Optimality Analyses of Operant Behavior and Their Relation to Optimal Foraging

J. E. R. Staddon

The ancient idea of hedonism, the pursuit of pleasure, comes in two forms. One may be termed philosophical or moral hedonism: that the aim of life ought to be the maximization of pleasure. Jeremy Bentham and the Utilitarians, although far from Epicurean in their personal habits, nevertheless endorsed a hedonistic ideal in their famous maxim of the "greatest good for the greatest number." In economics the Utilitarian ideal takes the form of Pareto optimality, that is, a state of the world in which no pair of individuals can become better off by exchanging something with each other.

Scientific hedonism is the idea that all (or a definable part) of behavior is guided by the pursuit of pleasure. Obviously scientific hedonism, if true, supersedes moral: if we are, in fact, guided by the pursuit of pleasure, it makes no difference whether we ought to be or not—-one does not debate the moral virtues of natural law.

The terms pleasure, good, and better off are undefined, as perhaps they should be. They are surplus terms that add vividness to what might otherwise be a dry topic, but their present scientific value is doubtful. The essence of scientific hedonism is not pleasure, a subjective quality, but consequences, an objective one (cf. Rachlin, this volume). Scientific hedonism is just the idea that behavior is guided by the outcomes it produces. This is of course the definition of operant behavior, so that reinforcement just becomes a label for those outcomes by which operant behavior is guided.

1 This work was supported by grants from the National Science Foundation to Duke University.
Much fruitless discussion has centered on the objective measurement of reinforcement. Economists have finessed the problem by means of the idea of revealed preference (e.g., Samuelson, 1965), which refers not, as might be supposed, to heavenly intervention, but to the outcome of choice: the better (more reinforcing, pleasurable, happier, etc.) state is simply the one chosen over alternatives. "The good," in consequence, is not assumed to have any intrinsic qualities. This is G. E. Moore's view: "I have maintained that very many things are good and evil in themselves, and that neither class of things possesses any other property which is both common to all its members and peculiar to them" (1903, p. x). Moore, of course, believed that there is just one set of good and evil things. Nowadays we are content to let each individual have his own set.

This admirably behavioristic view of utility has an exact parallel in the biological concept of Darwinian fitness: the fitter organism (genotype, gene) is simply the one that leaves more descendants. Again, no intrinsic properties are ascribed to fitness. The individual may be morally, physically, or aesthetically repugnant, but if it nevertheless leaves more descendants than others, it is by definition fitter.

There is, apparently, a degree of incompatibility between the demands of scientific objectivity, on the one hand, and the requirement that a theory predict, on the other. Objectivity requires that reinforcement be measurable. The easiest way to ensure this is simply to deline a reinforcer in terms of the organism's behavior—which leads at once to revealed preference and the empirical law of reinforcement (i.e., "a reinforcer is something that reinforces"). Predictiveness requires just the opposite: that a goal be specified in advance of the behavior it is to explain. This minor paradox—the apparent circularity of principles such as the empirical law of reinforcement—that neither class of things possesses any other property which is both common to all its members and peculiar to them" (1903, p. x). Moore, of course, believed that there is just one set of good and evil things. Nowadays we are content to let each individual have his own set.

This admirably behavioristic view of utility has an exact parallel in the biological concept of Darwinian fitness: the fitter organism (genotype, gene) is simply the one that leaves more descendants. Again, no intrinsic properties are ascribed to fitness. The individual may be morally, physically, or aesthetically repugnant, but if it nevertheless leaves more descendants than others, it is by definition fitter.

There is, apparently, a degree of incompatibility between the demands of scientific objectivity, on the one hand, and the requirement that a theory predict, on the other. Objectivity requires that reinforcement be measurable. The easiest way to ensure this is simply to deline a reinforcer in terms of the organism's behavior—which leads at once to revealed preference and the empirical law of reinforcement (i.e., "a reinforcer is something that reinforces"). Predictiveness requires just the opposite: that a goal be specified in advance of the behavior it is to explain. This minor paradox—the apparent circularity of principles such as the empirical law of reinforcement—that neither class of things possesses any other property which is both common to all its members and peculiar to them" (1903, p. x). Moore, of course, believed that there is just one set of good and evil things. Nowadays we are content to let each individual have his own set.

The economist's idea of a preference structure is more comprehensive than the parallel psychological concepts of motive, reward, or reinforcement because it allows for the resolution of conflicting goals. If a man eats Chinese food for lunch one day and American the next, it will not do to attribute this change of taste to a change in the relative reinforcing power of the different cuisines since there is no way in which such an account can be falsified—it would serve equally well to account for a sudden craving for dog food. But if Chinese and American foods are imperfect substitutes, then his change in behavior may very well be consistent with an invariant preference structure
that also precludes an attachment for dog food under any but the most dire circumstances. Without such a structure it is all too easy to account for apparent failure of one reinforcer by invoking another.

I will not attempt comprehensive description of optimality theory in the abstract here. Excellent accounts appear in several standard texts (e.g., Baumol, 1977), and good discussions of the strengths and weaknesses of this approach are also available (Lewontin, 1979; Maynard Smith, 1978; Oster & Wilson, 1978; Rapport & Turner, 1977). It is sufficient to note that in the theory of operant behavior the negative features of the approach have appeared decisive until recently. Either from a suspicion of formal analysis (e.g., Skinner, 1966) or from philosophical attachments to mechanistic (causal) theory, most "behavioral" psychologists have eschewed optimality analysis as teleological—despite the fact that variational methods have been a useful tool for theoretical physics since the time of Lagrange, and, with their imperfections, remain an essential part of evolutionary biology.

The aim of this chapter is to provide a selective survey of optimality analysis of steady-state operant behavior—to raise, if not to answer, the question: to what extent can operant behavior be explained in optimality terms? My objective is not to argue that optimality analysis is the only, or even the best, approach to these problems. Rather it is an attempt to redress the historical imbalance against it. In relation to the complexity of the problem, relatively little is known about the normative theory of operant conditioning procedures. In consequence, it is not possible to be sure about how useful this approach will prove to be. The usefulness of optimality methods in other areas gives grounds for modest optimism.

The methods of operant conditioning provide a laboratory model for both behavioral ecology and economics. The essential feature of operant conditioning is that access to a limited resource such as food (or the absence of electric shock) is made dependent on a specified behavior. The rule by which behavior (the operant response) and resource (reinforcement) are related is the schedule of reinforcement. For example, on a fixed-ratio schedule, a fixed number of bar presses (say) is necessary to procure each food pellet. This general arrangement is illustrated in Figure 4.1, which shows two functions: a feedback or schedule function, that describes the relation between behavior, \( x(t) \), and its effect on the environment, \( R(t) \), and a control function, \( x(R) \), that describes the relation between the environmental effect and the behavior made. Under laboratory conditions, the feedback function, \( R(x) \), is usually known; for the fixed-ratio schedule, for example, \( R(x) = x/m \), where \( m \) is the number of bar presses (responses) required for each pellet, \( x \) is the rate of responding, and \( R(x) \) the rate of pellet delivery (reinforcement). The objective of experiment is to learn something about the control function, \( x(R) \), that governs the behavior of the animal.
The virtue of this framework for both economics and behavioral ecology is that it makes visible the feedbacks inherent in any adaptive system. Under field conditions, little is usually known about the feedback function which is, in consequence, sometimes ignored. For example, in studies of foraging, the frequency with which a predator encounters different color or form variants (morphs) of the same prey species appears to affect its preference, commoner morphs often being taken disproportionately more frequently than rare ones (this is termed switching by ecologists, cf. Murdoch, 1969; Murdoch & Oaten, 1975). Obviously, interpretation of this finding depends critically on the search pattern adopted by the predator: if it looks in different places for different morphs (perhaps because different types have slightly different habitats), the relative encounter rates will be different than if it searches at random. In the former case, frequency-dependent predation may simply be a by-product of an optimal search strategy (look for the commoner type; Cornell, 1976; Manly, Miller, & Cook, 1972; Royama, 1970), whereas in the latter it may reflect constraints on the perceptual process—the predator’s inability to “look for” more than one prey type at a time (i.e., a search image: Tinbergen, 1960). To decide between these two interpretations it is essential to know both how the animal’s search pattern affects its exposure to different prey types (the feedback function), and how the spatial distribution of encountered prey types affects the animal’s search pattern (the control function).

Other chapters in this book trace the parallels between work on operant behavior and economic concepts. Therefore the emphasis in this chapter will be on the relations between operant conditioning and behavioral ecology.
No laboratory reinforcement schedule is an exact model of a natural situation, nor would such an identity be generally useful. The role of laboratory studies is to provide tests of behavioral mechanisms, based on hypotheses that may (although they need not) derive from the study of field data. Speculations about optimality criteria, such as energy intake per unit time, about temporal integration and forgetting, as in some interpretations of predator functional response (e.g., Holling, 1965; Hughes, 1979), and about constraints on perception or performance, are particularly amenable to test in this way.

**Optimal Choice and Allocation of Behavior to Patches**

The central evolutionary assumption is that organisms are adapted so as to maximize inclusive fitness (Maynard Smith, 1978). In studies of optimal foraging, the beginning simplification is that they do so by maximizing their rate of energy intake, expressed as energy acquired per unit time (cf. Krebs, 1978 for reviews; Pyke, Pulliam, & Charnov, 1977; Schoener, 1971). This assumption must, of course, be qualified when other requirements, such as safety or care of offspring, or differences in nutrient quality (imperfect substitutability—see below), become important. Limitations on the animal's knowledge of its environment, or on its performance or cognitive abilities must also be taken into account.

