Title

Models

Author

Winterhalder, Bruce

Publication Date

2002

Chapter 12

Models

Bruce Winterhalder

INTRODUCTION

Models represent observed or hypothesized relationships of structure and function in simplified or abstract form. They transform a reference situation, usually a complex system or process, in order to make it more accessible or tractable. Maps scale down a landscape in order to guide the newly arrived over unfamiliar terrain. The logistic equation isolates key demographic variables in order to guide the analysis of exponential population growth in a finite environment. Everything important about a model is shaped by its being an instrument of prediction or investigation. Because of this, the properties of models and the capacities expected of them are nearly as varied as are the goals in using them.

Models are ubiquitous in archaeology, biology, and related historicalevolutionary sciences. They are particularly useful in these fields because the subject matter is complex (which puts a premium on orderly techniques for simplification), multidisciplinary (which puts a premium on devices facilitating clear communication), and in the very early stages of scientific development (which puts a premium on instruments that balance the development of abstract ideas and empirical investigations of their explanatory potency). Models in these fields are non-denominational, in that their use is not restricted to particular theoretical approaches. To the degree we are clear in thinking and communicating about the practice of modeling itself, models bring order and rigor to analytic efforts. This claim recognizes that modeling is an activity (the common use of the gerund gives it away), not a concept, a problem, or a field of knowledge. Skill in modeling develops with experience, but it can be helped along by understanding the place of modeling in scientific investigation.

What follows is a schematic framework for thinking about the place and use of models in anthropology, archaeology, and cognate fields, particularly ecology. Consistent with the pedagogical objectives of an encyclopedia, it is a thin essay draped over a more substantial skeleton of definitions, concepts, taxonomy, and sources. I begin with a brief historical overview. The following sections discuss kinds of models, their uses, usefulness and drawbacks, testing, analytical compromises, and the relationship of models to narrative, metaphors, and reductionism. Three case studies are followed by some modest divination about the future uses of models in the fields mentioned above.

HISTORY

The beginnings of widespread use of explicit, quantitative models in ecology can be dated to the last 50 years. This short time span leaves us with a history that is largely biographical and informal. Excellent inside sources are Fretwell (1975), Golley (1993), and Schoener (1987). By contrast, Kingsland (1995) and Hagen (1992) write as professional historians of science. Because models are pervasive in science, a history of their use is potentially as complex as the intellectual and social histories of disciplines themselves. Nonetheless, in ecology, anthropology, and archaeology, modeling has followed largely dichotomous routes.

Tansley, Lindeman, Odum, and colleagues began the development of what have become ecosystem **simulation** models. These efforts drew on concepts from information theory, operations research, and cybernetics which flourished during and immediately after World War II. They initiated the work of giving empirical substance and conceptual form to Tansley's 1935 definition of the ecosystem concept. Ecosystem approaches adopted the premise that biotic communities and their physical substrates were organized according to system-level properties such as homeostasis. The goal was a structural-functional model that mimicked the processes observed, in particular, natural systems (typically biomes, such as temperate grasslands). Complex box-and-arrow flow diagrams were the visual expression of these models.

The second track developed somewhat later, in the 1960s. Ecologists led by Robert McArthur, Eric Pianka, and John Emlen challenged the reigning descriptive naturalism. They parted ways with ecosystem efforts by developing simple, frankly speculative, predictive models addressing such questions as feeding behavior, population regulation, and species diversity. This **evolutionary** ecology track used hypothetico-deductive methods, and problem-oriented fieldwork focused on adaptive design. It adopted the premise that natural selection would optimize or stabilize some ecological

property such as feeding efficiency, clutch size, or niche breadth, at the level of individuals. While the intent of many evolutionary ecology pioneers was understanding community structure, what they in fact created was the field of behavioral ecology. Behavioral ecology has since flourished despite routine rejection of early papers by major ecological journals (Schoener 1987).

Self-conscious attention to the use of models in archaeology lagged that in biology but has followed similarly divergent tracks. Clarke (1972) gives a thorough discussion of the kinds of archaeological models spawned by the "New Archaeology," categorized according to their alliance with morphological, anthropological, ecological, and geographical "paradigms." Three brief examples will establish the flavor of the early work with models. Using demographic and other assumptions, some drawn from ethnographic studies of extant foragers, as well as Monte Carlo techniques, Wobst (1974) simulated the MES (Minimal Equilibrium Size) of the social unit that would consistently provide suitable mates to maturing adults living in the lowdensity conditions assumed for the Pleistocene. He was then able to show how the size of this social unit, which varied between 175 and 475 persons, was affected by such factors as minimum band size (no effect), increased life expectancy (MES declines), or polygyny (MES increases).

Thomas (1972) used Steward's ethnographic reconstruction of Reese River (Great Basin) Shoshone life as the basis for a season-by-season simulation model of their subsistence-settlement system. His " systems analytic" (p. 691) model is represented as a multicomponent flow chart of male and female activities, lifezone locations, and branching points (pp. 679-680; Thomas, Figure 17.3). It draws on estimates of the productivity, localized year-to-year reliability, and spatial distribution of key resources (piiion nuts, Indian rice grass seed, and antelope). It has the objective of recreating the seasonal round in order to predict artifact distributions, and from that the goal of testing Steward's reconstruction against the archaeological record (Thomas 1973). Flannery (1968) developed an ecosystem approach based on cybernetic arguments to explore how seasonality and scheduling of an annual round might, when subject to small deviations such as fortuitous mutations in resource species, develop into early systems of plant cultivation in Mesoamerica.

Archaeologists and their anthropological colleagues in cultural ecology identified most readily with the system-simulation approach of biologists. However, the simulation approach did not flourish as Thomas (1972: 701) predicted. By contrast, use of evolutionary or behavioral ecology models of the type described more fully below appears to be growing (Winterhalder and Smith 2000). Bettinger (1991) gives the history for hunter-gatherer studies. Advocates of evolutionary ecology models in anthropology and archaeology generally followed a modeling strategy like that of their biological colleagues on the evolutionary ecology track (see O'Connell 1995; Smith and Winterhalder 1992; Winterhalder and Smith 1981).

A third and much newer development affecting archaeology, as well as other social sciences and the study of animal behavior, is agent-based modeling (Kohler 1999). These models typically simulate the outcome of interactions among a population of individuals who act on information collected from a particular environmental context. "Research using these models emphasizes dynamics rather than equilibria, distributed processes rather than systems-level phenomena, and patterns of relationships among agents rather than relationships among variables" (Kohler 1999: 2). An early example is the marvelously detailed and creative model by Flannery and Reynolds (Flannery 1986), analyzing the resource scheduling decisions of bands of foragers living in the prehistoric Oaxaca Valley. Similar applications can be found in Kohler and Gumerman (1999). In several respects-the scale of analysis, conceptual commitments, and degree of complexity-the agent-based approach is a potential bridge in archaeology between behavioral ecology and systems-level models.

