Sci & Educ
DOI 10.1007/s11191-011-9425-5

The Strategies of Modeling in Biology Education

Julia Svoboda · Cynthia Passmore

Abstract Modeling, like inquiry more generally, is not a single method, but rather a complex suite of strategies. Philosophers of biology, citing the diverse aims, interests, and disciplinary cultures of biologists, argue that modeling is best understood in the context of its epistemic aims and cognitive payoffs. In the science education literature, modeling has been discussed in a variety of ways, but often without explicit reference to the diversity of roles models play in scientific practice. We aim to expand and bring clarity to the myriad uses of models in science by presenting a framework from philosopher of biology Jay Odenbaugh that describes five pragmatic strategies of model use in the biological sciences. We then present illustrative examples of each of these roles from an empirical study of an undergraduate biological modeling curriculum, which highlight how students used models to help them frame their research question, explore ideas, and refine their conceptual understanding in an educational setting. Our aim is to begin to explicate the definition of modeling in science in a way that will allow educators and curriculum developers to make informed choices about how and for what purpose modeling enters science classrooms.

1 Introduction

Modeling is increasingly understood as an integral component of authentic scientific inquiry and, as a result, model-based curricula have been gaining traction in science education reform (NRC 2007). However, modeling, like inquiry more broadly, is a complex practice, which makes successfully translating modeling from the scientific community into science classrooms a significant challenge. One important element in addressing this challenge is to develop a better understanding of the reasoning and decision-making strategies that accompany authentic scientific practice. There are relatively few accounts of modeling in classrooms that do this in much detail. However, the philosophy of science, and, in particular, the philosophy of biology literature, contains several

---

J. Svoboda (✉) · C. Passmore
School of Education, University of California, Davis, Davis, CA 95616, USA
e-mail: jmsvoboda@ucdavis.edu

useful treatments of modeling that are applicable to developing a deeper sense of this practice in educational settings.

Historians and philosophers of science have documented and analyzed various ways that modeling is done in the sciences and described a range of activities that may all fall under the general heading of modeling. While there is some disagreement about how to define them, there is general agreement that models are central to scientific practice (e.g. Giere 1988; Morrison and Morgan 1999; Knuuttila 2005). However, because modeling is really a diverse set of practices, we feel that it is important to get more specific about the context and utility of models in science so that we can be more thoughtful about introducing science students to models and modeling. The diversity of uses for models is particularly apparent in the biological sciences where model use varies with the aims, interests, and disciplinary cultures of biologists. Detailed case-studies highlight how the role of modeling can differ across disciplinary contexts (e.g. Darden 1991; Lloyd 1994; Keller 2000) and have led philosophers of biology to challenge broad-brush accounts that describe science as simply "model-based" and to consider in more detail the roles that models play in scientific practice (Downes 1992; Godfrey-Smith 2006). The practice-based, or naturalistic turn in the philosophy of science has focused specifically on the activities rather than solely the products of science (Downes 1992, 142). In practice, "scientists use models to represent aspects of the world for various purposes" (Giere 2004, p. 743). Scientific modeling, like scientific inquiry more generally, is not a single method, but rather a complex suite of strategies. Models, of various forms and types can describe complex phenomena, represent core ideas about a system, be manipulated to explore the dynamics of a system, be used to make predictions about future events, suggest the need for empirical studies, and facilitate the communication of ideas. Each of these activities requires scientists to attend to different features of the model and make context-dependent decisions about how best to use and evaluate these models. There are real cognitive and epistemic differences in these different strategies of modeling.

We argue, in line with scholars such as Matthews (1994), that one of the main contributions philosophers of science can make to science education is to help explicate the diversity of strategies and motivations for model use in biology and science more broadly. In this article, we draw on the work of scholars such as Gregory Cooper and Jay Odenbaugh who have focused on uncovering *epistemic aims* and *cognitive payoffs* associated with different scientific methodologies, and biological modeling in particular (Cooper 2003; Odenbaugh 2005, 2009). Odenbaugh (2005) describes five important roles of modeling in biological research. In the body of this paper, we present his framework and then provide empirical examples drawn from an undergraduate biological modeling curriculum that serve to instantiate these roles. The examples that we chose illustrate what different modeling strategies can look like when enacted in an educational setting. In our analysis we describe how each strategy can support the cognitive, metacognitive and epistemic aims of science education. In sum, our primary aim is to begin to expand and explicate the definition of modeling in science in a way that will allow educators and curriculum developers to make informed choices about how and for what purpose modeling enters science classrooms.

1.1 Modeling in Science Education

A growing literature provides documentation of the learning benefits associated with engaging students in the practice of modeling. For example, researchers have shown that when students engage in modeling they develop a deep understanding of the content and an ability to solve novel problems (e.g. Wynne et al. 2001; Lehrer and Schauble 2005). Other studies have shown that modeling curricula can bring students into alignment with the epistemic aims of science and help them develop more sophisticated ideas about of the scientific enterprise (Schwarz and White 2005; Windschitl et al. 2008a). It is clear that modeling has the potential to support meaningful reform efforts in scientific education. However, as model-based curricula continue to proliferate, we argue for an increased need to examine in more detail how modeling is practiced in science and how these practices align with the aims of education.

We believe that, to a large extent, accounts of modeling in the science education literature are underspecified with regard to the means and purposes of modeling and the disciplinary variations that exist in this practice. For example, Schwarz et al. (2009, pp. 635–636) state that they have purposely limited their analysis to a general account of the practice of modeling and do not attempt to account for variations in modeling strategies. In their work, models are defined as simple, abstract representations that can be used to explain or predict. However, how the simplicity or abstractness of models might vary is not addressed, nor are the different aims of explanation and prediction explored in much detail. Other authors use "model" much more narrowly, as a simplified system that can be used to explore a more complex system (a model system) (e.g. Lehrer et al. 2008), a well established set of scientific ideas (consensus model) (e.g. Clement 2000); a representation of empirical observation (models of data) (e.g. Chinn and Brewer), or a set of ideas that explain some process or pattern (theoretical model) (e.g. White 1993). Collectively, these studies illustrate the range of aims and uses of modeling, but the relevant differences among them have not been made clear in the science education literature.

These different versions of models each have different meanings and uses and interface in different ways with other scientific practices. While these studies are productive starting points for introducing models into classrooms, we argue that without a clearer sense of the nature and purpose of models and modeling in science, there is a danger that modeling could go the way of more general calls for inquiry in science education, in which the meaning of the practice has become so general and diffuse that it has lost utility. The practices of scientific inquiry have often been enacted as over-simplified procedures that have provided limited opportunities for sophisticated scientific reasoning and even promoted anti-authentic images of the nature of science (Hodson 1996, 1998; Chinn and Malhotra 2002; Rudolph 2005). Like inquiry, for the most part, modeling has been discussed in the science education literature without reference to the different affordances and constraints of particular uses of models. If models are tools for thinking, we ask, what kinds of thinking can they be used to support, and to what ends? Too often, students are asked to consider what models are or how to go about modeling without a clear understanding of why modeling is a productive endeavor. Here, we bring ideas about the role of models from the philosophy of biology literature to bear on an analysis of a model-based educational experience for undergraduates in biology in an attempt to unpack the variety of, and relationships between, the cognitive strategies and epistemic aims that modeling has the potential to support.

1.2 The Cognitive Strategies and Epistemic Aims of Modeling in Biology

It is widely recognized that within the biological sciences, and in science generally, models are built for a range of different purposes and used in a variety of ways (Lloyd 1994; Smith et al. 1997; Morrison and Morgan 1999; Keller 2000; Cooper 2003; Knuuttila 2005; Nersessian 1999, 2002; Laubichler and Müller 2007; Odenbaugh 2009). In 2005, philosopher of biology Jay Odenbaugh introduced a framework that situates the practice of modeling as a collection of *cognitive strategies* that can be used to pursue particular aims of scientific inquiry. Odenbaugh's view emphasizes the pragmatic nature of modeling; depending on the objective, the form and function of modeling will vary. Similarly, philosopher Rom Harre describes how modeling strategies can vary both temporally (how far along is the inquiry?) and epistemologically (what kind of knowledge is the researcher after?). Attention to context is critically important in biology, a field that is dynamic, and inhabited by a diversity of scientists with very different kinds of aims (conservation biologists, biochemists, theoretical ecologists, computational genomicists etc.).

