were further coded for whether they specifically indicated that they had felt like a scientist in their IMCB laboratory course.

life. However, a significantly greater number of students within the AIM-Bio curriculum (44%) specifically reported having felt like scientists in IMCB compared with students within the unrevised curriculum (21%).

We were also interested in what made students feel like scientists. This reflects the students' views of what it means to participate in science and illuminates the curricular elements that supported students' identities as scientists in this context. We coded students' "Yes" responses for criteria for feeling like a scientist. Coded categories, examples of students' responses placed in these categories, and the number of students coded for each category are shown in Table 4. Interestingly, students' scientific identities were supported by both hands-on laboratory tech-

niques (pipetting, mixing chemicals, etc.) and more cognitively demanding tasks (testing hypotheses, designing experiments, drawing conclusions from results). Many students also reported autonomy/ ownership as a reason why they had felt like scientists. Consistent with our interpretation of the POS results, community college students specifically mentioned both the AIM-Bio units and the CURE unit when describing why they felt like scientists.

Does an AIM-Bio Curriculum Impact Students' Views of the Nature of Science and/or Scientific Skills?

Next, we wanted to measure how students' understanding of the nature of science and their skills for doing science changed over the course of the semester. We hypothesized that, through experiencing authentic scientific practices in the AIM-Bio curriculum, students would develop a more expert-like view of the scientific process as well as enhanced ability to solve scientific problems. To test this, we administered two pre- and postsemester assessments in the university AIM-Bio course: a written Nature of Science survey and the Classroom Test of Scientific Reasoning. A written survey was adapted from our pre/ post interview protocol based on analysis of those interviews (as described in the *Methods* section). Three of the prompts on the survey were designed to measure elements of students' understanding of the role of experimentation in science, experimental design, and what makes a theory. We referred to this part of the written survey as the Nature of Science survey. The adapted Classroom Test of Scientific Reasoning is intended to

TABLE 4. Two researchers independently applied coding categories to student responses for students who were coded as having felt like scientists (the coding scheme is available in the Supplemental Material)

Category	Representative student response	Number of coded responses (N = 104)[a]
Hands-on techniques	"When constructing procedures and running experiments (particularly on live organisms), it is very easy to feel 'like a scientist.' The lab coat helps, too!"	29
Hypothesis testing	"[I have felt like a scientist] because in this lab I have made hypotheses and created experiments to test them out."	20
Autonomy/ownership	"The biggest part was the way we had free access to use materials and practice on our own."	19
Experimental design/drawing conclusions	"Yes, I designed a lot of experiments in this lab. I constantly felt like I didn't have all the answers. Scientists probably feel like this often." "I feel like a scientist when I can come to my own conclusions based off an experiment I can see results from."	15
Critical thinking/problem solving	"Scientists solve everyday problems and I believe we all have the ability to do so!" "I feel like the more involved I get in science, the more I begin to think like a scientist. That is, more logically."	14
Content/understanding	"Coming to an understanding of science in a lab setting helps me feel like I'm in touch with how science works."	9
Explanation building	"When analyzing the results of our bacterial culture lab … I was close to the answer to make my model work all I needed was the results of another group in the class in a particular test … it was a eureka moment."	7
Asking questions	"A scientist is someone who is curious about how or why something works."	5

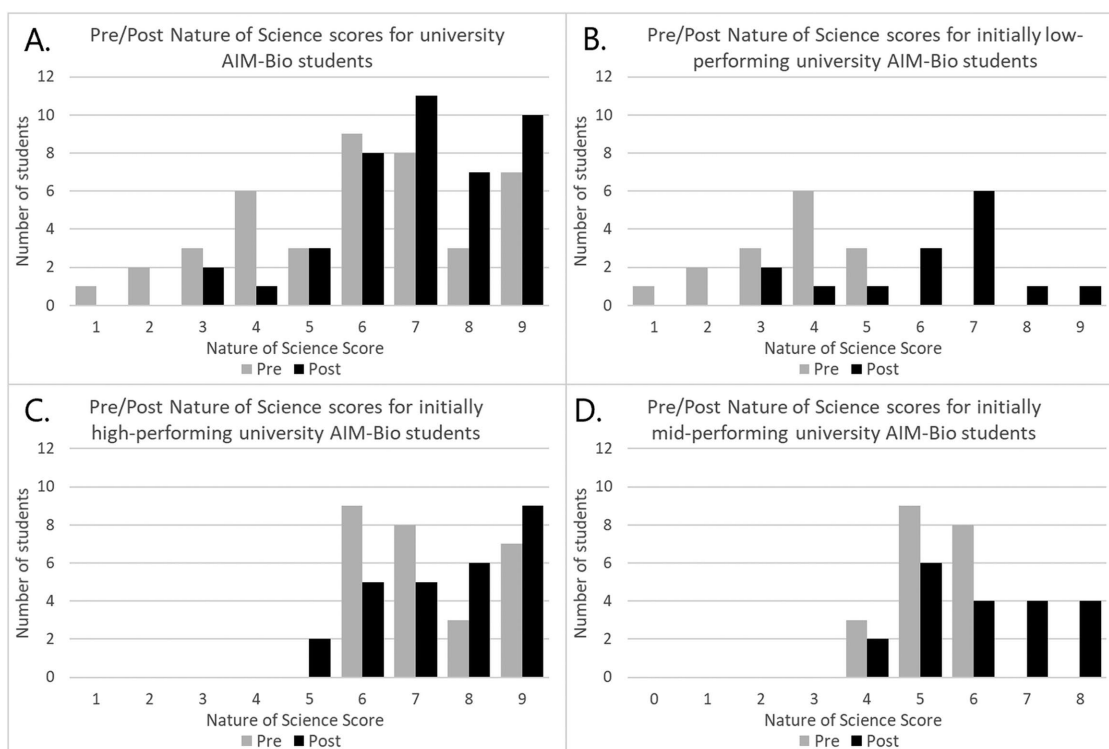[a]Each student response could be assigned multiple reason codes.

