

## FORAGING THEORY: UP, DOWN, AND SIDEWAYS

DAVID W. STEPHENS

*Abstract.* A large body of evidence is consistent with the idea that foragers tend to choose alternatives that yield more food in less time. But how do animals evaluate alternatives that vary both in time commitment and food gain? Two empirical trends, risk-sensitive foraging preferences and preference for immediacy, suggest that traditional models are incomplete because they ignore the temporal pattern of food acquisition.

Students of foraging theory are stepping *down* one level of organization by asking about the mechanisms of foraging behavior. I give two examples: an argument from foraging theory is used to evaluate the functional or adaptive significance of animal learning; and techniques from animal psychology are used to examine an issue—rules for patch leaving—that arose from arguments about foraging models.

Students of foraging theory are also stepping *up* one level of organization by addressing issues in population and community ecology. Although examples from functional response and resource partitioning show how this might proceed, advocates have yet to explicitly address its most fundamental issue: how accurate must a theory of feeding behavior be for it to be a useful building block of population models. If a precise theory is required, then work towards an accurate theory of behavior will contribute more to an understanding of population and community ecology than immediately applying current models to population processes.

### WHAT FORAGING THEORY IS AND IS NOT

In this paper I discuss current and future directions for foraging theory. I pay special attention to directions that empiricists might follow most profitably. By “foraging theory” I mean those models that are sometimes called by the unfortunate name “optimal foraging,” and their extensions and elaborations. The phrase, optimal foraging, is unfortunate for two reasons. First, although it is perfectly reasonable to try to distinguish foraging models that use maximization, minimization, or stability arguments from those that do not, “optimal foraging” is easily read to mean some claim about the single best way to forage. Foraging theory makes no such claim.

Second, even if you are an enthusiastic proponent of optimization models, many ideas—e.g., about perception, the development of behavior—must play important roles in any body of theory about foraging behavior and have little or nothing to do with optimization. Indeed, one source of the controversy surrounding optimization models of feeding behavior has been the absurd idea that “optimization” somehow summarizes everything anyone needs to know about foraging behavior.

This paper, reflecting my own biases and interests, is about how empiricists can most effectively influence foraging theory and foraging theorists. I would like to encourage others to do the kind of work that would tell foraging theorists what kinds of new models and new ideas are necessary to build a more accurate and general body of theory. My second motive is answering a question I am often asked: “Where is foraging theory going?” The title reflects the whimsical answer that I usually give to this question. I think

foraging theory is going in three directions—up, down, and sideways. I think students of foraging must go *sideways* by pursuing those questions they have traditionally asked: They must continue to ask evolutionarily motivated questions about the costs and benefits of, and constraints on, the foraging behavior of individuals. Students of foraging theory also find themselves stepping *down* one level of organization to ask questions about the mechanisms of foraging behavior. Moreover, many students of foraging theory have as their eventual goal stepping *up* one level of organization by using an understanding of foraging behavior to deduce things about the interactions of predators and their prey, or about population and community dynamics.

Following this logic, I have organized this paper into four sections. The first three sections correspond to my three directions: sideways, down, and up; while in the last section I discuss two components that make empirical studies influential.

### SIDEWAYS

In this section I outline the lessons that 20 years of foraging theory have taught, including lessons that encourage my own further interest in foraging models, and lessons that highlight the shortcomings of foraging theory. I am in a curious rhetorical dilemma: I want to convince the reader that foraging models have worked well enough to be worth further study, but no modeler wants to work in a field where all the problems are solved. As a consequence, I divide my review into two parts. First, I review the interaction between theory and data that encourages my further interest. Second, I discuss some more prob-

TABLE 1. ASSUMPTIONS AND PREDICTIONS OF BASIC FORAGING MODELS

---



---

Prey Model: Assumptions

Decision

- The set of probabilities of attack upon encounter for each prey type,  $p_i$  for prey type  $i$ .

Currency

- Maximization of the long-term average rate of net energy intake.

Constraints

- Searching for prey and handling prey are mutually exclusive activities.
- The forager encounters prey one after the other, and prey are encountered according to a Poisson process (a fine-grained environment).
- Three parameters—a net energy value  $e_i$ , a handling or involvement time  $h_i$ , and an encounter rate  $\lambda_i$ —can be associated with each prey type (e.g., with the  $i$ th prey type).
- Encounter without attack takes no time and causes no energy gains or losses.
- The forager is completely informed. It “knows” the model’s parameters, recognizes prey types upon encounter, and it does not use information it may obtain while foraging.

Prey Model: Predictions

- *Absolute preferences.* Prey types are either always taken upon encounter or never taken upon encounter (this is called the zero-one rule, because it is equivalent to saying that the optimal  $p_i$  must be either zero or one).
- Prey types are ranked by their profitabilities ( $e_i/h_i$ ), and types are added to the “diet” in rank-order.
- The “decision” to include a given prey type depends only on its own profitability and the profitabilities and encounter rates of higher ranked types. Specifically, inclusion should not be affected by a type’s own encounter rate.

Patch Model: Assumptions

Decision

- The set of patch residence times for each patch type,  $t_i$  for patch type  $i$ .

Currency

- Maximization of the long-term average rate of net energy intake.

Constraints

- Searching for patches and hunting within patches are mutually exclusive activities.
- The forager encounters patches one after the other, and patches are encountered according to a Poisson process (a fine-grained environment).
- Two things—a gain function  $g_i(t_i)$  that relates the time spent in a patch to the energy acquired there, and an encounter rate  $\lambda_i$ —can be associated with each patch type (e.g., with the  $i$ th patch type).
- The gain function has two important characteristics.