There are two broad approaches to the foraging problem. The most ambitious is the overall optimizing approach, that sets as its goal the derivation of a globally optimal strategy for all behavior, possibly even including details of behavior sequences (e.g., McFarland, 1977; Sibly & McFarland, 1976). The guiding assumption here is that natural selection has moulded animals that their behavior, at least in their "selection environment," conforms even in minute particulars to the requirements of overall maximization. At the other extreme are various piecemeal approaches that look at simplified aspects of the foraging problem. For example, foraging in patchy environments is accounted for by assuming that the animal will forage in a patch until its rate of intake drops below the level for the habitat as a whole (the so-called marginal-value theorem: Charnov, 1976). This idea can be applied by the ethologist as a formal analytic technique, or by the animal, in the form of an optimal "giving-up" time, as a decision rule or heuristic. If the feedback function for patch depletion is well behaved, and if the giving-up time chosen is appropriate, this rule suffices to optimize energy intake. In any case, equalization of giving-up time for all patches is a generally useful strategy that is an instance of satisficing (Simon, 1957), since it provides an approximation to optimal behavior that demands little computation by the
animal. The global approach subsumes these more limited, molecular analyses and, given the appropriate constraints, reduces to them. Emphasis on one or the other approach therefore boils down to one's intuition about scientific tactics: is the solution to the general problem best approached through comprehensive modeling of natural and quasi-natural situations, or through simplified models of constrained and perhaps artificial, but simpler, situations? The emphasis in this chapter will be on the analysis of behavior in relatively simple situations.

What may be termed the molar (global) and molecular (local) approaches have been developed to deal with the foraging behavior of numerous animal species in a variety of natural environments. Yet essentially the same dichotomy divides students of a much simpler system: the food-getting behavior of the white rat, and its avian counterpart, the White Carneau pigeon (Columba livia, var. Skinneriensis), in what must by now be its natural habitat, the Skinner box. The relevant issues are more salient on this smaller stage, and in most cases the experimental facts are of exemplary clarity. It may therefore be of interest to look at what is known of the determinants of behavior in these limited situations.

SIMULTANEOUS CHOICE: CONCURRENT SCHEDULES

A much-studied situation involves a hungry pigeon, trained to peck a response key for food and confronted with two choices. Pecks on each key occasionally yield brief access to food according to independent schedules of reinforcement. Although I later note some important differences, from an ecological point of view, this situation can be compared to two patches, varying in prey density (both schedules of the same type) or in terms of the function relating foraging effort to return (schedules of different types: there is no precise natural equivalent for this, although I discuss some possibilities below). The analogy to foraging in a patchy environment is further emphasized by the frequent use of a so-called changeover delay, which penalizes changes from one choice to the other (switches) by setting a time following a switch when responding will be ineffective. Changeover delay is thus loosely equivalent to travelling time between patches.

Two kinds of schedules have been studied most extensively, ratio and interval. Ratio schedules correspond approximately to variation in prey density: a variable-ratio m (VR m) schedule means that m key pecks (foraging movements) are required, on the average, for each brief access to food. Variable-interval schedules arrange for access to food to be jointly determined by time and by pecking: a variable-interval t-sec (VI t) schedule means that the first response after t sec since the preceding food delivery, on the average, will procure food. In principle, the distribution of intervals
or ratio runs can be determined by any distribution we choose. In practice, either fixed\(^2\) (ratios, times) or random distributions are most commonly used. In the random case, the probability that the food will become available for the next response, or during a brief increment of time, is constant. The implications of this are amplified below.

Concurrent schedules are obviously sufficiently similar to frequently encountered natural situations to be of more than passing interest to behavioral ecologists.\(^3\) Moreover, if behavior in these simple situations can be reduced to a general decision rule, foraging theories that assume such a rule will be on firmer ground.

Concurrent schedules have been looked at from three points of view: descriptive theory, molar (overall) maximizing, and local maximizing. Both local and molar maximizing are most readily illustrated by performance on concurrent ratio schedules. Descriptive theory, in the form of various matching relations, has been predominantly associated with studies of interval schedules. Common principles unify all these cases, however. It will be simplest to begin with molecular analysis of ratio schedules.

### Molecular Analysis\(^4\)

Obviously, if an animal is confronted with two choices, each paying off with constant probability (concurrent VR–VR schedule—see Herrnstein and Vaughan, this volume, for procedural details), then after an initial sampling period, nothing is gained by responding to the low-probability alternative. If the outcome for both choices is identical access to the same food source (perfect substitutability), then both local and molar maximizing predict eventual exclusive choice of the majority alternative, and this is the usual result with both pigeons and titmice (cf. Herrnstein & Loveland, 1975; Herrnstein & Vaughan, this volume; Krebs, Kacelnik, & Taylor, 1978).

This result strongly hints that the decision rule for choice in these experiments involves some form of maximizing. However, the concurrent VR–VR procedure provides too few degrees of freedom for fruitful analysis of the details. More useful are procedures that produce nonexclusive choice. All such procedures act by introducing a time-dependence into the feedback

---

\(^1\) In this case, they are, of course, termed fixed ratio or interval schedules.

\(^2\) It will not do to dismiss the Skinner box as an artificial contrivance, irrelevant to "natural" behavior. This apparatus underwent an evolution of its own, and there is much evidence that it constitutes something like an optimal environment for the production of pecking or lever pressing (cf. Atkin, 1977; Devenport, 1979; Staddon & Simmelhag, 1971). Its major limitation is that it provides no opportunity for animals to move from place to place in search of food, but this limitation is essential if the feedback function is to be fully under the control of the experimenter.

\(^3\) This section draws heavily on an unpublished paper by Staddon, Hinton, & Kram.
function for at least one alternative. Thus, in the concurrent VR–VR procedure, the feedback function for each alternative is constant, of the form

$$R(x_i) = x_i/m_i,$$

(1)

where $m_i$ is the average ratio value and $x_i$ the average response rate to the $i$th choice (see Glossary for these and subsequent symbols). If the distribution of response ratios is exponential, any response to an alternative has the same chance of procuring food. This time-independence can be abolished in two simple ways: either payoff probability increases with time (or successive responses), or it decreases. In practice, the decreasing case is the one that has been most studied, since it readily leads to stable, diversified behavior (i.e., nonexclusive choice). Real (this volume) has analyzed the reasons for this in connection with the formally similar problem of fitness maximization. In the context of choice, the basic ideas can most easily be introduced via a historical digression.

Students of comparative psychology will recognize the similarity between the concurrent VR–VR procedure and so-called probability learning experiments, in which animals are confronted on a trial-by-trial basis with two alternatives. On each trial, reward occurs with probability $p$ at one alternative and with complementary probability $1 - p$ at the other. A spirited debate continued for some years about supposed species differences in such procedures: rats and "higher" animals were supposed to maximize, that is, fixate on the majority alternative, whereas fish and, by some accounts, pigeons also, were said to probability match, allocating $p$ of their choices to one alternative and $1 - p$ to the other (cf. Bitterman, 1965; Brunswik, 1939; see Mackintosh, 1974, for a review). Yet, as we have already seen, pigeons settle for the majority alternative in the closely analogous concurrent VR–VR procedure. To what is this discrepancy attributable?

In the absence of additional experiments, no certain answer can be given. But the most likely culprit is the correction procedure frequently employed in these studies to ensure sampling of both alternatives. In its simplest form, this involves leaving an assigned reward until it is picked up, so that persistent, unreinforced choice of an alternative signals the probable availability of food for the other choice. It is easy to show that while the optimal strategy for the simple procedure is exclusive choice, optimal responding in the probability-learning-with-correction (PLC) procedure entails a probability of $1/(2 - p)$ of choosing the majority alternative, i.e., nonexclusive choice. Consequently, if either the animal or the experimenter fails to treat correction trials separately, nonexclusive choice is to be expected.\footnote{\textsuperscript{5} It may be that differences in steady-state PLC performance among species, to the extent that they exist, reflect differences in animals' ability to discriminate "correction" from "regular" trials, rather than any differences in the incremental effects of reward on behavior.}
The PLC procedure shows in extreme form the essential feature of procedures that produce nonexclusive choice: that the relative attractiveness of an alternative diminishes with continued sampling. In the PLC case probability of payoff for a choice declines to zero following one unsuccessful response, and at the same time increases to one for the other choice; hence a "lose-shift" strategy is optimal. Free-operant procedures such as concurrent variable-interval–variable-interval (VI–VI) and variable-interval–variable-ratio (VI–VR) share these negative feedback characteristics in a more subtle form.

The distinctive thing about free-operant procedures is that a response can occur at any time—time takes the place of trials. On random VI schedules, the probability that reinforcement will "set up" (be available for the next response) during any brief interval of time is constant. This means, of course, that as time goes by without a response, the probability that reinforcement has set up increases. Although not designed to model any natural situation, the feedback function for VI resembles many natural processes of replenishment following depletion: amount of nectar available since the last nectar-feeder visit (the similarity to temporal schedules is further emphasized by the circadian periodicities in nectar secretion of many flowers), grass or foliage recovery since the last grazing visit, and so on.

For the random case, where the probability of setup is constant over time, the probability of setup, \( P(R|t_i) \), for the \( i \)th choice is given by

\[
P(R|t_i) = 1 - \exp(-\lambda_i t_i), \tag{2}
\]

where \( t_i \) is the time since the last \( i \) choice, \( \lambda_i \) is the average programmed rate of setup for that choice, and \( \exp \) is the exponential function (cf. Feller, 1957; Staddon, 1977). Equation (2) completely defines the contingencies for concurrent VI schedules. The predictions about choice to be derived from it depends upon two things: the constraints—of memory, rate of response, ability to time, and so on—assumed to act on the animal, and the critical quantity that is believed to determine operant choice. I defer the problem of constraints until molar optimizing is taken up later.

Two quantities have been considered as moment-by-moment guides to choice: probability of reinforcement and rate of reinforcement. The two are obviously closely related, since if the probability-density function describing a stochastic process is known, the rate of occurrence of that event is also determined. For example, if an event, \( x \), occurs according to a stationary, random process (the most frequent case considered in this chapter), then the probability density function is \( f(t) = \lambda \exp(-\lambda t) \), where \( \lambda \) is equal to the average rate of occurrence of the event. The probability such an event has occurred after a time \( t \) is given by Eq. (2), whence by taking logarithms and rearranging \( \lambda = (\ln(1/P(x))/t) \). The relation is even simpler if \( t \) is small
so that terms in $L^2 t^2$ and above can be neglected, when $P(t) = \lambda t$, so that probability is directly proportional to rate. Obviously, if $P(x_i | x_i)$ are equated for all $x_i$, the expected values of rates, $\lambda_i$, are also equated. If the event of interest is not governed by a stationary process, or if rate, as estimated by the animal, involves time-dependent averaging, this simple relation between rate and probability will not hold, although there will still be a definite dependence of one on the other.