One way to parse this variation is to recognize, as ecologist and philosopher of science Richard Levins did in his seminal work, *The Strategy of Model Building in Population Biology*, that modelers face inherent epistemic trade-offs (Levins 1966). Levins argued that for both cognitive and methodological reasons, modelers must often choose among the desirable, but often conflicting, epistemic aims of realism, precision and generality. For example, a fisheries biologist interested in population projections of a species of interest might choose to sacrifice generality in order to construct a model that can generate accurate predictions of population fluctuations; while, an ecologist, like Levins himself, might forgo predictive precision in the interest of general explanatory power (ibid.). The main point to take away from Levins' argument is that modelers will and should build different models depending on their particular aims. The implication is that there is not a single type of model or modeling that can address all biological problems equally well; depending on the question at hand, a biologist will want to choose the model that is the best tool for the job.

To some extent, differences in epistemic aims among biological disciplines will be visible in the types of models used. Consider for example how modeling practices might differ between molecular and evolutionary biologists. Philosophers Lindley Darden and Evelyn Fox-Keller describe molecular biology as an enterprise whose aim is to elucidate molecular mechanisms at a high degree of detail (Darden 1991, 2002; Keller 2000). Darden describes how molecular biologists sought after the detailed series of steps that now fully describe protein transcription (1991). Similarly, Keller describes how developmental biologists used models along with carefully crafted experiments to specify the details of developmental pathways (2000). The aim of such scientists, says Darden, is to "fill in the gaps" in our understanding of the mechanisms at work at the molecular level (Darden 2002, p. S363), an aim, using Levins' terminology, that places a high value on realism. In contrast, Elisabeth Lloyd describes the importance for evolutionary biologists of relatively abstract mathematical models (1994). For example, population genetics models typically sacrifice reality (by, for example, modeling single-locus dynamics) in order to make generalizable predictions. These modelers have aimed for general explanatory power over detailed mechanistic descriptions.

One way to make sense of these differences is to recognize that different subfields within biology constitute different epistemic cultures (Knorr-Cetina 1999). Such cultures are not confined to traditional disciplinary boundaries, but instead refer to collections of practitioners that use common machineries to achieve similar epistemic aims (ibid). Importantly, these cultures can shift and evolve over time just as any cultural group would. This point is important because it once again emphasizes the need to consider the nature of modeling *in context*. The epistemic culture of molecular biology has shifted significantly in the past few decades. Darden describes how early attempts to articulate molecular processes of transcription and translation relied on abstract conceptual models compared to the detailed mechanistic models that are common in contemporary molecular biology research (2002). The utility of the notion of epistemic cultures is that it focuses on conversations about scientific practice on the aims of the practitioners and how such aims might shift over time and space. An appreciation for the socio-historical context of a discipline helps us understand why the detailed molecular pathway models are so different from theoretical population genetic models.

Odenbaugh, drawing inspiration from Levins, has argued that it is important to clarify the practice of modeling in terms of the cognitive benefits associated with different modeling strategies and how they can serve particular epistemic aims (2005). To this end, and drawing on the work of Levins (1966), Odenbaugh explores five major pragmatic uses for models in biology and their associated benefits: (1) simple, unrealistic models help scientists explore complex systems, (2), models can be used to explore unknown possibilities (3) models can lead to the development of conceptual frameworks, (4) models can make accurate predictions, (5) models can generate causal explanations. The focus of his argument, and a point we wish to present to the science education community, is that the first three roles of models have been underemphasized in comparison to the latter two. In addition to recognizing the role of models as tools that scientists use to make accurate predictions and build complete and accurate explanations of biological phenomena, it is important to recognize the ways in which models can support theoretical progress even when, and perhaps because they do not capture all the details of the system under study. Even when realism, precision, and generality are all fairly low, when models are "false", there is still important intellectual progress to be made towards "truer theories" (i.e. Harre 1986; Wimsatt 1987, 2002).

1.2.1 Using Simple Models to Understand Complex Phenomena

That models are simplified in comparison to the phenomena they represent is well understood. What is important is not simply that models are simplifications of reality, but that scientists purposely construct simplified models in order to serve specific ends. That is, model building is a subjective enterprise that reflects deliberate choices made by scientists about how much reality to include. The result, as Odenbaugh alludes to, is that there exists a range of models along a continuum from ridiculously simple to extremely complex. Neither approach is necessarily better than that other, for, as Levins, argues, "the difference between legitimate and illegitimate simplifications depends not only on the reality to be determined but also on the state of the science" (1966, pp. 421–422). This view emphasizes simplification not as a limitation of modeling, but as a deliberate and context-specific strategy.

Odenbaugh (2005) highlights several ways in which simplistic models can be productive tools for scientific reasoning. First, simple models can help scientists by allowing them to begin to unpack the reasons why a false model is wrong (see also Wimsatt 1987). Odenbaugh describes how a simple model can be compared to successively more complex models as a strategy for locating error. False simple models can be extremely important in guiding research and helping scientists formally describe what it is that they do not know. As the complexity of a model grows it becomes increasingly difficult to localize causal elements in the model. Increasing attention to detail can, somewhat counter-intuitively, impede scientists' ability to understand and explain how a system works. Starting with simple models and building upon them is a deliberate strategy that can help structure, in a systematic way, how scientists reason about complex biological phenomena. In particular, this can be achieved by helping to isolate the most important components out of the complexity. Thus one important feature of simple models is that they make reasoning about complex systems possible, and in doing so they must forgo realism.[1]

---

[1] Readers interested in an example from physics might want to consider Hughes' (1999) description of the Ising model, a model that is not faithful to reality but nevertheless has explanatory utility in physics.

1.2.2 Models as Exploratory

Role two goes beyond the first to suggest that models can help scientists explore alternative versions of reality. As Odenbaugh argues, part of the power of models lies in their ability to support and structure possibilities rather than actualities (2005).[2] That is, models can play an important exploratory role in science. Constructing a model, Odenbaugh argues, helps scientists organize and articulate their ideas about how a phenomenon might possibly work. This is one of the most creative and generative forms of scientific reasoning (Clement 1989; Nersessian 1992, 1999; Osbeck et al. 2010). The primary mode of reasoning during model construction is abductive: scientists draw on prior knowledge, reason by comparison or by analogy, generate novel visual representations and engage in thought experiments in an attempt to imagine which elements of a system may be critical and ultimately construct and choose among plausible explanatory hypotheses (Magnani et al. 1999; Morrison and Morgan 1999; Nersessian 2008).

The exploratory nature of modeling is also evident in the way conjectural models are analyzed. Once a modeler has made decisions about what to include in a model, he can manipulate the model to explore ideas about what might possibly be expected to happen over a range of different conditions. This kind of reasoning is particularly salient when computers are used to run models forward in time, but the same effect can be achieved (with less power) by thought experimentation (e.g. Hughes 1999; Nersessian 1999). In this way, the model can support reasoning that extends beyond what is immediately observable or measureable. This kind of exploratory analysis can help highlight areas of interest and open new potential areas to investigation. In sum, models can be used to explore and expand possibilities rather than simply as a way to rule out alternative hypotheses.

1.2.3 Using Modeling to Develop Concepts

While the first two roles focus on the model, this third role highlights the ways in which modeling can generate new ideas that have the potential to transcend the model. In this role the model is a tool that brings greater clarity to the nature of the phenomenon under study. As Odenbaugh describes, this may occur during the process of model construction as the modeler attempts to formally and clearly articulate ideas that were previously fuzzier. For example, the notion of a *species niche* was first developed as a component of a competition model and has subsequently become one of the most important ideas in biology (Cooper 2003, pp. 14–15).

Often opportunities for conceptual development take place during the process of formally representing model ideas by choosing a particular representational form. This process forces the modeler to make particular choices about how to represent model entities and how exactly to specify model relationships (Hughes 1999). Formal representation can often result in the generation of inscriptions that the modeler and others can interact with in productive ways (Latour 1990). Concepts may also emerge during model analysis as the model is manipulated and its behavior is characterized. For example, some parameters or groups of parameters that were included in the initial model can emerge as important enough to warrant naming them, and important conceptual work can occur in the translation of mathematical symbols back into biological concepts. Odenbaugh (2005) describes how biologist Robert May (1973) chose to represent the overall number and degree of

---

[2] See also Harre (1986) for his discussion of models that are used to explore “possibilities” and “impossibilities”.

interactions in an ecological community in terms of a “connectance” parameter C, which he defined as the proportion of all pairwise species interactions that were not equal to zero. May’s analysis suggested that C played a key role in the stability of the community over time. While May’s model was later criticized, Odenbaugh describes how his attempt to operationalize and interpret the role of C opened up a discussion in the ecological community surrounding the appropriate ways to conceptualize community complexity and stability. This analysis marked the beginning of a proliferation of ideas in the ecological community as well as a marked increase in experimental work in community ecology that extended well beyond the original model.

