Exploring the Unknown: Supporting Students' Navigation of Scientific Uncertainty With Coupled Methodologies

Julia Gouvea, Tufts University, Julia.Gouvea@tufts.edu
Aditi Wagh, Tufts University, Aditi.Wagh@tufts.edu

Abstract: Learning to make progress in science entails learning how to make progress in the face of theoretical and methodological uncertainties. In this paper we explore the potential for a coupled system that integrates computational and experimental methodologies to support learning to navigate scientific uncertainties. We present a case study of a group of undergraduate students in a 'hybrid' computational-experimental biology laboratory course as they investigated a complex biological system. We describe the nature of their progress: (1) mapping their theoretical understanding of the system and (2) planning and implementing their investigation. We then describe the role of the coupled methodologies in supporting their progress by (1) focusing the goals of their investigations, and (2) allowing for comparison and triangulation. Our results suggest that coupled methodological systems have the potential to provide a balance of resistances and footholds needed to support students in learning how to flexibly engage in scientific practice.

Keywords: uncertainty, scientific practice, computational modeling, epistemic agency

Introduction

> *"...that is what you are actually supposed to be doing in science...you are supposed to be trying to find things that are unknown, not known." – Nick, undergraduate student*

Learning how to do science means learning how to make progress exploring the unknown. In this exploration scientists encounter uncertainties – about what they know as well as about how to proceed. Scholars in science education have argued that students need practice navigating scientific uncertainties: to learn how to identify and characterize new problems, plan and design experiments, interpret complex datasets, and make and defend theoretical claims (Ford, 2005; Manz, 2015; Metz, 2004; Phillips et al., 2017; Reiser, 2004).

When students have opportunities to grapple with uncertainties they can learn scientific practices, and perhaps more importantly, they can learn when, how, and for what purposes to use them. For instance, Ford (2005) compared students who were taught the structure of experimental design with those who had to figure out how to design their own measures. Students who designed measures learned both the basic structure of a valid experimental test (controlling the variables) and how to evaluate experimental designs. When an elementary class studied by Manz (2015) encountered uncertainties in their investigations into a backyard ecosystem, it led to goal-oriented engagement with scientific practices. However, when some of their results were straightforward to interpret, the class engaged more superficially, ignoring potentially important variation in the data. In both of these examples, encountering and responding to uncertainty created a need for students to engage in scientific practices to attempt to understand something they did not already know.

Creating learning environments that allow students to learn how to navigate scientific uncertainties is a design challenge. In particular, it involves balancing opportunities for independent exploration with resources and guidance required to make meaningful progress (Engle & Conant, 2002; Jacobsen & Wilensky, 2006; Manz, 2015; Reiser, 2004). We have taken up this challenge in the context of an undergraduate introductory biology laboratory course. Many undergraduate biology lab courses emphasize learning procedures and demonstrating pre-defined outcomes (Beck et al., 2014). We aimed to create opportunities for students to learn how to conduct scientific inquiries into complex biological systems (Jacobsen & Wilensky, 2006). In considering this design challenge we drew on recent work in science studies that describes the potential for a coupled methodological approach to support progress on the frontiers of science (MacLeod & Nersessian, 2013). In this paper we report on how a coupled system of computational and experimental methods supported a group of students in making progress navigating a complex biological system.

Coupled methodological systems and progress on the frontiers of science

Systems biology is a frontier science in which researchers face multiple forms of uncertainty; the field currently lacks well-established theoretical models, standard experimental protocols, and complete datasets. Nersessian and colleagues have studied systems biology research laboratories in order to understand how scientists manage complexity and make progress on the frontiers of science. They note that some research labs specialize in either experimental or computational methodologies, while others take an integrative approach.

MacLeod and Nersessian (2013) conducted a case study of a graduate student researcher working in an integrative lab. They describe her use of computational modeling and experimental methods as *coupled* because investigations in each guided and constrained her progress in the other. They claimed that working within this coupled system allowed the student to make progress in two ways. First, having access to two modes of investigation, and therefore two sources of information, allowed this researcher to map her degree of certainty in her emerging understanding of the system. She was able to use output from each method to triangulate against the other, filling in her growing understanding of the system as well as locating gaps and uncertainties. Second, the two systems helped her identify, plan and implement next steps in her investigations. Because her research problem was open-ended there were multiple open questions and multiple possible approaches to addressing those questions. Working both with models and experiments helped her make choices about what to do next, with activity in one realm often suggesting avenues of inquiry to pursue in the other. And importantly, having access to both methods afforded her the flexibility to respond fluidly, changing course when she got stuck or taking the opportunity following unexpected, but potentially fruitful leads.