I consider first reinforcement probability as a guide to choice, leaving until second the more complicated problem of reinforcement rate.

If choice is determined by immediate probability of reinforcement (momentary maximizing; Shimp, 1966, 1969), then whenever the animal makes a choice, it should always be of the alternative with the highest probability of payoff. Shimp has provided a momentary-maximizing analysis of discrete-trial choice and a simplified analysis of concurrent VI–VI schedules. I now present a somewhat more general analysis of both concurrent VI–VI and VI–VR schedules. The momentary-maximizing condition is illustrated in Fig. 4.2, for a two-choice concurrent VI–VI situation, where responses to the two alternatives are denoted by $i$ and $j$. The top panel shows a sequence of two $i$ and two $j$ choices. $t_i$ and $t_j$ are the times elapsed since the last $i$ or $j$ choice. At the time of the second $i$ choice, $t_i = a$ and $t_j = b$; at the time of the second $j$ choice, $t_i = a'$ and $t_j = b'$. These values of $t_i$ and $t_j$ are plotted in the bottom panel of Figure 4.2. Thus, the coordinates of point $i$ (the second $i$ response) are $(a, b)$, and of point $j$, (the second $j$ response) are $(a', b')$. The ray from the origin that divides this clock space into two regions is derived.
from Eq. (2) as follows. The limiting case for choice is when \( P(R|t_i) = P(R|t_j) \) which, from Eq. (2), after rearrangement yields the relation

\[
t_j = \frac{\lambda_i t_i}{\lambda_j},
\]

(3)

where \( \lambda_i \) and \( \lambda_j \) are, as before, the average rates of programmed VI reinforcement for the two alternatives. Choice conforms to momentary maximizing if all \( j \) responses lie above and to the left of the switching line defined by Eq. (3) (i.e., \( P(R|t_j) > P(R|t_i) \)), and all \( i \) responses lie below and to the right of the line. The choices in Figure 4.2 are optimal in this sense.

The switching-line momentary-maximizing analysis can readily be applied to concurrent VI-VR schedules. In this case, the probability of reinforcement on the ratio alternative is constant and equal to the reciprocal of the ratio value, \( m \). The switching line can be derived from Eqs. (1) and (2), thus

\[
\frac{1}{m} = 1 - \exp(-\lambda_j t_j),
\]

(4)

where \( i \) is the ratio and \( j \) the interval alternative. By conventional manipulations, this yields the relation

\[
t_j = \frac{\ln[m/(m - 1)]}{\lambda_j}
\]

(5)

as the equation for the switching line. Since \( \lambda_j \) and \( m \) are constants, this describes a straight line parallel to the \( t_i \) axis.

The switching line in the clock space of Figure 4.2 constitutes a decision rule that tells the animal which alternative to respond to at any time, but not when to respond. Nevertheless, some general predictions can be derived from it. For example, in the concurrent VI-VR procedure, if a choice is made shortly after an interval response has occurred, it should always be to the ratio alternative. This follows from the property of interval schedules that the probability of reinforcement for a response decreases to zero after each response. Hence a further interval response should not be made until the probability of reinforcement [described by Eq. (2)] rises above the fixed value for responding to the ratio alternative. Given almost any aperiodic distribution of choices, momentary maximizing implies that ratio inter-response times will be smaller than interval ones, a universal empirical result. The choice patterns implied by some aperiodic and periodic choice distributions are derived below.

Reinforcement rate is the variable most often implicated as the major determinant of the "strength" of an operant response, hence of choice. In the form of rate of energy intake per unit time, it is also the major independent variable used in optimal foraging theory (e.g., Charnov, 1976; Krebs, 1978; MacArthur & Pianka, 1966; Parker & Stuart, 1976; Schoener, 1971). The
operation of delivering a reinforcer, such as a food pellet or a sip of water, like prey capture is a discrete event. Consequently, rate of reinforcement must be computed by counting instances and dividing the sum by some specified time interval. It is rarely obvious what this interval should be—and nor is the problem a trivial one that can be settled by convention. An animal such as a shrew that must eat every few minutes or starve to death obviously has little interest in average food intake over a period of even a few hours, much less the period of a week or two that would suit a large python. There is also evidence that rates of eating and drinking are variables that many animals regulate over periods considerably shorter than those forced by physiological limitations (cf. Staddon, 1979b), although the two periods are probably not unrelated. For all these reasons, reinforcement rate is less tractable as a molecular (as opposed to molar) theoretical variable than reinforcement probability.

Momentary maximizing is primarily a normative rather than a process theory, since it specifies what the animal must do to maximize probability of reinforcement, but says nothing about how it estimates the necessary time-dependent probabilities. Nevertheless, there is ample evidence that animals are sensitive to temporal contingencies of this sort. For example, pigeons will space their responses in time if only interresponse times greater than some fixed value lead to food (e.g., Staddon, 1965). They are also sensitive to the more subtle proportional relation between interresponse time and payoff probability implicit in interval schedules (Platt, 1979).

However, there is also evidence that animals are sensitive to more than just the immediate probability of payoff, although their adaptation is often imperfect. For example, pigeons respond quite adequately in choice procedures with a changeover delay, even though each changeover response has zero probability of reinforcement. But they adapt poorly to chained fixed-interval schedules, that is, schedules in which responses in the presence of the mth post-food stimulus, $S_m$, are reinforced by the appearance of stimulus $S_{m+1}$, with food reinforcement occurring only at the end of stimulus $S_k$. When $k$ is greater than 4 or 5, responding in $S_1$ is subject to long delays, so that overall rate of access to food is much reduced below its maximum value—and the value attained in the absence of distinguishing stimuli for each fixed-interval “link” (cf. Gollub, 1977). Evidently pigeons have some sensitivity to overall reinforcement rate, but their ability to adapt is limited by a short time-horizon and stimulus-specific effects. Competition between interim and terminal activities serves as a mechanism of intertemporal utility maximization in successive-comparison procedures of this sort in ways that are still not fully understood (cf. Hinson & Staddon, 1978; Staddon & Simmelhag, 1971). To go into these questions would take us
beyond the confines of this chapter. Consequently, I shall not further explore expected reinforcement rate as an optimization strategy.

There is, however, a simple alternative that is very similar to the marginal-value rule for patch sampling. It requires only that the animal learn the overall (molar) average reinforcement rate associated with a given situation and keep track of local reinforcement rate history. Recall that in optimal foraging theories (e.g., Charnov, 1976; Krebs, Ryan, & Charnov, 1974), the time between reinforcements (prey captures) is often taken as the relevant variable that determines choice. When this variable exceeds a critical "giving-up time," the animal switches to another patch. Giving-up time, and interreinforcement interval (the inverse of rate), are easy measures to apply in most natural situations, where several prey items are available in most patches, and travel time between patches is substantial. Animals will not usually leave a patch until at least one prey item has been obtained, so that it is possible to obtain a giving-up time for each patch and every switch. On concurrent VI–VI schedules, however, there are only two patches and each is sampled repeatedly between infrequent food deliveries. Giving-up time is therefore not a useful measure.

However, there is no reason why reinforcement rate for each alternative must be estimated interreinforcement interval by interreinforcement interval. It is at least as likely that the animal employs a time-dependent averaging process that provides a separate running average for each alternative. This is a process of local maximizing (termed melioration by Herrnstein and Vaughan, this volume) that allows the animal to track the alternative with the most recent history of payoff. It differs from overall maximization (discussed below) in that it is a process rather than a normative model; and like any process, will not invariably maximize overall payoff.

It may be instructive to consider how such a process might work. An initial sampling period will produce some reinforcements for both alternatives. For simplicity we may suppose that reinforcement rate is computed by the animal according to a weighting function in which more recent reinforcements count more heavily. Perhaps the simplest such averager is the following discrete-time difference equation:

\[ F_i(n) = a_0 F_i(n) + b_1 F_i(n - 1), \]

where \( F_i(n) \) is the value of reinforcement \((0 \text{ or } 1)\) during the \( n \)th brief time interval \( \Delta t \), \( F_i(k) \) is the averaged value of reinforcement for choice \( i \) on the \( k \)th time interval, and \( a_0 \) and \( b_1 \) are constants related to species and reinforcer (see Figure 4.3). Because Eq. (6) is linear, it is easy to show that if the probability of reinforcement \((i.e., \text{ of } 1 \text{ rather than } 0)\) during any \( \Delta t \) is fixed at \( p(F) \), then after a period of time sufficient to reach equilibrium, the
mean value of $F_i$ is given by

$$F_i = a_0 P(F)/(1 - b_i), \quad b_i < 1. \tag{7}$$

That is, the averaged rate is simply proportional to the actual average reinforcement rate, $P(F)$. Hence, the usual molar analyses (discussed below) which use actual reinforcement rates are unaffected. In addition to this useful property, Eq. (6) is also consistent with the exponential form typical of extinction and transitional-choice behavior (cf. Killeen, 1979; Myerson & Miezin, 1980).

Interpretation of Eq. (6) is relatively straightforward when only one alternative is being considered. When there is more than one, reinforcement and time intervals must be “tagged” with the name of the choice alternative to which they apply, so that $F_i(k)$ becomes a moving estimate of local reinforcement rate, equivalent to the slope of the total energy intake versus time-in-a-patch curves discussed by Charnov (1976). The discrete-time variable is then incremented only as long as the animal is attending to that alternative.

If the animal switches from alternative $i$ to alternative $j$ when the averaged reinforcement rate, $F_i$, for alternative $i$ falls below the (remembered) average for the situation as a whole, this process is formally identical to the giving-up-time rule prescribed by marginal-value theory. Although I can present no proof, it seems likely that it will approximate optimal behavior under many conditions, particularly symmetrical ones. Moreover, it makes specific predictions, most notably that time to a switch should be systematically related to the recent history of reinforcement for an alternative: As Figure 4.3 illustrates, successive reinforcements increase $F_i$ and should therefore delay switching to the other alternative. Changeover data of the necessary sort do not appear to be available, but there are confirming data on relative response rate. For example, Menlove (1975) reports a “positive recency effect” of reinforcement on subsequent responding: “the relative frequency of responding on a key was highest immediately following a reinforcement on that key [p. 318]” and similar, but less detailed, results have been reported.
by others. The model also makes the appropriate prediction about the effect of changeover delay on rate of switching, namely that changeover rate should be inversely related to changeover delay (cf. Hunter & Davison, 1978; Stubbs & Pliskoff, 1969).

