

Seeking Diversity in Mathematics Education: Mathematical Modeling in the Practice of Biologists and Mathematicians

ERICK SMITH¹, SHAWN HAARER² and JERE CONFREY³

¹ College of Education MIC 147, University of Illinois at Chicago, 1040 W. Harrison, Chicago, IL 60607, U.S.A.; ² 197 Hillside Avenue, Rochester, NY 14610, U.S.A.; ³ Department of Education, Kennedy Hall, Cornell University, Ithaca, NY 14853, U.S.A.

ABSTRACT. Although reform efforts in mathematics education have called for more diverse views of mathematics, there have been few studies of how mathematics is used and takes form in practices outside of mathematics itself. Thus legitimate diverse models have largely been missing in education. This study attempts to broaden our understanding of mathematics by investigating how applied mathematicians and biologists, working together to construct dynamic population models, understand these models within the framework of their perspective practices, that is how these models take on a role as 'boundary objects' between the two practices. By coming to understand how these models function within the practice of biology, the paper suggests that mathematics educators have the opportunity both to reevaluate their own assumptions about modeling and to build an understanding of the dialectic process necessary for these models to develop an epistemological basis that is shared across practices. Investigating this dialectic process is both important and missing in most mathematical classrooms.¹

INTRODUCTION

The rationale for investigating mathematics across disciplines is rooted in the primary constructivist contention that we build our mathematical understandings and meanings through our experiences and the resolution of problematic situations. This claim is fundamental to claims for mathematical diversity for it suggests that as we look at the experiences and problems of different disciplines, we should find different ways of seeing, doing, and understanding mathematics, that is we should find opportunities for seeing legitimately new mathematics. This opportunity is heightened when we can observe collaborative efforts across disciplines for the differing perspectives serve to make negotiations of the uses and meanings of mathematics more visible and open for discussion. Too often there is a tendency to see mathematics as an independent body of knowledge which can simply be applied in different situations, but which is independent of the problems and issues of any particular situation. We suggest that looking at the activities of mathematicians and biologists while working collaboratively on developing biological models provides an appropriate lens for seeing the differing developments of mathematics ideas in the two practices even though they develop and use the 'same' models in their collaborative efforts.

This perspective is explored by looking at the mathematical activity

involved in building models of population dynamics in a particular ecosystem. We focus on an advanced graduate-level biological modeling class at Cornell University which was taught by Dr. Carlos Castillo-Chavez (Carlos) and Dr. Sara Via (Sara).² Students in the course came from the fields of both biology and applied mathematics. Working in groups, they developed various population models in relation to a given ecosystem. This paper follows two groups of students in the development of their models, discusses the procedures by which their models were developed, and reflects on the reactions of individual group members in relation to their various practices. We then examine the perspectives of Carlos and Sara as a lens for discussing these processes.

One of the arguments we develop in the paper is that the intended role of these models in the practice of biology is usually *not* to be a 'reflection of reality'. Thus their role is different from what we think of in physics, for example, where the models and the science often seem inseparable, or in mathematics classrooms where the goal is typically to create a mathematical form (i.e., equation, table, and/or graph) which is a representation of the problem statement. Instead, we argue that the relationship between the mathematical models and biological knowledge is much more dialectic with each acting as a pointer or a lens for viewing and informing the other. In the final section we discuss possible implications of this view of mathematics for alternative approaches in the mathematics classroom.

WHAT IS A MODEL?

A common conception of a mathematical model is as a mathematical system that in some sense mimics some natural system. One expects that the output of the model will, within reasonable limits, be representative of actual events in the system. Doucet & Sloep (1992) offer a theoretical and practical description of what is involved in creating such a model, using what they call a 'semantic view of theories' as a framework for describing models. Within this framework, three components are required to establish a complete model. These are: 1) a formal or theoretical model; 2) a natural system; and 3) a theoretical hypothesis – a statement that 'claims that the theoretical model that we have defined actually *is* a model of the natural system of interest' (p. 282, authors' italics). Using an example of a chemical spill in a lake, the authors lead readers through a quite elaborate argument which helps establish the idea that there can never be a complete fit between model and system. In a later section, the authors describe the complexities of creating a model of the growth of a particular animal. Their discussion of this process reflected in many ways the concerns of the biologists in this study, for they quickly show that the complexity of even such a single process requires the modeller to make trade-offs that almost assure that the model will not accurately mirror the actual processes in nature. They describe four characteristics of a good

model, precision, realism, generality, and mathematical simplicity. However, they note: 'Realistic models are hardly ever simple; a high degree of precision often precludes generality; etc.' (p. 305). One makes choices based on one's purposes: 'models aimed at explanation tend to be the small ones, whereas models aimed at prediction tend to be the larger ones' (p. 306).

Although I have left out much of the detail of their arguments, they present a compelling case that actual models will always be based on the purposes and intentions of the modeler and that the process of creating a model suited to these intentions involves a matter of trade-offs. However, the goal is to have the output of the model be as close as possible to the natural system. This is stated in their definition of the theoretical hypothesis and this theme repeats itself at various places. For example, the discussion of how one makes decisions about the value of the model is based on how well the outcomes predicted by the model fit with the natural system being modeled. Thus despite the problems inherent in the process, a good model is, for them, a model of some system in the natural world. This is, to a large extent, compatible with our own conception of a model as we began this project.

Kinds of Models

In our initial meetings with Carlos, he described several ways to classify models. Two distinctions are important in understanding the notion of a 'good' biological model:

The first distinction is between *descriptive models* and *theoretical models*. A descriptive model, according to Carlos, attempts to describe 'everything' in a biological system, it attempts to 'mimic' reality within a restricted biological framework. A theoretical model, on the other hand, is related to a specific question and the modeler/theoretician will attempt to disregard all those things not important to the question.³

The second distinction is between *ad hoc models* and *first principles models*. Ad hoc models are based on a phenomenological approach and are built from observation/curve fitting/shape matching. As an example, he suggested logistic models. (The logistic produces the characteristic S-shaped curve of exponential population growth approaching a carrying capacity.) Carlos claimed that there were generally no principles by which one could argue for logistic growth even though it produces sensibly shaped curves which often are a reasonable fit for the data.

Another example of an ad hoc procedure, directly relevant to this study, is what is known as the 'mass action' term. If individuals from two distinct populations, P and Q are moving randomly within a confined area (such as the molecules of two gasses in a closed jar), the expected number of (instantaneous) contacts between individuals over a given time period is $\beta P_n Q_n$, where β is a fixed parameter. This 'mass action term' is also commonly used in epidemiological models to predict the number of new

infections of some disease, given a population, P , of susceptible individuals and a population, Q , of infectious individuals. Carlos provided several reasons why this kind of use is questionable: individuals don't move randomly, contacts are not instantaneous, and most seriously, under certain conditions the total population, $P + Q$ does not remain constant in the model. In relation to their use in this class, Carlos stated:

I argued strongly in class about the inadequacies of mass action, but at the end it becomes simple to handle mathematically and that sometimes dominates the way people think. One looks for justifications for the mass action law and people would argue that there are no justifications, but people will use them over and over because they know something about them.

First principles models on the other hand attempt to directly tie a mathematical form to a biological understanding of a given phenomena. Similar to the relationship of a 'prototypic function' to the actions one takes in a specific situation (Confrey & Smith 1991), a first principles model attempts to tie the biological action to the mathematical form. An important consequence, according to Carlos, is that all individuals are tracked and quantities are preserved.⁴

SCIENCE AS PRACTICE

The idea of science as practice is as old as the idea of science itself. Many of us were introduced to science in terms of the ideal scientific practice which involved practicing some version of the scientific method: observation – hypothesis – testing – theory. It was strict adherence to this model of knowledge production, coupled with the assumption that the knowledge produced was objective and independent of both the scientist and the larger socio-cultural setting, that separated scientific practice from other activities. However recent work in a number of fields has suggested that both the practice and the product (knowledge) of scientific practice can be viewed as social constructs. Andrew Pickering, for example, describes the basic premise of the relatively new field known as the sociology of scientific knowledge (SSK): 'First, as its name proclaimed, SSK insisted that science was interestingly and constitutively social all the way into its technical core: scientific knowledge itself had to be understood as a social product' (Pickering & Stephanides 1992, p. 1).

Lave and Wenger argue further that it is the act of participation in a community of practice that makes knowledge possible at all:

A community of practice is a set of relations among persons, activity, and world, over time and in relation with other tangential and overlapping communities of practice. A community is an intrinsic condition for the existence of knowledge, not least because it provides the interpretive support necessary for making sense of its heritage. Thus participation in the cultural practice in which any knowledge exists is an epistemological principle of learning (Lave & Wenger 1991, p. 98).

Pickering suggests that this 'interpretive support' can be viewed both through the culture of science, which 'denotes the field of resources that scientists draw upon in their work' and the activity or 'practice' which 'refers to the act of making (and unmaking) that they perform in that field' (Pickering & Stephanides 1992, p. 3, footnote 1). Smith sets this activity, or making in a goal-directed framework: '. . . "practice" is intended to capture our sense of both belonging and of purpose . . . knowing is related to action and participation; and practice only exists through active participation. Thus practice serves to locate and define the relationship between knowing and action' (Smith 1993, p. 123).