Our interest in this case was in considering how designing a learning environment that coupled these two methods could help students learn to make similar kinds of scientific progress, expanding their theoretical understanding and planning their investigations, in the face of uncertainties.

Design rationale: Coupling computational modeling and experimentation

This study is part of a design-based research project (Cobb et al., 2003) called 'Hybrid Labs' to emphasize the integration of computational and experimental approaches.

Drawing on prior research, we conjectured that introducing computational simulations into a biology laboratory course would enhance students' theoretical exploration of complex systems. By manipulating system parameters students could test and map their understanding of how the system could possibly behave (Svoboda & Passmore, 2013). Dynamic output from simulation runs could draw students' attention to how system outputs evolve over time as multiple system components interact simultaneously (Hogan & Thomas, 2001). In addition, computational models necessarily entail simplifying assumptions (Wilensky & Reisman, 2006), which we expected could create source of productive tension when compared with experimental systems (Manz, 2015).

We also centralized the role of experimental design, conjecturing that in developing their own method of measuring intended outcomes, students would have to consider the utility and meaning of their proposed measures (Ford, 2005). We further expected that in designing experiments and analyzing data from a complex biological system that students would encounter resistances (push-back from the material system), and that such resistances would create opportunities for students to accommodate their thinking and practices to interpret unexpected results and plan future investigations (Manz, 2015).

Most importantly, we conjectured that it would be critical for students to understand the two methodologies as linked and to use each as a tool to support progress towards their scientific aims. We wanted to avoid unproductive epistemic tensions between the two methods. For example, Smith et al. (1997) described how early in their learning to apply mathematical modeling to biology, some groups of college students constructed models that ignored important aspects of real systems and dismissed such mismatches as unproblematic. Other groups discounted the model, framing it as subordinate to more realistic experimental systems and claiming that the lack of match with reality made modeling useless.

We explicitly designed activities to emphasize interaction between the two methodologies. In this we drew on prior work in K-12 science education by Blikstein and colleagues (2016) who designed activities in which students simultaneously designed experiments and constructed models. They found that these *bifocal modeling* activities allowed students to identify and attempt to reconcile discrepancies between model output and experimental data (Blikstein et al., 2016). We also instructed students to compare output from simulations with experimental data. In addition, we instructed students to use the two methods as tools to plan steps in their inquiry, using the model to direct their experiments and using experimental data to structure how they interrogated the model.

Research aim

Our aim in this study was to explore how the coupled system could support progress in the face of scientific uncertainties that are part of investigating complex biological systems. We did so by identifying a focal group of students who made good progress and (1) describing the nature of their scientific progress, and (2) characterizing how the coupled methodological system supported their progress.

Methodological approach

We studied a group of four first-year undergraduate students, Nick, quoted above, and his group mates Abram, Damian and Walt, chosen because initial observations showed them using the computer simulation and experiment in coordination. For example, the group spontaneously turned to the simulation to overcome an obstacle in their experimental design. Our analysis of this group as a case study contributes an understanding of how a coupled methodological system can possibly support scientific progress.

Study context: Investigating variable mutation rates in bacteria

We studied the activity of the focal group during the first unit in a three-unit sequence of hybrid labs. In this unit students investigated the question, *under what conditions is it advantageous to have a higher or lower mutation rate?* This question was chosen because the relative advantage to organisms with different mutation rates is a complex outcome determined by organism-level parameters, population-level dynamics and environmental conditions, ensuring that the space of possible outcomes would be large enough that students would not be able to predict outcomes with certainty.

Students could investigate this question using a NetLogo computational model in which they could manipulate the relative mutation rates of two simulated bacterial strains, probabilities of different mutation effects (e.g. lethal vs. beneficial), and environmental conditions (e.g. presence/absence of antibiotic) and observe effects at the population level. Students could also design and conduct experiments with real bacteria by growing high and low mutating strains of *E. coli* in various growth media.