Reinforcement tracking implies that the allocation of behavior between two alternatives is solely determined by the animal’s immediate past history of reinforcement. Hence in a test where no reinforcements can occur, choice should soon converge on an alternating pattern in which neither side is favored. This is contrary to fact: After adequate choice training, animals consistently respond more on the richer alternative, even before the first reinforcement of the experimental session. Clearly, discriminative factors, including both external stimuli and, perhaps, time (as implied by momentary maximizing), are involved in choice in ways that go beyond a simple tracking process.

Obviously much is left unstated in this account. Tracking represents a more ambitious level of theorizing than optimality analysis, since it attempts to get at the process involved. The factors affecting the constants in Eq. (6) need to be spelled out, and assumptions about specific responses (as well as switching) need to be added. To explore this avenue in detail would take us beyond the scope of this paper, but, despite its obvious incompleteness, it is a promising one that relates naturally to the dynamic models recently proposed to account for some of these effects (McDowell & Kessel, 1979; Myerson & Miezin, 1980; Staddon, 1977).

**Empirical Results: Matching**

Experimental results from operant choice experiments provide a convenient way of summarizing the allocation of behavior across a pair of alternatives. If, for an experimental session of fixed length, the ratio of total number of responses to each choice is compared to the corresponding ratio of reinforcements obtained, a simple positive monotonic relation is invariably found (e.g., Herrnstein, 1961; Staddon, 1968; Baum, 1974; Hunter & Davison, 1978). This relation is well fitted by a power function, so that

\[ y_i/y_j = \alpha (R(y_i)/R(y_j))^\beta, \]

where \( y_k \) are the response rates and \( R(y_k) \) the associated reinforcement rates actually obtained. The constant \( \alpha \) provides a measure of the bias in favor of one or other alternatives: \( \alpha \) is equal to the ratio of choices when the reinforcements obtained for them are equal, hence, a value of unity signifies unbiased responding. The exponent \( \beta \) provides a measure of the diversity of behavior: if \( \beta = 1 \), animals match response ratios to reinforcement ratios; this is the typical result on concurrent VI–VI schedules (de Villers, 1976; Herrnstein,
1961). If \( \beta < 1 \), behavior deviates from matching in the direction of simple alternation; this is known as undermatching. Conversely, if \( \beta > 1 \), behavior deviates from matching in the direction of exclusive choice; this is overmatching, and represents minimal diversity. It is clear that Eq. (8) is widely employed precisely because it is capable of summarizing the full range of choice patterns.

What do the molecular-maximizing models just described predict about the form of the matching relation?

The switching-line analysis requires that some assumption be made about the temporal distribution of choices before predictions can be derived. The diagram in Figure 4.2 can then be used to compute the optimal sequence of choices. Considering concurrent VI–VI first, if choices are made at fixed intervals (which was approximated in experiment 1 of Nevin’s, 1969, discrete-trial study), the sequence of optimal choices will converge on a closed path, indicating a fixed sequence. Figure 4.4 shows an example, the trajectory of optimal choices when \( \lambda_i = 3\lambda_j \). The initial response occurs by convention at \( t_i = t_j = 1 \), and subsequent responses occur at unit time intervals. The horizontal and vertical dotted lines indicate that the occurrence of an \( i \) or a \( j \) choice sets \( t_i \) or \( t_j \) to zero; the 45° dotted lines indicate that during the time between choices, \( t_i \) and \( t_j \) grow at the same rate. The line of slope 3:1 through the origin is the switching line that determines the momentary-maximizing

![Figure 4.4](image-url)
choice at any instant. This example illustrates an ambiguity characteristic of all cases where $\lambda_i/\lambda_j$ is equal to an integer. If $\lambda_i/\lambda_j = N$, then after $N - 1$ occurrences of the majority response, the value of $P(\mathcal{F}|t)$ is the same for both choices; hence there are two possible momentary-maximizing sequences, depending on how this ambiguity is resolved. For the example in Figure 4.4, either iii or iij is a momentary-maximizing sequence.

It is simpler to derive predictions for concurrent VI–VR, since the temporal distribution of choices affects only the VI alternative. For example, suppose that a VI response is made only when the probability of payoff for it equals the fixed VR payoff probability $1/m$. Then, from Eq. (4), $1/m = 1 - \exp(-\lambda t_i)$, where 1 is the VI alternative and 2 is the VR alternative. This can immediately be translated into a relation between molar response and reinforcement rates. From Eq. (1), $1/m = R(y_2)/y_2$, and from Eq. (2), $P(R|t_i) = 1 - \exp(-\lambda t_i)$; but if VI responding occurs at intervals of duration $t_i$, then the mean VI interreinforcement time is just $t_i/P(R|t_i) = 1/R(y_1)$. Since $t_1 = 1/y_1$, it follows that $1 - \exp(-\lambda t_i) = R(y_1)/y_1$. Substitution at once yields the unbiased matching relation

$$R(y_2)/y_2 = R(y_1)/y_1.$$ 

This is a perfectly general result: if animals contrive invariably to match local reinforcement probabilities, then it follows that they will also match overall response and reinforcement ratios. This is theoretically possible on concurrent VI–VR, because $P(R|t)$ is constant for the VR alternative and the VI interresponse time can be chosen to match it. It is not possible on concurrent VI–VI: the geometry of Figure 4.2 shows that, in general, an animal cannot respond to both alternatives and place every response on the switching line.

The assumption that animals on concurrent VI–VR switch to the VI the instant that $P(R|t_i) > 1/m$ is obviously unrealistic. It is more likely that there is a switching threshold, so that $P(R|t_i) - 1/m$ must exceed some value, $\Delta$, before a VI response occurs; that is, $\Delta + 1/m = P(R|t_i)$ is the switching condition. Assuming a sort of Fechner's law for probability, $\Delta$ should be proportional to $1/m$, so we can write $\alpha/m + 1/m = P(R|t_i)$, $0 < \alpha < 1$. Substituting and rearranging yields the relation $y_2/y_1 = (1 + \alpha)R(y_2)/R(y_1)$, which is biased matching favoring the ratio alternative—which is the usual empirical result (Herrnstein & Helyman, 1979).

If choices are made at irregular intervals, the trajectory in the $t_i - t_j$ clock plane is not closed and there is no unique maximizing sequence. The payoff probability for each choice must then be individually computed and the totals summed to estimate the effectiveness of various choice strategies and their conformity to matching. Since animals never make choices at fixed intervals on concurrent schedules, any realistic attempt to evaluate the
The momentary-maximizing approach must make some other assumption about the temporal distribution of choices.

We explored the implications of the switching-line analysis of concurrent VI–VI and VI–VR schedules with the aid of a computer simulation which assumed a random temporal distribution of choices. Thus, during each program cycle, the probability of a choice was $P_i$, which varied from .15 through .75. The probability of a reinforcement being “set up” for each alternative was $p_i$ and $p_j$, where these two values were constrained such that $p_i + p_j = .003$. A setup was held either for the single program cycle when it occurred (VR; alternative $j$), or until the next response (VI; alternative $i$). Seven values of $p_i$, from .0001 through .0018, were examined at each $P_i$ value. When a choice occurred, it was allocated to $i$ or $j$ according to the switching-line analysis [Eqs. (3) and (5)], where $t_i$ and $t_j$ (see Figure 4.2) were measured in program-cycle units. Results reflect 100,000 choices at each $p_i$ value.

For each value of $P_i$, the number of $i$ and $j$ choices, $y_i$ and $y_j$, and the number of $i$ and $j$ reinforcements obtained, $R(y_i)$ and $R(y_j)$, were counted and the values of the ratios $y_i/y_j$ and $R(y_i)/R(y_j)$ at each value of $p_i$ noted. In every case, these pairs of ratios were well fitted by Eq. (8), with values of $\sigma$ in every case greater than .97, and in most cases between .99 and 1.00. The values of $\alpha$ and $\beta$ so obtained are shown in Table 4.1. For the VI–VI case, the value of $\alpha$ is invariably close to unity (unbiased responding). The value of $\beta$ is less than one (undermatching); $\beta$ increases as the value of $p_i$ (i.e., the absolute rate of responding) increases. For the VI–VR case, $\alpha$ is always less than one (bias in favor of VR) and $\beta$ is again directly related to the absolute response rate. The major difference between the VI–VI and VI–VR simulations is in the bias parameter; the slopes (exponents) of the fitted functions vary over a similar range in both cases.

| Table 4.1 |
| Parameters of Simulated Matching Functions |

<table>
<thead>
<tr>
<th></th>
<th>VI–VI</th>
<th></th>
<th>VI–VR</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_i$</td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\alpha$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>.05</td>
<td>.97</td>
<td>.60</td>
<td>.73</td>
<td>.57</td>
</tr>
<tr>
<td>.10</td>
<td>.96</td>
<td>.61</td>
<td>.64</td>
<td>.69</td>
</tr>
<tr>
<td>.15</td>
<td>.98</td>
<td>.66</td>
<td>.53</td>
<td>.70</td>
</tr>
<tr>
<td>.25</td>
<td>.96</td>
<td>.71</td>
<td>.53</td>
<td>.76</td>
</tr>
<tr>
<td>.75</td>
<td>.92</td>
<td>.80</td>
<td>.58</td>
<td>.78</td>
</tr>
</tbody>
</table>
Summary

In summarizing a number of optimal foraging theories, Hassell (1978) writes, "The tendency of all these models is always towards the reduction of all patches to the same rate of prey capture [p. 58]." This result is, of course, simple matching of marginal time ratios to reinforcement ratios. It will lead to overall matching under certain not-too-restrictive conditions described in the Appendix. In patch models it comes about because a given patch yields less and less as time, and foraging, continues. In concurrent schedules, the same result is a consequence of the growing attractiveness of alternatives other than the one being sampled. Both cases are instances of negative feedback, granted that choice is continually guided by the alternative offering the highest payoff, that is, some form of local maximizing.