FIGURE 9. (A) We observed a positive shift in the distribution of Nature of Science scores for university AIM-Bio students. The difference between the pre and post score distributions was confirmed with a Wilcoxon signed-rank test ($p = 0.00044$). To better understand how this shift is distributed in the population, we also visualized the pre/post distributions for (B) initially low-performing students (those scoring between 0 and 4 on the pre survey, 36% of students) and (C) the remaining, initially relatively high-performing students (those scoring between 5 and 8 on the pre survey, 64% of students). (D) We also visualized the pre/post distributions for the middle-performing roughly half of students (those scoring between 4 and 6 on the pre survey, 48% of students).

measure students' general skills for doing science (Benford and Lawson, 2001, adapted from Lawson, 1978).

We predicted that individual students' views of science may move along a gradient of increasing sophistication over the semester. To test this prediction, we assigned pre/post Nature of Science scores to each student according to his or her responses on the pre/post Nature of Science survey (details of this analysis are described in the *Methods* section). Possible scores range from 0 to 9, reflecting levels of sophistication of nature of science understanding evident in the survey responses. As predicted, we observed a positive shift in the distribution of Nature of Science scores from pre to post (Figure 9A). The difference between the pre and post score distributions was confirmed with a Wilcoxon signed-rank test ($p = 0.00044$). There was a positive shift in the Nature of Science score for 60% (25/42) of students, and a negative shift for only 12% (5/42) of students. To further analyze how these shifts were distributed in the population, we separated students into those who initially performed in roughly the bottom third (those who scored between 0 and 4, 36% of students; Figure 9B) from those who were higher scoring (those who scored higher than 4, 64% of students; Figure 9C). Students with relatively high initial scores experienced very modest gains, if any. This is possibly due to a "ceiling effect" on the ability of high scores to increase. There is, however, a clear and substantial shift toward higher scores for those with low initial scores. Because we were interested in the shift likely to be experienced by a "typical" student, we also looked at the pre/post

distributions for students who initially scored in roughly the middle half (the 48% of students with an initial score between 4 and 6; Figure 9D). These students' scores also shift toward higher scores, though the change is less marked than for those with the lowest initial scores. These results support the view that, although there was much variation in students' views of the nature of science both before and after the course, individual students shifted toward having more sophisticated views of the nature of science over the course of the semester.

The adapted Classroom Test of Scientific Reasoning (Benford and Lawson, 2001) is a validated 24-item, multiple-choice assessment intended to measure students' scientific reasoning abilities outside any particular disciplinary context. We administered the assessment online at the beginning and end of the Fall 2017 semester to students in AIM-Bio laboratory sections. We saw a statistically significant increase in the mean AIM-Bio score from pre to post, from 66% correct pre to 72% post (confirmed with a Welch's two-tailed, paired t test, $p = 0.032$). Most of these gains were concentrated on seven questions (questions 5, 7, 8, 9, 10, 21, and 22) assessing proportional reasoning and drawing inferences from data. These gains are consistent with skills practiced by students in the AIM-Bio laboratory course.

DISCUSSION

In this paper, we present data in the context of a novel, model based–inquiry curriculum for an undergraduate biology

laboratory course. Analysis of classroom artifacts and student interviews suggested that this curriculum prompted students to generate diverse hypotheses and models. In addition, the curriculum required students to test their hypotheses and revise their models in response to experimental data. In this way, the theoretical aspects of scientific practice (hypothesizing, modeling, explanation construction) were juxtaposed with more empirical practices (experimental design, data collection, data interpretation). Moreover, evidence supports that experiencing a level of scientific freedom was important for students. Both interviews and surveys indicated that students felt an enhanced sense of agency and identity as a scientist within the AIM-Bio curriculum as compared with the unrevised curriculum for the course. Finally, written survey and assessment data indicate that students increased in their understanding of the nature of science and their general scientific skills. In our view, inclusion (and integration) of both theoretical and empirical practices was important in making the curriculum an opportunity for students to learn about authentic scientific inquiry.

AIM-Bio Supports Integration of Science Content and Scientific Practices

The curricular approach we describe here is designed to move students away from learning explanations, which we see as an approach that focuses primarily on science content. Instead students are asked to build scientific explanations, an approach that requires integration of science content and scientific practices. This form of integration brings authenticity to science learning (Osborne, 2014; Passmore *et al.*, 2014). Further, we argue that learning science ideas in the context of doing science is likely to positively affect the learning process.