---

TABLE 1. CONTINUED

---



---

- It starts at zero ( $g(0) = 0$ ), spending zero time yields zero energy.
- It is initially increasing ( $g'(0) > 0$ ) and it eventually bends down ( $g''(t) < 0$  for all  $t$  values greater than some fixed  $t$  value).

- The forager is completely informed. It “knows” the model’s parameters, recognizes patch types upon encounter, and it does not use information it may obtain while foraging.

Patch Model: Predictions

- The patch-residence time should be chosen so that the instantaneous rate ( $g'(t)$ ) of gain at leaving equals the average rate of gain in the whole habitat. (Notice that this is an abstract mathematical condition, it is *not* the same as the leaving rule: “measure the instantaneous rate of gain and leave when it equals the habitat rate of gain.”) This condition has a number of interesting implications:
  - Whatever leaving rule the forager adopts, it should be one such that the instantaneous rates of gain at leaving are the same in all patch types.
  - If the time required to travel between patches increases, the rate-maximizing patch residence will increase. (For some degenerate gain functions it can stay the same.)

---

lematic (and hence more exciting) issues that have arisen during foraging theory’s development.

#### THE ENCOURAGEMENT

Broadly speaking, two models of foraging have been studied widely enough to allow discussion at a general level. These are the “prey model” (sometimes called the diet model) and the “patch model” (sometimes called the marginal-value theorem). Both models take the familiar form of optimality models; i.e., they make some assumption about what is maximized, a currency assumption; they make another assumption about what is controlled, a decision assumption; and they make assumptions about the things that place limits on the decision and currency, the constraint assumptions (see Stephens and Krebs 1986 for detailed discussion). The two models make identical currency assumptions and similar constraint assumptions; but they make very different decision assumptions (Table 1). These models are extremely simple, but they can in principle make detailed quantitative predictions and somewhat weaker qualitative predictions.

Kamil et al.’s (1987) recently published collection of papers on foraging behavior begins with two papers that evaluate the success of these models. The first of these (Schoener 1987:48)

concludes that foraging theory "... has often been verified with tests and therefore it *should* be pursued further" (emphasis Schoener's). Although the second (Gray 1987:95) concludes that foraging models (together with all optimality models) are such dismal failures that they "... could be said to weigh like a nightmare on the brain of the living." Krebs and I have addressed the difference between Gray's and Schoener's conclusions (Stephens and Krebs 1986), and while I agree more closely with Schoener than with Gray, I recommend that readers compare both papers and form their own opinions.

### *Empirical lessons*

By and large, the *quantitative* predictions of the patch and prey models have not fared well, with some exceptions. Stephens and Krebs (1986) found only 11 unambiguous, quantitative fits in our tabulation of 125 studies. (We took the authors' interpretations at face value, so even some of these quantitative fits have been criticized; but, on the plus side, many [about 64%] of these studies were not designed to test quantitative predictions.)

However, two astonishingly consistent *qualitative* trends are evident. The first is predicted by the patch model: the time spent exploiting a depleting patch should increase as the time required to travel between patches increases. This prediction has been found to hold practically everywhere it has been studied (the only disagreements I know of are a case in which exploitation time was unaffected by travel time [Waage 1979] and another in which the effect persisted when it should not have [Kacelnik and Cuthill 1987]). Indeed, I think this may be the most general empirical trend to emerge, not just from foraging theory but from the spate of modeling in behavioral and evolutionary ecology that began in the late 1960s.

The second qualitative success of these models is almost as universal: as predicted by the prey model, foragers selectively attack prey items that are most profitable (they have the highest ratio of "energy available,"  $e$ , to "time required for handling and consumption,"  $h$ , in symbols  $e/h$ ). Even Gray (1987) acknowledges the pervasiveness of this trend, and his tabulation shows that this prediction was supported in over 75% of relevant studies. Gray dismisses this by arguing that this prediction is trivially obvious, but Schoener (1987) points out that this obviousness is not reflected in the pre-foraging-theory literature. In fact, ecologists before the advent of foraging theory mainly argued about whether animals were selective at all (references in Schoener 1987). Gray did not review tests of the patch model.

To be sure, these models also have their qualitative failures. The prey model's prediction of absolute preferences (the idea that a given prey type should always be ignored or always be accepted) has been consistently rejected. The prey model's other main prediction (that a type's inclusion "in the diet" does not depend on its own abundance) has sometimes been supported and sometimes rejected. I think this prediction does pretty well, if one considers the relative quality of studies supporting and rejecting it, but at face value the results are clearly mixed.

How, then, can I be encouraged? The answer comes from knowing something about the models behind the predictions, and especially behind modifications of those models. While empiricists have been comparing the models to reality, modelers have been trying to improve them logically, either by making them more general or by making them more appropriate for particular empirical situations. These modeling efforts show us that the two empirically confirmed trends (the patch model's travel time-exploitation time correlation, and the prey model's preference for more profitable prey) are also the two trends that, on *a priori* grounds, we would expect to be the most robust. For example, McNamara (1982) has persuasively argued that the "travel time-patch exploitation time" correlation would be predicted by *any* rate-maximizing model; although the details of rate-maximizing patch-leaving behavior can vary widely, this simple trend should remain. To take an example from the other side of the coin, modelers have also shown that the absolute-preferences prediction is very fragile; indeed a modification as mild as allowing choice behavior to have a variance greater than zero makes the prediction of absolute preferences evaporate (Stephens 1985).