CONTEMPORARY USES Kinds of Models

A useful taxonomy separates scale models from heuristic models, and within the latter category distinguishes among theoretical, analogic, and empirical models. Brief descriptions follow.

A topographic map or architectural miniature is a very basic form of **scale model.** Both omit detail and reduce size but maintain fidelity of relative position and distance among selected features such as rivers and towns, or walls and windows. Some virtual (computerized) models are a variant, but are designed to give the perceptual illusion of realistic scales. An example is the house plan or archaeological site that one can "enter" visually in three dimensions on a computer screen. Geographic information systems (GIS) combine a scale model with visualization and advanced (database) abilities to analyze the elements that scale models best preserve: spatial correspondences based on relative position and distance.

Scale models represent relationships known to their makers. By contrast, heuristic models are problem-solving devices. Their usefulness lies in furthering the analysis of untested but hypothesized relationships. Heuristic models come in several forms, given further variety by the assumptions they make and the techniques used to solve them. In one basic form, heuristic models explore the implications of general theory. Such theoretical models serve as an operational bridge between abstract beliefs and concrete manifestations that are more directly amenable to observation and experimentation. Evolutionary ecology models, for instance, facilitate investigation of

the observable consequences of natural selection for the physiology and behavior of organisms.

Analogic models derive their heuristic power from comparison with phenomena believed to share one or more essential features but to differ in others. Paleontological reconstruction of hominid locomotion or diet based on the structural-functional anatomy of living species is a classic example. Another is the claim that we can learn about the behavior of Pleistocene hominids by thinking of them as if they were social carnivores, or a particular species of primate. Analogic models depend on a justification for the choice of referent, and in this they usually are implicitly theoretical. In the instances cited, evolutionary theory warrants the claim that the comparisons are meaningful. But theory seldom is prominent. Instead, the referent itself, social carnivores, serves as a source of ideas for postulating generalized features of the analog hominids. Tooby and DeVore (1987; see also Clarke 1972: 40-42) list reasons to prefer theoretical over analogic, or in their terms "conceptual" over "referential," models, especially for behavioral studies.

Empirical models are derived from the referent phenomenon itself. Thus, an empirical model of an ecosystem might represent its constituent species as trophic components in order to replicate and analyze nutrient or hydrological flows. Empirical models simplify by collapsing detail into conceptual categories. In ecological studies these are usually structural or functional features. Successful empirical models attempt to generalize beyond the observed features of the system: What will happen to nutrient dynamics and runoff if forest cover in this ecosystem is reduced by 40%? Empirical models do not require a strong theoretical association. Consequently, they are most valuable and apt to be used when appropriate theory is lacking or is only weakly developed.

The three types of heuristic models differ primarily in the referent that serves as their creative starting point. They move from the theoretical, to the comparative, to the empirical. The categories partition a gradient of analytical inspiration, but the boundaries between them are not sharp. The theoretical, comparative, and empirical can mingle to varying degrees in specific applications. Likewise, the categories are not of equal status. Theory is implicit in analog models and theory building is presumably a desired outcome of empirical models.

A large number of dichotomies are used to characterize specific modeling techniques. For instance, models can be **static,** without an inherent temporal component or **dynamic** and thus time-dependent. They can be **deterministic,** using fixed or average expected values for parameters and input variables, or stochastic. In a **stochastic model** one or more parameters is characterized by a distribution for which individual values are unpredictable. Variance itself becomes an important part of the analysis. Solutions to a model may be **mathematical** (also called **analytical),** either in the form

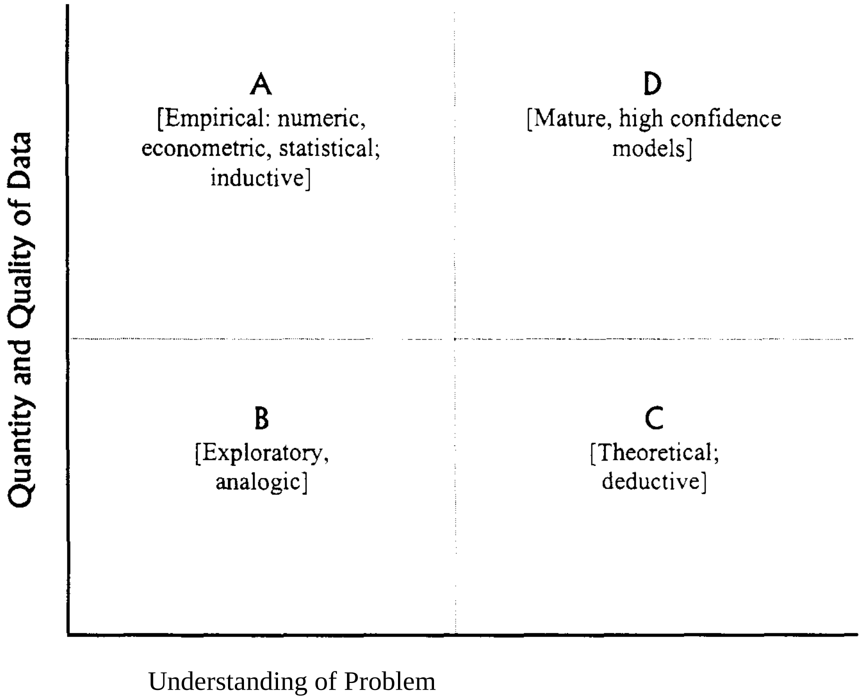

Figure 12.1. A model for classifying types and uses of models. This model emphasizes generality over realism and sacrifices precision to both (modified from Holling 1978: 67-69).

of equations or graphs; or they may be achieved indirectly through simulation, in which the model takes the form of computational rules executed with software. Many further distinctions can be found in the literature.

Uses, Usefulness, and Shortcomings of Models

Models can be used at all stages of scientific investigation. They can help to (1) define and isolate a problem, (2) organize thought, (3) advance understanding of data or direct attention to relevant data yet to be gathered, (4) communicate ideas, (5) devise hypotheses and tests, and (6) make predictions (Starfield and Bleloch 1986: 1). Ideally, they reveal emergent or novel phenomena not anticipated in their design, a quality Thomas (1972: 690) calls "serendipity," but which is a predictable benefit when models are used with skill.

Figure 12.1 (adapted from Holling 1978: 67-69) organizes uses and types of models according to the state of knowledge in a subject. The quantity and quality of data available on a problem are depicted on the y-axis;

the degree of understanding in conceptual or theoretical terms is measured on the x-axis.