The AIM-Bio laboratory curriculum is based on learning objectives that integrate biological concepts with science practices. For example, one learning objective from the bacterial growth unit is, "Apply ideas about metabolism and enzymes to construct explanations for bacterial growth." To meet this objective, students must use the scientific practice of explanation construction, but they also must use or learn specific biological concepts. Students combine their prior knowledge with ideas they are learning in the lecture and laboratory courses to construct and test explanations for the target phenomenon and thus further their understanding of bacterial growth. This integrated instructional approach requires that instructors provide scientific information (i.e., content) to students to facilitate their development of explanations for target phenomena. In the AIM-Bio curriculum, content was shared with students through whole-class instruction and readings as well as individual instruction that was required as students explored diverse ideas and explanations.

Asking AIM-Bio students to construct explanations situates their learning in the context of a problem. They may solve this problem through use of authentic scientific practices such as modeling, experimental design, and data interpretation. We did not directly compare how well students learned biological concepts through AIM-Bio versus the unrevised laboratory curriculum. However, previous work suggests that instructional practices that require students to construct meaning within the context of a problem or investigation (specifically, problem-based and inquiry approaches) can have a positive impact on learning (Hmelo-Silver *et al.*, 2007). For example, a well-controlled study within the context of a master's of business administration course directly compared learning through a traditional or problem-based approach and found a positive effect of the latter on students' ability to produce their own explanations of concepts during the end-of-semester exam (Capon and Kuhn, 2004). Although learning scientists have not yet established why learning within the context of a problem is beneficial, there are some hypotheses to suggest why this may be the case. First, when asked to solve a problem, students are likely to activate their prior knowledge and connect this knowledge to the new ideas they are learning. This could increase the number of connections to the learned concept and impact future retrieval (Schwartz and Bransford, 1998; Capon and Kuhn, 2004). In addition, there is evidence to suggest that the process of constructing one's own explanation (which can be viewed as a "problem" in a science classroom) has a positive impact on learning (Chi *et al.*, 1989, 1994).

Models Act as Powerful Thinking Tools in Science and in the Classroom

A key aspect of our instructional design is the focus on modeling as an organizing theme. In the AIM-Bio curriculum, model drawings play a role as representations of the current student explanations present in the classroom. The process of making model drawings encouraged students to discuss their ideas within their groups and commit to particular explanatory elements. These drawings made student thinking visible to instructors as well as other students. This focus on models drew explicit attention to students' cognition about relevant biological concepts. Importantly, the AIM-Bio curriculum treats models as flexible, able to change in response to current experimental data or new ideas. In addition to prompting students to draw and refine models, we asked them to reflect during individual lab reports on their process of model change. This pushed students to be metacognitive about their own scientific process.

The way that we used models in the AIM-Bio curriculum assumed that students would come to use models to support generative reasoning in the way that scientists do. Specifically, models have been shown to allow simulative and causal reasoning when thinking about scientific explanations (Nersessian, 2008). Further, models serve as the central organizing feature for the scientific process of making and testing predictions and constructing explanations from scientific data (Passsmore *et al.,* 2009). We found that model drawings did serve students in these ways within the AIM-Bio curriculum, but not without prompting and scaffolding by instructors. These scaffolds included encouraging students to include visual representations and icons, as well as tentative ideas and hypotheses, in their model drawings. The need for such scaffolding is consistent with prior research demonstrating the high cognitive demand of converting verbal information to visual (Van Meter *et al.*, 2006) and previous work suggesting developing expertise in modeling includes learning to view models as flexible representations that may change as new data become available (Schwartz *et al.*, 2009; Quillin and Thomas, 2015). Previous instructional design studies supporting MBI among middle school students also underscored the need for instructors to provide scaffolding for students in using evidence to evaluate scientific models (Rinehart *et al.*, 2014). We chose student lab

reports as the place to provide instructional scaffolds and feedback to support development of this skill.

AIM-Bio Supports Student Agency and Self-Efficacy

A central design goal in building our curriculum was to engage students in authentic scientific inquiry. This supported students in developing a sense of agency in the classroom, as it gave them control over many of their physical experiences in the classroom (e.g., which experiments they performed and how) and their cognitive engagement (e.g., how they built explanations of biological phenomena). This sense of agency played an important role in supporting students' identities as scientists. Scientific inquiry is well suited to allow a high level of physical and cognitive agency, as it assumes that different students will pursue different possible explanations (cognitive agency) through different means (physical agency). Importantly, we found that students' sense of control over their physical actions encouraged them to take more control and invest more in their explanations—having the power to determine how they were going to explore their ideas encouraged them to invest in those ideas. Moreover, the aim of the inquiry was to arrive at a biologically defensible explanation that was supported by the available data, rather than a single, predetermined (and likely pre-known) answer. We believe that this emphasis on *building* explanations, rather than learning explanations that were presented to them, also contributed to students' sense of cognitive agency.