To position the two approaches as coupled we structured the lab activities so that students moved back and forth between the two methods multiple times: first conducting a whole class experiment, then exploring the simulation, then designing their own experiments and analyzing their results, and finally returning to the simulation again. Transitions between the two methods were explicitly framed to support links. For example, the instructor, a member of our design team, encouraged students to use the computer simulation to inform their experimental design and specifically highlighted how they could "manipulate things in the simulation that you can't manipulate in the experiment." During analysis of experimental results, he instructed the class to "try to model something that is relatively close to the experiment that you did." And again the instructor focused their attention on how the simulation could allow them to "watch patterns for longer" and "relate that back to their experimental results." Finally, for their final presentations and reports, students were instructed to include representations of both simulation output and experimental data and discuss consistencies or inconsistencies between the two.

Data collection and reduction

We collected video of the focal group working on all lab activities, including screen capture of their activity in the computer simulation. We also collected copies of students' three writing assignments for this lab unit, assigned weekly. Finally, we conducted interviews with three of the four focal students after the three-week unit had ended. In these interviews we asked students to reflect on their experience in the lab, specifically asking them to describe moments of challenge, surprise and engagement.

For this analysis we selected two central episodes because they featured interactions between the computational model and experiment. *Episode one* begins at the start of the second week of the unit and covers the time the group spent designing their experiment, including their spontaneous revisiting of the simulation (2 hours of activity). *Episode two* takes place in week 3 and includes the group's analysis and presentation of simulation output and experimental results (1 hour of activity). We focused on these in-class episodes, written work and interviews in our analysis.

Analysis

In the first phase of analysis we characterized the nature of the progress the group made in investigating the biological system. Drawing from descriptions of the progress in MacLeod and Nersessian's (2013) study and emergent patterns in the data, we defined two broad dimensions of progress. The first concerned how students mapped their theoretical understanding of system behavior. Progress in this dimension included: identifying parameters or conditions that could affect outcomes, making specific claims about the possible causes of system behavior, and characterizing the relative certainty of such claims. The second concerned the group's progress in their investigative actions. Progress in this dimension entailed making decisions about directions of inquiry to pursue as well as articulating the purpose and expected epistemic value of such decisions.

In the second phase of analysis we focused on understanding the role of the coupled methodological system in this progress. Within the two episodes, we identified moments when students were discussing or working with both methodologies. We then characterized how students were using the two (e.g. to compare patterns, identify discrepancies, frame decisions in the other) as well as how they talked about the relative certainty of the knowledge they gained from each.

Findings

Progress during episode one: Articulating a question

In this section we describe how the focal group makes progress mapping their understanding of the system as well as in planning their investigation as they work to articulate a question to investigate in their experiment.

Mapping the theoretical space

Prior to any investigation into the simulation or experimental design, the students participated in a whole class discussion that elicited initial ideas about the conditions under which high and low mutators would be expected to have an advantage. For homework they individually wrote a paragraph on these initial ideas. An idea that arose from both was the expectation that high mutators would have an advantage in unstable, novel or stressful environments, whereas low mutators would have the advantage in stable environments. This idea was present in all four of the initial assignments for this focal group of students.

By the end of this episode the theoretical space that the group was considering had expanded (Table 1). First, they had differentiated between two different types of "novel" environment: one in which the effects on bacteria are negative (antibiotic) and one in which the effects are potentially positive (lactose, a new nutrient). They described their understanding of the negative antibiotic environment as more certain: According to Damian, in antibiotic environments, "you mutate or you die." However, they were less certain about the role of mutation in an environment with a novel secondary nutrient. In interaction with the instructor, Damian reported that he was, "not sure if will impact it that much, because there's not like negative force acting to reduce the population."

Part of their progress involved characterizing this uncertainty. One possiblity, offered by Damian was that the higher mutator might have an advantage if it could "use the lactose that nobody else is using." But Nick pointed out that they, "don't know how much they're going to compete for glucose. 'Cause like it seems like there should be enough glucose to go around." By the end, Nick summarized that they should expect the high mutator to have an advantage *only if* the population is at "carrying capacity" where a large population size would mean competition for resources. In unpacking their uncertainty they articulated the role of resource dynamics and competition in considerations of when a mutation can be considered beneficial (Table 1).