My conclusion is that a surprising amount of data from a wide range of taxa are consistent with the simple notion of rate maximizing. Long-term, average rate maximizing is, of course, a specific way to combine less time and more energy (or simply more food if nutrients other than energy are important). A critic might argue, and I would agree, that many models that somehow place value on options that provide more food in less time would be consistent with these qualitative trends. (Below I will explain why one cannot dogmatically assert rate maximizing.) So, we have as a minimal and conservative conclusion that foraging animals act economically, in the sense that they tend to choose alternatives that yield more food in less time.

### *Some conceptual lessons*

Foraging theory has not only had some empirical successes, but it has also had some im-

portant conceptual successes, because it has changed the way students think about this subject. For example, the prey model shows that the choice of a diet (a list of the things an animal eats, and sometimes the proportions of these things) is a consequence of two types of behavioral choices that are logically different. The prey model predicts which items should be attacked, given a fixed and well-defined process of encounters with prey; this *encounter process* characterizes which prey are encountered during search and how frequently. But nothing in theory or logic says that a forager cannot also make decisions that change the encounter process. Foragers might change it by doing obvious things, such as moving from one part of their habitat to another, or by doing subtle things, such as looking up instead of down. Hence, the diet is determined at a minimum by (1) choices that determine the parameters of the encounter process, and (2) the choice of which items to attack and ignore. This simple separation of choices casts doubt on stomach contents studies of selectivity and choice; rather, it suggests that watching foraging behavior may be the most informative way to study diet choice.

A related idea is the concept that a forager's perceptive abilities define what a "prey type" is; to paraphrase a clever phrase maker: animals do not eat Latin binomials. Two species may form a single type, as they do in model-mimic systems, or (more commonly) a single species may form many types; small *Genera generalis* caterpillars and large *Genera generalis* caterpillars may well be different types from a forager's perspective (see Getty 1985 for a sophisticated discussion).

The patch model also has its conceptual successes, but because the phenomenon of patch tenacity is really new, these successes cannot be contrasted with older approaches. Two generations of modeling this problem have suggested two quite different economic reasons to move on to a new patch. The first (originally proposed by Charnov 1976) is that patches usually decline in quality as the forager exploits them; one reason to leave is simply because things are getting worse. The second reason (Oaten 1977) is that experience gained while exploiting a patch may tell the forager that this patch is an inferior one and hence not worth further effort. Both of these reasons sometimes apply (e.g., Lima 1983 for the patch assessment case, and Cowie 1977 for patch depression), but we do not know much about their relative importance in nature.

#### THE DISCOURAGEMENT: WHAT NEXT?

Many things are wrong with foraging theory as presently constituted; most are aspects of foraging behavior that have been left out of tradi-

tional foraging models. A list of aspects that need to be included in a more general theory of feeding can be found in the chapter headings of Stephens and Krebs (1986): *Incomplete information*, including problems of resource assessment and the abilities of foragers to recognize and discriminate prey and other resources; *Tradeoffs*, including tradeoffs between energy and other "nutrients" (including toxins) between foraging and predation, or foraging and reproduction; *Risk-sensitivity*, including general questions about the pattern of resource acquisition in time and how foragers value different "patterns"; *Dynamic tactics*, the problem of allowing decision variables to be functions of other "state variables," so that one can solve for the best trajectory of decisions instead of the single best decision. Rather than discuss each of these here, I will discuss a particular empirical issue that I think addresses some fundamental flaws in the traditional assumptions. I would like to explain why I retreated from stridently advocating rate maximizing to the milder position that foragers value less time and more food (or energy) in some vague way. One might deny strict rate maximizing, because it ignores complementary nutrients and the threat of predation. Indeed these are limitations, but rate maximizing cannot be generally correct even in conditions in which time and energy alone are important because (1) it ignores the variability in food gain (or *risk*), and (2) because it ignores the importance of immediacy in food gain.

#### *Risk*

Conventional foraging models were built on the premise of maximizing the "long-term average rate of energy intake," which is a very specific and potentially restrictive assumption. Consider the difficulties inherent in limiting our attention to *averages*. Suppose that a forager can choose between two alternatives. Alternative A provides a mean food gain of 10 joules and standard deviation of 10 joules in a period of 1 min, while alternative B provides a mean food intake of 10 joules and standard deviation of 1000 joules also in a period of 1 min. Because both alternatives take the same time, they obviously provide the same average rate of energy intake. A model based on long-term, average rate maximization would provide no basis for preferring one of these alternatives; instead, any such model predicts that foragers should be indifferent between the high and low variance choices. But a real forager would hardly be indifferent between these two choices that vary so much in variance or risk? Animals have consistent preferences when presented with alternatives that vary only in their degree of riskiness, even when means do not vary (Caraco et al. 1980, Real 1981, Real et al. 1982,

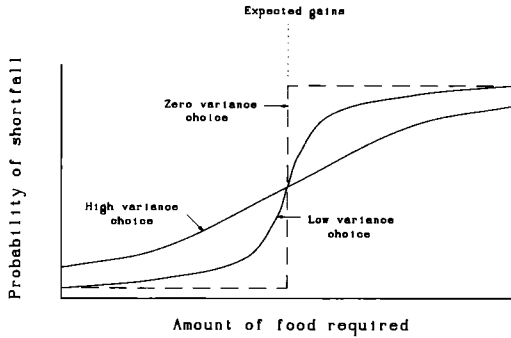


FIGURE 1. How minimizing the probability of an energetic shortfall predicts changing risk-sensitivity. The two curves represent the probability that a forager will fall short as a function of the amount of food required. The solid curve plots the probability of a shortfall for a normal distribution with low variance, and the dotted curve plots the probability of a shortfall for a normal distribution with high variance. The low and high variance distributions have the same mean (= expected gains). If the food requirement is greater than expected gains, then the high variance distribution yields a lower probability of a shortfall, but if food requirements are less than expected gains, the low variance distribution yields the lowest probability of a shortfall (after Pulliam and Millikan 1982).