The development of a sense of agency is tied to one's sense of self-efficacy in a particular domain. Given that self-efficacy is not a general trait of an individual—it varies from field to field, and even from task to task (Bandura, 2012)—encouraging students' self-efficacy in science courses is important. The literature describing and investigating self-efficacy in general and in educational settings has identified "enactive mastery experiences" as the primary way in which people build a strong and enduring sense of self-efficacy (e.g., Bandura, 1997; Palmer, 2006; van Dinther *et al.*, 2011). Characteristics of these experiences are that they are perceived as authentic, challenging, and successful after a period of struggle (Bandura, 1997). These characteristics align well with the goals and design principles that we used to design the AIM-Bio curriculum: we engaged students in authentic scientific reasoning tasks, paying close attention to the task difficulty level and providing scaffolds to support student success without infringing on their autonomy. It is relevant to note that in studies identifying factors that contribute to students' self-efficacy in higher education settings, practical experience alone—even when the students are successful—has not been found to be sufficient to increase students' self-efficacy (van Dinther *et al.*, 2011). Instead, grappling with and ultimately succeeding at a difficult task that does not initially seem within an individual's ability to accomplish sends a signal to that individual about his or her ability to productively work through challenges in that domain.

Given the emphasis on students' agency in AIM-Bio, it is important to note that the AIM-Bio classroom is not a free-for-all; there are many instructional scaffolds to support students' success (as is the case for effective inquiry instruction more generally; e.g., Hmelo-Silver *et al.*, 2007). We compare our roles as instructors of an AIM-Bio laboratory course to that of a faculty mentor for an undergraduate researcher working on his or her own project: the students are given the role as the drivers of the work, but we are there to provide resources, constructive feedback, and encouragement. Moreover, it is not the case that, in an AIM-Bio classroom, all ideas are equal. Although students are free to speculate widely, they are also responsible for creating explanations that are consistent with biological principles and currently available data. At the end of the day, students are held to account for the reasonableness of their explanations. A key difference between an AIM-Bio and a more traditional laboratory classroom, however, are the sources of accountability. In a traditional laboratory class, students are held to account for a specific, provided explanation—if they repeat the explanation imperfectly, they receive feedback from the instructor to that effect. In an AIM-Bio classroom, on the other hand, students' ideas receive pushback from the physical world—their explanations either are or are not consistent with data about how the system behaves.

AIM-Bio in the Context of Undergraduate Laboratory Course Reform

As the spaces where students encounter some of the tools and procedures for doing science and engage in hands-on activities, undergraduate laboratory courses offer an opportunity to introduce students to what it means to do science. It is broadly recognized, however, that traditional "cookbook" laboratory courses fail to engage students in authentic scientific experiences (e.g., Germann *et al.*, 1996; Brownell *et al.*, 2012; President's Council of Advisors on Science and Technology, 2012). By focusing on performing confirmatory experiments, such courses do not expose students to the roles that community, creativity, and uncertainty play in science, or even to the central purpose of science—to generate and vet novel ideas about how the world works. Failing to provide students with authentic experiences of science has a variety of negative consequences. For one, students do not get opportunities to develop the cognitive skills necessary for doing science in an authentic environment—for example, it is a very different experience to analyze data in light of a flexible model than to do so in the context of confirming a known principle. For another, by propagating a false view of what is involved in "doing science," confirmatory laboratory activities may alienate students who would thrive in a more authentic setting and discourage them from persisting in science programs (Graham *et al.*, 2013).

For these reasons, undergraduate laboratory courses have long been the focus of reform efforts. Many of the curricula developed through these efforts have been characterized as "inquiry." There is great diversity, however, among curricula labeled with this term. Typically, inquiry instruction involves students investigating a question with an unknown (to the students) answer by engaging in some subset of the following activities: making observations, asking questions, proposing explanations, designing experiments, conducting experiments, analyzing data, and revising explanations in light of data (NRC, 2000; Wallace *et al.*, 2003; Weaver *et al.*, 2008, Alkaher and Dolan, 2011). A given inquiry curriculum may give students primary control over more or fewer of these elements. Thus, inquiry instruction falls along a continuum from more instructor-guided to more open-ended (Windschitl, 2002; Bell *et al.*, 2005; Weaver *et al.*, 2008). Although some have conflated

inquiry and unscaffolded "discovery learning" (Kirschner *et al.*, 2006), even very open-ended inquiry is often highly scaffolded to support student success (e.g., Hmelo-Silver *et al.*, 2007). In the AIM-Bio curriculum described here, we began with a unit (membrane transport) that was heavily guided by the instructors in order to give students experience with the novel practices of building, discussing, and revising models and to introduce students to the idea that even models in which they have high confidence must be flexible in light of data that challenge those models. Throughout the semester, we moved toward more open-inquiry units, giving students greater control over aspects of the inquiry and fading the scaffolding for particular tasks as students gained competence and confidence. By the time students encountered the *Chlamydomonas* phototaxis unit, in which they took primary control over all aspects of the inquiry, they had received significant support, instruction, and feedback on how to conduct a productive investigation.