In addition, they began to consider how underlying parameters – the probability that a mutation has a harmful versus beneficial effect on the organism – could influence the relative costs and benefits to high and low mutators. By the end of the episode they put a tentative stake in the ground: if the probability that a mutation results in a metabolic benefit is high, then the higher mutator should have the advantage. However, they also expressed uncertainty about this possibility. For example, Damian wrote in his second assignment,

> the rate of deleterious mutations compared to the beneficial metabolic mutations in the E. coli is unknown. Therefore, it is possible that the rate of deleterious mutations could be higher which would result in more [of the high mutating] E. coli dying than gaining the potential benefits from the Lactose metabolizing mutation.

Planning their investigation

Students had the choice of designing an experiment that involved growing bacteria in media containing either an antibiotic or a novel nutrient, lactose. At first the group leaned toward an experiment involving lactose, in part because they wanted to do something different from the whole class experiment they just completed (which involved an antibiotic). Early on, Damian expressed concerns that their experiment was "vague" and lacking a "good theory," and Nick threw his paper and put his head on the desk. Their discomfort seemed related to uncertainty about the outcome of their experiment described above: They were unsure if the high mutator would actually benefit from being grown in lactose.

By the end of this episode the group made progress in planning their investigation and convincing themselves of its value (Table 1). This is evident from their shifting descriptions of their research question. Early iterations of the question were phrased in terms of desired output. They asked whether or not the higher mutator would have an advantage and seemed frustrated when they were unsure about this outcome. Later the group shifted to problematizing the role of lactose in the system. First, they asked whether or not the ability to digest lactose should be considered a benefit. Second, if digesting lactose were a benefit, would it be enough of a benefit to allow the high mutator to outcompete the low mutator? This more elaborated research question seemed to shift the group from a focus on outcomes to understanding what their experiment might tell them about the system. In writing down this final version of the question, Damian explicitly expressed his confidence in the research question: “I understand what we are doing. It's just…is it going to create a difference? But that's like part of the experiment I guess.”

Table 1. Summary of Progress in Episode One

Progress	Beginning of Episode	End of Episode
Mapping the Theoretical Space	• Predictions based on broad categories (unstable vs. stable) • Adaptation by high mutators as the primary mechanism of advantage	• Differentiate between novel environments with positive (new resource) and negative (antibiotic) types of selective pressures • Identify importance of and uncertainty around resource dynamics and competition • Begin to attend to role of parameter values
Investigative Actions	• Initial question is too “vague” • Research question as confirming advantage of high mutator	• Research question problematizes role of lactose in benefiting the higher mutator • Confident that question is “good”

Role of the coupled methodological system in episode one

During this first episode, mapping between the entities and relationships represented in each system helped the group locate and characterize the problem they wanted to investigate. The computational model did not explicitly represent nutrient resources, while in the experimental system the group had to choose a nutrient medium in which to grow the bacteria. This mismatch was due to a simplifying assumption made in the computational environment that “metabolic benefit” mutations gave all bacteria an energetic benefit regardless of resource levels (which were not represented in the simulation). In the experimental system, this benefit would depend on the relative abundance of real nutrients.

Prior to experimental design the group had set the probability of metabolically beneficial mutations to high in the simulation and observed that the high mutator had an advantage. This outcome matched their expectation that high mutators should have an advantage in a “novel” environment. Their uncertainty about the role of lactose in the experimental system caused a temporary disruption during which Damian raised concerns about the experiment being “too vague.” In the midst of this frustration Damian suggested that they revisit the simulation. Once again the simulation with metabolic benefit set to high showed them the high mutator winning out, but this time the group questioned that outcome. Watching the graph, Nick wondered, “But is being able to digest lactose a metabolic benefit?” and Damian responded, “That is our question!”

Now the group was using the simulation to mark one potential outcome of the experiment – what they would expect to see if they assumed lactose did provide a benefit to the high mutator. Whether or not they would actually see this result would depend upon whether or not the experiment was “like their simulation” and specifically whether or not the assumption that digesting lactose was beneficial held in the experimental setup. This comparison helped them see their question as both viable and interesting to explore.