Caraco 1983, Barnard and Brown 1985, Stephens and Paton 1986, Wunderle and O'Brien 1986). In my view, conventional foraging models fail because they require that foragers be indifferent over risk (or variance). This requirement is not only counter-intuitive, but it is also an empirical failure.

The work of Caraco and his colleagues provides several important examples of risk-sensitive preferences. Caraco et al. (1980) showed that Yellow-eyed Juncos (*Junco phaeonotus*) kept on positive energy budgets (i.e., fed at a rate that allowed them to maintain their *ad libitum* feeding weight) preferred certain alternatives, whereas those maintained on negative energy budgets preferred variable alternatives. This may mean that juncos are not maximizing the long-term, average rate of gain, but instead are minimizing the chance of falling short of some critical amount of food (see Fig. 1) at some critical time. The presence of risk sensitivity hints that part of the problem with long-term rate maximizing is its failure to consider details of the temporal pattern of food acquisition.

### Immediacy

Consider another set of hypothetical alternatives. Suppose that every 2 min a forager is offered two alternatives. Alternative  $\alpha$  leads to 1 joule of food delayed by 30 s and alternative  $\beta$

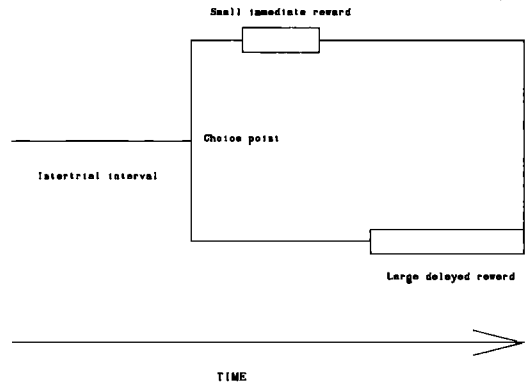


FIGURE 2. The contingencies for a typical experiment (in this case Green et al. 1981) showing preference for immediacy. After a fixed intertrial interval, the forager is presented with two mutually exclusive alternatives. One leads to a small, relatively immediate reward; the other leads to a larger delayed reward. In the experiment shown, reward size was controlled by controlling duration of food access, so a large reward takes more time. Notice that the total "cycle time" is the same regardless of the subject's choice.

leads to 10 joules of food delayed by 1 min. Because the time between offers is fixed, the forager must wait 1.5 min from being fed until the next offer if it chooses alternative  $\alpha$ , but it must wait only 1 min if it chooses alternative  $\beta$  (Fig. 2). If the long term is all that is important, then these alternatives amount to nothing more than 1 joule in 2 min, and 10 joules in 2 min, and  $\beta$  must be a much better choice. Annoyingly, real animals do not always agree. It is easy to find instances in which foragers prefer smaller but more immediate gains, even when they could do better in the long term by waiting for larger gains. This phenomenon is well known among animal psychologists (e.g., Green et al. 1981). Behavioral ecologists are just beginning to investigate this phenomenon in animals other than rats and pigeons, and at least one such study agrees with the psychological results (e.g., Barkan and Withiam [in press]).

Preference for immediacy is a vexing problem and few attempts have been made to explain it, compared to the number of attempts to explain risk sensitivity. The most convincing explanation is that foragers expect to be interrupted (by conspecifics or predators), so that the immediate small thing may actually be better than the delayed large thing (Kagel et al. 1986, McNamara and Houston 1987a). While this is the most reasonable explanation available, I think the effect is too strong and persistent to be explained completely thusly because none of the experiments have included any interruptions. So proponents

must argue that animals are “hard-wired” to expect interruptions (or that something external to the experiment itself has created such an expectation). This is possible, but animals react in other ways to the presence and absence of potential interrupters.

Like risk sensitivity, preference for immediacy points out that long-term, average rate maximization ignores some important features of the flow or pattern of food acquisition. Taken together, preference for immediacy and risk sensitivity punch a sizable hole in long-term, average rate maximizing.

Moreover, there are reasons to think that these two phenomena are related. To represent the decrease of food value with delay, suppose that a delay of  $\delta$  seconds means that an amount of food  $A$  is really worth only  $Af(\delta)$  (Fig. 3A), where  $f(\delta)$  is a *discounting function* that represents the fraction of  $A$ 's value that remains after a delay of  $\delta$ . Figure 3B shows that this positively accelerated discounting function also predicts that a forager should prefer a probability distribution of delays to a certain delay with the same mean (for example, the risky choice  $\Pr(\delta_1) = \Pr(\delta_2) = 1/2$  should be preferred to the certain choice  $\Pr(\frac{\delta_1 + \delta_2}{2}) = 1$ ). Indeed, this trend has been widely observed; rats and pigeons prefer alternatives with variable delays before reward to alternatives with fixed delays (see Hamm and Shettleworth 1987).

*A biological time-energy problem.* Both risk sensitivity and preference for immediacy suggest that something about the pattern of food acquisition is important and neglected. An enormous body of evidence is consistent with the “more food in less time” postulate, but both risk sensitivity and preference for immediacy show that, when it comes to details, we do not know how or why animals evaluate combinations of time commitment and food gain as they do. The possibility of a link between these two phenomena makes me hopeful that some crucial piece of the puzzle may click into place at any moment. I think this is the most fundamental “sideways” problem in foraging theory, because a solution would change our view of every aspect of feeding behavior.