The details of the control process are still obscure. In terms of molecular maximizing, the decision rule represented by the switching-line analysis is consistent with matching results on concurrent VI–VI and VI–VR schedules. A related possibility is some form of local reinforcement rate maximization process in which running averages of local reinforcement rates for each alternative (equivalent to the giving-up times of marginal-value theory) are compared moment by moment (cf. Herrnstein & Vaughan, this volume). Additional experimental results on post-reinforcement behavior as well as local switching patterns are needed to evaluate these possibilities. Such results would also bear on dynamic models that have recently been developed from which both matching (in concurrent VI–VI and VI–VR) and exclusive choice (in concurrent VR–VR) can be derived (e.g., Myerson & Miez, 1980; Staddon, 1977).

There has obviously been a substantial convergence between operant choice theory and optimal foraging theory. Local reinforcement rate maximization appears to differ only trivially from an optimal decision rule in terms of giving-up time. However, momentary maximizing with its emphasis on mechanisms of temporal discrimination, has no direct counterpart in optimal foraging theory. On present evidence, both processes may be involved in operant choice.

Negative feedback at the molecular level corresponds to diminishing returns at the molar level: increments in time or effort expended on an alternative become relatively less and less attractive as the absolute amount of time or effort increases. Diminishing returns shows itself in two ways: as a negatively accelerated feedback function, and as imperfect substitutability of different reinforcers. I turn now to these two topics.

Molar Analysis

There has been a flurry of interest in this type of analysis recently, and many more published theoretical papers are available on molar than on
molecular optimizing. This section is therefore just an overview of some of the more important ideas and experimental results.

Molecular optimality theories such as local reinforcement rate maximization often have something to say about the processes by which maximization is achieved, about decision rules. Molar maximizing theories tend to be strictly normative, however, saying nothing about process. It is important to keep this distinction in mind when tempted to compare the two types directly.

Molar optimality analysis derives directly from the framework illustrated in the block diagram of Figure 4.1. The argument generally proceeds in four steps. First, the constraints on responding are identified. Two cases commonly considered are specification of the distribution of responses to each choice, and incorporation of a time-allocation constraint, which requires that an increase in the level of one response entails a decrease in some others. Second, the form of the feedback relation, \( R(x) \), between responding and reinforcement is defined. Third, the objective function, which identifies the quantity to be maximized or minimized as a function of choice responding, is defined. Fourth, the extremum of this function (subject to the specified constraints) is derived.

These steps are best conveyed by example. Perhaps the simplest is the analysis of concurrent VI-VI by Staddon and Motheral (1978). Independent, random responding to each alternative is assumed (cf. Staddon & Motheral, 1979). This means that the expected time between reinforcements for responding to a given alternative is \( E_i = 1/a_i + 1/y_i \), where \( 1/a_i \) is the scheduled minimum average interreinforcement interval, and \( y_i \) is the response rate to the \( i \)th choice. If we let the expected rate of reinforcement \( R(y_i) = 1/E_i \), then the feedback function is

\[
R(y_i) = a_i y_i / (a_i + y_i),
\]

where \( a_i \) is the scheduled reinforcement rate (= 1/\( T_i \)) for the \( i \)th choice. If the quantity to be maximized is total expected reinforcement rate, then for two alternatives the objective function is

\[
R_T = a_1 y_1 / (a_1 + y_1) + a_2 y_2 / (a_2 + y_2).
\]

The matching relation to be expected from maximizing Eq. (11) can be derived very simply by applying the marginality condition, that is, for any pair of choices, the maximum of \( R_T \) must satisfy the condition that

\[
\partial R(y_1)/\partial y_1 = \partial R(y_2)/\partial y_2.
\]

But \( \partial R(y_1)/\partial y_1 = a_1^2 / (a_1 + y_1)^2 \) which, from Eq. (10), equals \( [R(y_1)/y_1]^2 \). Substituting in Eq. (12) at once yields \( y_1/y_2 = R(y_1)/R(y_2) \), which is simple matching of response and reinforcement ratios. The relation between re-
spouse and reinforcement rates to each alternative, considered separately, can be derived once the form of the time-allocation constraint is known. Staddon and Motheral dealt with the simplest case, a linear constraint of the form $r_1 + cr_2 = 1$. When the slope constant, $c$, is equal to unity the result is the hyperbolic relation between response and reinforcement rates proposed for VI responding by Herrnstein (1970).

Rachlin, Kagel, and their associates (this volume; Rachlin, 1978; Rachlin & Burkhard, 1978; Rachlin, Green, Kagel, & Battalio, 1976; Rachlin, Kagel, & Battalio, in press; see also Becker, 1971, for a related approach) have pursued a complementary approach using an econometric framework. Again, an example provides the most direct illustration. Consider two mutually exclusive and exhaustive activities that occur for times $T_1$ and $T_2$ in an experimental session of unit length, so that $T_1 + T_2 = 1$. Furthermore, suppose that the amount of payoff for spending time at each alternative is directly related to the time spent, according to a negatively accelerated function (such as might be enforced by a VI schedule). A convenient form for such a feedback function is the power relation
\[ R(T_t) = a_i T_i^{\beta_i}, \quad 0 \leq \beta_i \leq 1, \]
where $R(T_t)$ is the total amount of reinforcement obtained in the session by spending a time $T_t$ at the alternative, $a_i$ is the maximum possible payoff that can be obtained (when $T_t = 1$), and $\beta_i$ is a parameter that describes the diminishing returns between time spent and amount of reinforcement obtained. If $\beta_i = 0$, the feedback function corresponds to intermittent response-independent reinforcement (a variable-time schedule); when $\beta_i = 1$, Eq. (13) is the feedback function for a ratio schedule. At some intermediate value (Rachlin, 1978, has suggested .2) the function provides a reasonable approximation to the properties of VI schedules.

As before, the matching relation can be found by equating partial derivatives of the feedback function for all alternatives:
\[ \frac{\partial R(T_t)}{\partial T_t} = a_i \beta_i T_i^{\beta_i - 1}, \]
which, from Eq. (13), is equal to
\[ R(T_t) \beta_i / T_t. \]
Hence, for any pair of alternatives
\[ R(T_1) / R(T_2) = T_1 \beta_1 / T_2 \beta_2. \]
For schedules of the same type, it is perhaps reasonable to assume that $\beta_1 = \beta_2$, whereupon Eq. (14) is again the matching relation.

This analysis assumes perfect substitutability between the reinforcements for responses $T_1$ and $T_2$, which is reasonable if they are in fact identical. In
Equation (17) can be rewritten in terms of obtained reinforcement rates: thus, from Eq. (13)

\[(K_1[R(T_1)]^{b_1})/T_1 = (K_2[R(T_2)]^{b_2})/T_2,\] 

where \(K_i\) represents the lumped constants of Eq. (17). Clearly optimal choice implies linear matching of obtained reinforcement ratios to response (time) ratios only if reinforcers are perfectly substitutable \((b_1 = b_2 = 1)\). Otherwise (so long as \(b_1 = b_2\)) power-ratio matching, described by Eq. (8), is to be expected. If \(b_1 < 1\), as expected from the diminishing-returns assumption, the result will be undermatching. However, the feedback function of Eq. (13) means that linear matching will be obtained on all schedule comparisons with perfectly substitutable reinforcers where at least one feedback function has \(b_i \neq 1\). If the exponents are unequal, as in concurrent VI–VR, the matching will be biased, with time allocation favoring the alternative with the higher feedback exponent—in this case the ratio alternative.

The economic account of operant choice is largely descriptive, in the sense that it has not so far emphasized the basis for the diminishing returns assumption, nor has it provided interpretations for the parameters involved. It is explanatory in the sense that it assumes behavior is consistent with an invariant value or preference structure. It is, therefore, a substantial advance over most previous reinforcement theories, which had no natural way of relating the effects of one source of reinforcement to the effects of another, competing one. It has the advantage over the simple statement of matching (in whatever form) in that it begins to account for the different effects of different reinforcement schedules. Nevertheless, the economic view takes the animal’s preference structure as given. It would be nice to have a basis for preference that goes beyond the analytic convenience of the power-function utility curve, and some explanatory attempts have been made.

Several convergent explanatory schemes have been independently developed to deal with behavioral allocation under constraint (Houston & McFarland, this volume; Rachlin & Burkhard, 1978; Sibly & McFarland, 1976; Staddon, 1979a). All conform to the following general description. A state space is defined, with “value” isoclines within it defined according to some general rule. Environmental, and perhaps internal, constraints such as time limitations and reinforcement schedules can be represented in the space and act to displace the point representing the animal’s current state away from the region of highest value to one of lower value. Optimal behavior then consists of a reallocation of behavior so that the representative point resides in a region of as high value as possible, consistent with the imposed constraints.

The chief differences between the various approaches are in the rule by which the isoclines are derived, and in the identification of the axes of the
space. McFarland's approach is the most general, but also the hardest to apply to particular cases. His state space has axes that are the essential physiological variables to be regulated by the animal. Observed behaviors then map, via their motivational consequences, on to these axes. Staddon's (1979a) model takes as the axes of the space the (molar) rates of observed behaviors, with "reinforcement" considered simply as an activity (e.g., eating, but also such things as running and grooming). His scheme is a formalization of Timberlake and Allison's (1974) response-deprivation analysis of Premack's well-known (1965) results on reinforcement relativity. The approach of Rachlin and Burkhard also takes molar response rates as defining the state space. Both McFarland and Staddon define isoclines in terms of a quadratic (Euclidean) cost function, of the form \( C(x_i) = A_i(x_i - B_i)^2 \), where \( A_i \) is a cost parameter and \( B_i \) is the level of \( x_i \) in the absence of a schedule constraint (in Staddon's proposal) or zero (in McFarland's analysis). The quadratic model implies circular indifference contours centered on the point of zero cost (the "bliss point," with coordinates \([B_1, \ldots, B_n]\) for \( n \) behavior axes). If the costs of different activities are not equal, the indifference contours become elliptical. If the bliss point has a positive value on a given pair of behavior axes, the corresponding indifference contours will be convex to the origin over a moderate range, as assumed by standard economic theory. Rachlin's and Staddon's approaches are molar and static. McFarland's can in principle embrace both aggregate behavior, and molecular changes, since the Pontryagin method he advocates deals not just with equilibrium values, but also with the optimal trajectory through the state space. The analytic difficulties of the method make this generality hard to realize in practice, however (McCleery, 1978).