Statistical or other forms of empirical or numerical models are most useful if data are plentiful but theoretical understanding limited (cell A; Figure 12.1). Here analysis is exploratory. It is intended to reveal regularities, order data, and organize thought. Much of human demography and econometrics operates in this realm. By contrast, in cell D we have plentiful data and understanding to match. Here models take up the job of representing and applying relatively mature scientific knowledge. Models are used primarily to communicate and generate predictions. Some population genetic models and many models of engineering and physics have these qualities. Maps, of course, excel in them. Simplifying models can provide relief from the paralyzing effects of having too much knowledge to fully comprehend, that is, to keep usefully "in mind." Simplification can be important because overly complex models have a suite of liabilities (Richerson and Boyd 1987: 33-34): they can be difficult to understand and analyze, and they may offer few advantages in predictive ability. The importance of simple models in this cell reminds us that models are not the temporary tools of nascent or emerging knowledge, dispensable as a science matures. Their forms and uses may shift even as their usefulness endures.

Many models concerned with either human behavior or the historical (evolutionary) sciences operate in the cells B and C. With limited data and understanding (cell B), models necessarily are speculative and provisional. Here models help to define and isolate problems. They facilitate preliminary analyses. Evolutionary archaeology and some life history and behavioral ecology models sit here. As theoretical understanding increases (cell C), models can be used in an interpretive or inferential fashion, even if data are limited (cell C). Models in cells B and C commonly direct attention to novel data, facilitating efforts to advance empirically, in a positive direction along the y-axis.

Models in cell A commonly are inductive; those in cell C are deductive. Analogic models usually are confined to cell B. While scientific advance consists most obviously in moving from cells A, B, and C as far as possible into cell D, this most desirable of outcomes may not be feasible, especially in the short term. Game managers face population-ecology decisions that cannot wait decades for data collection. For many subjects of high interest, archaeology and related historical sciences will always operate with much less data than we would like. Extraordinarily creative and time-consuming effort may be required of prehistorians hoping to nudge their subjects of investigation from cells B or C small distances toward A and D. Taphonomic studies are an example.

However, even if data are constrained, there are potential gains to knowledge in striving to move from cell B toward C. Interpretive models well verified in other contexts may be used to make sound inferences about

phenomena only incompletely substantiated with data (viz., ethnoarchaeological studies; see O'Connell 1995). When insufficient time or empirical hurdles are a problem, theoretical models advance the pragmatic aim of working as effectively as we can with limited information. By specifying and narrowing the analytical quest, they help to identify what kinds of data are needed in order to perform a compelling test, a special advantage when data are difficult to come by.

In practice, a model might be applied in three different modes: the evaluative, interpretive, and normative. The evaluative mode is focused on testing the model against observations. Evaluative statements commonly take the form: *If* the assumptions of the model accurately reflect the referent situation, *then* we expect the following observations. Here models provide provisional hypotheses and the analytical work lies in testing those hypotheses (see "Testing," below). The model is made subordinate to the empirical evidence.

The evaluative mode grades into the interpretive mode. In this mode one offers a plausible covering argument that the model applies to a situation; its congruence is accepted on this argument rather than by direct test. The consistency of the model with available observations is then used to interpret or extend them. Statements take the form: Because of its theoretical generality for social foragers, the producer-scrounger game likely characterized hominid hunter-gatherers, with the consequences x, y, and z for subsistence and social behavior. Interpretive use implies a model that is only weakly validated in either the theoretical or the empirical sense in that situation. Our willingness to engage in the interpretive use of a model depends on: (1) general confidence that the model captures the relationships it claims to represent, that is, on prior testing in related contexts; (2) on the covering argument for suitability in this context; and (3) on inability to make a more direct, evaluative use of the model in this context. The first two criteria recognize that models might be quite useful in a provisional way, even in situations in which they may eventually be tested directly. The third criterion recognizes their utility for the historical sciences (archaeology, history, geology, evolutionary biology) in which direct testing may not be feasible for any of a variety of reasons.

Purely evaluative use of social science models is rare because they are difficult to test thoroughly. An analyst willing to pursue evaluation is seldom so disinterested as to eschew completely extrapolating to interpretive uses. It likewise is the case that interpretive uses gain their plausibility by appealing to various indirect sources of evaluation. As a consequence it can be difficult to discern the relative balance of evaluation and interpretation in a particular application. Judgments are further complicated if authors are not explicit. It is tempting for the advocates of a model to slip unintentionally from evaluative to interpretive modes with few discernable signals that they have done so. Critics, on the other hand, are tempted to misread interpretive claims as evaluative. By associating the interpretive mode with the harder, scientific claims of the evaluative, they make their critical task easier.

The normative mode of applying models is more easily recognized. Here a model becomes the basis for a recommendation. The normative mode expresses confidence in the merit and applicability of the model, whether or not it conforms to empirical observations in the particular case. If the model has been tested and found to fit the case, it can serve its normative role, mainly that of reassurance, with high confidence: not only is behavior *x* observed when predicted, it is the behavior that *should* be observed. However, even if x is not observed, there may remain reason to assign a model a normative role. Although individuals are doing otherwise, x is what they should do if they are to most effectively realize their goals. The normative stance has an applied or practical flavor evident in microeconomic models. Normative use of models comes to the fore when pragmatic decisions must be made in the absence of exhaustive data, in fields such as conservation biology (Starfield and Bleloch 1986), economic development (Krugman 1995), or environmental assessment (Holling 1978).

The most immediate costs or dangers of modeling are simplification and reification. Models do simplify, selectively. In this, their most attractive feature (Boyer 1995) is also, for some, their greatest liability, lack of "realism." No one will easily confuse a map for a landscape. But many other models are less clearly distinguishable from the phenomena they represent, making it easy to mistake the model for the phenomena (reification). The selectivity accompanying the construction of a model means that information has been discarded, perhaps at a high cost. Models impose certain kinds of blindness. Problems may also arise if the potential arbitrariness of the selectivity is neglected. A model may neglect essential features of the problem or grant undue importance to features of little consequence. A more remote danger is the possibility that insistence on modeling will lead to the neglect or disparaging of subjects not, or not yet, amenable to formal treatment (Krugman 1995). Two preventative measures address these problems. First, users of models need to remind themselves periodically that models are instruments; they are the tools, not the object, of scientific workmanship. Second is thorough testing (see below).

> [We] ... build models to *explore the consequences of* what we believe to be true. ... Because we have so little data ... we learn by living with our models, by exercising them, manipulating them, questioning their relevance, and comparing their behavior with what we know (or think we know) about the real world. (Starfield and Bleloch 1986: 3) (emphasis in original)

However models are applied, their ultimate scientific value depends on the skill and thoroughness with which they have been tested. Appraisal is a significantly more general process than assessing the empirical fit of hypotheses with observations (Maynard Smith 1978). Four general classes of tests can be recognized:

(1) Is a model strictly faithful to its antecedent theory, and is that theory sound? These are not the straightforward questions they may seem. Some theories can be accepted with greater confidence than others. No theory, or more than one theory, may be perceived as applicable to the problem. And, even if a single choice is obvious, broad and compelling theories sometimes generate multiple, competing mandates concerning their most appropriate uses. The "pluralism" debates within neo-Darwinism attest to this (Ayala 1982; Gould 1983), as do debates over the best way to do ecology (Kingsland 1995) or evolutionary archaeology. Different schools of thought may make conflicting and hard-to-assess claims about which is the more legitimate heir to a particular theory.