I have talked about *the* solution of the time-energy problem, but many solutions might exist. Different individuals or species may value time-amount combinations differently at different seasons or phases of their life history. I might hope for a general solution, but I certainly do not insist on one. In fact, there may well be ecological correlations that would be exciting discoveries if they hold up. For example, while birds (pigeons and chickadees, for example) seem to have strong preferences for immediacy, my own work with

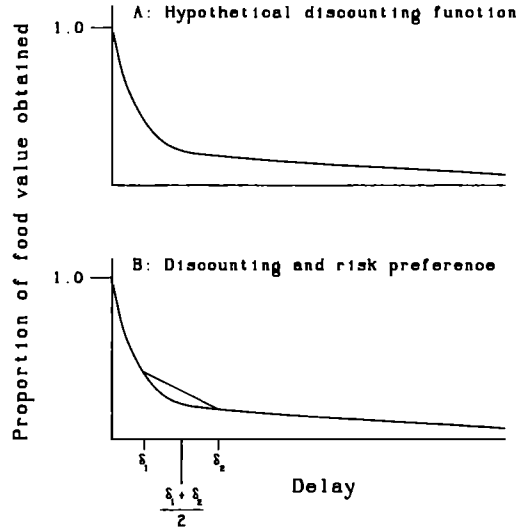


FIGURE 3. (A) A hypothetical relationship between delay and food value: food item that would have value  $A$  if obtained immediately will only have value  $Af(\delta)$  if it is delayed by  $\delta$  s. The “discounting” function  $f(\delta)$  is shown in this panel. (B) If the “discounting” function is curved as shown here (positively accelerated), then this predicts risk preference over delays. Suppose that the forager can choose between (1) an alternative that yields  $\delta_1$  half the time and  $\delta_2$  half the time, and (2) an alternative that always yields  $\frac{\delta_1 + \delta_2}{2}$ . The expected food value from the risky alternative will lie halfway along the line segment that connects the points  $(\delta_1, f(\delta_1))$  and  $(\delta_2, f(\delta_2))$ . The expected value of the risky alternative is higher than the expected value of the certain alternative, because this line is always above the curve.

honeybees suggests that the preferences of these social, dawn-to-dusk foraging machines are consistent with long-term rate maximizing, because honeybees will pass up immediate gains to increase longer term gains (Stephens et al. 1986). A similar point can be made about risk sensitivity, and how general the switch from risk preferring to risk avoiding shown by Caraco’s juncos may be. Early indications suggest that it may not work for Bananaquits, *Coereba flaveola* (Wunderle et al. 1987).

## DOWN

In the last few years, many behavioral ecologists have begun an exciting collaboration with animal psychologists. Why should psychologists care about the evolutionary function of behavior? Stephens and Krebs (1986) answer this question using a slide rule as an example. *The mechanism* of a slide rule is defined in terms of how it accomplishes the function of multiplication. You would interpret this differently if you thought

it was a device for digging holes. In more general terms, function is implicit in most arguments about mechanism. It is always a good idea to bring hidden assumptions into the light (see also Kamil and Yoerg 1982, Shettleworth 1983).

A traditional psychologist might agree, but argue that general models of function are sufficient to interpret mechanisms, such as "the function of feeding behavior is to acquire food." Yet, foraging theory has discovered new phenomena such as the "travel time-patch exploitation time" correlation that were not evident from more general arguments. Indeed, the concept of patch-exploitation tactics has fostered new research on psychological questions about rules for patch leaving (see below).

What does animal psychology offer, and why should foraging ecologists care about mechanisms? Simply, more details lead to better models. I can make this point somewhat more formally by returning to the elements of optimality models: currency, constraint, and decision. Early foraging models make unrestrictive and general assumptions about constraints (e.g., foragers cannot search and eat at the same time), which results from psychology can make more sophisticated. For example, Getty and his colleagues (Getty 1985, Getty and Krebs 1985, Getty et al. 1987) have taken some models derived from sensory psychology (signal-detection theory, Egan 1975) to derive predictions about the detection of cryptic prey. The results are impressive; they have refined foraging theory's view of what a prey type is (see Stephens and Krebs 1986, Chapter 3).

A subtle variant on the same theme is what I call the feasibility-of-mechanisms problem. A traditional foraging model can work only if there is some decision mechanism, which in the patch model must allow a link between travel time and patch-leaving decisions. Traditional models ignore these mechanisms, by assuming that one mechanism can be implemented as easily (and as cheaply) as any other. This cannot be generally correct, and may be an issue that psychologists might help resolve.

I think most students of feeding behavior would agree that both laboratory and field studies have something to offer, but persons tend to specialize. Kamil (1988) has addressed this dichotomy by defining the different goals of laboratory and field work. He argues that all studies should have two goals: high internal validity (such issues as the repeatability of results, and avoiding confounding variables) and high external validity (how readily one can generalize from the situation studied to others). These goals are usually in conflict: an operational decision that increases internal validity will often decrease external validity and *vice versa*. Laboratory work tends to have high internal validity but compromised external

validity, while the reverse is true of field work. Behavioral ecology is a complex subject; it is probably expecting too much for a single study to establish the *general* validity of a result.

#### TWO EXAMPLES ON THE WAY DOWN

Below I give two examples of important "down" questions. One is a case in which foraging theory seems to say something new about a field in the traditional domain of animal learning. The second is an instance in which a mechanistic perspective is making inroads into a traditional question in behavioral ecology.

##### *Information acquisition and animal learning*

An animal's experience often changes its behavior. Animal learning is a central topic in animal psychology, in which psychologists have focused on mechanisms. Recently, behavioral ecologists and students of foraging behavior have begun to look at functional aspects.