In this study we focus on the practices of biology and mathematics. Each are assumed to have a field of *resources*, including certain kinds of problems and means of solutions which lead to identifiable goals and actions related to 'making and unmaking' knowledge claims in each field. In mathematics this includes not only the classical deductivist methods of theorem and proof, but also, as Lakatos has shown (1976), the informal methods of 'conjectures and refutations' and the kinds of actions and goals that arise within different mathematical systems and different branches of mathematics. Likewise the biologists draw upon the traditional resources of biological classification, the forces and dynamics involved in genetic change and evolution, and, more recently, the dynamic interactions of ecosystems. These descriptions are not intended to fully characterize either practice, rather to point at some of the important resources that will give rise to the view of what constitutes a problem in each practice and thus the kinds of goals and actions which a practitioner might undertake in relation to these problems.

Model as 'Boundary Objects'

When biologists and mathematicians work together to create a mathematical model of a biological system, it must be assumed that both bring with them the tools and perspectives of their practice. What is created, the model, becomes a product of both practices, that is it sits between the two practices playing a mathematical role for the mathematicians and a biological role for the biologists. Thus we suggest a model becomes what Star & Griesimer (1990) have called a '*boundary object*'. Star and Griesimer were particularly interested in how different social worlds interact and developed the idea of boundary object to 'facilitate the multiple translations needed to engineer agreements among multiple social worlds' (p. 393). They provide the following description:

Boundary objects both inhabit several intersecting worlds . . . and satisfy the informational requirements of each of them. Boundary objects are objects which are both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common use, and become strongly structured in individual use. They have different meanings in different social worlds but their structure is common enough to more than one world to make them recognizable, a means of translation (p. 393).

A question we investigate, then, is how it is that the models that are developed come to be both plastic enough to suit the needs of the biologists and the mathematicians and yet 'robust enough to maintain a common use'. A second question we investigate, particularly from a biological perspective, is what the 'different meanings' these models have in the two practices. For this we adopt the notion of what Confrey has termed an '*epistemological object*' (personal communication), conceptual objects which inform one's ways of knowing within the context of a particular setting or practice. Referring to a conceptual object as a boundary object focuses on its role between practices, whereas referring to it as an epistemological object focuses on its epistemological role within a practice. In a case where a particular boundary object may be related to accepted facts which extend across practices, viewing it as an epistemological object allows one to distinguish why those facts are relevant and accepted within the frameworks of the individual practices. We focus on how mathematical models come to play the role of boundary objects between the disciplines with a particular focus on the roles they play as epistemological objects within the practice of theoretical biology.

DYNAMIC POPULATION MODELS

For our purposes, 'modeling' refers specifically to the construction of mathematical population models relevant to some biological system. When we refer to a model, we include a number of components, including certain assumed mathematical relationships, schematic drawings, sets of equations, and results obtained from the model either through numerical or analytic methods. The models discussed are called 'dynamic population models', that is they are models designed to describe changes in one or more populations over time. Thus the 'set of equations' of the model are created as either differential or difference equations.⁵ In our interactions with participants in the class, it was not uncommon to use 'model' only in reference to that specific set of (either difference or differential) equations which had been developed. For example, one might have many representations of a population which was growing exponentially at an annual rate of 5%. However, it was common to refer to the 'model' of such a population as either the differential equation: $dP/dt = 0.05P$ ⁶ or as the difference equation $\Delta P/\Delta t = 0.05P$. For clarity, we will use the phrase, 'algebraic model' to refer to the set of difference/differential equations which are part of a model.

LAYING THE GROUNDWORK: THE BIOLOGIST AND HER SYSTEM⁷

Sara played a unique role in the course in two respects: First she was able to bring her own field data and accumulated knowledge on a species of aphids to the class. This aphid population was the system for those models

created in the class which we followed. Second she brought a theoretical biologist's perspective to the class – both in terms of the data she had collected and in terms of the particular questions of interest to her in relation to the evolution of predator resistance among insect species. We will return to the role of her questions below. First we turn to a brief description of the ecological system which formed the subject matter for the models.

The basic system on which Sara and her graduate students had collected field data consists of a species of pea aphids, its host plants (clover and alfalfa), a predatory wasp, and a fungal pathogen in agricultural fields in upstate New York. The aphids hatch from eggs in the early spring and reproduce clonally⁸ over the spring and summer. In the fall the aphids mate once and the resulting eggs lay dormant over the winter before hatching in the spring. It is quite common for the host plants in the field to be harvested at approximately 45 day intervals. Harvest is followed by a catastrophic drop in the aphid population.

Wasps seek out individual aphids, attach themselves to the body and deposit a single egg inside the body cavity (oviposition). The wasp larva grows inside the parasitized aphid, killing the aphid after eight days, upon which an adult wasp emerges.

When a fungal spore contacts an aphid, the fungus begins its growth inside the aphid. After four days, the aphid dies, while attached to a leaf. The aphid carcass then emits spores, dropping a large (uncountable) number of spores over a fixed area until exhausted.

CONSTRUCTING THE MODELS

Thirteen students in the class indicated they were interested in working on the aphid-wasp-fungus model. They were divided into three groups such that each group included both biologists and mathematicians. One group of three students intended to spend much of the semester away from campus, thus our study followed only the work of the other two groups. The members of these groups are as follows:

Group 1	Group 2
Alan: grad student in ecology	Karen: post doctorate in applied math
Mary: grad student in agronomy	Jack: grad student in mathematical biology
David: grad student in applied math	Cynthia: grad student in ecology
Yvonne: grad student in biology	Mack: grad student in statistics
George: grad student in applied math	Sam: grad student in biology

Each group was initially asked to either create several questions which

would provide the focus for their modeling work or to choose some questions from a list which was provided to the class. For most of the remainder of the course, the groups met separately, occasionally meeting with either Sara or Carlos when questions arose. The first author also attended group meetings and interviewed students individually. Each group was expected to derive a model of the aphid/wasp/fungus populations relevant to their questions and to prepare a final paper presenting the results of their inquiry. Students had access to several computer-based modeling programs on both work stations and micro-computer systems on campus. Class periods during this time were devoted to learning to use the software, guest lectures on related research, and group presentations on work to date.

THE INITIAL BOUNDARY OBJECTS

Given the basics of the biological system, we can start to describe some of what we saw as the boundary objects in the modeling process. Two basic conditions seemed initially to be accepted among both mathematicians and biologists. We provide both a narrative and symbolic description of each:

1. The change in population over any given time period is equal to births minus deaths plus immigration minus emigration, or $\Delta A/\Delta t = B - D + I - E$.
2. For any given time period, the number of births and the number of deaths are each proportional to the total population, or $B = f_b(x_1, x_2, x_3, \dots, x_n) \times A$ where f_b is a function of some number of variables which may include A among other factors, and likewise: $D = f_d(x_1, x_2, x_3, \dots, x_n) \times A$.

The first of these relationships is a global claim that seems to underlie all population modeling. The second is more specifically related to our intuitive understanding of population change and in the modeling process becomes a heuristic device. That is, since the proportionality factor itself is variable from one time period to the next, this claim becomes a framework for creating the basic quantitative relationships of the model, i.e., algebraic models are initiated in the form:

$$\Delta A/\Delta t = f_b(x_1, x_2, x_3, \dots, x_n) \times A - f_d(x_1, x_2, x_3, \dots, x_n) \times A.$$

Immigration and emigration were assumed to be zero in all the models considered in the class.

Although only two of the three models prepared in the class included a schematic model, these two were essentially identical and would presumably also have been accepted by the third group. Thus a third boundary object could be the schematic model shown below.⁹

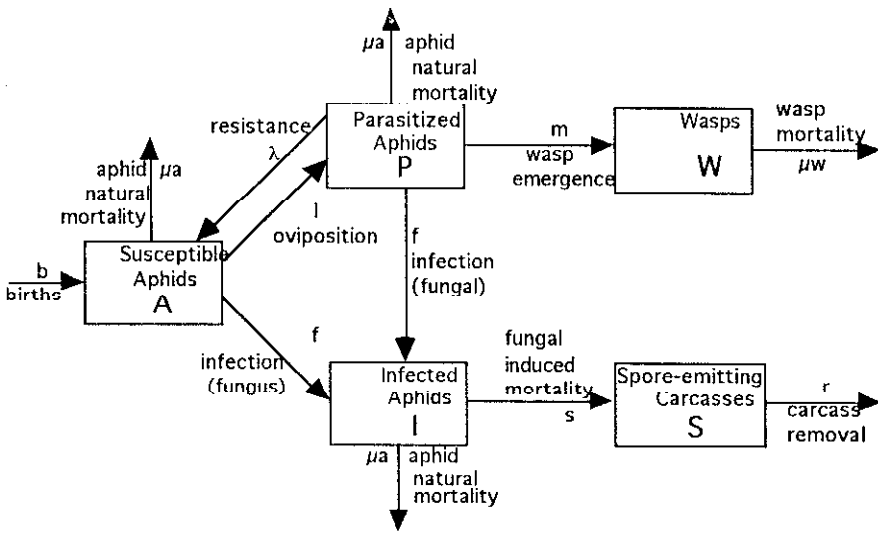


Fig. 1. Basic schematic model.

FROM SCHEMATIC TO ALGEBRAIC REPRESENTATIONS

The schematic model can be viewed as playing a critical role in the modeling process of the two groups as it was, in both cases, the only major boundary object specific to the aphid system which was accepted across the groups. We will see that while each group did create an algebraic model and derive various results based on their model, there were individuals, primarily biologists, in both groups who never accepted these algebraic models. More specifically, whereas the schematic model was 'plastic enough' to satisfy the needs of the biologists, the algebraic models, for the most part, were not. This became evident as the groups worked to create their individual algebraic models.