Lately, there has been much focus on CUREs as a way to address the shortcomings of traditional laboratory instruction. These courses aim to scale up the experience of a mentored research internship to involve an entire class of students in ongoing scientific research (Auchincloss *et al.*, 2014). Characteristics of a CURE, identified by a working group of the Course-Based Undergraduate Research Experiences Network (CUREnet), are use of multiple scientific practices, discovery (in that it is unknown to both the students and the instructor what they will find), broadly relevant or important work, collaboration, and iteration (Auchincloss *et al.* 2014). Those researching CUREs have reported positive outcomes that address many of the criticisms of traditional laboratory curricula. These include increased content knowledge, skills for doing science, self-efficacy, persistence in science, increased project ownership, and increased science identity (reviewed in Corwin *et al.*, 2015).

The AIM-Bio curriculum presented here has characteristics and outcomes in common with CUREs. In both CUREs and the AIM-Bio curriculum, students engage in multiple scientific practices beyond those typically exercised in traditional laboratory courses; the outcomes of students' investigations are unknown to the students; students collaborate with one another as peers and with the instructor as a mentor; and iteration is built into the process. AIM-Bio students experienced a sense of project ownership similar to that reported for CURE students (Hanauer and Dolan, 2014). This is consistent with the recent findings by Corwin *et al.* (2018) that collaboration and iteration contribute significantly to students' feelings of ownership in a course and by Ballen *et al.* (2018) that students in a nonmajors inquiry course experienced a sense of ownership similar to students in a comparable CURE course despite the absence of novel discovery and broad relevance. Both CUREs and the AIM-Bio curriculum have resulted in increases in students' skills for doing science. Because assessment measures differ between studies and often rely on self-reporting by students or faculty, however, it is difficult to compare the relative impact of CUREs and the AIM-BIO curriculum on specific scientific skills. It is likely that students in the different curricula develop different but overlapping skill sets, depending on the emphasis of each course. Similarly, students in both CUREs and the AIM-Bio curriculum have experienced increases in their understanding of the nature of science, but these results cannot be directly compared, as they were assessed differently. We have not directly assessed the impact of the AIM-Bio curriculum on increases in students' conceptual knowledge or student persistence in STEM programs—these are potential avenues for future research.

Key differences between CUREs and the AIM-Bio curriculum stem from the fact that, in CUREs, students are investigating questions with answers unknown by and relevant to the broader scientific community, whereas in the AIM-Bio curriculum, the answers are known to the community and the instructors, though unknown to the students. Each has its advantages. Because CUREs involve students in addressing questions that are relevant to the broader community, they promote students' sense that they are making a meaningful contribution and potentially provide avenues for students' ongoing involvement in the greater community. On the other hand, the goal of generating data that are useful to the broader scientific community may constrain the types of data collected and therefore encourage greater designer/instructor control over, for example, which questions are asked and how experiments are designed in order to productively address those questions. When evidence is being collected strictly for classroom use, there is little opportunity cost to giving students greater control over which questions are asked, which data are generated, and how those data are generated. By their nature, CUREs require significant ongoing input by research faculty to remain current and relevant. Those with a large benchwork component are also likely to require more state-of-the-art equipment and expensive reagents than are typically used in classroom settings. Where resources permit, these are advantages. They can, however, represent significant hurdles for implementation and maintenance of CURE curricula at institutions with fewer resources and/or large numbers of students, particularly at the introductory level. An advantage of the AIM-Bio curriculum is that it was deliberately designed to be relatively inexpensive to initiate and run, while still incorporating authentic benchwork skills.

Challenges and Limitations

There are several challenges to designing an AIM-Bio–style curriculum. As described in the *Instructional Design* section, the choice of phenomena is central to its success. Crucially, a target phenomenon for an AIM-Bio unit must have multiple biologically plausible potential explanations that are accessible to the students. Moreover, tools must be available to test the variety of hypotheses that students generate. Because traditional lab curricula tend to rely on confirmation experiments with results that are (by design) already known to the students, they are not suitable for "retrofitting" with the AIM-Bio model—a developer must identify and develop curriculum around new phenomena and tools, many of which require adaptation and troubleshooting. Identifying appropriate phenomena and developing them for laboratory classroom use is thus no easy task. Another challenge, typical for implementation of an inquiry curriculum, is deciding when to include curricular scaffolds to provide students with the resources to productively engage without sacrificing the mystery motivating that engagement. Finally, most institutions have limited resources to devote to undergraduate labs. This places some limits on the techniques available for testing students' ideas, which constrains the target phenomena and tools that may be included in the curriculum.

Implementing an AIM-Bio curriculum also presents challenges. Instructors as well as students must shift their focus

from students' understanding of provided explanations to students' development of the skills necessary for building explanations. This is bound to place many instructors and students outside their comfort zones. From an instructor's perspective, teaching an AIM-Bio course requires constraint—resisting the temptation to "nudge" students toward correct ideas in favor of allowing students to develop their own ideas. Providing feedback also becomes more challenging. The goal is not to guide students toward an understanding of a particular explanation, but rather to challenge students—and encourage them to challenge one another—to defend their ideas in light of biological principles and available data. For students to benefit from this shift in instructional focus, a supportive classroom culture must be established: students must feel safe presenting ideas but accepting of the reality that not all ideas will survive the process. In an effort to establish such an environment, instructors in our implementation demonstrated enthusiasm and curiosity about students' ideas. At the same time, we worked with students to help them think through and develop their ideas and experimental approaches. Given these supports, the challenges of grappling with the unknown and generating their own answers, rather than having answers provided them, supported students in feeling like scientists (e.g., as seen in the quotes under "Experimental Design/Drawing Conclusions" in Table 4).