Older models supposed that foragers were completely informed and did not need to use experience to improve their foraging decisions. Modelers were initially attracted to learning simply because they wanted to improve their models by allowing foragers to use new information. Because the approach of these "learning" models has been functional, conditions seem ripe for the kind of "function-mechanism" interaction that I advocated above.

With a few exceptions (Hollis 1982, Johnston 1982, Kamil and Yoerg 1982, Staddon 1983), psychologists seldom discuss functional significance of learning. Johnston (1982:74) concludes that "the ability to learn . . . has as its primary selective benefit that it permits adaptation to ecological factors that vary over periods that are short in comparison with the lifetime of an individual." This is a sensible idea; however, recent analyses of "incomplete information" problems in foraging theory show that it is only a part of the functional story.

When I began to model how to track a changing environment, I thought that the inclusion of a term that represented how frequently the environment changed would allow me to make a more quantitative statement than Johnston's. In rough outline, I made the following assumptions (see Stephens 1987 for details). Some varying resource always looks the same even though it can actually be in one of two states, good or bad. Although these states look the same, the forager can easily tell the difference when it exploits the resource. There is an alternative, stable resource whose quality is mediocre. I represented the *persistence* of the varying resource by a conditional probability; hence, persistence is the probability that a good state will be immediately followed by another good state, or that a bad state will be

immediately followed by another bad. When my hypothetical forager experiences a bad state, it can switch to exploiting the stable-but-mediocre alternative, but should occasionally return to check out the varying resource.

My hypothetical forager was free to use its experience or to ignore it by adopting an “averaging” tactic—attacking only the varying resource or only the mediocre-stable resource, whichever had the highest average quality. Hence, I was able to look for conditions in which learning was worthwhile.

The first thing I discovered surprised me: when the varying resource has no persistence, an averaging tactic is best. Superficially this seems like a counter-example to Johnston’s assertion, a case in which an environmental feature changes over a period that is much shorter than an individual’s life, but learning is not an economically sound policy. At first I dismissed this as a special case, believing that if I looked at the whole range of persistences I would find that learning paid off most at some intermediate level of variability. Instead I found that the payoff increased continually, and that the longer states persist the more worthwhile it is to learn about them.

This suggests that there is more to the relationship between environmental variability and the value of learning than Johnston’s statement implies. This apparent paradox can be resolved by thinking of two kinds of predictability. Johnston’s statement deals with the ability of the previous generation to predict the environment of its offspring (when this predictive link is weak, learning is favored), while my argument has to do with the ability of an individual’s experience to predict the future states of its own environment (when this predictive link is strong, learning is favored). This opens up an enormous number of new and fascinating questions about how these two kinds of “predictability” may or may not be related, and how these relationships may affect the value of learning.

Finally, these arguments about the functional significance of learning provide a more serious example of my “slide rule” point. Behaviorists often want to make statements about the presumably mechanistic limitations of what can be learned: A stimulus of type A can be associated with food, but a stimulus of type B cannot. I have concluded that learning may not be worthwhile in some situations. There is a big difference between something that is not learned because of a mechanistic limitation and something that is not learned because it is not worthwhile.

#### *Rules for patch departure*

Can animals count, keep track of the time between two events, or integrate information about

time and number? These are the kinds of questions that psychologists study.

Recently such questions have become important, for purely theoretical reasons, in foraging theory. Early workers (see Charnov 1976) on patch-leaving models seemed to suggest that a general rule for patch departure has the form: leave the patch when the instantaneous rate of gain drops to some critical value. While this rule may work sometimes, its generality has been widely criticized (Oaten 1977; Green 1980, 1984; Iwasa et al. 1981; McNair 1982). Four types of patch-leaving rules have been presented: (1) a fixed-number rule: leave after finding  $n$  prey; (2) a fixed-time rule: leave after spending  $t$  seconds in the patch; (3) a giving-up time or run-of-bad-luck rule: leave  $\tau$  seconds after the last prey capture; (4) a rate rule: leave when the “instantaneous” rate of prey capture drops to some critical rate. Iwasa et al. (1981) have shown that different rules work in different situations. For example, if prey are captured at random intervals and all patches have the same number of prey, then a fixed-number rule makes sense. However, if the number of prey per patch is highly variable, a run-of-bad-luck rule makes more sense. If the number of prey per patch follows a Poisson distribution, then a fixed-time rule works well (Iwasa et al. 1981). Hence, an esoteric argument about models of patch leaving has helped to place some issues from animal psychology (such as counting and timing) in ecological perspective.

More importantly, some of my colleagues at the University of Massachusetts have performed an experiment designed to deduce what kinds of patch-leaving rules animals actually use. Kamil, Yoerg, and Clements (in press) presented feeding Blue Jays (*Cyanocitta cristata*) with a simple patch-leaving problem. Two patchy resources were available to an individual jay. One resource was initially of high quality but eventually depleting, while the other was of low but constant quality. To simplify matters, Kamil et al.’s depleting patch depleted suddenly; hence up to  $n$  prey were delivered at a high fixed rate in the depleting patch, but no prey were delivered after the  $n$ th. Each bird was exposed to three treatments  $n =$  three, six, and nine prey; and each jay experienced a single treatment (two patches per day) for a very long time (often up to two months).