For the mathematicians, the schematic model seemed to play the role of an alternative representation for a set of differential equations. While one of the groups did not present a schematic in their final presentation or paper, they did acknowledge that the schematics of other groups were representative of their algebraic models. In the other group, David initially created and presented to his group an early version of the schematic above and a corresponding set of differential equations. As they worked together over time, David continued the process of using the suggestions of the biologists in the group to create [revise?] a series of schematics and related algebraic models. This was not, however, a simple one to one relationship, as he would often propose several alternative algebraic models in relation to the same schematic. For the mathematicians, then, the move from schematic to algebraic representations was not a fixed move in that there

was not a one to one correspondence between representations. However, the movement between representations was not completely flexible either because the number of corresponding algebraic and schematic representations was limited.

For the biologists, however, the schematic seemed to play a much more flexible role. It represented a set of *possible outcomes* in an extremely complex system. Whereas the schematic does provide a pictorial representation of the set of all possibilities in the life of an aphid and the connecting arrows do suggest that an explanatory hypothesis might be created for any particular movement from one stage to another, the immense complexity of the biological world is most certainly not captured in such a diagram. For example, not all aphids are the same, not all wasps are the same, birth and death rates might depend on any number of factors, the process of infection by fungal spores is largely unknown, and so on. Thus while the mathematician might see the schematic model as simply an alternative representation for an algebraic model, for the biologist it is a simplified set of relationships representing a complex system, and the movement from the schematic to the algebraic representations was not at all straightforward. The following narrative offers a sequential interpretation of the work of the two groups.

Group 1

In the work of the first group, this distinction between the perspectives of the mathematicians and the biologists became apparent at two points. At the initial meeting, Alan took the lead, listing several questions he wanted to pursue. These questions were centered on the issue of modeling the resistance of the aphids to both fungus and wasp under the assumption that there would be a cost to resistance and thus a trade-off among clonal types between resistance to the wasp and resistance to the fungus. By the second meeting the other biologists in the group added to Alan's list and a tentative plan of action was developed for constructing the desired model(s). This plan was structured around the idea that an initially simplified model should be created which would contain the relevant biological parameters. Complexity relevant to the specific questions would be added one step at a time. In this initial plan, seven increasingly complex models were proposed where the final model would include all of the following:

1. Searching and handling time of wasp for fungus.
2. Delays in wasp emergence to account for larval stages.
3. Seasonal cuttings of host plants.
4. Different reproduction rates for different clonal types.
5. Resistance to wasp, fungus and a mechanism for trade-off.
6. Seasonally variable infection rates for the fungus.
7. Genetic recombination during sexual reproduction over the winter.

It seemed that the working assumption of the group was that a simplified

model of the system could be developed that would be reasonably close to the actual behavior in the field. The group would then be able to add complexity in a goal-oriented manner aimed at answering the specific questions that had been raised. Thus this approach depended on the initial development of a working and 'realistic' model.

At this second meeting, the biologists made a list of the biological parameters which they considered important for the model and divided up responsibilities for searching for appropriate values in the literature.

As the biologists worked on collecting this data over the next few weeks, the first problem began to appear. To the frustration of the biologists, much of the biological data they needed for their approach was either unavailable or reported in the literature with a disturbingly large range of values (e.g., birth rates ranging from 2 to 10 per day). Thus the sense that each of the arrows in the schematic model could be replaced with a realistic quantitative relationship was cast into serious doubt.

The second problem for this group arose about four weeks into the project when it had been agreed that the best biological data possible had been collected and the whole group met to do some initial simulation runs on the computer. Although David had shown the group various possible algebraic models before this meeting, none of the biologists had apparently taken the general forms of the models he proposed as necessarily prescribing limits for the complexity of what could be included in the algebraic model. At this meeting the biologists became aware, apparently for the first time, that the algebraic model simply could not incorporate much of the biological data that had been gathered. For the mathematicians, an important aspect of constructing an algebraic model was that it be *mathematically analyzable*, that is that such things as equilibrium points, saddle points, etc., could be found. The mathematicians stated flatly that any system containing more than five equations with five unknowns was not analyzable. For the biologists, this seemed to be nearly a death blow – they seemed to feel that not only were they unable to gather biological data with the precision they would have liked, but the algebraic models would not even allow them to incorporate much of the data they had. The models which were ultimately developed were not mathematically analyzable, but were instead run numerically on a more powerful computer capable of running the simulations and displaying the results graphically. Thus one can begin to see a distinction between the interests of the biologists and mathematicians. Although the biologists would probably not, at this point, have used these terms, one could see that many of their questions could potentially be addressed through the patterns developed in numerical approaches, whereas the interest for the mathematicians was focused more on an analytic approach providing mathematically verifiable solutions.

The basic algebraic model which this group used and was included in their final paper is shown below:

1. $\frac{dA}{dt} = bA - lAW - fAS - \mu_A A + \lambda P$ A = aphid population
2. $\frac{dP}{dt} = lAW - \lambda P - fPS - mP - \mu_A P$ W = wasp population
 S = spore-emitting carcasses
 P = parasitized aphids
3. $\frac{dW}{dt} = mP - \mu_W W$ I = infected aphids
4. $\frac{dI}{dt} = f(A + P)S - \mu_A I - dI$
5. $\frac{dS}{dt} = dI - rS$

Notice that the terms, lAW , fAS , and fPS , are mass action terms, which from Carlos' perspective are effectively ad hoc terms in the model. This issue will be important in discussing the reactions of both Sara and Carlos to these models. The models used by the second group also used mass action terms to evaluate the rate of parasitism, lAW , and of infection, fAS , in their model.

Using this model and various modifications, they did investigate three of their six original questions (3, 5, and 6). However, the sense of disappointment and lack of enthusiasm of the group seemed to be reflected in the paper. For example, in most of the results they reported, the aphid population started out at a certain level, immediately dropped rapidly to near zero, then went into a period of wide cycles (in conjunction with either the wasp or the fungus). By varying the rate of infection (f) and the rate of parasitism (l), they were able to make either the wasp or the fungus become the dominant predator with the other eventually dying out. However as they stated in the paper: 'We were not able to find values for l and f which would allow all three species to coexist for long periods of time' (p. 6). Likewise, neither the initial rapid drop nor the continuing wide cycles were evident in the field data on the aphids. Thus they had a schematic model which all agreed was reasonable, an algebraic model which the mathematicians presented as derivative from the schematic, and results which did not seem to fit the data in a reasonably close fashion. What this seemed to lead to was a rising sense of skepticism on the part of the biologists, as suggested by this comment in an interview with Alan a few weeks before final papers were due:

For me and people I work with, unless you have a really good idea what these parameters are, which most times you don't, and then, in my own mind, I'm not real sure what you can get out of the types of models we're doing and relate it directly to what, for example, to what Sara's trying to do, to what she sees in the field. She knows the populations grow exponentially. She knows the predators don't control them - maybe they do on occasions. If you have a specific goal in mind, like for me with genetics, and how it might evolve over time, things like that interest me. But I don't see the value of doing these kinds of

model of the system could be developed that would be reasonably close to the actual behavior in the field. The group would then be able to add complexity in a goal-oriented manner aimed at answering the specific questions that had been raised. Thus this approach depended on the initial development of a working and 'realistic' model.

At this second meeting, the biologists made a list of the biological parameters which they considered important for the model and divided up responsibilities for searching for appropriate values in the literature.

As the biologists worked on collecting this data over the next few weeks, the first problem began to appear. To the frustration of the biologists, much of the biological data they needed for their approach was either unavailable or reported in the literature with a disturbingly large range of values (e.g., birth rates ranging from 2 to 10 per day). Thus the sense that each of the arrows in the schematic model could be replaced with a realistic quantitative relationship was cast into serious doubt.

The second problem for this group arose about four weeks into the project when it had been agreed that the best biological data possible had been collected and the whole group met to do some initial simulation runs on the computer. Although David had shown the group various possible algebraic models before this meeting, none of the biologists had apparently taken the general forms of the models he proposed as necessarily prescribing limits for the complexity of what could be included in the algebraic model. At this meeting the biologists became aware, apparently for the first time, that the algebraic model simply could not incorporate much of the biological data that had been gathered. For the mathematicians, an important aspect of constructing an algebraic model was that it be *mathematically analyzable*, that is that such things as equilibrium points, saddle points, etc., could be found. The mathematicians stated flatly that any system containing more than five equations with five unknowns was not analyzable. For the biologists, this seemed to be nearly a death blow – they seemed to feel that not only were they unable to gather biological data with the precision they would have liked, but the algebraic models would not even allow them to incorporate much of the data they had. The models which were ultimately developed were not mathematically analyzable, but were instead run numerically on a more powerful computer capable of running the simulations and displaying the results graphically. Thus one can begin to see a distinction between the interests of the biologists and mathematicians. Although the biologists would probably not, at this point, have used these terms, one could see that many of their questions could potentially be addressed through the patterns developed in numerical approaches, whereas the interest for the mathematicians was focused more on an analytic approach providing mathematically verifiable solutions.