1.2.4 Prediction and Explanation

We discuss the final two roles of modeling together, as does Odenbaugh, because they both concern uses of models that require a higher degree of realism. In some contexts, the ultimate test of the utility of a model is that it is able to make accurate predictions or generate plausible explanations for data patterns. Odenbaugh does not debate the need for models to speak to empirical data, and neither do we. We do, however, agree with Odenbaugh’s argument that progress in science results from a dialogue between generating new conjectures and, when possible, subjecting those conjectures to empirical testing. Modeling is important in both of these contexts; models can direct empirical work by generating testable predictions, and empirical testing can lend support to the hypotheses generated by the model or identify gaps in our understanding that could benefit from a new theoretical lens. It is worth noting however, that while they both rely to some extent on matching empirical data, explanation and prediction are two different kinds of criteria for evaluating a model. Explanation refers to the ability of a model to account, in sufficient detail, for the underlying causal mechanisms that produce some observed outcome or phenomenon (MDC 2000; Bechtel and Abrahamsen 2005). Prediction, in contrast, refers only to a model’s ability to predict, with a high degree of accuracy, what, given some specified initial conditions, will happen in the future.[3] A disease model that can predict how fast a disease will spread through a population, without specifying in any detail the underlying mechanisms of such spread satisfies this aim, while a detailed account of the molecular mechanisms that underlie circadian rhythms is an example of a model with high explanatory power (Bechtel and Abrahamsen 2010). Over time and with iteration, a good model will ideally be able to both explain and make accurate predictions. At this stage in inquiry the model has transitioned from acting as a tool used to represent tentative ideas to one that represents a well-supported consensus view of how some phenomenon really works. Models of this sort have a different epistemic status—they represent the current consensus about what is “known” and as such they play less of a role in supporting new inquiries.

In sum, the essence of Odenbaugh’s argument is that matching reality is not the only role for models in biology. We can still value models that are able to make predictions and explanations without demanding that *all* models must do so. Odenbaugh wants to ensure that the utility of the exploratory role of models in generating new ideas and new ways of thinking is recognized as well. His aim is to argue for a plurality of uses for models, and

---

[3] Prediction in this context is used to refer to the practice of predicting future events with some accuracy. This is different from the reasoning strategy of imagining the implications if a given model were true, often also referred to as the predictions of a model. This type of reasoning is more closely related to exploring possibilities as discussed in Sect. 1.2.2.

like Levins, to recognize that not all models need to be precise, realistic and general, but that instead we should acknowledge that models can be used in different ways depending on the goals of the modeler. An understanding that models are not blunt instruments but rather a suite of tools that can be used to organize and conceptualize ideas, explore possibilities, generate new questions, as well as serve as formal testable hypotheses, will serve science educators well as they craft lessons to engage students in particular types of reasoning.

2 Supporting the Plurality of Roles for Models in Science Education: A Case Example

In the remainder of the paper, we present an account of model-based inquiry that features a group of seven undergraduate mathematics and biology majors. These students were participants in an NSF-sponsored traineeship called Collaborative Learning at the Interface of Mathematics and Biology (CLIMB). The central focus of this traineeship was to introduce students to the strategies of modeling in biology. During the first half of the year, the CLIMB cohort worked with faculty from a range of biological disciplines to solve different modeling problems. They then worked as a group to conduct their own research project over the spring and summer. Throughout this project the group received advice from a number of different scientists who guided them through the process of constructing and analyzing a scientific model.

What was unique about this context is that we, as researchers, had the opportunity to observe the ways in which scientists guided students in the context of a truly authentic inquiry. What is evident from our analysis of what the faculty said and did, is that without explicitly planning to do so, they helped support various uses of modeling for the CLIMB students, and in the process supported a range of different kinds of scientific reasoning. Below we use Odenbaugh's five strategies to frame our presentation of illustrative empirical examples that demonstrate how the CLIMB program supported the different cognitive uses and epistemic aims of modeling to help the students make progress in their inquiry. These examples are drawn from data collected during an ethnographic study of the 2008–2009 CLIMB cohort. In this analysis we draw on detailed field notes of student-faculty meetings and transcripts of audio-recorded student presentations that were collected over 5 months, from the beginning stages of project formulation in April 2009 to the final presentation of results at the CLIMB conference in September 2009. In addition, we draw on transcripts of audiotaped interviews with CLIMB students.

What is clear from these examples is that much of the productive reasoning took place while students were using fairly simple, unrealistic models that did not match the "real" empirical data, and that despite or perhaps because of this, students were able to engage in creative, generative modes of scientific reasoning.

2.1 Overview of the CLIMB Summer Project

The CLIMB cohort consisted of seven undergraduate students: five biology majors, one mathematics major, and one physics major (who later switched to an independent major in biophysics). The students consulted regularly with three scientists: Professor R, Professor S, and Dr. Marcia. They also had the support of an advanced graduate student teaching assistant (GTA), who had a background in mathematical modeling.

The goal for the CLIMB summer project was to choose a research question in biology that could be addressed using a mathematical model. The specific question was largely left open to the students to choose. The group considered ideas related to the evolution of gene regulation, altruism and cooperation, the role of macrophage proliferation in cancer treatments and evolutionary models of phenotypic plasticity before they ultimately converged on a project inspired by an article that described how media attention about a possible link between autism and the measles, mumps, rubella (MMR) vaccine led to reduced vaccination and recent disease outbreaks in countries that have voluntary vaccination strategies.

What was especially interesting about this project was that the nature of the question that the group was attempting to answer shifted continually throughout the 5-month duration of their research. Even in the final days of their analysis, the group was still refining their research question. As will be seen in the empirical examples below, these shifts occurred for a variety of reasons: changing interests of the group, suggestions by mentors, access to new information, and in response to emerging results obtained from their modeling activities. In this way the students' inquiry into vaccination-disease dynamics reflected the idiosyncratic and often opportunistic nature of authentic inquiry. The students never had a firm hypothesis that they were attempting to "prove" or "disprove". Rather, the group was engaged in a creative and dynamic process of trying to make sense of a complex phenomenon. Odenbaugh's framework has allowed us to select instances from this modeling experience that were particularly productive in terms of the depth and diversity of reasoning strategies they supported.

2.2 Strategically Using Simple Models

The importance of working with a simplified model was emphasized by various CLIMB mentors throughout the project, but for different reasons at different points. Early on, when the CLIMB students were still trying to decide what kind of question they wanted to ask about disease dynamics, the mentors encouraged them to first understand the simplest possible model, the SIR model, which conceptualizes disease spread in terms of transitions among three classes of individuals: those that are susceptible to infection (S), those that are already infected (I) and those that have recovered (R). This exercise had two goals, to familiarize the students with a basic starting point used by disease modelers and to present to them the strategy of localizing gaps in current knowledge by examining the assumptions made by the simplest model (field notes, May 8).

> Professor R: What is the SIR model trying to predict?
> Rose[4]: The dynamics of the disease.
> Professor R: Exactly, and if SIR does not predict the disease dynamics then you know that either your parameters are crappy or you left some important variable out of your model. Maybe there is something important biologically that we are missing. So you add, subtract, whatever. You take something as simple as SIR and you ask why it does or does not work.

This strategy then guided students in the construction of their own model. They spent several days creating a list of possible factors that could influence disease dynamics that were not present in the basic SIR model. Once the group had created a list of possible factors, they were instructed to narrow their focus. The quote below from the graduate TA

---

[4] All names are pseudonyms.

echoes the advice from the faculty mentors about why the CLIMB students needed to keep their model simple if they wanted to understand anything about the complex system (transcribed audio, June 2, 2009).

> GTA: So let me just ask you guys, because you are going over a couple things here: you are saying voluntary vaccination, age structure, then you are talking about this information. You know, you have all these things into one model. Those are quite a few things. You want to know which ones of these things is actually doing anything. How are you going to be able to discern which one is having the effects on it?

Here, the GTA explicitly articulated simplicity as a strategy for making conceptual progress on their problem. Using this strategy would of course mean that they would be leaving many elements of what is really happening biologically behind.