(2) Does the model display internal logical consistency? There are two conditions here: (a) the conclusions should follow from the stated assumptions, and (b) they should follow from *only* those assumptions. Errors affecting the first condition are relatively uncommon. Deficiencies pertaining to the second condition are, unfortunately, a routine affliction of modeling. It commonly happens that the predictions of a model may depend on unrecognized assumptions, with the consequence that assessments about applicability and generality are misled. Relethford *(1999)* gives an example from models of hominid paleontology. A model may unwittingly assume what properly should be the object of inquiry. Krugman (1995: 53) states that von Thünen's central place model "assumes the [very] thing you want to understand: the existence of a central urban market."

In pursuit of logical consistency, good models bare their features forthrightly. Even good modelers sometimes neglect this transparency obligation, perhaps because the generality of a model is enhanced if it appears to depend on fewer rather than more assumptions. Nonetheless, those who use models regularly testify that the act of converting intuition or verbal reasoning into numerical, graphical, or mathematical form reveals assumptions, logical connections, and outcomes that otherwise would have been missed.

(3) Empirical evaluation of a model may entail direct testing of its assumptions. For instance, the encounter-contingent, resource selection model of foraging theory assumes that resource types are encountered randomly, according to a Poisson distribution. Although the model is used widely, to my knowledge this assumption has never been field tested. Direct testing of assumptions is important because assumptions may in part be dictated

by the availability of techniques rather than suitability to problem. For instance, early foraging theory models assumed deterministic input variables because stochastic or risk-sensitive techniques were not yet available.

(4) The final way of testing a model is a direct comparison of predictions with observations. Fit constitutes an indirect test of the model's assumptions and, working backward through this list of assessments, its internal consistency and antecedent theory. This is true *provided that* the predictions are unique to the model, the model unique to the theory, and the theory uniquely applicable to the problem. If we can derive the same predictions from theory or a model with different and incompatible assumptions, as happens with some frequency, then we are handicapped in choosing between them. In practice, when the predicted result is observed, it typically is taken as conferring a limited, favorable appraisal on higher levels of testing even when uniqueness is not assured. When fit is not observed, a search for the reason is initiated at those same levels. Was there an error in the derivation of the hypothesis from the model? Were the assumptions misplaced? Was the theory inappropriate? Hypotheses

fail for distinctive and, when they can be discovered, usually informative reasons.

The most basic form of empirical test is a single circumstance or noncomparative hypothesis, as in the following example: Its socioecological conditions are such that we would predict polygyny in species (or society) x. More satisfying are multicircumstance, or comparative, tests of the form: given their different socioecological conditions, we would predict that species (or society) x is polygynous and related species or society y is monogamous. Or, given seasonal changes in the density, costs, and benefits of the available resources, we can predict that the forager's resource selection will show pattern z through the year. Multicircumstance tests require that we find natural or experimental situations that exhibit the required variability of input conditions. This can be difficult, even though the opportunities for comparative testing are diverse (Maynard Smith 1978: 43). One can make comparisons among entities (species, societies or populations, individuals), or track comparative differences as any of these entities experiences changing conditions (seasonal, migratory, life cycle, etc.). Simulation can be used to replicate natural or experimental tests that cannot be conducted in actuality. If simulation increases the predictive range and specificity of models, it may thereby make them easier to test against the limited data sets available (Kohler 1999).

Few are free of the temptation to base allegiance to a model on the creative effort that went into its making, rather than the skeptical effort that may go into its evaluation. Nonetheless, confidence in a model grows as the number and detail of its tested linkages with the phenomena increase. Confidence is enhanced if the parameter values in the model are estimated independently of the observations that are used to assess hypotheses. It always is worthwhile to push as far in the direction of quantitative tests as

is possible. Qualitative tests can establish the necessity of an adaptive explanation, but not its sufficiency (Orzack and Sober 1993). Noting that hypothesis testing can only invalidate a model, Holling (1978: 95) argues that we are seeking *degree of belief* in the efficacy of a model. The more ways in which the credibility of a model survives being put at risk, the greater our justification for confidence in it.

Analytic Compromises

Levins (1966) argued that models can be classified according to three independent and sometimes competing qualities: generality, realism, and precision. Good maps have virtually no generality, but score well on precision and to a lesser extent on realism. The Hardy-Weinberg law has a high degree of generality, low realism, and moderate precision. Ecosystem models give priority to realism; precision is secondary and generality of lesser concern. Evolutionary ecology models, by contrast, emphasize generality; realism and precision are secondary objectives. Agent-based models tilt that configuration of qualities toward realism.

Levins states that because it is impossible to maximize all three of these qualities at once, trade-offs are central to effective model-building. He also argues that it takes a family of interrelated and partially overlapping models to address any particular subject, such as population biology. The simplifications necessary to make models analytically tractable mean that each can represent only a partial view of the whole. A family of models is bound together by their antecedent theory, and by the ways in which they overlap in addressing the phenomena. Thus, the evolution of life history characteristics cites neo-Darwinian theory and involves models of clutch size, parental investment, altruism and helping behavior, senescence, and so on. Finally, Levins suggests that individual models could be assessed by a desirable quality, their *robustness*. Robustness appraises the degree to which the predictions of a model are independent of the details of its construction (e.g., its assumptions). In practice, it means achieving the same result from partially independent starting points. For instance, it is said that the encounter-contingent model for resource selection has six separate derivations (Schoener 1987).

Levins' views on modeling have been extraordinarily influential, despite some deficiencies. He was not specific about the meanings of generality, realism, and precision. He did not provide explicit arguments that this was an exhaustive list of desirable model qualities, or that an increase in one or two of them by necessity comes at the expense of another. He did not provide a specific protocol for assessing robustness. Nonetheless, these shortcomings appear not to have diminished widespread citation and use of Levins' analysis, perhaps because his description resonates so well with the experience of those who create and use models.

The disparity between the substance and the influence of Levins' brief for modeling tempts philosophizing, and Orzack and Sober (1993) have taken up the shortcomings listed in the previous paragraph. They try to sharpen Levins' terminology, for instance, by defining generality and realism as relative or comparative terms, applicable only within the context of models addressing like phenomena. They use examples to show that generality, realism, and precision are not independent of one another and do not always combine to generate a stable taxonomy of models. For example, the same model can be classified differently depending on the degree to which its parameter values are made quantitatively specific, or "instantiated," in their terminology. They reject as vague the argument that models can be appraised by their robustness.