The basic algebraic model which this group used and was included in their final paper is shown below:

1. $\frac{dA}{dt} = bA - lAW - fAS - \mu_A A + \lambda P$ $A = \text{aphid population}$
 $W = \text{wasp population}$
2. $\frac{dP}{dt} = lAW - \lambda P - fPS - mP - \mu_A P$ $S = \text{spore-emitting carcasses}$
 $P = \text{parasitized aphids}$
3. $\frac{dW}{dt} = mP - \mu_W W$ $I = \text{infected aphids}$
4. $\frac{dI}{dt} = f(A + P)S - \mu_A I - dI$
5. $\frac{dS}{dt} = dI - rS$

Notice that the terms, lAW , fAS , and fPS , are mass action terms, which from Carlos' perspective are effectively ad hoc terms in the model. This issue will be important in discussing the reactions of both Sara and Carlos to these models. The models used by the second group also used mass action terms to evaluate the rate of parasitism, lAW , and of infection, fAS , in their model.

Using this model and various modifications, they did investigate three of their six original questions (3, 5, and 6). However, the sense of disappointment and lack of enthusiasm of the group seemed to be reflected in the paper. For example, in most of the results they reported, the aphid population started out at a certain level, immediately dropped rapidly to near zero, then went into a period of wide cycles (in conjunction with either the wasp or the fungus). By varying the rate of infection (f) and the rate of parasitism (l), they were able to make either the wasp or the fungus become the dominant predator with the other eventually dying out. However as they stated in the paper: 'We were not able to find values for l and f which would allow all three species to coexist for long periods of time' (p. 6). Likewise, neither the initial rapid drop nor the continuing wide cycles were evident in the field data on the aphids. Thus they had a schematic model which all agreed was reasonable, an algebraic model which the mathematicians presented as derivative from the schematic, and results which did not seem to fit the data in a reasonably close fashion. What this seemed to lead to was a rising sense of skepticism on the part of the biologists, as suggested by this comment in an interview with Alan a few weeks before final papers were due:

For me and people I work with, unless you have a really good idea what these parameters are, which most times you don't, and then, in my own mind, I'm not real sure what you can get out of the types of models we're doing and relate it directly to what, for example, to what Sara's trying to do, to what she sees in the field. She knows the populations grow exponentially. She knows the predators don't control them - maybe they do on occasions. If you have a specific goal in mind, like for me with genetics, and how it might evolve over time, things like that interest me. But I don't see the value of doing these kinds of

things where you make a model, you plug parameters in and you see which parameters produce stability which, in the real world, I don't know if stability has any biological relevance. . . .

[later]. Carlos is trying to predict the future. I want to understand my system.

This suggests that for Alan, and possibly the other biologists in this group, the algebraic models did not play a significant epistemological role in relation to their practice, that is they were not legitimate epistemological objects and never became part of what Alan saw as the practice of biology. Instead he possibly saw this work as remaining in mathematics where the presumed interest was more in seeking parameters which produce stability. However, it also did not appear that Alan faulted the mathematicians. The last sentence seems to suggest that Alan simply sees the goals of modeling as separate from his goals which are to understand a particular ecological system.

Group 2

Whereas Alan and the other biologists in group 1 provided the initial direction for the activities of their group, the second group had quite a different orientation from the start. This group was organized by a mathematician, Karen, and a mathematical biologist, Jack, who had met together prior to the first group meeting and formulated several questions which interested them. Both had worked on similar kinds of models before and mutually agreed that the primary effort should be on the creation of mathematical models specifically relevant to their questions. Two issues are particularly important in understanding the progress of their group. First, each of them had a primary question revolving around the role of genetic resistance of the aphids to the wasp and fungus and the conditions under which all three species are able to exist. Karen wanted to investigate how introducing a 'cost of resistance' would effect these conditions where this cost was realized as either a lower birth rate or a higher death rate. Jack was interested in the role of the variance in the resistance within the aphid population, particularly in allowing the variance itself to vary as a function of the level of resistance. Neither of these questions seemed to be an obvious issue in the actual aphid populations observed in the field as Sara had experimentally measured predator resistance and variation and had found neither a cost of resistance or any significant variation in the variance of the resistance over time. Thus, the second issue is that from the beginning, Karen and Jack considered this primarily a theoretical project rather than one directly related to actual aphid populations. In the beginning of their paper they state:

We developed several mathematical models to explore ecological and genetic questions concerning a prey species interacting with various natural enemies. Although our explorations were general in nature, our models were conceptually based on an ecological system involving pea aphids, . . . , a fungal pathogen and a parasitic wasp (p. 1).

This initial joint decision to focus on these two specific conceptual issues had a marked influence on the interactions of the group. At the first meeting, which Sam, a biologist, did not attend, Jack and Karen explained their goals to the others. Although Mack was apparently interested in their questions, as a statistician, he also had an interest in developing a stochastic model. The model he developed was presented in a separate section of the group's final paper and will not be discussed here.

Sam came to the second meeting with a list of parameters he had found in the literature for the aphid, wasp, and fungi species which were in Sara's data. He began writing them on the board. Jack almost immediately stopped him with the statement: 'We don't care about reality' (a phrase made only partly in jest which he was fond of repeating). Karen then proceeded to put her algebraic model which included a cost of resistance on the board and the remainder of the discussion at that meeting revolved around the appropriateness of her model to the question she was asking and whether it was mathematically equivalent to an alternative algebraic model proposed by Jack. Later, when I asked Karen about the interaction with Sam, she replied:

Part of the problem is [Sam] missed the first meeting too. So, the other four of us had agreed on a sort of approach we were going to take and where we going to draw the line for reality and what kind of models we were going to look at. And so when he came in and wrote all that stuff up, the problem was he didn't ask us first what we were doing. So he wrote all that stuff up and it was useless to us because we already agreed that, see if you are looking at a continuous time model then things, then it can be argued that things like generation time don't mean an awful lot to you because there's just a constant, there's a rate of birth rate, right? It doesn't matter if right now the guys that are born now aren't going to start having little aphid babies for another 8 days. It doesn't matter because somebody's having little aphid babies, right? So when he came in and wrote all that up, the problem was that he didn't stop to ask us. Having missed the first meeting, he didn't realize that we had already settled these things . . .

For the remainder of the course, there was evidence of tension between Sam and Karen/Jack, presumably as a result of this particular incident and the related implication that he (Sam) would have little or no say in the overall direction of the group. Although Sam remained in the group, he tended to play an independent role and also created his own model which was a separate section in the group's final paper.

As opposed to Sam, Karen described Cynthia's role in the initial meeting as:

. . . of all of us the person with the absolute least amount of [mathematical] background is [Cynthia] and so I don't think that's really what the issue is. I think, you know, with [Cynthia], she was put in our group and we talked to her and said we should warn you that we have pretty much been campaigning to go this route. And she said well that sounds really interesting, you know here's what I can bring to it.

Initially, at least from Karen's perspective, Cynthia was a willing participant. When asked what roles the various group members were playing in relation to the overall project, Karen acknowledged that she and Jack

were doing most of the mathematical analysis and that she, at least, did not need much input in the way of biological data. She described Cynthia's role as:

Well, I think Cynthia is sort of keeping us all in line and saying, 'OK I'm doing a lot of literature searching' and saying here are some things I learned and here are some parameters and she has actually helped out Mack a lot because he's doing stochastic simulations.

When we interviewed Cynthia after the completion of the course, she offered an alternative perspective which, in many ways, was very similar to Alan's:

I really wanted to learn about modeling because modeling has a definite place in biology because sometimes you can't go out and sample 1000 generations . . . but . . . it seems that [models] are built for no reasons and they mean nothing and people say we'll just set this up and we'll make up the parameters and we'll find out where the equilibrium point is and that makes no sense . . . I want to know why, what this means in real life.

This suggest that for Cynthia, as for Alan, the models constructed in her group did not come to play the role of epistemological objects in relation to her own practice of biology.

Parallels between the groups

Even though we have used Alan and Cynthia as representing a biological perspective and implied that the perspective of Karen and Jack was more mathematical, it is clear from Karen and Jack's comments that both of them consider their work to be biological and to offer insights into biological knowledge. However, this view of biological knowledge may be quite different from that of Alan and Cynthia. For example, one of the conclusions offered in the paper was:

A striking result of the model is that variable genetic resistance with tradeoffs can allow two predators to coexist on the same resource. Although the ecological equations of the model, if used alone, would result in competitive exclusion and the survival of only one predator, when genetic variation of resistance is added, the predators are able to coexist over a wide range of parameter values (p. 2).

The first part of this result seems self-evident as both predators do, in fact, exist in the field. When asked about this, Jack pointed out that the importance of the result was not in the biological result itself but in the creation of a mathematical system that supported that result, a relatively rare event. This allows one to use the mathematics to the range and types of conditions under which this coexistence might be possible. Cynthia, on the other hand, had another reaction:

We see that if we increase the variance, the rate of evolution is faster, big surprise. I learned a lot of mathematics, but learned nothing or almost nothing about the system.

Likewise, Karen described her own relationship to biology as follows:

I'm not and probably never will be an experimental biologist but part of what I hope to

learn talking with Sara for starters is how to interpret a lot of what the experimental biologists do. And how to hopefully work with an experimentalist without coming to blows over where reality is going to leave off. That would be an important thing . . . You know I'm kind of split. I do, I want to do mathematics in the mathematical biology that I do, the math is very important to me as math but the mathematics in the mathematical biology is very important to me that I not be doing a problem, that I not be saying oh, well we can sort of justify this as biologically interesting let's do some math. It's very important to me that if I'm going to call it a mathematical biology problem I better have my biology right.