Later this strategy became more specific. The group began to more systematically evaluate other disease models in order to decide whether they could be transferred to the case of measles and MMR. They found, for example, a paper that described how the dynamics of smallpox are influenced when vaccination is voluntary. This journal article spurred a discussion about the extent to which the network structure of the smallpox model was appropriate for a model of measles (field notes, May 22, 2009).

> Kevin: I like the idea of adding a network thingy – a non-Poisson, small world network.
> Sean: We could focus on changing the network structure.
> Lillian: But measles is not a social contact disease.
> Romy: Measles is airborne…transmissibility is high.
> Professor S: Is that critical? The social interactions are in the network. The only difference is the transmissibility constant, not the nature of the network.
> Lillian: But smallpox has a smaller neighborhood. Less people get sick who you contact.
> Professor S: Oh, I see, my measles network would be larger.
> …
> Rose: If we assume children are transmitting the disease – let's face it, kids touch everything.

Several students, Sean and Kevin, are excited about the possibility of modeling disease transmission using a network structure instead of the simpler random-mixing model. Three other students, Lillian, Romy and Rose argue that in the case of measles a more complex model of transmission is not justified. Given what they understand about measles (it is airborne), as well as their intuitions (kids touch everything), these three group members are able to convince the others that keeping disease transmission simple is a justifiable decision for this particular disease and that the group should retain the random-mixing assumption of the SIR model.

The above examples demonstrate two ways in which reasoning with simple models can be an effective modeling strategy. First, the students had to consider to what extent the simple SIR model was a relevant starting point for a model of measles spread, an example of reasoning by localizing gaps in existing disease models. For the CLIMB students, the SIR model was a set of ideas that they could build upon by asking themselves "what's missing" from this model. As the group explored increasingly complex models they engaged in comparative analogical reasoning in which they made decisions about the applicability of analogous disease models to their problem. Exercises such as these led to productive brainstorming sessions in which students considered the possible factors that might influence how diseases spread such as voluntary vaccination or adding different age classes to the model.

Second, in the face of multiple factors, the group relied on simplicity to keep their model cognitively tractable. Had they added all of their ideas into their model they would have had a potentially more realistic model, but they would have limited power to explain which of those factors were most important in the dynamics. Working with simple, unrealistic models helped the students productively constrain their inquiry by helping them narrow a complex problem. The group decided that adding network structure to their model would not be a relevant or useful model feature, and they justified their decision to treat transmission as random mixing. Rose's final comment provided an intuitive justification for this decision. If transmission among children with measles is high, then random mixing is a reasonable assumption. In discussions like these the CLIMB students knowingly sacrificed reality for tractability, but also did so in ways that shaped their aim. The exclusion of network structure meant that spatial dynamics was not part of the question they wanted to ask about measles spread.

2.3 Using Models to Explore What is Possible

After spending several weeks reviewing and analyzing the literature, the group began to explore their ideas more systematically. The CLIMB mentors asked the students to choose among the many factors that they thought might be important in the human-measles system and to justify those choices. Once again, the students were engaged in selective abduction. To help them make these choices, the faculty regularly encouraged the students to employ the strategy of thought-experiment; they asked the students to imagine what effect a variable might have on the system. If they had reason to believe that it could significantly influence disease dynamics, then it might be a variable worth exploring. In the excerpt below the students were prompted to reason about how they might build age sensitivity into their model (Field notes, June 30, 2009).

> Professor R: You need to think about what you think is the biology. For example, the risk [of vaccination] for a newborn may be higher. So you might reduce the likelihood of aversion [to vaccination] as the kid gets older.
>
> …
>
> Then consider the age dependence of a kid encountering another kid who is infected. This might be small for young kids, then increase, and that will be affected by prevalence.

Professor R asked the students to consider whether differences among different age classes could possibly impact patterns of disease spread. He asked them to draw on their prior knowledge of biology to decide whether adding age structure to the model is something worth exploring. This same idea was echoed the next day by the graduate student TA, who led the students through a consideration of *how possibly* various model components could matter (field notes, July 1, 2009):

> GTA: People have looked at immigration, age-structure, and seasonality.
> If you are going to add voluntary vaccination, think about intuitively, is it going to change something?
> Lillian: It makes epidemics easier to spread?
> Sean: It will depend on the age-dependent component of the vaccination very heavily, if people who are older are more likely to volunteer to vaccinate.
> GTA: Are you going to consider what you do if you are under 18?
> Eve: I can't figure out what would be the results.
> GTA: It's okay, you don't have to come up with it right now; you just have to think about it.
> Professor R: Here's another way to think about it. What would I predict would happen that might be interesting if I added this or this to existing models? How would it change the dynamics?

What is notable about the strategies described above is that it that they are not motivated by the need to replicate reality; the students did not have data to suggest that transmission is significantly different for different age classes. They were asked to speculate on whether this model feature could possibly matter, and if so, could they imagine a plausible mechanism of action.

The CLIMB group decided that the two most important additions to the basic SIR model would be to include age-structured dynamics and the influence of human learning and decision-making on vaccination levels. Importantly, the latter decision was supported by the original article that spurred their interest. This article gave them reason to believe that human behavior could impact vaccination patterns, and hence disease spread in a significant way. The inclusion of age-structure however, was ultimately included because other modelers had included it. The students never really addressed Professor R's suggestion to do the thought experiment. Instead they relied on another modeler's contention that age-structure was an important aspect to include because it was able to generate the observed biennial dynamics. While this was a reasonable decision, it impacted the group's ability to make sense of their results because they were unable to explore the question of how or even whether different aged populations influenced disease spread in different ways.

Overall, this phase of model construction encouraged the CLIMB students to pursue a number of different strategies for reasoning creatively about how human behavior and human social organization might possibly influence disease dynamics. In making and defending their model building choices, the group was engaged in a degree of intellectual independence that is rare in school science even at the undergraduate level. The model that the group built was both novel and grounded in biologically plausible mechanisms and served as an important reasoning tool to structure the remainder of their inquiry. Nevertheless, the students did miss an opportunity to better understand the effects of age variation in their model. Instead of considering the possible mechanisms through which age could influence infection and spread, the group simply included age-structure into their model because it was more realistic.

2.4 Translating Mathematical Expressions into Biological Concepts

The process of generating formal inscriptions can create opportunities to both clarify and refine model ideas. For the CLIMB students, formally articulating their mathematical model and interpreting the results of their numerical analysis, afforded the students with the opportunity to engage in more careful, more specific reasoning strategies. As the students began to formally construct their model they were faced with the challenge of operationalizing some of the ideas that they had decided were important in terms of mathematical expressions. For example, the equation below was an early attempt at defining how exactly different forms of human learning could impact vaccination rates:

$$^{dx}/_{dt} = \mathrm{L_S}x(1-x)[\omega \mathrm{I} - 1] + \mathrm{L_E}(1 - {}^{x}/\mathrm{c})$$

This expression captures how the number of vaccinating individuals (x) changes over time as a function of two forms of learning: learning from others in dyadic social interactions at rate $\mathrm{L_S}$ and learning directly from the environment at rate $\mathrm{L_E}$. It also includes parameters that describe how sensitive individuals are to socially acquired information (ω) and how responsive they are to environmental information (c). In constructing this mathematical expression, the students referenced other similar expressions they had encountered in the literature, but tailored it so that it incorporated the ideas that they had agreed were most important. In constructing their full model, the CLIMB students had to decide both what biological ideas to include and how to represent them in a way that was both logical and resulted in a tractable expression, both forms of quite sophisticated meta-representational competence in the sense that the students had to link the symbols to their meanings (diSessa 2004).

In conducting their numerical analysis they needed to reverse this strategy to make biological sense of their model output. One of the strategies the students used in their analysis was to plot changes in an output variable (vaccination levels) over a range of different values of two input parameters: omega (ω), or the level of infection at which non-vaccinators are sensitive enough to information to switch to vaccinators, also called the sensitivity to infection, and the rate of social learning (L_S) the proportion of individuals who participate in social learning interactions. They visualized this information on a temperature plot with omega on the y-axis and L_S on the x-axis. The rate of vaccination was represented by color (in this case red corresponded to low levels of vaccination and blue to high levels.) Once again, an inscription was integral in allowing them to reason with their model. In the process of interpreting this plot, the parameter ω emerged as a predictor of vaccination levels. The CLIMB mentors instructed the students to spend some time making sense of this parameter so that they could explain exactly what this parameter meant conceptually. The following excerpt is an example of one such conversation (field notes, August 10, 2009).