These schemes further emphasize the similarities between reinforcement theory and economic theory. For example, economists usually distinguish between two components of the effect on commodity choice of a change (such as an increase) in the price of one of the commodities: an income effect and a substitution effect (cf. Kagel, Battalio, Green, & Rachlin, this volume). Staddon (1979a) has pointed out a parallel distinction in the analysis of optimal behavior under constraint. The income effect of a price change refers to the reduction of income caused by the increase in price of one commodity; if income is reduced for any reason (including a price increase), the distribution of commodities bought will shift just on that account. This shift can be distinguished from the effect on choice of the change in relative price caused by the increase: other commodities that satisfy some of the same needs as the now more expensive commodity are likely to be purchased in larger quantities than before. In a similar way, the imposition of a reinforcement schedule automatically restricts the availability of the contingent response (i.e., the reinforcer), so that some of the changes in level of other activities may be
just because more time is now available for them. These are restriction effects. However, some of the changes will be related to the contingency (i.e., the schedule) that relates the instrumental response or responses to the contingent response: if lever pressing is required to obtain food, lever pressing is likely to show an especially large increase; this is a contingent effect. These effects can be predicted once the preference structure (isoclines) is known.

**Empirical Results**

Optimality models are in some cases distressingly flexible (cf. Lewontin, 1979; Staddon, 1979a). Even when the quantity to be maximized can be narrowly defined, as in single-reinforcer studies of simple choice, for example, failure to support a particular model may just mean that the constraints have been imperfectly identified. When the allocation of time and energy among many different activities is to be explained, there is liable to be uncertainty about the cost/utility function as well as the constraint set. Nevertheless, it is possible to reach conclusions about some concurrent-schedule situations. First, despite some disagreement (Heyman, 1979; Heyman & Luce, 1979), it is clear that simple matching on concurrent VI–VI schedules is consistent with constrained overall (i.e., molar) maximizing of reinforcement rate (Rachlin, 1978; Rachlin et al., 1976; Staddon & Motheral, 1978, 1979). Even critics of overall maximization as a general principle can point to only small discrepancies between the optimal allocation and matching on concurrent VI–VI.

The behavior at least of pigeons on concurrent VI–VR schedules does not appear to conform to molar reinforcement maximization. Staddon and Motheral (1978) show that a maximizing model that works well for concurrent VI–VI fails to predict the biased matching that is reliably found on concurrent VI–VR. Herrnstein and Heyman (1979) have shown that pigeons suffer a substantial loss of reinforcements by matching on VI–VR. Rachlin’s (1978) approach predicts the biased matching found on VI–VR, but also predicts much greater allocation of time to the ratio alternative, which is opposite to the experimental results. None of these molar models yet provides a natural account of the known difference in molecular structure of responding on ratio and interval schedules (animals respond much faster for the same payoff frequency on ratios), although Staddon and Motheral’s approach can incorporate this limitation in the form of the slope of the linear response constraint. Molecular maximization does much better here.

Pigeons trained to switch between two signal stimuli, each associated with the delivery of response-independent food at a different rate (a concurrent variable-time, variable-time schedule, e.g., Brownstein & Pliskoff, 1968), settle down at a ratio of times in the presence of each stimulus that matches
the ratio of reinforcements obtained (see also Baum & Rachlin, 1969). It appears that this result, also, is not easily deducible from overall reinforcement rate maximization (Heyman & Luce, 1979). Several other, more specialized, procedures are well-known to produce maladaptive behavior: omission training of an autosshaped response (Williams & Williams, 1969) and behavior maintained by the production of electric shock (e.g., Morse & Kelleher, 1977) are obvious examples. These serve as reminders of the provisional status of optimizing theories.

The models of Rachlin and Burkhard (1978), and Staddon (1979a) have been quite successful in accounting for the differences between schedules in single-response situations: for example, for the generally higher rates on ratio as opposed to interval schedules, for the bitonic form of relation between response and reinforcement rates on both schedules, and for the prevailing negative tilt of the function for ratios versus the prevailing positive tilt of the function for variable-interval schedules. The general form of responding on essentially all standard single-response schedules—avoidance, punishment, spaced-responding, etc.—also follows directly from these molar models.

Summary

Molar and molecular maximizing models differ in their time scale, and also in the clues they provide to optimizing processes: molecular models suggest decision rules to guide behavior on a moment by moment basis; molar models tend to be strictly normative. The time-scale difference is not merely arbitrary. When a hungry animal is allowed to choose between two disparate pieces of food, it will almost invariably pick the larger or the tastier one. There is little doubt that when confronted with simple, composite alternatives, animals behave like maximizers. But when reinforcers are separated in time, when reinforcement contingencies are complex, and when current choice limits future options, the required optimal strategy is not so simple. The strategy required of an animal that acts to maximize reinforcement rate across an entire experimental session may sometimes be quite involved. For example, the true optimal strategy for concurrent VI−VI, given a fixed number of responses to “spend” in a fixed-duration experimental session, is a trajectory through the clock space illustrated in Figures

7 This prediction does not contradict my earlier statement that these models do not account for molecular rate differences. The prediction of higher ratio response rates refers to a comparison between two single-response experiments, not to differences in a single, concurrent-schedule experiment. Since these molar models assume, for simplicity, a fixed trade off between time and responses, they have no way of dealing with local rate differences. Relaxation of this assumption (at some cost in mathematical tractability) permits the prediction of local rate differences.
4.2 and 4.4. The switching-line analysis discussed earlier is therefore an approximation—suitable for an optimizer constrained by a limited time horizon, but not infallible.

The evidence presently available suggests that animals in choice situations act as local maximizers of some sort, picking the alternative with the highest local rate or probability of payoff. They behave as if they are unable to look very far ahead. In simple, particularly symmetrical, situations, this strategy produces results not too different from overall maximization. In more complex, or asymmetrical, situations, the resulting allocation of behavior may depart substantially from the optimum. Since simple matching of response to reinforcement ratios is not always consistent with overall maximizing, it may provide clues to the mechanisms involved. Two processes that appear to generate matching under appropriate conditions are momentary maximizing, illustrated by the clock-space analysis, and the local-reinforcement-rate-tracking process termed “meltioration” by Herrnstein and Vaughan.

**Optimal Choice and Functional Response**

Current understanding of the adaptation of predators to variations in prey density is based on a series of classic papers by C. S. Holling (1959, 1965, 1966). Holling noted three types of functional response curves relating prey density to rate of prey capture: linear (Type I), hyperbolic (Type II), and sigmoidal (Type III). Type I is characteristic of organisms such as filter feeders that consume prey at a rate proportional to rate of encounter. Type II is derived from Type I when each prey item takes a certain handling time: as prey density rises, more and more time is taken up handling captured prey, so that foraging time is reduced and further increments in prey density yield smaller and smaller increments in capture rate.

Holling originally suggested that the Type II response is characteristic of invertebrate, and the Type III of vertebrate, predators. Recent evidence belies this, however, since Type III responses have been found in protozoa and arthropods, and Type II responses have been found in vertebrates (Hassell, 1978; Murdock and Oaten, 1975). Instead, the most important variable appears to be the availability of more than one prey type: “almost all . . . predators show type 2 responses when given only one prey species. This result applies to such a wide range of organisms that one might expect it is the basic and most widespread response [Murdock & Oaten, 1975, p. 38].” The Type II curve may therefore be taken as an appropriate feedback function (in the sense used earlier in this paper) for animals foraging for a single, randomly spatially distributed prey type.

The contingency assumed in Holling’s (1959) derivation of the “disc
equation’’

8

for the Type II curve is equivalent to a ratio schedule. Thus, if

R is the amount of prey taken in unit time, prey density is D, and “attack rate” is \( \rho \), then

\[ R = \rho D t_v, \]

where \( t_v \) is the available time. In turn, \( t_v \) is reduced by the handling time, \( h \), required by each prey item captured, so that

\[ t_v = 1 - R h. \]

Combining these two equations leads to Holling’s result:

\[ R = D \rho / (1 + Dp h), \] (19)

which is the feedback function for a ratio schedule where the time denominator for \( R \) is total session time (rather than session time less reinforcement time, which is more usual), \( h \) is reinforcement duration, \( p \) is response rate, and \( D \) is the reciprocal of ratio value (i.e., probability of reinforcement).

When two or more prey types are involved, many predators take disproportionately more of the higher density prey; this is termed switching (Murdoch, 1969). The disc equation for \( n \) equal-value prey types can easily be shown to be

\[ R_t = \rho D_t (1 - \sum_{i=1}^n R_i h_i). \]

For just two types this reduces to

\[ R_1 = \rho D_1 (1 + \rho_1 D_1 h_1 + \rho_2 D_2 h_2), \] (20)

which still corresponds to a Type II response, and yields linear matching between the expected ratios of obtained prey types and prey densities:

\[ R_1 / R_2 = \rho_1 D_1 / \rho_2 D_2. \]

Switching corresponds to a relation of the form

\[ R_1 / R_2 = \alpha (D_1 / D_2)^\beta, \]

where the exponent is greater than unity.8 Hence the occurrence of switching implies a change in the form of the functional response from Type II to Type III, and this is usually observed (cf. Holling, 1965; see Krebs, 1973; Real, 1977, 1979, for reviews).

The disc equation is the feedback function for a ratio schedule, and it will be profitable to pursue this analogy for a moment. In the standard, two-choice operant experiment, responses to the two alternatives are incompatible

8 So termed because it accurately modeled the behavior of Holling’s blindfolded secretary “foraging” for sandpaper discs randomly scattered on a tabletop.