Although both Jack and Karen do have a strong interest in the mathematical system, it would also seem to be true that they do not equate the rejection of 'reality' with a rejection of biology. In fact it would seem that from their perspective they did work to create and offer to the biologists algebraic models which could legitimately be considered 'boundary objects'. However, for the most part, the biologists rejected these models as only marginally related to the biological system. Even in the first group where the mathematicians seemed more inclined to try to follow the lead of the biologists, the algebraic models that were ultimately developed by the group were effectively rejected by the biologists. What this suggests is that if such models are to become legitimate boundary objects, an investigation at the level of 'epistemological object' may be called for. We need to better understand what epistemological role the biologists expected of the mathematical models and what epistemological role the mathematicians anticipated the models might play. In effect this becomes then a study of what it means to come to know through these models within the practice of biology.¹⁰

Before turning to the perspectives of Sara and Carlos, we mention one more model that was done in this class. This model was prepared by one student, George, who was originally a member of the first group but attended almost no meetings and had very little interaction with the group. He is a self-described 'loner' who has an undergraduate degree in biology and is currently a graduate student in applied mathematics. Of the three models prepared in the course (group 1, group 2, and George), his was considered most valuable by both Sara and Carlos. Two features of his model were considered to be particularly important. First, his was the only model that did not use a mass action term to describe the rate of parasitism by the wasps, and second his was the only model that made a serious attempt to address a problem that Sara considered to be of interest and important in relation to the genetic dynamics of the aphids. Assuming a negative correlation between susceptibility to the fungus and susceptibility to the wasp (a relationship which Sara had found experimentally in aphid populations), George investigated the genetic changes in susceptibility over the course of a season under the assumption of clonal reproduction. The importance of these two features will become more apparent as we discuss the reactions of Sara and Carlos to these three models.

THE PERSPECTIVE OF THE THEORETICAL BIOLOGIST

Sara agreed to teach the course with Carlos and share her work on aphids, wasps, and fungi, in order to both learn more herself about the usefulness of modeling to her own research and to provide the context of a specific biological system for the class modeling projects. Though she has had little modeling experience herself, she has a great deal of expertise in doing theoretical biology by investigating a particular biological system and hence brought an important perspective to the class.

Sara [distinctly] considers herself a theoretical biologist. Her primary interest is in the dynamics of evolutionary interactions. As she states: 'I am not all that interested in pea aphids themselves. They have nice features which make them useful in terms of getting some ideas about basic evolutionary interactions'. She is not particularly interested in quantitative models aimed at predicting actual aphid populations over time: 'Well a lot of people dream of it [quantitative models]. And applied people are very concerned about it, but . . . I would be more interested in the way in which particular factors interact with one another to affect the dynamics'. Thus it seems quite clear that we are amiss if we refer to a singular 'biological perspective'. Instead we have the particular perspective of a theoretical biologist with an interest in evolutionary genetics. It is from this perspective that we need to investigate how these models become possible epistemological objects for Sara. Sara occupies a critical point in this investigation for she seems to simultaneously question the fundamental value of models in relation to her work and at the same time be able to find value in certain aspects of the models that are related to her background and interests.

At one level, in a way reminiscent of Alan and Cynthia, Sara questions whether these models have any substantial biological value:

I have to say that relative to what we are doing this semester I am a little depressed about the ability to make very realistic models because the two papers, the one by [the first group] and the one by [the second group], really [make] the simplest of all possible assumptions about everything. You know, about the course of the aphids, about handling time interactions, the counter probabilities, etc. And they are so simple so as to really not to be able to be related to anything real. They don't make any predictions that are very realistic. So in a sense of connecting these models to my work, there is not a very explicit connection. And there is a lot of work that still has to be done if that connection were to be made.

Later in the interview she states:

. . . the mass action . . . didn't seem, to a biologist, as that big a difference and yet it made a gigantic difference in results. And so as I am talking to Carlos on Friday, I did become sort of depressed about it

. . . it's very hard to take these three papers, particularly you know [group two's paper], and come to any general conclusions that are very realistic with respect to the biology because if you can get that big of a difference between these . . . oscillations which eventually wipe them out and one epidemic per season just because of some little tweak

in the way you handle encounters, that . . . What about all the other things that are too simple about the models?

The side of Sara we are hearing in these two quotes seems to resonate with the concerns of Alan and Cynthia, and might be characterized as a fundamental concern that the complexity of the models is not even comparable with the complexity of any biological system. There is a sense that a mathematical model should actually *be* a model of the biological system and that that does not seem possible. This is a question we will return to later: In what sense is a 'good' biological model a model *of* a particular biological system.

A second area of concern to Sara revolves around what constitutes a good biological question or issue in the development of a biological model. She is critical of the focus questions articulated by both Karen (role of cost of resistance) and Jack (variation in the variance of resistance, but neglect of covariation). She states:

I was a little bit miffed at [Jack and Karen's] group because they were very set on modeling a situation where there was a metabolic cost to resistance and yet in my situation we didn't find that cost [in the field data] and in other situations other people haven't found a metabolic cost to resistance . . . so I felt that in fact that was an example of something you might think would be a specific parameter and yet including it or not including it could be important to a whole class of systems. Just like they didn't want to model the fact that we found that there was a negative genetic correlation and susceptibility to the two natural enemies. In other words, genotypes which were more susceptible to fungus were less susceptible to the parasitoid. [George] modeled that, but they weren't interested in modeling that so they cut out a whole class of systems in which that situation was the case and they did not have the potential to investigate the consequences of that on the dynamics of the model.

Sara seems to be making two arguments here. First since the experimental data did not show evidence of a cost of resistance, it was inappropriate to enter it as a parameter in the model rather than making it an object of investigation itself. Second was the issue of leaving out covariation which, in fact, had shown up in the data.

Both of these questions focus on the issue of what constitutes a legitimate biological question in the formation of these models, that is, how do they become legitimate epistemological objects for a biologist? The issue of framing a good question is closely related to the issue of the perspective of the modeler. Sara seems to attribute at least part of the problems she sees in Karen and Jack's approach to perspective:

I think there would be a big difference between the perspective of a biologist who is only interested in this one system that they were working on versus someone like me who is more interested in a general class of interactions overall. And the mathematicians, well, often they just want to do something that is mathematically interesting. So . . . that leaves out a large class of important biological problems which may involve mathematics which is not particularly new or interesting. I mean I think that's why [Karen] isn't continuing to work on this. She wants to get a job in the math department and so she has to do something that is mathematically interesting.

Again it would seem that Sara, like Alan and Cynthia, saw the models potentially as being more an epistemological object of mathematics than of biology, in effect questioning their legitimacy as boundary objects. However, given her background and experience, Sara was also able to articulate what she saw as legitimate roles for these models in her work much more than the less experienced graduate students. She has partially set the stage for these views in the above quote as she describes her interest in a general class of interactions, rather than in only one biological system. She articulates this view more fully below:

I don't really care about making this model a prediction, but in my experimental work on the genetics of insect predators, I think of the aphids as a model of the system . . . with which to pose the question.

. . . yet the question of what is the impact of variation on interspecific interaction is really at the core issue that we try to model and that is a very general question. So there is a whole body of ecological models that people have made, which are general. They don't really concern any particular system on predator-prey relations or parasitoid-host interactions. And I would be really interested in the effects of adding variation upon that class of models. Using the aphids is just sort of an example. So it didn't really bother me that the parameters weren't estimated all that great.

When asked what these are models of, she stated:

I would think they are models of . . . ecological and genetic interactions between species.

That is, models become arenas in which to examine dynamic changes of general systems under various prescribed conditions. Sara's particular interest is in the effect of introducing genetic variation into these models. Thus she valued George's model in particular because it explored how changes in the (negative) covariation effect the patterns of genetic change in the aphids over a season.

A second epistemological role which Sara describes for these models is as an experiment. Thus despite her criticisms of Karen for leaving covariation out of her model and for putting in a cost of resistance, Sara is intrigued by the stability of the three-species system in Karen's results. She mentions two other studies reported in the literature with similar results and states:

Those are the only three models I know of with predator and prey or parasitoids and host that have incorporated genetics, and all three have found that in some sense stability is enhanced with increased variability. So that seems that maybe we are working towards getting some kind of a general picture, because all three of these models have some different elements and yet they come out with a similar sort of overall result.

From this she proposes her idea of model as experiment:

. . . and I suggested to Alex [another mathematical biologist at Cornell], well, Alex maybe these models are sort of like experiments. If you do an experiment on a single system, you can't necessarily say something general about all biology, but when you . . . read over a period of time in the literature and you see similar development in many biological systems you can start to derive general ideas. . . . I said maybe we just need a lot of different people with different points of view and different kinds of assumptions in

their models to incorporate variation and see whether over all these different kinds of models we find this result. He said, well, I never thought about modeling like that but that is sort of true because you can think of each model as an experiment with a different system and if you add variation to all of these different kinds of models and you kind of stabilize the system, well, that would make me feel that there might be something different, even though it might not make a mathematician feel [good] because he couldn't solve a single equation. . . .

Yeah, and what is general. It is a different sense of, I mean, I would accept that something is general if I see it in four or five different models with different simplifying assumptions thrown in as well.