> Nina: What is ω [omega]? ω is the sensitivity to infection right? Wait, that doesn't make sense. The less sensitive you are, the longer it takes to react?
> Rose: Pick points in the regions and move up the plot.
> Nina: As ω is increasing people are becoming more sensitive. I don't understand the red.
> Rose: Red shows that there is no vaccinators basically. L_S is the rate that people switch. It's κ [kappa] times p.
> Nina: p is the people who engage in pairwise interaction.
> Lillian: ω is the sensitivity of the proportion of those playing the social learning strategy.
> Nina: What?
> Romy: Of those who are doing social learning, how sensitive are they?
> Nina: The red part is saying people are not sensitive enough? That means they are not going to vaccinate?
> Rose: Yeah, so there are no vaccinators and you get Schenzle-like dynamics [oscillations].
> Romy: The reason the region is red is because of ω, not κ. The imitation dynamics do not matter. Any ω smaller than 5,000 should be red.

In the above excerpt the students are making sense of the concept of sensitivity to infection, which is represented in their model by the symbol ω and how it helps them understand their model output. As Romy articulates at the end of this discussion, this plot allowed them to see that this concept of sensitivity to information was an idea they needed to think more about. This plot became the focus of their analysis, and while it made no precise predictions, the faculty found it exciting because it suggested to them that this parameter ω, this concept of sensitivity of infection, might be important enough to inspire future investigations (field notes, August 10, 2009).

> Professor S: I am totally excited by that plot! You have gotten to the point where you can create that figure, and you even have a story about why you have it. That is what you are hunting for, the types of figures where there is a clear story. It is interesting because something happens at the critical point where ω is 5,000, and there is a reason for why it happens that you can examine.

In summary, conceptual and representational work was important both during model construction and model analysis. In building their model the students had to carefully define the elements they were including in their model and make decisions about how to represent relationships among model elements mathematically. Then, during analysis, the students had to interpret the mathematical results they generated in terms of biological concepts. In the process they identified omega, a parameter, but also an *idea*, or concept, that emerged from their work. The ideas that the students were able to draw out of the temperature plot formed the core of their results, yet these ideas were not specific predictions about how a measles infection could be expected to behave in a real population. Nor were these ideas sufficient to explain the specific data patterns of measles incidence that were the inspiration for their work. Rather, they were qualitative descriptions that identified conceptual entities that have the potential to play an important role in this complex system.

2.5 The Importance of Matching the Data

Early on in the project the CLIMB mentors highlighted the value of building a model that could speak to actual data. In May, the group was still deciding which disease they might want to model, and they were given the following advice (field notes, May 8).

> Professor R: In terms of the flu, spend some time thinking about what there is to model that would be gratifying. What are you trying to predict that you could use as the basis of your model? Is there some pattern that you are trying to explain?
> Professor R: Or look at the theory to see what people are trying to explain. If they are trying to make a prediction, then what is it they are trying to predict and why? There could be a previously documented case; there could be some data there. If I change this what is the impact? That is how you can use a model to do experiments. Make a list, and tell us what each of those papers is trying to explain, and what the paper did to try to improve on the power to explain some pattern. This is all about patterns in nature. What makes science, science is there is some pattern that we are trying to explain and understand. If you don't know anything about those patterns then it is hard to come up with a framework.

Here Professor R is advising the students to use data patterns as the inspiration for their modeling project. He wanted them to ground their inquiry in an empirical pattern that had already been identified. Ultimately, it was this discussion that led the students to abandon the idea of modeling the flu virus in favor of measles. As they explained to their mentors a few days later after looking for data patterns in the scientific literature (field notes, May 12, 2009):

> Romy: There were just too many variables for what causes flu. And the data are all local, not global.
> GTA: So the difficulty of the flu made you decide to go to MMR?
> Lillian: After Professor R picked it apart we needed a new question.
> Romy: He basically said, what if there is not a pattern?
> Rose: We found a survey on flu, and it said that of people that refuse vaccination, only 7% said it was because it didn't work last year.
> Romy: And the vaccination rate for flu is fairly constant - around 30%.
> Rose: The biggest reason people don't vaccinate was economic.

The CLIMB students had been interested in how vaccinator behavior in one year might influence flu dynamics in subsequent years, but this initial lack of data shifted the group's attention to measles. The group was able to find data from the United Kingdom that described how vaccination levels dropped in 2003, presumably as a result of a vaccine scare linking vaccination to autism. These data were what initially inspired the group to focus on the effects of voluntary vaccination on disease dynamics. However, as we have already described, the focus of the project shifted from explaining this graph to exploring ideas about human vaccination behavior and disease spread more generally.

As the project drew to a close, the students got anxious for more precise results, and they proposed to use their model to replicate the actual pattern of disease incidence in the UK as a result of the 2003 vaccine scare. They stated that they hoped they would be able to predict precisely how long it would take for the population to regain pre-scare vaccination levels. The students proposed that the answer to the question, how much time will pass after vaccine scare before we reach pre-scare levels, would be an important main result. But the faculty saw their project as productive in a different way (field notes, August 25, 2009):

> Professor R: This question strikes me, as it's stated, as almost trivial because we actually know how much time will pass. What motivated you to do this, it seems to me, is trying to understand the processes that are likely to influence what we are talking about. Age structure and how information about vaccine scares is transmitted absorbed and used - those are two likely factors as suggested from other models.

Professor R refocused their attention on their theoretical contribution, that of exploring a model that included both age-structure and voluntary vaccination. Still the students seemed anxious about this as an important result (field notes, August 25, 2009):

> Kevin: We are unsure of what to show. What is worth showing? I don't know if the bifurcation is worth showing.
> Professor S: Are you spending time reproducing the UK dynamic?
> Rose: It's been hard. And we recently found age-specific notification of the disease, and we tried to incorporate that as initial conditions. But we are still having problems with the vaccination being low and not that many infected.
> Professor S: If you can't replicate it, then you can't. Several articles couldn't do that. Focus on the bifurcation plot.

Ultimately, in the final presentation of their work, the students emphasized the importance of their model in exploring the possible roles of social and environmental learning on vaccination levels and disease dynamics and, for the most part, accepted that making accurate predictions was not an outcome of their modeling. As one student put it, when asked about the major contribution of their work: [Interview, September, 24, 2009]

> Romy: Honestly, I feel like even though we originally wanted to talk about vaccine scares and stuff, I think incorporating environmental learning and like saying like… there are different ways that people learn, that people haven't necessarily examined through modeling or you know empirical data about how people learn, or how vaccines…. So like here's two ways that can affect how people learn about vaccines and how it could affect the vaccination rates and infection rates.
> …
> Interviewer: So, what do you think are the strengths and weaknesses of the model that you ended up with?
> Romy: Mm strengths are it's, it obviously does, I don't know, that actually the model suggests something, like it works I guess, like it incorporates things that…both it incorporates really all the things relevant to measles and MMR vaccination, at least like empirically and what we think as scientists or whatever. But, and it seems to give qualitatively really interesting results. On the flip side of that, as a weakness it doesn't give quantitatively correct results.
> Interviewer: Mm hmm. So how problematic is the quantitative mismatch from your point of view?
> Romy: Um, it would have been nice, to get, you know, something that reflected real data. But at the same time,…I feel like it's something that can be tweaked with the parameter values of like our constant values. So maybe our constants are not necessarily what the contact rates are out there right now or something. So it's something that can change. I don't think it's a huge, huge issue.

The above examples show that the distinctions that philosophers of biology have chosen to highlight were relevant in this educational setting. Throughout their inquiry the CLIMB students used models to help them frame their research question, explore ideas, and refine their conceptual understanding. Their attempt to match their model output to the existing empirical data was arguably one of the least productive aspects of their project. We do not highlight this to suggest that evaluating models empirically is not an important part of science. Ultimately, models, including the model built by the CLIMB students, need to answer to empirical data. The point here is that given the exploratory nature of their project, matching to data did not need to be the highest priority. As the faculty mentors stated, predicting how long it would take for the vaccine to reach pre-scare levels was not something their model was built to do. If they had wanted to build a model that could make this sort of prediction they would have had to make different decisions in the construction of their model, namely focusing much more carefully attending to how time was represented in their model. In focusing on this match (or lack of match) the students were applying inappropriate epistemic criteria. Moreover, they were overlooking the potential of their model to introduce new concepts worthy of further exploration. To reiterate the point made by Levins: “modelers always must keep in mind that the utility of their construct depends on the particular purpose for which it was built” (1985, p. 8, as cited in Odenbaugh 2005, p. 253).