Limitations of our data collection prevent us from addressing certain questions. For logistical reasons and to avoid "survey fatigue," we did not administer the Classroom Test of Scientific Reasoning in unrevised sections of the university IMCB laboratory course. Thus, we cannot make comparisons between the unrevised and AIM-Bio courses for this instrument. Therefore, we cannot rule out the possibility that factors other than the AIM-Bio curriculum were responsible for student gains on the instrument. It has been demonstrated elsewhere, however, that students in traditional educational settings do not typically demonstrate statistically significant improvements on this and similar instruments (Johnson and Lawson, 1998; Benford and Lawson, 2001). At the community college and in the unrevised university lab course, the first two class periods focused on the nature of science and the scientific method. Because this occurred immediately before pre-instruction data collection, we did not deem it productive to administer the written Nature of Science survey to students at the community college or in the unrevised university course at that time. Thus, we are unable to make a pre/post comparison of these students' understanding of the nature of science in the same way that we did for the university AIM-Bio students. Finally, we did not administer a pre/post biology concept assessment, so we are currently unable to compare students' biology concept learning gains between unrevised and AIM-Bio sections of the course.

Because the community college course simultaneously incorporated both AIM-Bio and CURE units, we are unable to separate the impacts of the two. We can, however, report that community college students, as well as university students, productively engaged in generating, testing, and revising a diversity of explanatory models in the AIM-Bio units. Moreover, our POS results and community college students' responses to the written question "Have you ever felt like a scientist? (Yes/No) Please explain" are consistent with the

interpretation that the CURE unit primarily impacted students' sense of having made a real-world contribution, whereas the AIM-Bio units supported other aspects of autonomy and ownership.

We are currently limited in our ability to describe increases in students' skills for doing science. Although we did see an encouraging gain on the adapted Classroom Test of Scientific Reasoning, there is imperfect alignment between the instrument and the skills primarily targeted by the AIM-Bio curriculum. Students showed gains for items that most closely aligned with skills practiced in the class; more complex skills involved in explanation building, however, are not assessed by the instrument. Such skills are difficult to assess, particularly in a pre/post manner. Engaging in explanation building is a highly contextualized process: it is difficult to assess a student's explanation-building skills outside their content knowledge and conceptual understanding relevant to particular biological phenomena. This presents a challenge for measuring students' pre-instruction skills or making comparisons between different instructional contexts.

Future Directions

Our ongoing research is focusing on the mechanisms that support learning in the MBI curriculum. We are analyzing classroom audio recordings and artifacts to better understand the ways in which students develop the scientific practices of modeling, experimental design, and data interpretation. We are also investigating further potential links between students' scientific agency and their development and exercise of these practices. Furthermore, we are working to describe students' skill development through qualitative analysis of pre/post interviews in which students engage in think-aloud modeling and hypothesis-testing tasks. Future research will explore how best to scale up and disseminate this AIM-Bio curriculum for implementation in a variety of settings.

REFERENCES

Alkaher, I., & Dolan, E. (2011). Instructors' decisions that integrate inquiry teaching into undergraduate courses: How do I make this fit? *International Journal for the Scholarship of Teaching and Learning*, *5*(2), art9.

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action.* Washington, DC.

Auchincloss, L. C., Laursen, S. L., Branchaw, J., 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*, 29–40.

Ballen, C. J., Thompson, S. K., Blum, J. E., Newstrom, N. P., & Cotner, S. (2018). Discovery and broad relevance may be insignificant components of course-based undergraduate research experiences (CUREs) for non-biology majors. *Journal of Microbiology & Biology Education*, *19*(2), 19.2.63. https://doi.org/10.1128/jmbe.v19i2.1515

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: WH Freeman.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, *38*, 9–44.

Bangera, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE—Life Sciences Education*, *13*(4), 602–606.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.

Baze, C. L., & Gray, R. (2018). Modeling Tiktaalik: Using a model-based inquiry approach to engage community college students in the practices of science during an evolution unit. *Journal of College Science Teaching*, *47*(4).

Bell, R. L., Smetana, L., & Binns, I. (2005). Simplifying inquiry instruction. *Science Teacher*, *72*(7), 30–33.

Benford, R., & Lawson, A. E. (2001). Relationships between effective inquiry use and the development of scientific reasoning skills in college biology labs. Educational Resources Information Center (ERIC) accession no. ED456157.

Bierema, A. M. K., Schwarz, C. V., & Stoltzfus, J. R. (2017). Engaging undergraduate biology students in scientific modeling: Analysis of group interactions, sense-making, and justification. *CBE—Life Sciences Education*, *16*(4), ar68.

Brewe, E. (2008). Modeling theory applied: Modeling instruction in introductory physics. *American Journal of Physics*, *76*(12), 1155–1160.