In response, Levins (1993) states that the formal analytic methods used by Orzack and Sober have little or no bearing on the practice of modelbuilding, or indeed, on science more generally. Orzack and Sober set out detailed examples to show that Levins' analysis is afflicted by ambiguous definitions and permeable, shifting categories. Levins embraces many of their examples and adds some of his own, but he rejects the larger lesson. In his view, it is entirely appropriate that the terms of his characterization are relative and that useful taxonomies of models are situational. That they can be confounded by formal logic and carefully chosen cases does not make them less reasonable as guides to the practice of constructing and using evolutionary models. Levins notes that philosophy is not evolutionary biology, and a too-tidy or strict sense of formalization does not necessarily advance creative or sound scientific practice. While the Orzack and Sober (1993) critique introduces some laudable caution, it seems unlikely that it will reduce the intuitive appeal of Levins' description of the trade-offs entailed in model-building.

General terms of approbation, like robustness, are common in the literature on modeling. Starfield and Bleloch (1986: 6-8) use the term *resolution* to describe a model's scope, what is or is not included, and the detail of its components. Models with too fine a resolution can become burdened with distracting detail; those with too coarse a resolution will yield inadequate understanding. Aris and Penn (1980: 9-11) speak of *craft,* to emphasize that a well-constructed model requires the triumph of judgment and skill over the uncertainty of incomplete knowledge. Clarke (1972: 4) speaks of *power,* a composite judgment made up of the comprehensiveness, predictiveness, efficiency, and accuracy of a model. Richerson and Boyd (1987: 43-46) use the term *plausibility argument* to summarize the degree to which a family of models is (1) theoretically and logically sound, (2) consistent with data, and (3) separable from competing approaches to the same subject. The variety of these terms for assessing model validity presumably arises from the multiple kinds of compromise that accompany the use of models. The diverse ways in which models can be used productively may also mean that any one-dimensional scale for judging their utility is necessarily somewhat ambiguous.

Metaphors, Narratives, and Reductionism

Several issues of a more philosophical nature might benefit from brief comment. First, how do models differ from metaphors and narrative? And second, what is the relationship between models and reductionism?

Metaphors amend a definition or image, usually by shifting its context. The metaphor, "love is a prickly rose," communicates the bittersweet quality of close personal relationships. Like models, metaphors represent one thing in terms of another in order to reveal properties not otherwise apparent. Like models, they are meant to add to understanding. Their qualities, such as precision, realism, or generality, are balanced according to the goal in using them. Metaphors may trade on suggestive ambiguity, whereas most models seek clarity and discipline, but even in these features there may be overlap. In search of a distinction, it will not do simply to claim that metaphors belong in the realm of literature and models in science. Science, like literature, depends on expressiveness and metaphors excel at facilitating communication. Nonetheless, there is a difference worth noting. In using a heuristic model we ask that it reveal not just the existence of a relationship in the referent-love is associated with troublesome as well as positive emotions, as roses come with thorns-but something about the causal basis for the relationship. Love may occasionally suffer emotional turmoil, but not for the reason that roses have thorns. Even in the case of analogy, a kind of model only slightly removed from metaphor, this causal likeness is implied, if not stated (Hesse *1966).* To the extent that a metaphor has scientific aspirations, it is properly viewed as an analogy in disguise.

Terrell (1990) explores how the study of prehistory can take another literary form, that of **narrative** or story-telling. His examples come from archaeological accounts of the successive waves of migrants that settled the dispersed Pacific Islands. Terrell argues that such narratives have a number of liabilities as science, but are only one step removed from models, and can be valuable when formalized sufficiently that they can be compared and tested, used evaluatively.

It is also important to distinguish between simplification and reductionism. Both sometimes invite hostility. If they are confounded such criticism will be inappropriately magnified. "The complexity of real social and biological phenomena is compared to the toylike quality of the simple models used to analyze them, and their users charged with unwarranted reductionism or plain simplemindedness" (Richerson and Boyd 1987: 27). Models do simplify by selective neglect, but they are not thereby reductionist in ways that should invite automatic skepticism. Following Mayr (1988), simplification, or what he calls **constitutive reductionism,** is both appropriate and perhaps necessary for scientific analysis. Richerson and Boyd (1987: 40) call this **tactical reductionism.** By contrast, **explanatory reductionism** (the claim that higher-level phenomena are fully explained by subcomponent properties) and **theory reductionism** (the claim that any theory applied to a higher level is only a special case of theory appropriate for lower levels) are more problematic. Models usually practice constitutive reductionism, but they need not entail either of the second two types. In fact, by fostering explicit simplification by parts, models may help us to avoid explanatory or theory reductionism, even as they give us entry to complex phenomena like cultural evolution (Richerson and Boyd 1987). The facile objection that models threaten human dignity by their simplification or reductionism is not creditable.

CASE STUDIES

Three examples will illustrate the utility of modeling in archaeology. These are drawn mainly from behavioral ecology, for the simple personal reason that I know this literature best. However, they might as easily have come from new archaeology (Binford's forager-collector model [1980], or Gould's [1980] analogic models of hunter-gatherer adaptations), from evolutionary archaeology (Rindos' [1984] co-evolutionary model of plant domestication), or other archaeological approaches. Modeling is one of the most ecumenical of scientific commitments; it transcends even sharp rivalries of method, interpretation, school, and totemic theorist.

Residential and Field Processing

Interpretations of prehistory depend on knowing if there are patterned relationships among the procurement and transport of resources, residential geography, and the discard of archaeologically visible materials. An important aspect of this problem arises when resources have high and low utility components. Because low utility parts (e.g., husks, shell, bone) are more likely to preserve, it is useful to know if processing and discard occur in the field where the resource is first located or at the camp where it may be transported for consumption. Bettinger et al. (1997) take up this question with a model first used in this context by Metcalfe and Barlow (1992).