3 Discussion and Implications

The examples presented from the CLIMB program were chosen to highlight how attention to studies of scientific practice can inform our understanding of learning environments in science. More specifically, we argue that explicit attention to the different uses of models can engage students in meta-strategic reasoning. For example, the CLIMB group began their project by evaluating models in the scientific literature. They did so by both considering the kinds of data the models could explain (i.e. could the models replicate the kinds of biennial cycles typical of measles incidence?), and also by evaluating the theoretical contributions of each model (i.e. did the models include ideas about age-structured populations?). They used these models as a baseline against which to compare their own ideas. As a group they discussed both whether the models were able to account for the existing data patterns of disease dynamics as well as considering what possible mechanisms might be missing from these models.

The students’ analysis of existing models provided the backdrop against which they defined and refined their approach to model construction. Constructing a model required a great deal of creative thought. The students used analogical reasoning to incorporate and modify aspects of others’ models into their own. They also generated a list of candidate mechanisms and reasoned abductively to choose those that were both plausible and tractable. In formally articulating the relationships in their model, the students consciously left out realistic details in order to gain analytic power. Finally, in the analysis of their ultimate model, the students did not examine whether the model was able to fit the data. Rather, with guidance from the CLIMB mentors, they used their model output to explore qualitative trends that would help them begin to build a conceptual understanding of the potential relationships among model parameters. During this phase of their modeling project the CLIMB students designed inscriptions that allowed them to visually explore how various model components influenced the model output. The resulting temperature plots showed how values of key variables such as vaccination rate and disease incidence fluctuated over different parameter values and allowed the group to speculate on the relative importance of different forms of human learning on the spread of disease.

The CLIMB experience provided a context in which students were engaged in creative and productive forms of model-based reasoning, using the model to explore novel ideas and to consider as yet untested possibilities. The CLIMB mentors, who were biologists and modelers, were largely responsible for helping the students find and follow productive lines of inquiry. At times this meant abandoning particular aims (i.e. constructing a model that can replicate existing data patterns) in favor of new aims (i.e. exploring parameter values to consider their possible influence on the system). These shifts exemplify the pragmatic nature of scientific practice and how modeling is best conceived of as a set of strategies that can serve a variety of epistemic aims.

It was the objective of the CLIMB program to train future scientists in modeling so that they would be more likely to uptake and carry these strategies with them as they move into positions as professional biologists or mathematicians. It is our contention that students are more likely to do so if they have an appreciation for when, how and why modeling can inform scientific research. Smith et al. (1997) describe a similar program involving groups of graduate students engaged in mathematical modeling in the context of ecology. In reflecting on this experience and the successes and failures of the participants, the program mentors, a biologist and a modeler, articulated what they saw as the most productive approaches to using models in biology. The aims they list: exploring dynamic patterns, ruling out possibilities, suggesting new areas for research, and developing new knowledge (p. 467), closely mirror those presented by Odenbaugh, most notably in their emphasis on models as "epistemological objects", rather than representations that mimic reality. The reflections of these scientists as well as our observations of students engaging in modeling in practice further emphasize the importance of training future scientists, and future biologists in particular, to understand the aims and purposes of modeling so that the cognitive strategies of modeling will be relevant and available to them.

The CLIMB program was designed as a science traineeship, a context very different from science classrooms in many ways. We now turn our attention to the ways in which the lessons we have learned from this research might apply to science education in K-12 classrooms.

3.1 Contextualizing Modeling and Model-Based Reasoning in Classrooms

Understanding of the diversity of ways in which modeling supports scientific inquiry can help educators make decisions about how exactly to incorporate models into science curricula (c.f. Lehrer and Schauble 2006). Moreover, if educators make a conscious effort to cultivate a variety of uses for models in science, we believe that students will have at their disposal a greater number of tools for reasoning and a greater appreciation for the dynamic and context-dependent nature of scientific practice.

We build on the work of scholars like Schwarz et al. (2009), who have outlined a learning progression for teaching modeling that emphasizes the importance of integrating modeling practice with an understanding of the nature and purposes of modeling, which they refer to as *metamodeling* knowledge. We agree that this integration of practice and purpose is the key to successfully teaching modeling. However, we wish to add to the general account presented by Schwarz et al. an appreciation that modeling, as practiced by scientists, is always grounded in a specific context (Lehrer et al. 2008; Smith et al. 1997). If modeling is meant to reflect authentic scientific reasoning and authentic features of the nature of science, there needs to be more clarity surrounding both what it means to practice modeling and the specific purposes that modeling can support.

In the science education literature modeling tasks often emphasize the need for students to fit real systems or match some existing target model (e.g. Harrison and Treagust 2000; Gobert 2005; Schwarz and White 2005; Lehrer et al. 2008; Schwarz et al. 2009). Our contention is not that these are unimportant roles for models, but that by emphasizing the need for models to fit with reality they underemphasize other important forms of model-based reasoning. If modeling is understood *only* in relation to explaining and predicting (c.f. Koponen 2007), we believe that important aspects of productive scientific thought will be neglected. We saw some indications of this in the CLIMB project. When the CLIMB students were fixated on the idea that a model should be predictive, they began to try to fit prediction into their inquiry, which, in effect, distracted them from potentially more meaningful theoretical contributions of their work.

Part of the solution, as we see it, is to be explicit about the fact that modeling can play different roles in different contexts. We argue that a key aspect of metamodeling knowledge is being able to distinguish among different purposes and forms of modeling in a contextually relevant way. This point has been made about inquiry more generally (e.g. Hammer et al. 2008; Ford and Forman 2011). As notions of what counts as inquiry continue to expand (Grandy and Duschl 2007), it is important to clarify the context of their use. As alluded to by Tang et al. (2010), it is not enough to expand the definition of inquiry to include a long list of possible activities. Removed from the context of an actual problem or question, the reasons for choosing one activity over another are obscured—the same applies to the strategies of modeling.

As part of their learning progression, Schwarz et al. (2009) include as a goal that students recognize that different models have different strengths or weaknesses, and that sometimes multiple models might be needed to account for complex phenomena. They do not, however, specify when multiple models might be needed or what kinds of strengths and weaknesses such models might have. Ultimately, these kinds of decisions can only be made in the context of practice by choosing a model that can best support the relevant aims. Consider for example an excerpt from Schwartz et al. in which a group of students are discussing their ideas about the uses of models. Schwarz et al. (2009, p. 646) present this as an example of confusion about the use of models and write that,

> … sometimes students struggled with using a model to make predictions. For example, in one conversation, two students suggested that "your model can be your best guess," while others disagreed, stating that "to make it accurate… you can't really model it without knowing." Another student pointed out that "like your prediction," you can build a model representing "stuff that you think is true," and then "after you get the information you can make the model then" [Interview #7, Lesson 8].

Several different ideas about models are raised in this short excerpt, and while we cannot attribute a deep understanding of these ideas to any of these students, we view their statements as the fodder for a potentially sophisticated discussion about the multiplicity of roles that models can play. The idea that a model is a "best guess" emphasizes the important role that models play, often early in an inquiry, to articulate hypotheses and organize prior knowledge. Another student emphasizes the importance of a model holding "accurate" information, another legitimate role for models, in particular when a model is being used to explicate the details of how some process works. Another student emphasizes a slightly different role for models by suggesting that the model represents a testable hypothesis, "a prediction" or the "stuff you think is true". This stance is a potentially productive starting point for motivating the need for empirical studies or other kinds of model evaluation. While there are only fragments of ideas in this short excerpt, it hints that even middle school students, after working with models in their classrooms, are able to pick up on the different roles that models can play. This points to the potential of being explicit about encouraging and legitimizing these intuitive ideas of students in a rich modeling context.

3.2 Developing the Strategies of Modeling

CLIMB was an advanced program for undergraduates, and some might argue that younger students do not have the degree of epistemological sophistication to use models in a variety of different ways. For example, much has been made of Grosslight et al.'s (1991) characterization of students as epistemologically naïve about models. However, it is important to remember that these results represent students' unscaffolded, decontextualized ideas about models and do not reflect what students can do with models in more supportive or appropriately contextualized settings. Our interpretation of Grosslight et al.'s findings is that the student responses they collected reflect the simplistic definitions of models that are typically presented in school, but are not representative of what students can do when engrossed in more authentic modeling activities.