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*(4), 36.

Capon, N., & Kuhn, D. (2004). What's so good about problem-based learning? *Cognition and Instruction*, *22*(1), 61–79.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182.

Chi, M. T., De Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439–477.

Clement, J. (2008). The role of explanatory models in teaching for conceptual change. In Vosniadou, S. (Ed.), *International handbook of research on conceptual change* (pp. 417–452). New York: Routledge.

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.

Corwin, L. A., Runyon, C. R., Ghanem, E., Sandy, M., Clark, G., Palmer, G. C., ... Dolan, E. L. (2018). Effects of discovery, iteration, and collaboration in laboratory courses on undergraduates' research career intentions fully mediated by student ownership. *CBE—Life Sciences Education*, *17*(2), ar20. doi: 10.1187/cbe.17-07-0141

Craik, K. (1967). *The nature of explanation*. New York: Cambridge University Press.

Darden, L. (2002). Strategies for discovering mechanisms: Schema instantiation, modular subassembly, forward/backward chaining. *Philosophy of Science*, *69*(S3), S354–S365.

Dunbar, K. (1999). How scientists build models: In vivo science as a window on the scientific mind. In Magnani, L., Nersessian, N., & Thagard, P. (Eds.), *Model-based reasoning in scientific discovery* (p. 89). New York: Plenum Press.

Germann, P. J., Haskins, S., & Auls, S. (1996). Analysis of nine high school biology laboratory manuals: Promoting scientific inquiry. *Journal of Research in Science Teaching*, *33*, 475–499. doi: 10.1002/(SICI)1098-2736(199605)33:5<475::AID-TEA2>3.0.CO;2-O

Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A., & Handelsman, J. (2013). Increasing persistence of college students in STEM. *Science*, *341*(6153), 1455–1456. doi: 10.1126/science.1240487

Hanauer, D. I., & Dolan, E. L. (2014). The Project Ownership Survey: Measuring differences in scientific inquiry experiences. *CBE—Life Sciences Education*, *13*(1), 149–158.

Harrison, M., Dunbar, D., Ratmansky, L., Boyd, K., & Lopatto, D. (2011). Classroom-based science research at the introductory level: Changes in career choices and attitude. *CBE—Life Sciences Education*, *10*(3), 279–286.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, *8*(6), 280–285.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107.

Ippolito, M. F., & Tweney, R. D. (1995). *The inception of insight*. In Sternberg, R. J., & Davidson, J. E. (Eds.), *The nature of insight* (pp. 433–462). Cambridge, MA: The MIT Press.

Johnson, M. A., & Lawson, A. E. (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes? *Journal of Research in Science Teaching*, *35*(1), 89–103.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Jones, M. T., Barlow, A. E., & Villarejo, M. (2010). Importance of undergraduate research for minority persistence and achievement in biology. *Journal of Higher Education*, *81*(1), 82–115.

Khan, S. (2007). Model-based inquiries in chemistry. *Science Education*, *91*(6), 877–905.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. doi: 10.1207/s15326985ep4102_1

Lawson, A. E. (1978). he development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, *15*(1), 11–24.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, *39*(6), 497–521.

Lehrer, R., & Romberg, T. (1996). Exploring children's data modeling. *Cognition and Instruction*, *14*(1), 69–108.

Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. In Sawyer, R. K. (Ed.), *The Cambridge handbook of: The learning sciences* (pp. 371–387). New York: Cambridge University Press.

Lopatto, D. (2007). Undergraduate research experiences support science career decisions and active learning. *CBE—Life Sciences Education*, *6*(4), 297–306.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*(1), 1–25.

McComas, W. F., Clough, M. P., & Almazroa, H. (1998). The role and character of the nature of science in science education. In McComas, W. F. (Ed.), *The nature of science in science education* (pp. 3–39). Dordrecht, Netherlands: Springer.

National Research Council (NRC). (2000). *Inquiry and the national science education standards*. 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.

Nersessian, N. J. (1999). Model-based reasoning in conceptual change. In Magnani, L., Nersessian, N., & Thagard, P. (Eds.), *Model-based reasoning in scientific discovery* (pp. 5–22). Dordrecht, Netherlands: Springer.

Nersessian, N. J. (2002). The cognitive basis of model-based reasoning in science. In Carruthers, P., Stich, S., & Siegal, M. (Eds.), *Cognitive basis of science* (pp. 133–153). New York: Cambridge University Press.

Nersessian, N. J. (2008). *Creating scientific concepts*. Cambridge, MA: The MIT Press.

Odenbaugh, J. (2005). Idealized, inaccurate but successful: A pragmatic approach to evaluating models in theoretical ecology. *Biology and Philosophy*, *20*(2–3), 231–255.

Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education*, *25*(2), 177–196.

Palmer, D. H. (2006). Sources of self-efficacy in a science methods course for primary teacher education students. *Research in Science Education*, *36*(4), 337–353.

Passmore, C., Gouvea, J. S., & Giere, R. (2014). Models in science and in learning science: Focusing scientific practice on sense-making. In Matthews, M. (Ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1171–1202). Dordrecht, Netherlands: Springer.