Overall, the role of the simulation during this first episode was to represent what the group expected to see. In isolation it functioned to simply to confirm their intuitions; but in interaction with the experimental system, the simulation represented one possible outcome based on a particular set of assumptions. The role of the experiment was both as a source of resistance, causing the group to notice and consider the assumptions in the simulation, and as a promising method to make progress – the results of the experiment could potentially tell them something about the lactose problem they had identified.

Progress during episode two: Analysis and interpretation of results

In this episode the group makes tentative theoretical claims and identifies additional uncertainties as they interpret their results and consider possible next steps in their investigation.

Expanding the theoretical space

At the end of the first episode the group was wondering about resource dynamics. They had also begun to discuss the role of harmful mutations, but had not systematically discussed their impact on the outcome. During this episode they further expanded their map of the theoretical space (Table 2).

One way they did this was by articulating their lack of knowledge of the probabilities of different mutation effects. This is reflected in their final assignments. For example, Nick wrote that while they “could set the mutation type frequency sliders to whatever we wanted” in the simulation, “we had no way of knowing the actual mutation frequency for each different type of mutation in bacteria.” Similarly, Damian qualified claims about relative success in his report, adding that it would "depend on the ratio of the rate of lethal mutations to the rate of beneficial Lactose metabolizing mutations."

While they did not know the exact values, by the end of the episode, all members of the group made tentative claims about the relative impact of the parameters. In his final report, Walt argued that, "The lethal mutation rate has a much stronger impact and is a stronger indicator of the high mutator's likelihood to survive than the beneficial mutation rate." Abram wrote that, "a higher metabolic-benefit rate would not counteract the effects of the lethal component on the high mutator." Nick even used this argument about relative impacts to claim that it was unlikely that beneficial mutations are more frequent in the real system because otherwise they would have seen a signal of this in their experiment.

Finally, there was evidence, both during the group's interaction in the simulation and in their final reports, that they were beginning to attend to mutation dynamics – the rate of spread of mutations through the population. As Abram and Walt ran trials in the simulation, Walt noticed an unexpected result: the high mutator was doing well at first, but after some time the low mutator began to take over. Walt pointed out that this was happening as the percent of beneficial mutations "plateaued" for the higher mutator. In his final report, Abram described how despite the high mutator's early advantage, "eventually, the low mutator strain gained a beneficial mutation that allowed it survive while the high mutator strain died out." These observations demonstrate a beginning understanding that advantage can shift over time as different mutations sweep through populations at different rates.

Planning simulation trials and interpreting and revising experiments

During this episode, Abram and Walt spent 40 minutes in the simulation and conducted six trials (Figure 1). With the exception of one trial suggested by the TA, Abram and Walt negotiated and defined a rationale for each, using the experimental setup to guide them. Their focus was on comparing the relative impact of lethal and metabolic benefit mutations, which they did systematically. By trial three Walt proposed that, "it seems the low mutator does best when the lethal rate is equal to or higher than metabolic benefit." In the remaining trials they edged back towards the parameter values in the first trial, increasing the metabolic benefit mutations relative to lethal to find a combination that would allow the high mutator to have the advantage. At the end they summarized across the set, Walt claiming, "I guess that makes sense. If it's a high mutator, it needs a really low chance of a lethal mutation in order for it to really thrive," and Abram nodding in agreement.

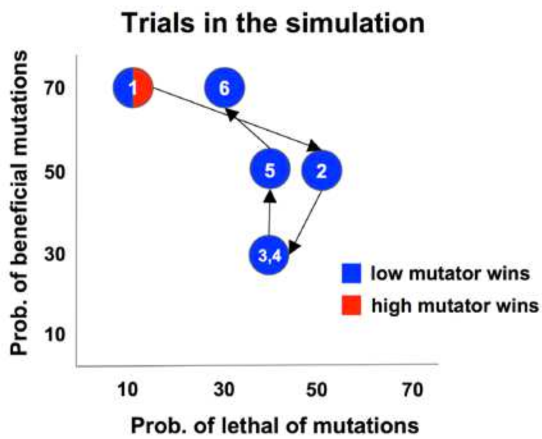

Figure 1. Path of students' six trials in the simulation plotted by the parameter combinations they chose.