A third role which Sara describes for models is as a guide to empirical inquiry:

Sure, the environment may change and you may start all over again but that could change some of the interactions, but the point is . . . it's one of those holding all other things constant, or all others being equal argument. What is the effect of this? I think that is very valid to look at that because that's what provides a context for empiricists to say, well, this term really makes a big difference. Now its worth it for me to go out and determine what the values for this are. [what] the ranges of this parameter might be in populations because I can see it causes a big difference in the ecology or genetics of the particular organism.

Sara has provided three possible roles for models as epistemological objects: an arena in which to explore patterns of dynamic change; an experiment; and a guide for empirical inquiry. With these three descriptions, Sara has set the stage for revisiting the question of what is a good mathematical model for a biologist. To help answer this question we return to Carlos with a particular interest in how he divides up the world of mathematical models and what, from his perspective, goes into the construction of a good model.

THE PERSPECTIVE OF THE MATHEMATICAL BIOLOGIST

The final interview in this project was a return visit to Carlos to discuss the various perspectives that we had seen in our study. Carlos started out with two issues which he sees as closely connected: What makes a good biological model and how does the perspective of a mathematician differ from that of a biologist in the process of making a model. Eventually this led to a fairly forceful claim of what is a good modeler. In reading Carlos' comments, it is important to remember that his educational degrees are in mathematics and applied mathematics, thus by training he would be called a mathematician.

Making a 'Good' Model

Above, we briefly discussed two distinctions in models that Carlos had described in an early meeting with him, descriptive vs. theoretical models

and ad hoc vs. first principles models. These distinctions play a central role in the current discussion.

Descriptive vs. Theoretical Models

A common view of a model is that it is a model of something, that it depicts or copies some external system. The model could be of a specific system, such as our solar system, or a general system, such as a general planetary model. Yet even with a general model, there is often an expectation that by plugging in the appropriate parameters one can model a specific system. For example, by plugging certain parameters into a general planetary model, one can create a model of our solar system. From this perspective the criteria to judge the value of a model is to determine how well it mimics the behavior of the 'real' thing. Carlos calls a model 'descriptive' if it attempts to describe everything within a given setting: '. . . if you want to have a descriptive model, right, essentially you are trying to mimic reality . . .'. Thus a 'descriptive model' corresponds loosely with the common view of models as a copy of a 'real system'.

In an epistemological sense, of course, you cannot model 'everything' – the modeler will have to make certain choices about areas of focus and to a certain extent about the phenomena of interest. However, when the goal is to make a model 'realistic' the modeler attempts to include as much as possible. Carlos stated:

A descriptive model is essentially when your whole objective is to reproduce the dynamics of what is going on. For example in this case you might want to have reproduced the patch dynamics. In other words, you might want to have divided alfalfa fields or whatever into patches of different sizes and study the local dynamics and you might want to have introduced temperature changes and humidity changes over time and things like that. And you might have modeled it that way and you might have modeled individuals, taking into account reproduction at different times, and then at different genetic times and then you make a very complex computer simulation and then you have your descriptive model.

One interpretation of group 1, then, is that their initial objective was to create a descriptive model. The tasks they undertook were to collect a wide range of biological data on the three species and put the data into their model as parameters with the expectation that this would lead to a model that mimicked the behavior of the aphids, wasps, and fungi. From this perspective, their list of seven objectives could be treated as seven experiments of increasing complexity. That is, once the model mimics the actual behavior, one can ask it questions. For example, one can insert seasonal cuttings or one can vary the infection rate of the fungus, etc., and observe the results. When asked about the appropriateness of a descriptive approach, Carlos replied:

But the word descriptive would really be using the wrong approach, because if you want to have a descriptive model, right, essentially you are trying to mimic reality and you might as well make some sort of simulation and that really disguises what's going on. In that sense what you're really trying to do is mimic reality and we will get a complex

simulation. You probably won't make too much sense out of this unless you really have a very good idea what the parameters are. The goal of this [class projects] was not really descriptive, and it should not really be descriptive at any level.

The high level of frustration which all three biologists in group 1 exhibited when they realized that their approach was not going to work, and the general frustrations expressed by Alan, Cynthia, and Sara may all be related to this background assumption that a good model should somehow be more descriptive. However, a second level of their frustration was related to their perception of the relationship between mathematics and biology. Alan and Cynthia both seemed to feel that their models had primarily turned into a mathematical exercise and Sara seemed a bit uneasy over whether Karen and Jack were more interested in biological or mathematical questions. This distinction was central to Carlos in our conversations as well. His criticisms of descriptive models was balanced by the value he placed in theoretical models.¹¹ When asked what makes a good model, Carlos replied that first it requires a 'good modeler' by which it seemed he meant primarily someone competent in mathematics. By this criteria, he called all the (mathematics) students in the class 'equally qualified'. What was of primary importance for him, however, was the issue of how a mathematician can address an appropriate biological question:

But what becomes central to me is how important it is for you to address a particular biological question. So you have to learn a lot of biology, you have to be a competent modeler, but then you really have to focus on what it is you want to use with this model for.

Many times people just write models, just to write them. You know we have this compartment and this compartment and this compartment and this compartment and this happens to this and to that. That's okay. But it's not an underlying question. You ask people why they're doing this and many people won't know exactly why they're doing it.

Carlos' description of what makes a good theoretical model and his distinction between a mathematical and a biological perspective are closely related. In our conversations, he provided several examples:

So, in biology when people ask what is the origin of life, in mathematics that would be a question that doesn't even make any sense. A good mathematical biologist would focus his research on the question asked, what is the origin of life, and then the question would be how can I contribute to the mathematics of that particular question. A mathematician per se would find that type of approach unthinkable. They would be more interested in saying I have to study the properties of this equation, this particular space, say continuous solution or something like that, what would be the structure of these solutions on a larger space. Or what would be the structure of solutions if I change this condition. And that person will do, you know, is capable of carrying on this kind of investigations regardless of the importance of the model.

Carlos emphasized his concern with issues of communication in his description of the importance of the modeler or mathematician listening to and hearing the biologist. For Carlos, theoretical biological modeling is applying math to address a biological question, and not modifying the

question to suit the mathematics (either for the sake of simplified analysis or for the sake of 'interesting' mathematics).

One also has to understand it has to do a lot with priorities. Mathematicians might claim they want to compromise between the math and the reality. But the fact is that many are just interested in the mathematics. Many times they will never listen to the biologist. . . .

In other words, if we really believe that mathematics is going to make a contribution to theoretical biology, we have to constantly listen to people like Sara Via and address their concerns or argue about their concerns until we reach a reasonable agreement because there is a tendency for mathematicians to minimize the knowledge that biologists have.

From his own work on modeling the spread of the HIV virus, Carlos gave an example of what it means to have a good biological question and of the importance of modifying the question over time as new biological insights are obtained.

One of the first questions I addressed when I got involved with the project is what is the role of a long period of infectiousness on the dynamics of it? That is a very specific question. I didn't ask a mathematical question. I asked a biological question. Given the fact that you have long and variable periods of incubation, how does that affect the dynamics of it? What we found out in that case is that you could put arbitrary incubation periods and you will get the same qualitative analysis, not the same quantitative result but the same qualitative analysis.

When asked what he meant by qualitative analysis, Carlos replied ' . . . some sort of broad description of what is going on in such a way that you will know the properties of this model for all possible parameters. You could never do that in a simulation. . . '. Qualitative analysis, for example, includes such things as the existence of steady states, saddle points, repeating cycles and so on which are independent of specific parameter values.

Carlos then pointed out that in this initial model, they had assumed that the infection rate was constant over the entire incubation period. However new information from field studies indicated that this might not be the case:

Then data keep coming up that individuals were very capable of passing the disease at the beginning of the disease and then they were not very capable for a long time and then they might be very capable before full-blown AIDS. This is a new feature that we didn't think about. And it is a feature that people have not thought about before in epidemiological models. And we put this feature into the same equations which make the system more complicated and then we found that it might make a difference. We still haven't been able to complete the analysis.

Although introducing the variable rate of infection has complicated the mathematics of the model to the point that they have not yet succeeded in fully analyzing the system, the indication that it 'might make a difference' is an important result for Carlos in relation to the effect it has on his own work and the work of others:

So in the future when people use epidemiological models hopefully they will remember this experience as something that is important. If they know that variable infectivity, we will call it, is important then they should not probably neglect this issue.

Carlos also mentioned the issue of the relation between complexity and simplicity. As opposed to a descriptive model, the best theoretical model is the simplest possible model that allows one to address the question. While Carlos (like Sara) saw this balance successfully struck in George's work, he found the models of the other groups to be over-simplified with respect to the questions they asked.

The key reason is that what appeared to dominate the discussion is let's look at the simplest possible model that can help us to reach decisions. But that's not what Einstein would say. You know we have to look at a simple model but not at a simpler model or something like that. In other words, I think that the approach they are using in the other papers is perhaps too simple for these questions.

Ad Hoc vs. First Principles Models

A second distinction Carlos has made in classifying models is between ad hoc and first principles models. As stated at the beginning of the paper, ad hoc models are based on a phenomenological approach and are built from observation/curve fitting/shape matching, whereas first principles models attempt to directly tie a mathematical form to a biological understanding of a given phenomena. As Carlos pointed out in an early interview, an important feature of first principles models is that all individuals are tracked and quantities are preserved. He provided an example from his work on sexually transmitted diseases. He explained that the use of mass-action law, which treats the population as a whole, to predict the rate of infection does not necessarily preserve the number of individuals of each sex. He suggested a variant based on first principles that assigns a probability of sexual contact to each individual. Carlos claimed that not only was such a modification closer to the observed behavior of individuals, but that by tracking each individual in the system it necessarily preserved quantities.