The central place foraging model assumes that (1) humans will optimize their delivery of useful material to a central place; (2) procurement occurs in a locale some distance from the residential site, imposing outward travel and return transport costs; and (3) the resource can be separated into high and low value components. The essential trade-off is represented in Figure 12.2. Field processing increases transport efficiency because loads are made up of high utility parts only. But it reduces harvesting efficiency by taking

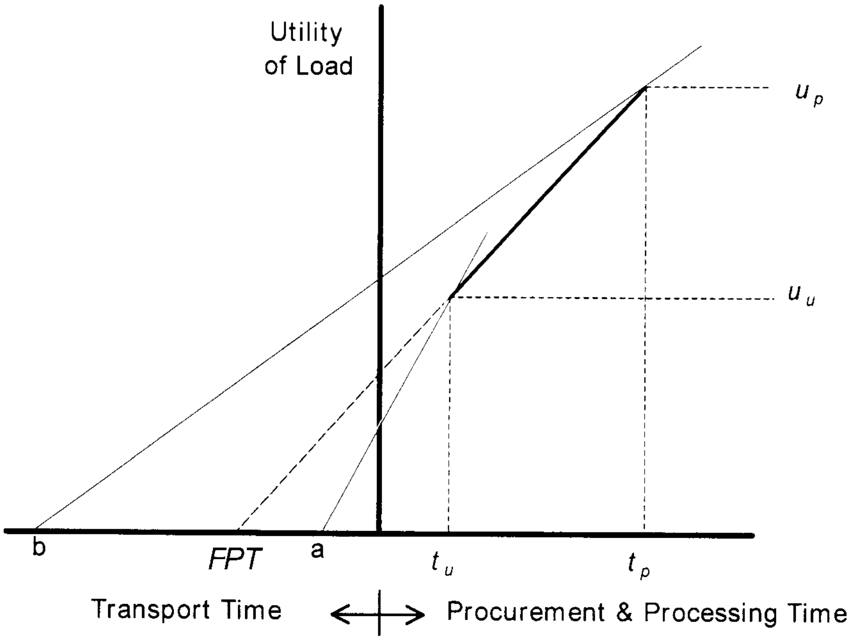

Figure 12.2. Central place foraging and field processing (after Metcalfe and Barlow 1992). This model applies to resources that can be processed to remove a bulky but low utility component (such as the shell of acorns or mussels), and that must be transported to a central place for consumption. The foraging costs and benefits of an unprocessed load of the resource (subscript u) are given by the procurement time (t_u) and resource utility (U_u). A load that has been field processed requires more work (t_p) to obtain and process, but has a higher utility (U_p). *FPT* is the Field Processing Threshold. If transport time is less than *the FPT* (point a), then hauling the resource unprocessed is the higher efficiency option. If transport time is greater than *FPT* (point b), then field processing before transport is a more efficient choice. Note that overall efficiency is given by utility divided by the time required for procurement, processing, and transport (or, the slope of the line segments drawn from "a" and "b"). A behavioral ecology model like this can be used to generate either qualitative or quantitative predictions by examining how changes in any of the input variables affect the relative efficiency of the two options.

field time that could be spent procuring more of the resource. Given measurement or estimates of activity costs and resource utility, processed and unprocessed, the model can be used to determine the travel costs (time or distance) at which field processing becomes more efficient; or, given a certain travel distance, it will show how much change in utility is required to shift from residential to field processing.

Bettinger and colleagues measure the volume of "burden baskets" used

by California Native American groups to estimate transport parameters (e.g., volume, weight). They survey ethnographic observations which allow them to estimate the costs and benefits of processing black oak acorns *(Quercus kelloggii)* and big mussel *(Mytilus californianus)* to various stages. The results show that even short travel times (45 minutes) make it advantageous to dry acorns in the field. However, subsequent stages of acorn processing (shucking, cleaning, pounding, etc.) almost certainly occurred residentially, as the minimum travel time associated with field processing is prohibitively long (25 hours, or a distance of about 80 miles). In the case of plucking big mussel from productive habitats, the travel threshold for switching between field and residential processing is well within a daily foraging radius of two hours. In this circumstance, model predictions may allow for more accurate interpretation of the archaeological record, even though they are highly dependent on local environmental details (Bettinger et al. 1997: 897).

The model employed by Bettinger et al. gives them a means of identifying "distinctive archaeological signatures of specific human behaviours with reference to the general principles that organize those behaviours" (1997: 887; see O'Connell 1995). It stimulated experimental and ethnoarchaeological collection of novel data. The model would be interesting enough for its qualitative applications (e.g., increased travel time makes field processing more likely). But even in prehistoric settings, it can be applied quantitatively (e.g., travel times greater than 25 hours are required to make field processing and discard of acorn shells effective).

Production Risk and Social Integration

Four hundred years of northern Anasazi prehistory is the setting for a model-based study of relationships between agricultural production and social adaptation. Kohler and Van West (1996) develop a model in which the sufficiency and variability of maize yields predict periods of social integration and disintegration. They test these predictions against markers of social structure at village and supra-village levels in a 1,500-squarekilometer region of southwest Colorado.

Utility theory and risk-sensitive techniques developed in economics and behavioral ecology make up the theoretical and conceptual basis of their model. The model assumes that: (1) the utility function relating the quantity of production and its value is sigmoidal; (2) households have the goal of optimizing utility; and (3), intra-annual (spatial) and interannual (temporal) correlation among household maize yields are < 1. A simple graphical formulation shows that in years of high yield a household does best by avoiding production variance. Household members would prefer to consume the average yield available if households pool and evenly redistribute their production. Further, the advantage of pooling mechanisms such as

village-level exchange and regional sociopolitical integration increases with production variance. By contrast, high variance and low average yields should lead to defection from such arrangements. Households can do best by "going it alone," and should withdraw from social interactions entailing food transfers. Through a stretch of variable but generally poor years, a household gives up more utility in its net donations to a pooling arrangement than it gets back, even if the quantities of maize given and received balance.

Empirical assessment of these predictions requires environmental measures of relative production level and variance, and archaeological measures of social integration. For the first, Kohler and Van West estimate relative corn yields for the A.D. 901-1300 period using paleoecological assessments of soil depth, elevation, soil moisture, and drought severity indices. Close dating is possible through tree-ring studies. They identify five unusually high, and an equal number of unusually low, production intervals, each lasting on the order of 10 to 50 years. This duration is presumed long enough for environmental conditions to have an impact on archaelogically visible social arrangements. The same data are used to calculate interannual and spatial variability of yields during each of the recognized intervals. The ten intervals are then rank-ordered from high production/high variance (greatest pressure for exchange) through low production/high variance (greatest pressure for defection), a deft translation of archaeological data into the input variables required by the model. For the second data set required, Kohler and Van West aggregate information from nine types of archaeological observation, each of which should correlate with increasing social integration (e.g., site size increase, roads). They use the break-up of aggregate sites as a measure of social disintegration.

An interval-by-interval comparison shows that "The general pattern of the [archaeological] record is strongly in the directions anticipated by the model" (1996: 183). For instance, the Chacoan build-up (A. D. 1100-1129) coincides with the highest period of expected cooperation, whereas the Chacoan break-up (A.D. 1130-1179) and regional abandonment (1270s1280s) coincide with low yield/high variance situations, when defection from social integration is expected. This success is somewhat tempered by the observation that predictions based on yield alone, and on variability alone, perform as well as the risk-sensitive model which combines these two variables. This occurs because of a particular limitation in the data set: level of production and production variability are highly correlated. In principle, with less covariant data it would be possible to distinguish among these alternatives.