9 This is of the same form as Eq. (8), and therefore suggests a direct comparison with descriptive matching theory. There is an important difference, however: the generalized matching “law” involves obtained, rather than scheduled, reinforcement rates, hence it is an equilibrium relation involving interdependent variables. But the variable \( D \) is a true independent variable. The emphasis in matching theory on obtained measures makes for regularity of functional relations, but is misleading because much of this regularity is an artifact of the constraint between response and reinforcement rates enforced by the feedback function. For example, in numerical simulations I have found that simple random variation in response ratios on concurrent VI – VI, given the feedback function of Eq. (10) and two fixed VI values, yields a matching relation very well fitted by Eq. (8) with an exponent of approximately .85.
with each other: responding to one alternative precludes responding to the other. This is most simply expressed in the form of the linear constraint:

\[ \rho_1 + \rho_2 = C, \]

where \( \rho_1 \) is the response (attack) rate to alternative \( i \) and \( C \) is a constant. Under these conditions (as we have already seen), animals follow the optimal strategy of responding exclusively to the higher probability alternative. This is switching in its most extreme form. Hence, it is obviously important to look into the factors that might introduce a constraint of this sort into the natural foraging situation.

Competitive constraints on attack rates to two or more prey types are of two kinds, external and internal. External constraints relate to the spatial distribution of prey; internal constraints relate to information processing and the notion of search image. I consider external constraints first.

The most important source of incompatibility is nonrandom spatial distribution of prey. If each prey type is preferentially associated with a different location, then foraging for one necessarily interferes with foraging for the other, the degree of interference depending upon the degree of differential association, the travel cost between locations and their spatial relationships. Hence, spatial inhomogeneity favors switching. Since distributions in nature are almost invariably nonrandom (cf. Taylor, Woiwod, & Perry, 1978)—seeds are found in the vicinity of the parent tree, each prey insect tends to have its own microhabitat, and so on—switching should perhaps be the rule. Indeed, the theory of optimal diets (e.g., Hughes, 1979; Macarthur & Pianka, 1966; Pulliam, 1974; Schoener, 1971) predicts exclusive choice of a limited number of prey types—just one in an experimental situation where abundance is not a problem.

However, there are at least two types of negative feedback that tend to weaken the exclusive selection of a single prey type implied by the disc equation and a patchy prey distribution. The first is simply depletion (cf. Charnov, Orians, & Hyatt, 1976; Manly et al., 1972). In nature, unlike the operant laboratory, continued foraging in a region usually diminishes the rate of payoff obtainable from it: the parameter \( D \) in Eq. (19) is not constant under these conditions. If rich and lean patches are spatially intermixed, and rich patches can be depleted—or if the animal must leave a patch for reasons unconnected with foraging—then it will often pay it to forage in a nearby lean patch rather than travel farther to another rich one. Space, as well as time, enters into the natural feedback function in ways that often make modeling difficult. Under realistic conditions, the marginal-value idea, taken in conjunction with typical search patterns, might well lead to matching, rather than exclusive choice, even with an initial prey distribution that favors specialization. Depending on the period over which prey choice is averaged, either extreme specialization or, perhaps, none at all, might be found.
The depletion factor greatly complicates the problem of interpreting the functional response shown by freely moving animals in a static environment. Strictly speaking, no interpretation can be given unless the feedback function is known, but under natural conditions this is often impossible.

The second negative feedback may act over a longer term than depletion; it is the substitutability of different prey types. Different foods satisfy overlapping, but not identical, nutritional needs, and animals are sensitive to this both in general, by tiring of a monotonous diet, and sometimes specifically, by acting so as to remedy particular dietary lacks (cf. Rozin & Kalat, 1971). When confronted with two ratio schedules for the same reinforcer, most animals settle for the higher-probability one. But Rachlin et al. (1976) have shown that when each schedule is for a different reinforcer—cherry cola versus tom collins mix—rats show nonexclusive choice, even if the ratio values for each choice are quite disparate. Different food types are rarely perfect substitutes for each other: after a month of gourmet French cooking, even the most avid Francophile might yearn for a hamburger. This factor will also tend to weaken any tendency to switch in the ecological sense of exclusive preference for the more profitable alternative (Pulliam, 1975).

Analysis of feeding tactics in laboratory situations has revealed other factors that make functional response curves difficult to interpret in terms of individual behavior. For example, rats and many other omnivorous species tend to avoid novel foods. They also sample foods in ways that allow them to detect the delayed effects of novel, poisonous foods (cf. Milgram, Kramer, & Alloway, 1977; Rozin & Kalat, 1971). Obviously, judgments about such animals’ adaptation to variations in prey density will be different depending upon the time period over which diet is assessed and the presence or absence of novel items. These complications may appear irrelevant to an ecologist interested in apostatic selection (say), since his concern is obviously with the longest possible time intervals, group, rather than individual, behavior, and feedback functions that mimic the natural environment as much as possible. Nevertheless, they must be confronted whenever laboratory results are extended to the natural situation. These difficulties are most acute when the laboratory experiment is used simply as a model for the natural situation. The proper objective for laboratory studies is to learn something about the control function in the feedback diagram of Figure 4.1, since this will certainly be invariant across environments.

INTERNAL CONSTRAINTS—SEARCH IMAGE

The search-image concept has had a checkered history. Tinbergen’s (1960) original definition was not notably precise, and had subjective overtones about “learning to see” (cf. Dawkins, 1971). This idea is introspectively
4. Optimality Analyses of Operant Behavior

reasonable, but has probably been something of a distraction in efforts to
give the concept a rigorous theoretical and empirical base.

The search-image idea appears to have two essential features, both of
which represent constraints on animals' ability to adapt to varied prey
distributions. The first is temporal: animals evidently take some time to
incorporate a newly abundant prey type into their diet, especially if it is
cryptic. Tinbergen interpreted this delay as owing to the time necessary to
form a "specific searching image" for the new prey. In addition to Tin-
bergen's own observations, there is evidence for his contention from both
animal and human experiments. For example, Dawkins (1971) has shown
that when confronted with colored rice grains against either cryptic or non-
cryptic backgrounds, chicks consume the noncryptic prey first and take the
cryptic grains only after some delay. People take longer to identify a target
stimulus surrounded by many "distractor" stimuli than one with few dis-
tractors or in plain view (see Rabbitt, 1978, for a review).

The second feature is that search images are assumed to be mutually
exclusive, or at least incompatible to some degree. Obviously, if there is no
limit to the number of search images, there is no reason why animals should
not form one for every prey species. There would then be no need for the
concept. Again, there is animal and human evidence in support. Pietrewicz
and Kamil (1979) have recently shown that bluejays are better able to identify
color slides of a given cryptic prey if it is presented alone, rather than inter-
mixed with slides of another cryptic prey. In human search experiments
where subjects are asked to make a yes-no response depending on the pres-
ence or absence in a display of a target item (e.g., word or letter), the number
of target types, termed memory-set size, corresponds to number of prey types.
Crypticity is analogous to the proportion of distractor items in the display
and their similarity to memory-set items. Under most conditions, reaction
time to identify a target item is directly related both to memory-set size
(number of prey types "looked for") and proportion of distractors in the display (crypticity). People take longer to learn to look for more than one
type of item, and, at any stage, take longer to identify particular items,
the more types they are looking for (Green & Swets, 1966).

Practice interacts with these effects in ways that are still under vigorous
study (cf. Rabbitt, 1978; Schneider & Shiffrin, 1977). For example, highly
trained human and animal (pigeons: Blough, 1979) subjects are able to
identify two or more target types almost as rapidly as one, although the
amount of additional practice required may be substantial. In nature, prey
abundances fluctuate quite rapidly, so it is likely that experimental results
from early in training are of more relevance than the "automatic search" char-
etic of highly overtrained subjects.

The effects of "learning" reflect the incompatibility of search images and
are apparent on quite a short time scale. For example, Pietrewicz and Kamil found an increase in the proportion of correct positive identifications within runs of a single prey type, but no improvement when a second prey type was intermixed. In human absolute identification experiments, accuracy in identifying a particular stimulus (even a simple one, such as a tone of a given loudness out of a set of 10 such tones) is reliably improved if the same stimulus was presented on the preceding trial (cf. Holland & Lockhead, 1968; Staddon, King, & Lockhead, 1979).

Although the processes involved in the formation of a search image are evidently quite complicated, the major implication for optimality analysis of foraging is simple: It generally takes longer to look for two cryptic prey types than for one. I now show that search-image behavior—specialization on a single prey (usually the most profitable in terms of density and energy value)—may represent an optimal response to this time difference. This view makes no assumptions whatever about hypothetical images and their supposed incompatibility. Nevertheless, what it lacks in psychological content is perhaps more than compensated for by its simplicity and testable consequences.

A simple model for the constraint imposed by the time to recognize cryptic prey can be derived in the following manner. Suppose first that the animal devotes a fixed time each day to foraging for two equal-value, cryptic prey types. Then, for simplicity, assume that foraging consists of repeated episodes of the form: “search, decide to attack (or not), handle and eat prey (or not), search, etc.” Let the probability of encountering the $i$th prey type at the end of the “search” phase be $p_i$, and further assume that the search time, $\tau$, is small enough that the probability of encountering more than one type can be neglected. Thus, $\sum p_i + p_0 = 1$, where $p_0$ is the probability of finding no prey at the end of a sequence. When each “search” episode is brief, the rate at which the predator encounters prey type $i$, while it is actually searching (i.e., excluding handling and decision time), will be equal to $p_i/\tau = \rho D_i$, that is, proportional to prey density. If the predator makes no errors (i.e., never attacks when no prey is present, never misses a prey item when it is), which appears to be characteristic of stable performance when it has been measured (e.g., Dawkins, 1971), then the expected time between eating episodes in a single-prey system is given by the recursive equation:

$$E(T_i) = \tau + p_i(h + \Delta) + (1 - p_i)[\epsilon + E(T_i)].$$

(21)

That is, the expected time between attacks is equal to the search time plus the probability a prey will be encountered times the sum of handling ($h$) and decision ($\Delta$) times, plus the complementary probability of not encountering prey, times a second decision time ($\epsilon$) plus the expected time to the next successful encounter ($E[T_i]$ again). Data from animal detection experiments
4. Optimality Analyses of Operant Behavior

(4.)0ltrittttlit.y

[229x612]4.