Due to the numerous shortcomings of a mass-action term in biological models, Carlos seems particularly concerned with their widespread use:

One looks underlying questions of mass action law and people would argue that there are no questions, but people will use them over and over because they know something about them.

In the later interview, Carlos described a second fault of mass action in relation to the use of time scales. In general, the actual scale for time in these models is not terribly important, particularly if one's focus is on a qualitative analysis. However, when a model includes more than one dynamic process, the coordination of the time scales across these processes is very important. In the aphid models, differences in time scales are not dramatic, but still potentially important, particularly in handling time, the amount of time it takes for a wasp to find and deposit an egg in an aphid

relative to other processes in the model. Using mass action to predict new wasp eggs in effect assumes instantaneous contacts. In the models from class, George was the only one who modeled this parasitism rate to include handling time. Because group 1 used mass action, Carlos questions some of their results.

Despite his overall concern, however, he does not totally dismiss mass action, arguing instead that one must evaluate its use in relation to the particular question being investigated:

It depends upon the question you want to answer. If you want just to address the question of invasion or you want to address just the question of coexistence, then it is okay. If you want to address another question, for example you want to address the importance of handling times . . . [then it is not].

Thus Carlos describes four features for the creation of a good theoretical model: 1) the presence of a good modeler/mathematician; 2) the focus on an appropriate biological question; 3) the development of the simplest model possible which is still complex enough to address the question; and 4) the model should be developed on a first principle basis wherever possible.

The Epistemological Role of Models in Theoretical Biology

Given these kinds of criteria for a good model, we now return to the issue of what epistemological role such models may take on in the practice of biology. How can and do these models function as epistemological objects for biologists? An initial answer to this question can be found by re-examining the components of what Carlos described as a good model. The first claim, and perhaps the most dramatic, is that these models are not intended to 'mimic reality'. For the most part, they are also not intended to provide predictive or quantitative results. Thus if one knows the number of aphids, wasps, and fungi in a field at a particular moment and uses these numbers to initialize a model, one does not necessarily expect the model to predict the actual numbers of each species at some future point in time. (In many ways this contradicts our sense of what a model should be—that is a model should be a model *of* some system in the sense that it quantitatively predicts the behavior of the system.) A second and related feature of these models is that they are not general models of a biological system, rather each is designed around a specific biological question. Thus the epistemological role of each individual model is necessarily in relation to a specific question. Finally, despite this epistemological role, they are generally not intended to provide an answer. That is, the results of analyzing and/or running a model typically do not, in any direct sense, provide the biologist with new biological knowledge. The question then is, what are they intended to provide?

From our conversations with Sara, we described three types of epistemological roles that these models might play. In our conversations with Carlos, he also provided several criteria which seem complementary to

the roles described by Sara. One characteristic of a theoretical model is that it acts like a *pointer*, leading one to possible conjectures. In describing what he considered to be a good model he had come across in a recent article, Carlos described it as a 'caricature' of the actual biological system:

Well, in this case it is a very rough caricature of what could possibly be going on and you can't use a caricature to predict what's going on. You can use a caricature to provide some possible explanation that then later you would evaluate.

One analogous example he provided was a Newtonian model of the interaction of multiple planets. Carlos says the system of equations required to accurately describe the interaction of more than two or three 'bodies' are too complex to be solved. Yet there is still a role for a simplified model in pointing to areas where further investigation might be fruitful:

But sometimes by observing small disturbances and some things like that with simpler models people might be able to predict the existence of another planet or the existence of another moon or the existence of things like that. Based on some information or some very simple models people might predict as to how the universe was started, was the big explosion, things like that

Likewise in the example from Carlos own work described above, one sees Carlos' suggestion that the value of his investigation into the impact of variability in infectious rates in HIV is in pointing at issues that should be taken into account in further investigations.

Models also serve biologists in *ruling out possibilities* and *generating hypotheses*, a pair of roles which Carlos grouped together. For example, he suggested that the role of mathematical models in such questions as 'What is the origin of life?' would fall in this category. He used Darwin's theory of natural selection as an analogy. Modern biologists are now questioning its adequacy and proposing alternatives. Thus, the question might be of what value was Darwin's theory.

. . . before that theory there was Lamarkian theory and there were many theories. Perhaps the only thing . . . that was accomplished . . . by the theory of natural selection is to show convincingly to people that all the other theories were useless.

Viewing natural selection as a model, its power to accomplish this, for Carlos, would seem to be that it is a good example of a first principles model:

And natural selection has a mechanism because it is based on Mendel's data and these kinds of things so it has some sort of formal mathematical basis, but even then its purpose is just making things clearer for everybody else. . . . What I would argue is that whatever we do in any of these areas of biology, as long as they are useful to address some specific uses that we can actually go and check, that is good. In other words, natural selection might not be the mechanism that drives evolution, but it has allowed us to explain a lot of phenomena to a first approximation. It has helped us to understand a lot of issues related to biology and it has helped science to progress.

Finally, Sara describes a role for mathematical models as experiments

in providing evidence for theoretical claims. We discussed this role with Carlos and he framed the development of evidence from a model as an interactive process between the biologist, the modeler, and the model.

Well I think what happens is that, biologists worked with this system for a long time. They have very little sort of real data, but they have an incredible knowledge of the system. In other words, there are some answers that she already has, okay. She has very strong intuition about what is going on and she is probably correct. She is trying to gather some sort of use of the models as evidence for her own intuition and if the models do not support her intuition she is skeptical and she should be. And only after she fully understands what is going on, she might change her mind or she might reject the model.

Thus from Carlos, as from Sara, we see multiple roles for mathematical models in biology: as a 'caricature' of a biological system which can both suggest areas of future investigation and suggest what kinds of data biologists might need to collect in relation to a particular question; as 'ruling out possibilities', and 'generating hypotheses', where first principles models in particular allow one to seriously question alternative explanatory frameworks; and finally the sense that both these processes depend on an on-going interaction between biologist, model, and modeler (mathematician).

MODELS AS EPISTEMOLOGICAL OBJECTS

Carlos and Sara have suggested a number of possible roles for models as 'epistemological objects' in the practice of biology. We summarize these below:

1. Model as 'caricature' of a biological system, or as 'arena' in which to explore dynamic patterns of interaction.
2. Model as 'experiment' in which evidence from a variety of models accumulates.
3. Model as a 'first principle' where the relationship between biological dynamics and the mathematical structure creates opportunities to 'rule out' possibilities.
4. Model as 'empiricist' where the model suggests what data should be collected in relation to a particular problem.
5. Model as 'dialectic' where the interaction between results from the model and the knowledge of the biologist forms the basis for the development of new biological knowledge.

For us, this story is both revealing and informative. It is revealing in the sense that it causes us to step back and reevaluate many of our assumptions about how a model is constructed in science, how it is used, and how it relates the body of knowledge associated with that science. Using a Kuhnian analogy, we might see this as a scientific revolution. The 'underlying assumptions' of the dominant paradigm might be that one gathers all the relevant information, creates the appropriate mathematical relationships, enters the data, runs the model, and then learns from the

results. Even though the semantic view, described above by Doucet & Sloep (1992), allows one to see the folly of imagining that the process could be quite so straight-forward, the epistemological role of a model, from that perspective, is as model *of* a real-world system, however imperfect any particular model may be. What we have heard from Carlos and Sara suggests a significant shift in the nature of this epistemological object. When one first begins to realize that these models cannot 'mimic reality', the sense of discomfort suggested in the comments of Alan, Cynthia, and, to a lesser extent, Sara sets in. However as one comes to realize that there is an alternative in which the models are not *intended* to play that role, new possibilities arise, as suggested in the set of possible epistemological roles described by both Carlos and Sara. For us this is the informative part. Sara suggests three possibilities for the role of models, as suggesting patterns of interaction and change, as experiments, and as guides to empirical inquiry. Carlos adds the idea of a model as a 'caricature' pointing towards possible knowledge claims and the dialectical interaction between biologist, model and modeler/mathematician, emphasizing the central importance of a good biological question in the development of any model. His emphasis on using first principals may be particularly noteworthy, for it offers the possibility of changing the process of evaluating a model from an emphasis on the fit between model output and empirical data, as suggested by Doucet, to an emphasis on modeling one's understanding of the processes involved in some system. It is the fit between the processes of the mathematical model and one's understanding of the biological processes involved that is significant, more so than the fit between model output and empirical data.

Taking these possibilities together suggests that models are a part of a diversity of ways of constructing biological knowledge. As 'epistemological objects', these models can better be viewed in terms of 'threads interwoven in a tapestry' (Confrey 1992, 1993) than as containers of isolated pieces of knowledge.

We believe that this view also suggests an interactive relationship between models and the practices of biology and of mathematics. As described earlier in the paper, we do not view practices as fixed modes of social behavior but as essentially guided in relation to the problems and goals of a social group. Thus Carlos' dictum that a good modeler/mathematician is one who starts by listening closely to the biologists and focusing on a *biological* question suggests that the practice of mathematics will change under these conditions but does not suggest that the mathematicians will become biologists. The distinct interests of mathematicians in the analysis of mathematical systems remains, yet is funneled towards biological issues. For example in focusing on the question of the origins of life, mathematical systems might be developed with the specific goal of ruling out particular possibilities. However, it also suggests that mathematics is not just a useful tool in the practice of biology, but is stretched and shaped by the biologists such that legitimately new mathematics is created.