The Kohler and Van West model produces an important counterintuitive prediction: a modest quantitative increase in environmental variance, or a modest drop in yield, could select for a qualitative reversal in the adaptive trend: village-level and regional social integration will switch to disintegration. A single model shows how household self-interest can move social structure in either direction, and exactly what environmental conditions make the critical difference.

Paleolithic Population Ecology

Stiner et al. (2000) take up the implications of resource selection for human population dynamics in the Upper Paleolithic of southern Europe and western Asia. Their study gives an old idea (the Broad Spectrum Revolution, or BSR) a new and more powerful analytic purchase by allying it with models in population ecology and foraging theory. As developed by Lewis Binford and Kent Flannery, the BSR states that a combination of late Paleolithic population growth and/or habitat shifts led to intensified exploitation of a wide range of resources, especially small game and low value, difficult-to-process plant foods, thereby setting the stage for domestication and the origins of agriculture.

Stiner and colleagues document the pattern of small-game use in northern Israel and western Italy during the early Middle Paleolithic through EpiPaleolithic periods-roughly 200,000 to 9,000 years B.P., with more precise dating varying by site-using archaeofaunal data. The percentage of small game in the overall diet and the variety of species taken are steady through this interval, but there are clear shifts in the type of organism that predominates in the small-game portion of the harvest. Relatively sessile, easy-tocapture and easily depleted species predominate early, but decline in importance late in the sequence. Edible shellfish and tortoises are key examples in Italy; tortoises, ostrich eggs, and slow-moving reptiles in Israel. In complementary fashion, relatively mobile, hard-to-capture and difficultto-overexploit organisms increase in importance late in the period under scrutiny. Italian instances are birds of several types, hare, and rabbits; in Israel, instances are birds, hare, and squirrels. Some of the less mobile and more vulnerable species such as limpet and tortoise also decrease in size in the latter half of this period, an indicator of heavy exploitation.

Easily captured, small-game species do not go unused in the Upper and Epi-Paleolithic periods, but they are supplanted in importance by species more difficult to capture. Because these changes occur against a background of stable biotic communities, an environmental explanation for the shift is unlikely. Rather, Stiner et al. (2000) suggest that human foragers met subsistence needs from highly ranked (easily caught) species in the Middle Paleolithic. Because these hunter-gatherers were low-density, dispersed, and perhaps highly mobile predators, they did not seriously deplete these readily captured species, despite their vulnerability to overexploitation. The Upper Paleolithic shift of the small-game component of the diet to low ranked and hard-to-catch species shows that exploitation pressure had grown. High ranked small game continued to be captured on encounter, but their diminished numbers and size also reflect more intense harvesting. This interpretation is consistent with predictions of the diet breadth model of foraging theory, given an increase in human population density and greater economic pressure on faunal resources.

This explanatory scenario depends on a clear economic separation between small game that are easy to catch and easy to overexploit, and those that are more difficult to catch but also more resistant to exploitation. To test their idea, Stiner and colleagues develop a population ecology model which simulates the effect of different exploitation levels on species like those in their scenario. They draw life history parameters such as age at first reproduction, litter size, life span, and so on, from wildlife biology studies of modern analog species. To assess the sensitivity of their approach to the obvious approximations of this procedure, they calculate the results for one standard deviation above and below mean values. The desired analytic output is assessments of potential yield and resilience to exploitation; the calculations are done by a computer spreadsheet. The results show that hare and partridge will sustain harvests of 7 to 10 times those of tortoises before exploitation results in local extinction. This difference is of a magnitude sufficient to account for the archaeological pattern.

In their conclusions, Stiner et al. (2000) argue that foraging theory and population ecology models reveal, better than simple diversity measures, the subsistence "signature" that should be associated with intensification like that envisioned in the BSR. The life history variability of small game, and the scatter of small-game species across the lower end of the resource ranking scale, make them sensitive indicators of prehistoric human demography and subsistence stress.

Summary

Each of these studies is packed with archaeological data and each draws on models from economics and behavioral ecology. None of these studies sets out to directly test the model it employs. Rather they attempt to make interpretive sense of the archaeological record in terms of the model. Inasmuch as the empirical record or theory behind the models have independent credibility, and the archaeological record is subject to no competing and superior explanatory contenders, the consistency of model and data lends sound inferential substance to explanation.

FUTURE IMPORTANCE

The explicit use of models is an act of cognitive humility, admission that raw intuition-which usually means the implicit, unexamined use of models-is quite often a fallible guide to causal relationships in complex systems. Models bring discipline to private comprehension by fusing creativity

with formal procedures. They facilitate clear and accurate communication. They advance efforts to identify and make use of critical data and to assess provocative ideas. It is hard to imagine a science without models of some form. Krugman (1995: 5) states that an idea will not be taken seriously in economics unless it can take the form of a simple model. It is a defensible claim that a science matures in parallel with its ability to develop the use of models.

An archaeologist attempting to assess progress in that respect might well watch the course of four efforts. First, to what extent can prehistorians recover data in forms that allow direct testing or evaluative use of models, or failing that, increasingly sound interpretive use of them? This is a y-axis question (Figure 12.1). Second, to what extent will simple, uni-topical models in fact coalesce into families predicted by Levins (1966), creating theory that is more synthetic in scope? This is an x-axis question. Third, will it be possible to reconcile use of optimization and equilibrium models with the role of agency and historical contingency in evolutionary processes? Ross (1999) discusses parallel issues between optimizing and historical approaches to adaptation. Broughton and O'Connell (1998) draw attention to the distinction in their review of evolutionary approaches in archaeology. Fourth, will it be possible to productively reconcile different schools of modeling (e.g., simple evolutionary ecology models based on individuallevel, optimization assumptions and analytic solutions, with system-level approaches based in equilibrium assumptions and simulation solutions, with dynamic, agent-based simulations)? Said differently, will the generality of marginal value theorem (Charnov 1976) and the realistic intricacies of Guilá Naquitz (Flannery 1986) ever come to rest in the same analytical effort, and will the result be enlightening?

ACKNOWLEDGMENTS

For encouragement, ideas, and useful, critical comments, I would like to thank Robert Bettinger, Tim Earle, Sheryl Gerety, Tim Kohler, Paul Leslie, Hector Neff, Mitch Renkow, John Speth, Mary Stiner, and Brain Tucker.

REFERENCES

Aris, R., and M. Penn. *(1980).* The mere notion of a model. *Mathematical Modeling 1: 1-12.*

Ayala, F. J. *(1982).* Beyond Darwinism? The challenge of macroevolution to the synthetic theory of evolution. *Philosophy of Science Association 2: 275-291.* Bettinger, R. *L. (1991). Hunter-Gatherers: Archaeological and Evolutionary Theory.* New York: Plenum.