Emerging evidence suggests that even young children have the capacity to engage in abstract and theoretical discussions, but that the consistency with which they are able to do so is largely context-dependent (e.g. Hammer et al. 2008 and references therein). A study by Metz (2010) investigated second grade students' success at independently developing and implementing an inquiry project. Metz's analysis shows that these students were able to engage in fairly sophisticated epistemological themes. Metz concluded that,

> … under particular instructional conditions, even second graders can take their investigation—including the plan they devised to investigate their question, the implementation of the plan, and the inferences they based on their data—as an object of critical reflection. In other words, they viewed their investigation as a tool, problematic in specific ways, to try to address their question (p. 283).

However, Metz's analysis revealed that while children were able to identify theoretical uncertainties in their investigations, they had few, if any, strategies for dealing with that uncertainty. She suggests that this was due in part to the lack of curricular support for developing theoretical understanding (Metz 2010, p. 267). Like Metz, we believe that the root of this challenge is more likely to be instructional rather than developmental.

A recent study by Pluta et al. (2011) provides further evidence that is specific to the context of modeling. Without significant instruction, middle school students were able to generate a variety of epistemic criteria for evaluating scientific models. When prompted to list the features of a good model, students responded with epistemic criteria such as communication, explanatory power, and fit to data. However, the most common responses had to do with the amount of detailed information presented in the models, suggesting that students were thinking of models primarily as useful for conveying information, much as a textbook diagram would. One weakness of this study is that this model evaluation task was framed without reference to a particular problem, question or aim. It is difficult to tell how students would have interpreted the task of "making a list of the most important characteristics of good (rather than bad) models" without a specific purpose in mind. Looking closer at the nature of that task, students were asked to evaluate a variety of static representations of models including diagrams, pictures and text similar to what they might see in textbooks (p. 500). Given that in the context of this task, students were interacting with final form models, it is not surprising that many students describe models as tools that help communicate ideas, rather than objects to support scientific inquiry. Nevertheless, this study suggests that students do have some resources for thinking about using models in a variety of ways and supports the argument that the aim of instruction should be to reinforce and refine these ideas with reference to particular scientific aims. We caution against teaching epistemic criteria to students as a normative list of characteristics of "good models." Such criteria cannot be universally applied to all models, and while we agree that ultimately model accuracy is a primary aim in science, it is not always the most appropriate or relevant aim for scientific reasoning.

We envision metamodeling knowledge as an understanding of the variety of aims that models can serve in scientific practice and an ability to apply the epistemic criteria that are relevant to those aims. In this way our version of metamodeling knowledge is similar to Ford's notion of a "grasp of practice" (2008; Ford and Forman 2011). Developing a grasp of practice means that students begin to know the criteria that can be used to evaluate knowledge claims, use those criteria in the context of their own practice and understand how those claims are used in the scientific community (Ford 2008). Like Ford, we are advocating for a more flexible understanding of scientific practice and a version of science learning that helps students understand and reflect on why they are doing what they are doing rather than learning a normative set of rules.

We acknowledge that developing a meta-understanding of the roles of modeling in scientific practice will require repeated and extended opportunities to practice modeling and reflect on that practice. We therefore argue that it is important to develop curricula that challenge students to consider the purpose of the modeling activity they are engaged in for a range of grade-levels. The kinds of challenges that different students might face are open to investigation, since as far as we know students, especially younger students, are rarely asked to consider the epistemic aims of their scientific practice. With this in mind, we offer some general suggestions for how we might begin to introduce these ideas into science classrooms over a broad range of grade-levels, and acknowledge that this is an open area for research.

3.3 Implications for Classroom Enactments of Modeling

We suggest that curriculum developers begin by considering the kinds of reasoning they want to support and then designing modeling tasks that will best support that kind of reasoning. At times this goal might be to teach students how to generate and test accurate explanations and predictions. At other times, modeling can be used to guide the initial phases of developing a research question, identifying knowledge gaps or locating problems of interest. In this sense our conception of modeling is similar to Windschitl et al.'s (2008b) definition of model-based inquiry—a term that suggests that modeling, is not simply a method in science, but rather a collection of methods that permeate every stage of the inquiry process from organizing and understanding a problem to testing and evaluating explanatory models, and proposing arguments about the adequacy of a particular model and how it might be revised, expanded and improved. Designing curricula that embody the complexities of modeling as practiced by scientists is a significant challenge. It is much easier to implement simple algorithms than support the contingent and unpredictable nature of authentic model-based inquiry. However, oversimplified, epistemically inaccurate, and reasoning-poor conceptions of scientific practice have dominated science education for too long. We maintain that we can no longer afford to shy away from this challenge, and we see several reasons to be hopeful that such change is possible.

In the interest of sparking a conversation about instructional implementation, Windschitl et al. (2008a) presented a framework for supporting model-based inquiry that revolves around a series of four conversations that they argue should form the backbone of classroom activity. They are: *1. Organizing What We Know and What We Want to Know, 2. Generating Testable Hypotheses, 3. Seeking Evidence, 4. Constructing an Argument.* For each of these conversations the authors propose a list of potential "starter-questions" that could form the basis of classroom discussions that accompany modeling activities. Many of these questions are aimed at supporting critical reflection about the use of modeling in the classroom, and we encourage readers to consult the original paper for a more detailed discussion of their use (Windschitl et al. 2008a, p. 17–19). For our purposes, we explore a few examples that have the potential to help instructors and students begin to address some of the epistemological themes relevant to those we have raised in this article.

What are we leaving out of our model and why? (p. 17) is a question that could form the basis for the kinds of analogical and abductive reasoning we saw CLIMB students engage in as they struggled to make decisions about what to include and what to leave out of their model and to justify those decisions. The question provides a way to go beyond stating that models are simplifications to support important conversations about *when* simplifications are appropriate and *why*. This question suggests the need for conversations about using simple models in strategic ways. We imagine that a follow-on to this question might be: *Why, and under what circumstances is it okay to leave things out of a model?*

What do we already think we know about this situation, process, or event? Which of the elements of what we know are based on observations? Inference? Other sources of information? (p. 17.) These questions focus attention on the formation of *how possibly* explanations and form the basis for conjectures that inform model construction. In the context of these sorts of questions the model is a tool for organizing what is known, what is possible, and for localizing gaps in need of further inquiry. Just as the CLIMB students relied on their evolving model to hold their ideas, model building in the classroom can support discussions that are explicitly epistemological in nature. Such discussions can potentially make use of students' prior knowledge by converting it into a potential object of investigation.

When we look at our tentative model and consider the question we want to ask, what would our models predict? (p. 18.) To this we add: *Do we expect our model to produce accurate predictions? Why or why not?* These questions are necessary precursors to asking students to make predictions. They first ask students to consider what those predictions might look like, and then ask them to consider the accuracy of those predictions. Such questions could support thought experiments much like the ones the CLIMB mentors asked the CLIMB students to conduct, which were crucial to the decision making processes involved in model construction. Part of choosing how to construct a model involves imagining how the output of the model might change as a result of the choices made during construction. Such conversations could also spur discussions about what "fitting the data" really means. *How close must the fit be? Is the lack of fit due to our model or our experimental procedures?*

What does it mean if our model cannot produce accurate predictions? How consistent and coherent is our final explanation with the phenomenon of interest? (p. 19.) Finally, it is important that instead of telling students that models are predictive and explanatory, that we engage them in a discussion of what predictive accuracy and explanatory power really mean. As the CLIMB project showed us, predictive models are not necessarily always very informative. Our model may not be completely consistent with the details of reality, but *can it still help us make better sense of the world?* Questions such of these can stimulate explicit discussions about the nature of prediction and explanation, when the two are in line, and what it means when they are not.

While these kinds of conversations can be productive starting points, ultimately the success of model-based instruction hinges on embedding these discussions within authentic and well-designed tasks—tasks in which the model plays a key role in helping students reason scientifically and in which that role has been made clear for the students. The strategies of modeling in science are most powerful when linked to an authentic scientific aim.

4 Conclusion

The aim of this paper has been to expand the conversation about the nature of modeling in science education by drawing on the insights of philosophers of biology. Our hope is that what we add to the conversation is a sensitivity to the context of model use both in terms of the discipline and the phenomena under investigation. We believe our analysis of the work of the CLIMB students using Odenbaugh's framework makes a strong argument for the need to engage students in extended authentic modeling experiences, which have the potential to provide the opportunity for students to encounter a range of scientific reasoning strategies situated within the context of specific epistemic aims. We further believe that attention to the relevant details of context can help scaffold teachers and students in deciding when and why it is appropriate to engage in the various different practices that constitute modeling or model-based inquiry. In our future work we plan to develop these ideas in ways that can help us translate the insights of philosophers, historians and psychologists of science into productive conversations and practices in science classrooms. We hope that others will take up this challenge as well.