On the other hand, some of the interaction patterns observed between mathematicians and biologists suggest that the 'authority' of mathematics may sometimes exclude the biologists and their essential and legitimate insights. Thus for the biologists, this evolution of practice in relation to modeling may be even more dramatic, for, as Sara and Carlos both suggest, it is the biologists who must become the weaver of the tapestry and to do so, they must have the control to be able to pull all the threads into the process as needed. Sara, by her own description, was in the process of learning and, as might be expected, was expressing a certain measure of skepticism yet also proposing new possibilities in the design of the weave.

MODELS, REFLECTIVE KNOWLEDGE, AND MATHEMATICS EDUCATION

Ole Skovsmose (1990) describes three kinds of knowledge involved in mathematical modeling:

- a) Mathematical knowledge itself.
- b) Technical knowledge, which in this context is knowledge about how to build and how to use a mathematical model. we may also call it pragmatic knowledge.
- c) Reflective knowledge, to be interpreted as a more general conceptual framework, or metaknowledge, for discussing the nature of models and the criteria used in their constructions, applications and evaluations (p. 767).

Skovsmose argues that mathematical knowledge itself is not sufficient for developing technical competence, that is to learn about modeling, 'the solution must be to engage students in modelling processes. . . . (Like-wise,) Reflective knowledge cannot be reduced to technological knowledge. . . This idea is equivalent to the core idea of the technological society: the knowledge by which we create technologies is insufficient as a foundation for an evaluation of the technologies created. an additional interpretation of our constructions is necessary' (p. 768).

Since models never are nor should be considered models of reality, the importance of reflective knowledge, for Skovsmose, lies in the ways it can both point at the conceptual 'preunderstanding' that underlies the creation of a model and the purposes and intentions for which a model is built. In many ways this seems remarkably similar to the emphasis which Carlos placed on both the use of first principles and the necessity of working from a good biological question. However, whereas Carlos' interest is more specifically on how a mathematical modeler goes about the process of creating and using a model in a biological context, Skovsmose' interest is the context of mathematics education. He poses the following questions, relative to the current emphasis on a more applied approach to mathematics education:

Does the trend imply that the students will be able to understand the social practice of mathematics? Understand the potentials of mathematical model-building activity in society? Understand the nature of problems connected to the use of formal methods such as problem-solving technique? Is it at all possible to develop such an understanding in a mathematical education (p. 766)?

The issues Skovsmose is raising in relation to the social context of mathematics education seem strikingly parallel to the issues being raised by Carlos and Sara in relation to mathematics (education) in the context of science. Both seem to be suggesting that a treatment of mathematics that does not go beyond the pragmatic/technical is inadequate. In the context of this study, the mathematicians' primary interest in modeling was as an epistemological object in relation to interesting mathematical questions. This seemed to have little to offer Sara and the other biologists. Likewise when the biologists viewed the epistemological role of a model as a picture of the world (a biological system), they found actual models to have little relevance to their own practice. It was through the discussions with Sara and Carlos that we began to understand how alternative views of the modeling process and of the model itself (as caricature, arena, experiment, first principle, empiricist, and/or dialectic) supported the possibility of the model becoming a genuine boundary object in relation to the two practices.

We believe that this has significant implications for the practice of mathematics education. In most classrooms, even those in which problem-solving plays a central role, finding an answer to the mathematical problem is the solution and thus the end of the problem-solving process. What we also find, however, is that more and more students find mathematics distasteful and/or irrelevant as they progress through their schooling. One interpretation of this is that the epistemological role of the mathematics they are learning in school has little relevance to the complexity of their lives. School mathematics fails to become a boundary object between the practice of school and the practice of life outside of school and, like many of the biologists in this study, they reject mathematics as irrelevant. This is where Skovsmose' last question becomes important: 'Is it at all possible to develop such an understanding in a mathematical education?'

Although there certainly are pressures not to extend mathematics education beyond what Skovsmose called mathematical and technical knowledge, we also think that there are potential links available should we choose to go in this direction. One of the most powerful, we believe, is suggested in the idea of an 'epistemology of multiple representations' (Confrey 1992; Confrey & Smith 1991), which claims that even within the practice of mathematics, knowledge is built up through a coordination of actions and patterns across representations and that our claims to know become more certain as we find ways to confirm them across representations. It would seem that this view of what it means to come to know is readily extended to the role of models as described by Carlos and Sara. The dialectical relationship between the model and one's biological exper-

tise grows towards resolution through a coordination of the representations of knowledge across these domains, and as in the relationship between mathematical models, each informs the other. What this potentially creates for students is a much more flexible role for mathematics which is both less alienating and more relevant to the experiences of their daily lives.

Notes

1. The authors express their thanks to the participants in the Biomathematics Seminar offered at Cornell University during the fall of 1993, particularly to Dr. Carlos Castillo-Chavez and Dr. Sara Via for their cooperation without which this project would not have been possible.
2. Dr. Castillo-Chavez and Dr. Via are identified by their actual names. All other names are pseudonyms. Dr. Castillo-Chavez specializes in the development of epidemiological models, particularly in the spread of sexually-transmitted diseases. He describes as a primary focus of his work the importance of bringing a biological perspective to the development of mathematical models. Dr. Via is a theoretical biologist with a particular interest in quantitative genetics, particularly in the issue of changes in the genetic characteristics of a population over time in response to environmental pressures. Although she has a substantial background in statistics, she said that her direct work with models had been limited to her experience as a post-doctoral associate prior to coming to Cornell.
3. For Doucet & Sloep (1992), a theoretical model is any mathematical model within this framework. This is somewhat more general than Carlos' description of a theoretical model being based on a specific question. What Carlos calls a descriptive model would be a theoretical model within this framework if all three components were present.
4. Under certain conditions, using a mass action relationship to depict the number of new infectious individuals from a given population of susceptible individuals and infectious individuals could result in an absolute increase or decrease in total individuals, independently of births, deaths, and migration.
5. In fact the criteria as to whether difference or a differential equations were appropriate for a given situation was a subject that often came up in our discussions. The complexity of this issue prohibits a full discussion.
6. Technically, this equation should be: $dP/dt = \ln(1.05)P$, but the functional difference is insignificant.
7. A paper describing their work with this system is currently under preparation: Via, S., Webb, C. & Laude, R., 'Coevolution, Quantitative Characters and Costs in a Host-Parasitoid System'.
8. Clonal reproduction is asexual reproduction in which all offspring are clonal copies of the adult.
9. This schematic is almost a direct copy of the initial schematic created by group 1.
10. Thus we are intentionally neglecting for the most part the role of these boundary objects in the practice of mathematicians. However, we believe this is less in need of articulation for these algebraic models lend themselves to traditional mathematical analysis in terms of determining steady states, saddle points, and other traditional objects of inquiry within the practice of mathematicians. In doing so, we do not want to underemphasize either the potential of biology to suggest new methods of mathematical inquiry or the potentially powerful role of graphical simulation programs in mathematical practice. However we consider such topics to be beyond the scope of this paper.
11. Carlos' criticism of descriptive models was not so much a general criticism as a criticism in relation to the goals of this course. Thus he acknowledged the value of large predictive

models in certain situations as being able to predict the (short term) future course of events. However, the complexity of such models detracted from their ability to assist in answering specific theoretical biological questions.

REFERENCES

- Confrey, J.: 1992, 'Using Computers to Promote Students' Inventions on the Function Concept', in S. Malcom, L. Roberts & K. Sheingold (eds.) *This Year in School Science 1991: Technology for Teaching and Learning*, American Association for the Advancement of Science, Washington, DC.
- Confrey, J.: 1993, 'Voice and Perspective: Hearing Epistemological Innovation in Students' Words', *Revue des Sciences de l'Éducation*, January, 1993.
- Confrey, J. & Smith, E.: 1991, 'A Framework for Functions: Prototypes, Multiple Representations, and Transformations', *Proceedings of the 13th Annual Meeting of PME-NA*, Virginia Polytechnical University, Blacksburg, VA.
- Doucet, P. & Sloep, B.P.: 1992, *Mathematical Modelling in the Life Sciences*, Ellis Horwood Limited, London.
- Lakatos, I.: 1976, *Proofs and Refutations: The Logic of Mathematical Discovery*, Cambridge University Press, Cambridge.
- Lave, J., Wenger, E.: 1991, *Situated Learning: Legitimate Peripheral Participation*, Cambridge University Press, Cambridge.
- Pickering, A., Stephanides, A.: 1992, 'Constructing Quaternions: On the Analysis of Conceptual Practice', in A. Pickering (ed.) *Science as Practice and Culture*, University of Chicago Press, Chicago, IL.
- Skovsmose, O.: 1990, 'Reflective Knowledge: Its Relation to the Mathematical Modelling Process', *International Journal of Mathematics, Education, Science and Technology* **21**(5), 765-779.
- Smith, E.: 1993, *Practice in a Radical Constructivist Setting: The Role of Virtues and Activities in Mathematical Knowing*, Unpublished Doctoral Thesis, Cornell University, Ithaca, NY.
- Star, S. & Griesimer, J.R.: 1990, 'Institutional Ecology, "Translations" and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology', *Social Studies in Science* **19**, 387-420.