Since in any given treatment there is a fixed number of prey in the good patch, the best rule is obviously a fixed-number rule: exploit the high-quality depleting patch until it has provided all  $n$  prey, then switch to the constant, mediocre patch. Kamil et al. examined patch-leaving rules by calculating the relative frequency of patch-leaving events that were preceded by all possible numbers of prey captures. They found that a fixed-number rule did not completely explain the



data; instead they found that the birds seemed to use a combination of a number rule and a run-of-bad-luck rule. Kamil et al. argue that a mixed rule is a safe but flexible policy. A jay adopting a fixed-number rule would do very badly if the number of prey in the depleting patch suddenly went up (since it would miss all prey arriving after the old  $n$  value), although a pure run-of-bad-luck rule may be prone to early leaving mistakes. Real animals seem to have outsmarted the theoreticians.

Studies of "rules of thumb" are a growing and important part of foraging theory. Behavioral ecologists and animal psychologists will have much to say to one another about this search for good but simple rules for implementing foraging tactics.

#### UP

One of the goals of early foraging theorists was to build models of feeding behavior that would help them understand issues in population and community ecology. This approach follows the reductionist view. Many ecological interactions are feeding interactions: for example, a predator consumes its prey, and competitors may compete for food. Hence, the logic goes, a correct theory of feeding behavior is a precursor to deducing a correct theory of population- and community-level phenomena. This argument is sound, but requires two caveats. First, even a perfect theory of feeding will need to incorporate facts about reproduction, physiology, and morphology before it can be extended to population- and community-level interactions. Second, it remains to be determined how correct must such a theory be. Some ecologists feel that current theory is adequate, while I and others do not. I would like to hear more explicit discussion of how accurate a theory of feeding must be to serve as an adequate building block of population processes (see Comins and Hassell 1979 for discussion of this question).

#### TWO WAYS TO COMBINE FORAGING THEORY AND ECOLOGY

Foraging theory has been combined with population- and community-level models in two ways. The first is a kind of separate but equal approach, in which it is used to explain and refine ideas from existing ecological theory. The second is the seldom tried approach of a completely integrated and simultaneous study of foraging behavior and population-level processes.

##### *Separate but equal*

Any model of predator-prey interactions makes assumptions about the predator's feeding behavior, such as each predator eats  $x$  prey per day. Population modelers try to make these assump-

tions fairly general and *descriptive*. Foraging theory, on the other hand, is fundamentally an *explanatory* approach, which considers why a predator eats only  $x$  prey per day in terms of such economic variables as the abundance of alternative prey and the predator's hunting abilities. A population modeler may be justified in ignoring such explanatory questions, so long as a good estimate of  $x$  is available.

These contrasting perspectives make sense because the two approaches have different goals, but they also complement each other. Specifically, explanatory models can help population modelers to refine their assumptions. For example, simple predator-prey models usually assume that all members of the predator species are equally effective at consuming all members of the prey species. A population modeler might use arguments from foraging theory, such as size selectivity, to make some members of the prey species more vulnerable than others.

Conversely, the descriptive assumptions of population modelers show, in a general way, which feeding phenomena are critical to population processes. For example, population models show that a predator's functional response can have an important effect on the stability of predator-prey interactions. Holling's (1959b) demonstration that predators can become saturated at high prey density worried population modelers, because it meant that such predators were not very effective at "stabilizing" prey populations, at least at high prey densities. This destabilization can be partially overcome, even for functional responses that saturate, if the functional response is positively accelerated at lower prey densities (a sigmoid or type III functional response; see May 1976). Hence, an obvious way to apply foraging theory is to try to explain this positively accelerated piece of the functional response.

A naive interpretation of the prey-choice model might lead one to think that there should not be a functional response, since this model predicts that a prey *type's* inclusion in the diet is independent of its own abundance (see Table 1). However, as I pointed out above, the prey *species* of a population model may not correspond to the prey *type* of a foraging model. Krebs et al. (1983) have discussed one implication of this for functional response. More commonly, the prey model's rather strict assumption of fine-grained search and encounter will be violated (e.g., prey may be clumped), so that a functional response may result because clumps of prey become more valuable as prey become more abundant.

Getty (1985) has presented a particularly fascinating study of functional response by considering what a "prey type" really is, and what would happen if the degree of resemblance between two

kinds of prey varied. Suppose that two prey species resemble each other perfectly, as they might in a model-mimic system. One might find a functional response for the high quality (mimic) species even if the restrictive assumptions of the prey model hold, because an increase in the abundance of the high quality prey species would increase the average value of the prey type (the set of indistinguishable models and mimics). Getty uses models from signal detection theory to take the analysis further, which allows him to consider intermediate degrees of resemblance between the variants of prey type. He finds that relatively close resemblance can bring the stabilizing effects of a sigmoid (type III) functional response into play; however, when “good” and “bad” types do not resemble each other closely, destabilizing negatively accelerated (type II) functional responses may result.

This approach of using explanatory arguments from foraging theory to guide and refine the descriptive assumptions of population biology is one in which population biology and foraging theory still go separate ways, although Maurer (this volume) has advanced an integrated approach.

#### *The integrated approach: a case study*

Can a case be made for a more closely integrated study of foraging and population processes? Perhaps the most focused attempt to accomplish this is the work of Earl Werner and his colleagues, who used foraging theory to study how centrarchid fishes partition the resources in small, freshwater lakes in Michigan. Werner and Hall (1974) performed a careful series of laboratory experiments to test and refine the prey model specifically for each species. After satisfying themselves that these models were reasonable, they argued that two species of centrarchids—Bluegill Sunfish (*Lepomis macrochirus*) and Largemouth Bass (*Micropterus salmoides*)—partition food resources by prey size, because the morphology of bluegills means that their most profitable prey are plankton, while the morphology of bass means that their most profitable prey are small fishes. It would pay neither bass nor bluegill to specialize on the other's food supply (Werner 1977).