Acknowledgments This research was supported by the National Science Foundation Interdisciplinary Training for Undergraduates in Biological and Mathematical Sciences (UBM) program under Grant No. 0531935. We would like to acknowledge CLIMB mentors and the students of the 2008–2009 CLIMB cohort for allowing us to bare witness to their learning process. We also thank Richard Lehrer and four anonymous reviewers for their helpful suggestions for improving this manuscript.

References

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36*(2), 421–441.

Bechtel, W., & Abrahamsen, A. (2010). Complex biological mechanisms: Cyclic, oscillatory, and autonomous. *Handbook of the Philosophy of Complex Systems, 10*, 1–26.

Chinn, C., & Malhotra, B. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*(2), 175–218.

Clement, J. (1989). Learning via model construction and criticism. In G. Glover, R. Ronning, & C. Reynolds (Eds.), *Handbook of creativity: Assessment, theory and research* (pp. 341–381). New York: Plenum Publishers.

Clement, J. C. (2000). Model based learning as a key research area for science education. *International Journal of Science Education, 22*(9), 1041–1053.

Cooper, G. J. (2003). *The science of the struggle for existence: On the foundations of ecology*. Cambridge: Cambridge University Press.

Darden, L. (1991). *Theory change in science: Strategies from Mendelian genetics*. New York: Oxford University Press.

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

diSessa, A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction, 22*(3), 293–331.

Downes, S. (1992). The importance of models in theorizing: A deflationary semantic view. In *PSA: Proceedings of the biennial meeting of the philosophy of science association* (Vol. 1, pp. 142–153). Chicago: The University of Chicago Press.

Ford, M. J. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education, 92*(3), 404–423.

Ford, M. J., & Forman, E. A. (2011). Redefining disciplinary learning in classroom contexts. *Educational Research, 30*(2006), 1–32.

Giere, R. N. (1988). *Explaining science: A cognitive approach*. Chicago: University of Chicago Press.

Giere, R. N. (2004). How models are used to represent reality. *Philosophy of Science, 71*, 742–752.

Gobert, J. D. (2005). The effects of different learning tasks on model-building in plate tectonics: Diagramming versus explaining. *Journal of Geoscience Education, 53*(4), 444–455.

Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy, 21*(5), 725–740.

Grandy, R., & Duschl, R. (2007). Reconsidering the character and role of inquiry in school science: Analysis of a conference. *Science & Education, 16*, 141–166.

Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching, 28*, 799–822.

Hammer, D., Russ, R., Mikeska, J., & Scherr, R. (2008). Identifying inquiry and conceptualizing students' abilities. In R. Duschl & R. Grandy (Eds.), *Teaching scientific inquiry* (pp. 138–156). Rotterdam: Sense Publishers.

Harre, R. (1986). *Varieties of realism: A rationale for the natural sciences*. New York: Blackwell.

Harrison, A., & Treagust, D. (2000). A typology of school science models. *International Journal of Science Education, 22*(9), 1011–1026.

Hodson, D. (1996). Laboratory work as scientific method: Three decades of confusion and distortion. *Journal of Curriculum Studies, 28*(2), 115–135.

Hodson, D. (1998). Science fiction: The continuing misrepresentation of science in the school curriculum. *Pedagogy, Culture and Society, 6*(2), 191–216.

Hughes, R. I. G. (1999). The Ising model, computer simulation, and universal physics. In M. Morrison & M. S. Morgan (Eds.), *Models as mediators: Perspectives on natural and social science* (pp. 97–145). Cambridge: Cambridge University Press.

Keller, E. F. (2000). Models of and models for: Theory and practice in contemporary biology. *Philosophy of Science, 67*(S1), S72–S86.

Knorr-Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge: Harvard University Press.

Knuuttila, T. (2005). Models, representation, and mediation. *Philosophy of Science, 72*(5), 1260–1271.

Koponen, I. (2007). Models and modelling in physics education: A critical re-analysis of philosophical underpinnings and suggestions for revisions. *Science & Education, 16*(7), 751–773.

Latour, B. (1990). Drawing things together. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 19–68). Cambridge, MA: MIT Press.

Laubichler, M. D., & Müller, G. B. (2007). *Modeling biology: Structures, behavior, evolution*. Cambridge: MIT Press.

Lehrer, R., & Schauble, L. (2005). Developing modeling and argument in the elementary grades. In T. A. Rombert, T. P. Carpenter, & F. Dremock (Eds.), *Understanding mathematics and science matters (Part II: Learning with understanding)*. Mahway, NJ: Lawrence Erlbaum Associates.

Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 371–388). New York: Cambridge University Press.

Lehrer, R., Schauble, L., & Lucas, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development, 23*, 512–529.

Levins, R. (1966). The strategy of model building in population biology. *American Scientist, 54*(4), 421–431.

Lloyd, E. A. (1994). *The structure and confirmation of evolutionary theory*. Princeton: Princeton University Press.

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

Magnani, L., Nersessian, N. J., & Thagard, P. (Eds.). (1999). *Model-based reasoning in scientific discovery*. New York: Kluwer.

Matthews, M. (1994). *Science teaching: The role of history and philosophy of science*. New York: Routledge.

May, R. (1973). *The stability and complexity of model ecosystems*. Princeton: Princeton University Press.

Metz, K. (2010). Children's understanding of scientific inquiry: Their conceptualization of uncertainty in investigations of their own design. *Cognition and Instruction, 22*(2), 219–290.

Morrison, M., & Morgan, M. (1999). Models as mediating instruments. In M. Morrison & M. Morgan (Eds.), *Models as mediators: Perspectives on natural and social science* (pp. 10–37). Cambridge: Cambridge University Press.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. N. Giere (Ed.), *Cognitive models of science* (pp. 3–44). Minneapolis: University of Minnesota Press.

Nersessian, N. J. (1999). Model-based reasoning in conceptual change. In L. Magnani, N. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery*. New York: Kluwer/Plenum Publishers.

Nersessian, N. J. (2002). The cognitive basis of model-based reasoning. *The cognitive basis of science* (pp. 133–153). Cambridge: Cambridge University Press.

Nersessian, N. J. (2008). Model-based reasoning in scientific practice. In R. Duschl & R. Grandy (Eds.), *Teaching Scientific Inquiry: Recommendations for Research and Implementation* (pp. 57–79). Rotterdam, the Netherlands: Sense Publishers.

NRC. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies 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.

Odenbaugh, J. (2009). Models in biology. In E. Craig (Ed.), *Routledge encyclopedia of philosophy*. London: Routledge.

Osbeck, L., Nersessian, N. J., Malone, K. R., & Newstetter, W. (2010). *Science as psychology: Sense-making and identity in science practice*. New York: Cambridge University Press.

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching, 48*(5), 486–511.

Rudolph, J. (2005). Epistemology for the masses: The origins of "The Scientific Method" in American schools. *History of Education Quarterly, 45*(3), 341–376.

Schwarz, C., Reiser, B., Davis, E., Kenyon, L., Acher, A., Fortus, D., et al. (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.

Schwarz, C., & White, B. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction, 23*(2), 165–205.

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.

Tang, X., Coffey, J., Elby, A., & Levin, D. (2010). The scientific method and scientific inquiry: Tensions in teaching and learning. *Science Education, 94*(1), 29–47.

White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and instruction, 10*(1), 1–100.

Wimsatt, W. C. (1987). False models as means to truer theories. In M. Nitecki (Ed.), *Neutral models in biology* (pp. 23–55). New York: Oxford University Press.

Wimsatt, W. C. (2002). Using false models to elaborate constraints on processes: Blending inheritance in organic and cultural evolution. *Philosophy of Science, 69*(s3), S12–S24.

Windschitl, M., Thompson, J., & Braaten, M. (2008a). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, 1–27. doi: 10.1002/sce.

Windschitl, M., Thompson, J., & Braaten, M. (2008b). How novice science teachers appropriate epistemic discourses around model-based inquiry for use in classrooms. *Cognition and Instruction, 26*(3), 310–378.

Wynne, C., Stewart, J., & Passmore, C. (2001). High school students' use of meiosis when solving genetics problems. *International Journal of Science Education, 23*(5), 501–515.