Bettinger, *R. L.,* R. Malhi, and H. McCarthy. *(1997).* Central place models of acorn and mussel processing. *Journal of Archaeological Science 24: 887-899.*

Binford, *L.* R. *(1980). Willow* smoke and dogs' tails: Hunter-gatherer settlement systems and archaeological site formation. *American Antiquity 45: 4-20.*

Boyer, P. (1995). *Ceteris paribus* (All else being equal). In J. Brockman and K. Matson (eds.), *How Things Are: A Science Tool-Kit for the Mind.* New York: William Morrow and Co., pp. 169-175.

Broughton, J. M., and J. F. O'Connell. (1998). On evolutionary ecology, selectionist archaeology, and behavioral archaeology. *American Antiquity 64:* 153-165.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology 9:* 129-136.

Clarke, D. L. (1972). Models and paradigms in contemporary archaeology. In D. L. Clarke (ed.), *Models in Archaeology.* London: Methuen & Co., pp. 1-60.

Flannery, K. V. (1968). Archeological systems theory and early Mesoamerica. In B. J. Meggers (ed.), *Anthropological Archeology in the Americas.* Washington, DC: Anthropological Society of Washington, pp. 67-87.

Flannery, K. V. (ed.). (1986). *Guilá Naquitz: Archaic Foraging and Early Agriculture in Oaxaca, Mexico.* Orlando, FL: Academic Press.

Fretwell, S. D. (1975). The impact of Robert MacArthur on ecology. *Annual Review of Ecology and Systematics 6:* 1-13.

Golley, F. B. (1993). *A History of the Ecosystem Concept in Ecology: More Than the Sum of the Parts.* New Haven, CT: Yale University Press.

Gould, R. A. (1980). *Living Archaeology.* Cambridge: Cambridge University Press.

Gould, S. J. (1983). The hardening of the modern synthesis. In M. Grene (ed.), *Dimensions of Darwinism: Themes and Counterthemes in TwentiethCentury Evolutionary Theory.* Cambridge: Cambridge University Press, pp. 71-93.

Hagen, J. B. (1992). *An Entangled Bank: The Origins of Ecosystem Ecology.* New Brunswick, NJ: Rutgers University Press.

Hesse, M. B. (1966). *Models and Analogies in Science.* Notre Dame, IN: University of Notre Dame Press.

Holling, C. S. (1978). *Adaptive Environmental Assessment and Management.* Chichester: John Wiley & Sons.

Kingsland, S. E. (1995). *Modeling Nature: Episodes in the History of Population Ecology,* 2nd ed. Chicago: University of Chicago Press.

Kohler, T. A. (1999). Putting social sciences together again: An introduction to the volume. In T. Kohler and G. Gumerman (eds.), *Dynamics in Human and Primate Societies.* Oxford: Oxford University Press, pp. 1-18.

Kohler, T., and G. Gumerman (eds.). (1999). *Dynamics in Human and Primate Societies.* Oxford: Oxford University Press.

Kohler, T. A., and C. R. Van West. (1996). The calculus of self-interest in the development of cooperation: Sociopolitical development and risk among the Northern Anasazi. In J. A. Tainter and B. B. Tainter (eds.), *Evolving Complexity and Environmental Risk in the Prehistoric Southwest.* Reading, MA: Addison-Wesley, pp. *169-196.*

Krugman, Paul. (1995). *Development, Geography, and Economic Theory.* Cambridge, MA: MIT Press.

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

Levins, R. (1993). A response to Orzack and Sober: Formal analysis and the fluidity of science. *The Quarterly Review of Biology 68:* 547-555.

Maynard Smith, J. (1978). Optimization theory in evolution. *Annual Review of Ecology and Systematics 9:* 31-56.

Mayr, E. (1988). Is biology an autonomous science? In E. Mayr (ed.), *Toward a New Philosophy o f Biology: Observations o f an Evolutionist.* Cambridge, MA: Harvard University Press, pp. 8-23.

Metcalfe, D., and K. R. Barlow. (1992). A model for exploring the optimal tradeoff between field processing and transport. *American Anthropologist 94:* 340-356.

O'Connell, J. F. (1995). Ethnoarchaeology needs a general theory of behavior. *Journal o f Archaeological Research 3:* 205-255.

Orzack, S. H., and E. Sober. (1993). A critical assessment of Levins's *The Strategy of Model Building in Population Biology* (1966). *The Quarterly Review of Biology* 68: 533-546.

Relethford, J. H. (1999). Models, predictions, and the fossil record of modern human origins. *Evolutionary Anthropology 8:* 7-10.

Richerson, P. J., and R. Boyd. (1987). Simple models of complex phenomena: The case of cultural evolution. In J. Dupré (ed.), *The Latest on the Best: Essays on Evolution and Optimality.* Cambridge, MA: MIT Press, pp. 27-52.

Rindos, D. (1984). *The Origins of Agriculture: An Evolutionary Perspective.* Orlando, FL: Academic Press.

Ross, C. F. *(1999).* How to carry out functional morphology. *Evolutionary Anthropology 7:* 217-222.

Schoener, T. W. (1987). A brief history of optimal foraging ecology. In A. C. Kamil, J. R. Krebs, and H. R. Pulliam (eds.), *Foraging Behavior.* New York: Plenum, pp. 5-67.

Smith, E. A., and B. Winterhalder (eds.). (1992). *Evolutionary Ecology and Human Behavior.* New York: Aldine de Gruyter.

Starfield, A. M., and A. L. Bleloch. (1986). *Building Models for Conservation and Wildlife Management.* New York: Macmillan.

Stiner, M. C., N. D. Munro, and T. A. Surovell. (2000). The tortoise and the hare: Small game use, the broad spectrum revolution, and paleolithic demography. *Current Anthropology 41:* 39-73.

Terrell, J. (1990). Storytelling and prehistory. *Archaeological Method and Theory* 2: 1-29.

Thomas, D. H. (1972). A computer simulation model of Great Basin Shoshonean subsistence and settlement patterns. In D. L. Clarke (ed.), *Models in Archaeology.* London: Methuen & Co., pp. 671-704.

Thomas, D. H. (1973). An empirical test for Steward's model of Great Basin settlement patterns. *American Antiquity* 38: 155-176.

Tooby, J., and I. DeVore. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. In W. G. Kinzey (ed.), *The Evolution of Human Behavior: Primate Models.* Albany: State University of New York Press, pp. 183-237.

Winterhalder, B., and E. A. Smith. (2000). Analyzing adaptive processes: Human behavioral ecology at twenty-five. *Evolutionary Anthropology 9:* 51-72.

Winterhalder, B., and E. A. Smith (eds.). (1981). *Hunter-Gatherer Foraging Strategies.* Chicago: University of Chicago Press.

Wobst, H. Martin. (1974). Boundary conditions for Paleolithic social systems: A simulation approach. *American Antiquity* 39: 147-178.