Passmore, C., Stewart, J., & Cartier, J. (2009). Model-based inquiry and school science: Creating connections. *School Science and Mathematics*, *109*(7), 394–402.

Pease, M. A., & Kuhn, D. (2011). Experimental analysis of the effective components of problem-based learning. *Science Education*, *95*(1), 57–86.

Penner, D. E., Giles, N. D., Lehrer, R., & Schauble, L. (1997). Building functional models: Designing an elbow. *Journal of Research in Science Teaching*, *34*(2), 125–143.

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.

Quillin, K., & Thomas, S. (2015). Drawing-to-learn: A framework for using drawings to promote model-based reasoning in biology. *CBE—Life Sciences Education*, *14*(1), es2.

Rinehart, R. W., Duncan, R. G., & Chinn, C. A. (2014). A scaffolding suite to support evidence-based modeling and argumentation. *Science Scope*, *38*(4), 70.

Rodenbusch, S. E., Hernandez, P. R., Simmons, S. L., & Dolan, E. L. (2016). Early engagement in course-based research increases graduation rates and completion of science, engineering, and mathematics degrees. *CBE—Life Sciences Education*, *15*(2). ar20.

Russell, C. B., & Weaver, G. C. (2011). A comparative study of traditional, inquiry-based, and research-based laboratory curricula: Impacts on understanding of the nature of science. *Chemistry Education Research and Practice*, *12*(1), 57–67.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475–5223.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*(6), 632–654.

Simon, R. A., Aulls, M. W., Dedic, H., Hubbard, K., & Hall, N. C. (2015). Exploring student persistence in STEM programs: A motivational model. *Canadian Journal of Education*, *38*(1), 1.

Songer, C. J., & Mintzes, J. J. (1994). Understanding cellular respiration: An analysis of conceptual change in college biology. *Journal of Research in Science Teaching*, *31*(6), 621–637.

Southard, K., Wince, T., Meddleton, S., & Bolger, M. S. (2016). Features of knowledge building in biology: Understanding undergraduate students' ideas about molecular mechanisms. *CBE—Life Sciences Education*, *15*(1), ar7.

Southard, K. M., Espindola, M. R., Zaepfel, S. D., & Bolger, M. S. (2017). Generative mechanistic explanation building in undergraduate molecular and cellular biology. *International Journal of Science Education*, *39*(13), 1795–1829.

Speth, E. B., Shaw, N., Momsen, J., Reinagel, A., Le, P., Taqieddin, R., & Long, T. (2014). Introductory biology students' conceptual models and explanations of the origin of variation. *CBE—Life Sciences Education*, *13*(3), 529–539.

Stewart, J., Cartier, J. L., & Passmore, C. M. (2005). Developing understanding through model-based inquiry. In Donovan, M. S., & Bransford, J. D. (Eds.), *How students learn* (pp. 515–565). Washington, DC: National Academies Press.

Stratford, S. J., Krajcik, J., & Soloway, E. (1998). Secondary students' dynamic modeling processes: Analyzing, reasoning about, synthesizing, and testing models of stream ecosystems. *Journal of Science Education and Technology*, *7*(3), 215–234.

Svoboda, J., & Passmore, C. (2010). Evaluating a modeling curriculum by using heuristics for productive disciplinary engagement. *CBE—Life Sciences Education*, *9*(3), 266–276.

Tien, L. T., Teichert, M. A., & Rickey, D. (2007). Effectiveness of a MORE laboratory module in prompting students to revise their molecular-level ideas about solutions. *Journal of Chemical Education*, *84*(1), 175.

van Dinther, M., Dochy, F., & Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educational Research and Reviews*, *6*(2), 95–108. doi: 10.1016/j.edurev.2010.10.003.

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, *31*(2), 142–166.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review*, *17*(4), 285–325.

Walker, J. P., Sampson, V., & Zimmerman, C. O. (2011). Argument-driven inquiry: An introduction to a new instructional model for use in undergraduate chemistry labs. *Journal of Chemical Education*, *88*(8), 1048–1056.

Wallace, C. S., Tsoi, M. Y., Calkin, J., & Darley, M. (2003). Learning from inquiry-based laboratories in nonmajor biology: An interpretive study of the relationships among inquiry experience, epistemologies, and conceptual growth. *Journal of Research in Science Teaching*, *40*, 986–1024. doi: 10.1002/tea.10127

Weaver G. C., Russell C. B., & Wink D. J. (2008). 2008. Inquiry-based and research-based laboratory pedagogies in undergraduate science. *Nature Chemical Biology*, *4*, 577–580.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81.

Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research*, *72*(2), 131–175.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, *92*(5), 941–967.

Zagallo, P., Meddleton, S., & Bolger, M. S. (2016). Teaching real data interpretation with models (TRIM): Analysis of student dialogue in a large-enrollment cell and developmental biology course. *CBE—Life Sciences Education*, *15*(2), ar17.

Zwickl, B. M., Finkelstein, N., & Lewandowski, H. J. (2014). Incorporating learning goals about modeling into an upper-division physics laboratory experiment. *American Journal of Physics*, *82*(9), 876–882.