This seems to hold up for large bluegills and bass, where the interaction is primarily competitive. However, some of the small fish that bass find so profitable are small bluegills. This means that these two species have a mixed relationship; one size class of bluegills competes with bass, but another is a bass prey item. Werner and his colleagues began to think of bluegills not as pure competitors or pure prey items, as simple two-species-interaction models would, but as a size-structured population that may be prey or com-

petitors. They discovered that small bluegills feed in the weeds but large bluegills feed in the open water. This might mean: (1) the weeds may provide cover; or (2) the weeds might provide better small bluegill feeding (i.e., small bluegills and large bluegills might partition resources); or (3) some combination. Using foraging theory they were able to calculate that all size classes of bluegills do best when they feed in the open water. So the predator avoidance explanation of habitat choice by small bluegills seemed satisfactory. Werner et al. (1983a) nailed down this explanation in their celebrated split-pond experiment. A man-made circular pond was divided in half by a screen, each was given the same size distribution of bluegills but one side also had eight largemouth bass: the small bluegills on the predator-free side of the pond ate more or less the same open-water diet as their larger conspecifics, but the small bluegills on the predator side of the pond took prey primarily from the weeds.

Werner et al. (1983b) found that in ponds without predators they could predict the bluegills' seasonal shifts in habitat use simply on the basis of where the highest rate of food acquisition could be found, but they needed a model that incorporated predation and feeding to predict the size-dependent habitat shifts that they observed in ponds with predators. Gilliam (1982) and Werner and Gilliam (1984) were able to build such a model, that also predicts and explains situations in which a mixed relationship exists between interacting species (e.g., predator-prey at one ontogenetic stage, but competitive at another).

#### *Prognosis: continued slow but fruitful interaction.*

The success of Werner and his colleagues makes me wonder whether my cautious approach to the interaction between foraging theory and population ecology is warranted. Their work shows that ecologists can fruitfully use foraging theory as a tool to understand predator-prey and competitive interactions. However, ecologists must have the patience to build the necessary behavioral tools that Werner's group has demonstrated.

The extent to which foraging theory and population ecology can be profitably combined depends on how one answers the question, “How accurate must a theory of feeding be to be a useful building block of models of higher level processes?” If a precise theory is required, then the “sideways” questions I have mentioned might be the most helpful “up” questions. If only vague generalities about feeding are required, then an adequate theory probably exists. However, if, as seems most likely, population modelers must understand some aspects of feeding behavior pre-

cisely, but can get by with a vague understanding of others, then the future of this interaction will depend on which aspects of feeding behavior are deemed to be critical.

### EMPIRICAL DIRECTIONS

In my opinion, the most influential studies are those that force modelers to change their assumptions at the most fundamental level, such as Caraco et al.'s (1980) demonstration that the preferences of juncos are sensitive to variance, and that this preference can change direction with the junco's state of hunger. This study challenges fundamental assumptions, because it casts doubt on the traditional assumption of *average* rate maximizing. Moreover, it provides an interesting new observation, the switch in risk sensitivity, for which alternative models would have to account. Caraco et al. were fortunate that juncos are sensitive to variance, since they might have confirmed the traditional assumption instead of overthrowing it, and that people were beginning to think about variance sensitivity, so that their work got immediate attention. But, two features of this study cannot be ascribed to luck. Caraco et al. tested explicit and meaningful alternatives, and they chose these alternatives and the conditions of the experiment with an understanding of their larger theoretical significance. Unfortunately, many tests of foraging models do not meet these two criteria.

#### *Understanding the theory*

Many would-be testers do not understand the theory well enough to perform meaningful tests. For example, the patch model has been tested in cases in which there is no patch depletion (see Dunning, this volume), and the prey-choice model has actually been applied to situations in which the two prey types of interest were found in different parts of the forager's habitat. No one is interested in the "discovery" that a model does not work when it would not be expected to. Understanding the theory pays empirical dividends in another more important way, because it focuses attention on important problems. For example, an attack on an assumption made in 50 foraging models will be more influential than an attack on an assumption made in one model. Caraco et al.'s (1980) demonstration of risk sen-

sitivity challenged an assumption made by practically all models before 1980.

#### *Testing meaningful alternatives*

Many earlier tests of foraging models were essentially confirmatory tests. Because they tested only one model, if they compared it to any alternative it was the alternative of some random model. While this approach may once have been justified, that time is past. Alternative models are available, and the failure to use them makes tests of foraging theory difficult to evaluate. The ideal alternatives are contrasting views already present in the literature. See Kamil (1988) for a full discussion of this and other empirical issues in ornithology.

### CONCLUSIONS

I think that foraging theory has helped to refine our ideas about feeding behavior, and that its future will be more interesting than its past. It faces some fundamental challenges about how animals evaluate alternatives that vary in both time and amount, about the mechanisms that govern feeding behavior, and about what the answers to such questions will tell us about ecological processes. Each of these directions represents exciting empirical and theoretical opportunities.

Foraging theory is nothing if not controversial. The critics have helped theorists to see the limitations of their approaches, producing a more productive and cautious discipline. But they have failed to convince some that foraging theory is "a complete waste of time" (Pierce and Ollason 1987). In my view foraging theory may be limping along, but it is moving, and its critics have failed to make clear what alternative research programs would provide a more productive approach. Until critics meet this challenge, foraging theorists can take comfort from a north African proverb: "The dogs may bark but the camel train goes on" (quoted by Murray 1981).

### ACKNOWLEDGMENTS

I thank Al Kamil, Ted Sargent, and Sonja Yoerg for their comments. A number of the ideas expressed here were formulated and refined in conversations with Tom Getty, whom I thank for patience and sound advice.