A second type of progress was in interpreting their experimental results, which was difficult due to low numbers and an inability to distinguish between strains on some of the plates. Still, when asked by another student during their final presentation if they still believed their initial predictions, they were able to make some tentative claims. Damian responded, "If you look at the simulation, it shows that we saw that the metabolic mutations didn't have a huge impact on the colonies. So I think if we did the experiment again we wouldn't get numbers that showed that the lactose actually helped the high mutating strain."

Finally, there was progress in identifying new avenues for research. In interviews at the end of the unit, Abram, Damian and Nick all expressed a continued curiosity about the system. Unprompted, Nick articulated two potential ideas for future experiments. He described wanting to "try an experiment in which the glucose was scarce, as that may be a better way to create a selective pressure for lactose digestion where we would more likely see [the higher mutator] benefiting." Then, in his final report, he proposed an experiment in which he would incubate the bacteria for longer, noting that in the simulation it took hundreds of generations for the high mutator to benefit, whereas they only let their experiment run for 30 generations.

Table 2: Summary of Progress in Second Episode

Progress	Beginning of Episode	End of Episode
Mapping the Theoretical Space	• Begin to attend to role of parameter values • Uncertainty around resource dynamics and competition	• Problematize parameter values in experiment • Make claims about relative effects of different parameter values (lethal stronger than beneficial) • Attend to mutation dynamics
Investigative Actions	• Use simulation to confirm one expected result • Uncertainty over how to interpret results	• Use simulation to systematically explore parameter space • Make prediction about what results would have been (without experimental error) • Propose future experiments

Role of coupled methodological system in episode two

During this second episode we identified three ways in which the simulation and experiment interacted to support progress. First, Abram and Walt used the experiment to frame their investigation in the simulation. As Abram described in his final report,

> In the simulations we ran, we made changes to two conditions that we thought would tell us more about the results of our experiment: metabolic-benefit rate and lethal rate. Since we were not entirely sure whether the high metabolic-benefit rate actually aided the high mutator in our actual experiment (due to not being able to differentiate between the two strains), we decided to run the experiment again.

Abram described thinking that the simulation could give them information relevant to their experiment (to essentially allow them to "run it again"). Further, it focused their choice of parameters to manipulate, ultimately allowing them to claim that the impact of lethal mutations was stronger than the impact beneficial mutations.

Second, the ability to explore parameter space in the simulation caused some group members to see the whole system as more complex. Abram had initially felt relatively certain that the high mutator would win in the lactose experiment. In an interview at the end of the unit, Abram described how, "the fact that we got so many outcomes from the one that we predicted kind of just contradicts our original predictions and kind of shows that …we were really fishing for a specific thing, when it was likely that a lot of possible outcomes could have happened instead." This is different from how the simulation functioned in episode one to represent simple outcomes. Now both the simulation and the experiment emphasized that the outcomes of the system can vary depending on the conditions and parameter values.

Third, there is evidence that mapping between simulation and experiment influenced at least one of the future experiments proposed by Nick. In his interview Nick described noticing that the simulation ran through hundreds of generations compared with 30 generations in the experiment. Nick noticed that in the simulation sometimes the dynamics would "switch" if run for long enough. These considerations seem to have inspired Nick's idea to run the experiment for longer, which he explicitly connects to his wondering about whether more time was needed for the beneficial mutation to sweep through the high mutating population. In this example, the simulation once again functioned to open up the space of possible ways the experimental system might behave, this time drawing attention to the importance of mutation dynamics.

Discussion

Designers of authentic scientific learning environments have recognized the need to allow students to explore unknown terrain while at the same time pointing to footholds that will help them make progress. This design tension is often discussed in terms of balancing opportunities for students to encounter problems and resistances on the one hand and access to resources (guides, information, instructor support) on the other (e.g., Engle & Conant, 2002; Reiser, 2004). In this case, each approach – computational simulation and experiment – functioned at times as a source of knowledge and at other times as a source of uncertainty. While the computer simulation was initially used to represent straightforward theoretical predictions, it later became a method for exploring how dynamics changed over parameter space. In the experimental system, the group was initially uncertain about the resource dynamics, but became more certain, despite small numbers, that the experimental system would not favor the high mutator as they had initially expected. This suggests that an alternative to providing students with targeted help is to provide them with alternative avenues for continuing their investigations.

The progress made by this group was clearly linked to their taking up the idea that the two methods can and should interact; but not all groups understood the system as coupled. This raises a new tension in our design: How much to structure the two methods as interacting as opposed to providing students with access to both methods to use flexibly. In this case, moves between experiment and simulation were primarily guided by the activity structure and framed by the instructor. One exception was the group's spontaneous move to revisit the simulation during their experimental design. In this moment they demonstrated the kind of flexible use of scientific tools seen in the systems biology PhD student who described the autonomy and confidence she developed in her work: "I like the idea that I'm building my model things are popping up in my head oh wow this would be a good experiment. I plan out the experiment myself and then go into the lab and I do it." (MacLeod & Nersessian, 2013). Ultimately, the ability to plan and make goal-directed moves in the face of uncertainty is the kind of expertise we want to develop in science students, and coupled methodologies may be a fruitful way to structure opportunities for students to practice this.

References

Beck, C. W., Butler, A., & Burke da Silva, K. (2014). Promoting Inquiry-Based Teaching in Laboratory Courses: Are We Meeting the Grade? *CBE—Life Sciences Education, 13*(3), 444–452.

Blikstein, P., Fuhrmann, T., & Salehi, S. (2016). Using the Bifocal Modeling Framework to Resolve "Discrepant Events" Between Physical Experiments and Virtual Models in Biology. *Journal of Science Education and Technology, 25*(4), 513–526.

Cobb, P., Confrey, J., DiSessa, A. A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher, 32*(1), 9–13.

Engle, R. A., & Conant, F. R. (2002). Guiding Principles for Fostering Productive Disciplinary Engagement: Explaining an Emergent Argument in a Community of Learners Classroom. *Cognition and Instruction, 20*(4), 399–483.

Ford, M. J. (2005). The Game, the Pieces, and the Players: Generative Resources From Two Instructional Portrayals of Experimentation. *Journal of the Learning Sciences*, *14*(4), 449–487.

Hogan, K., & Thomas, D. (2001). Cognitive comparisons of students' systems modeling in ecology. *Journal of Science Education and Technology, 10*(4), 319-345.

Jacobson, M., & Wilensky, U. J. (2006). Complex Systems in Education: Scientific and Educational Importance and Implications for the Learning Sciences. *Journal of the Learning Sciences, 15*(1), 11–34.

Macleod, M., & Nersessian, N. J. (2013). Coupling simulation and experiment: The bimodal strategy in integrative systems biology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 1–13.

Manz, E. (2015). Resistance and the Development of Scientific Practice: Designing the Mangle Into Science Instruction. *Cognition and Instruction, 33*(2), 89–124.

Metz, K. E. (2004). Children's Understanding of Scientific Inquiry: Their Conceptualization of Uncertainty in Investigations of Their Own Design. *Cognition and Instruction*, *22*(2), 219–290.

Phillips, A. M., Watkins, J. E., & Hammer, D. (2017). Problematizing as a scientific endeavor. *Physical Review Physics Education Research*, *13*(2), 20107.

Reiser, B. J. (2004). Scaffolding Complex Learning: The Mechanisms of Structuring and Problematizing Student Work. *Journal of the Learning Sciences, 13*(3), 273–304.

Smith, E., Haarer, S., & Confrey, J. (1997). Seeking diversity in mathematics education: Mathematical modeling in the practice of biologists and mathematicians. *Science & Education, 6*, 441–472.

Svoboda, J., & Passmore, C. (2013). The strategies of modeling in biology education. *Science & Education*, *22*(1), 119–142.

Wilensky, U. J., & Reisman, K. (2006). Thinking Like a Wolf, a Sheep, or a Firefly: Learning Biology Through Constructing and Testing Computational Theories: An Embodied Modeling Approach. *Cognition and Instruction, 24*(2), 171–209.

Acknowledgments

We gratefully acknowledge financial support from the Davis Educational Foundation established by Stanton and Elizabeth Davis after Mr. Davis's retirement as chairman of Shaw's Supermarkets, Inc. to J. Gouvea and the Gordon and Betty Moore Foundation through Grant No. GBMF3475 to D. Hammer, Tufts University.