[240x612]0ltrirtttlit.y

[282x612]/noll,.st,s

[316x589]ol

Opcront

IJclmtior

ll,1

(101x589)(notably

[139x589]llloLrgh,

[174x588]1979)
suggest

that at asymptote $\Delta = \epsilon$, which I shall assume,

writing this as $\Delta_1$ to indicate the decision time when looking for a single prey type. Thus, Equation (21) reduces to:

$$ E(T_1) = h + (\tau + \Delta_1)/p_1. $$

(22)

If the animal is looking for two prey types, identical save for appearance (e.g., two color morphs of the same species), then an additional outcome is added, which yields the recursive equation:

$$ E(T_{1.2}) = \tau + p_1(h + \Delta_2) + p_2(h + \Delta_3) + (1 - p_1 - p_2)(\Delta_2 + E(T_{1.2})) $$

(23)

given that the search time, $\tau$, is unchanged. Equation (23) reduces to:

$$ E(T_{1.2}) = h + (\tau + \Delta_2)/(p_1 + p_2). $$

(24)

The search-image hypothesis implies that $\Delta_2$ will generally be greater than $\Delta_1$, that is, that it takes the animal longer to decide whether prey is present or absent when looking for two types than when looking for only one. The condition for search-image behavior (i.e., specializing on one prey type) to be optimal is $E(T_1) < E(T_{1.2})$, which yields the simple relation:

$$ \frac{\tau + \Delta_1}{\tau + \Delta_2} < \frac{p_1}{p_1 + p_2}. $$

(25)

Several things are noteworthy about this relation. First, if handling time for both prey types is the same, the optimal choice does not depend on it. Second, if $\Delta_1 = \Delta_2$, as might be expected for noncryptic prey, then the optimal behavior is always to generalize. Third, for a given search time, $\tau$, the optimal choice depends on the relative, rather than the absolute, abundance of the more abundant prey type, as students of predator "switching" have proposed (e.g., Murdoch & Oaten, 1975). Fourth, and almost obviously, search-image behavior is favored by large differences between $\Delta_2$ and $\Delta_1$, that is, by cryptis.

Equations (21) and (23) can be generalized in obvious ways to accommodate differences in handling and decision times for different prey types, and in the absence of prey. Differences in nutritional value could also be accommodated in straightforward fashion. However, if there is a tradeoff between decision time and probability of encounter—for example, if the animal can reduce decision time substantially by permitting a few "misses" and "false alarms"—then the simple analysis will not be sufficient. Its virtue is that it shows the effects of decision time on optimal choice in perhaps the simplest possible way. Nothing more elaborate is necessary to establish the point that a competition for information-processing resources (cf. Norman & Bobrow, 1978) that shows itself as a longer decision time when looking for
two prey types rather than one may be sufficient to outweigh the encounter-time advantage of generalizing.

The general conclusion of this section is that either external or internal competitive constraints will favor exclusive choice, given a ratio-like (positive) feedback function. When matching, rather than exclusive choice, is found either competitive constraints are minimal, or negative feedbacks due to depletion (resource depression) or non-substitutability of prey may be responsible.

**Conclusion**

The principles that govern choice by animals, in the field and in the laboratory, turn out to be quite simple, at least in outline. The critical factors are the form of the feedback relation that relates value obtained to responses made, and constraints on responding. “Value” is usually, but not invariably, measured as rate of payoff, in energy per unit time. The constraints usually involve time: responding to one alternative takes up time that could have been spent responding to the other.

Behavioral diversity (nonexclusive choice) is favored by negatively accelerated feedback functions (diminishing returns, negative feedback) and weak response-competition constraints, exclusive choice by linear or positively accelerated functions and strong constraints. Negative feedback can be contributed either by reinforcement schedules of the interval type, or by internal processes that make an outcome relatively less valuable as more of it is obtained (nonsubstitutability)—for example, dietary diversity is favored by processes which ensure that one food type becomes less attractive relative to another as more of it is eaten. Constraints can either take the form of physical or temporal separation of choice alternatives, or information-processing limitations which make recognition of two items more time consuming than recognition of a single item.

These molar, static variables have their counterparts at the level of molecular, dynamic processes. The negatively accelerated interval-schedule feedback function corresponds to the negatively accelerated growth of reinforcement probability with time following each choice response. Presumably similar, internal processes are responsible for the maintenance of dietary diversity under conditions of abundance: the attractiveness of the chosen food decreasing with time and, perhaps, the attractiveness of alternative foods increasing with time. Self-inhibition at the molecular level is the counterpart of diminishing returns at the molar level.

Optimality theory takes the fun out of hedonism, but in return it gives a
picture of how things fit together. In simple situations, animals behave “rationally” as maximizers; only in situations that are complicated, or dissociate features that are reliably associated in the natural environment do they sometimes behave in strikingly nonoptimal ways. The difference is not that animals are motivated differently in simple and complex situations, but that their internal constraints allow optimal behavior only under a limited range of conditions. The two constraints I have considered here are limited information-processing capacity and a short time horizon. In turn, these constraints provide clues to the underlying processes. Little can be said with certainty on this topic, but it is encouraging to note that quite simple dynamic models have recently been shown to account for a considerable range of quantitative choice results.

This chapter has provided a review of optimality theory and animal choice that is intentionally rather uncritical. The approach is still relatively untried in behavioral psychology and few experiments have yet been done explicitly from this point of view. It is easy to find fault with most of the simple approaches so far suggested. For example, neither the hyperbolic nor the power-function feedback functions are likely to be exact representations of interval schedules (Heyman & Luce, 1979). No complete model of local reinforcement rate maximization has been proposed, much less tested. Convenient assumptions about exponentially distributed choice responses are almost invariably wrong in detail, and so on. Neverthe less, it seems obvious that the approach has considerable potential, although until recently it has been a stunted plant. The present best course may be just to give it some water and sunlight, rather than critical comments on its spindly aspect and lack of blooms. Future growth will provide ample opportunity for pruning and weeding out.

Appendix

MATCHING AND MARGINAL VALUE

The momentary-maximizing and local-rate maximizing ideas are very similar to the marginal-value approach to optimal foraging in a patchy environment. It is of some interest, therefore, to see under what conditions the marginal-value idea will predict matching of overall numbers of prey items obtained per patch to amount of time spent in a patch.

First, notice that matching of ratios of numbers of prey items (or total amount of food or energy obtained) per patch to times spent implies that the number of items obtained per patch is just proportional to the time spent in the patch. Thus, if the total time available is $T$ and the number of prey
items obtained in patch $i$ is $N_i$ for a sampling time $T_i$, then overall matching means that $N_i/N_j = T_i/T_j$, so that $N_i/T_i$, the local rate of prey capture, is constant for all patches.

This result is similar, but not identical, to the marginal-value condition that the marginal rates of capture be equal—to each other and to the average rate of capture for the habitat as a whole (Charnov, 1976; Parker & Stuart, 1976). The condition under which the two are the same is illustrated in Figure 4.A1, which shows cumulative food intake plotted as a function of time in a patch for two patches of different richesses. Suppose that the marginal value condition defines point $D$ on the higher curve as the appropriate foraging time for a patch of that richness. If time in the patch is equal to $B$, the local rate of prey capture is obviously the ratio $BD/OB$, which is the slope of line OED. Consequently, the matching point on the lower curve is at $E$, the point of intersection of line OED with that curve. In turn, point $E$ is only consistent with marginal value if the tangent at that point has the same slope as the tangent at point $D$. This will obviously be true of some classes of curve and not others, so that matching will sometimes be expected from a marginal-value analyses and sometimes not.

Analytically, a condition for matching to be consistent with marginal-value optimizing is that for any prey-capture function, $F(x)$,

$$\frac{dF(x)}{dx} = G(F(x)/x), \quad (A1)$$

where $G$ is a fixed function. That is, the derivative of the function with respect to $x$ should be expressible solely as a function of the ratio $F(x)/x$.

![Figure 4.A1.](image)

Figure 4.A1. The relation between marginal and overall matching in optimal foraging.
4. Optimality Analyses of Operant Behavior

Both the power function, \( F(x) = ax^b \), and the hyperbolic function, \( F(x) = \frac{ax}{(b + x)} \), where \( a \) and \( b \) are constants, have this property, which is why they show matching and maximizing to be equivalent when used as feedback functions in molar models of concurrent schedules.

Glossary of Symbols

- \( x_i, y_i \): rates of the \( i \)th response
- \( R_i, R(x_i), R(y_i) \): obtained reinforcement rate for the \( i \)th response
- \( R(x) \): feedback function: \( R \) as a function of \( x \)
- \( R_T \): total reinforcement rate
- \( t_i \): time since the last occurrence of response \( i \)
- \( T_i \): time spent in the \( i \)th activity
- \( a, b \): constants
- \( p \): probability
- \( \lambda_i \): rate constant of an exponential (random) process; equal to average rate of the event
- \( F_i(n) \): value of reinforcement (0 or 1) during the \( n \)th discrete-time interval
- \( F_i(n) \): running average value of reinforcement rate during the \( n \)th interval
- \( p(F) \): average probability of reinforcement during a discrete-time interval
- \( a_i \): scheduled VI reinforcement rate for alternative \( i \)
- \( U(T_i, T_j) \): subjective utility of time expenditures \( T_i \) and \( T_j \)
- \( \rho \): attack rate for a foraging predator
- \( D \): prey density
- \( h \): handling time of prey
- \( \tau \): search time—time between the end of a "decide, attack (or go on)" sequence and the next sequence
- \( \Delta_i \): decision time when looking for \( i \) prey types—time to decide whether a prey is present or absent at the end of a search sequence
- \( E(T_i) \): expected time between eating episodes

Acknowledgments

I am grateful to Robert Dale, Rob Gendron, Richard Herrnstein, Howard Rachlin, Will Vaughan, and John Vaughan for comments on an earlier version.
References


Dawkins, M. Perceptual changes in chicks, another look at the 'search image' concept. *Animal Behaviour*, 1971, 19, 566–574.


Heyman, G. M., & Luce, R. D. Operant matching is not a necessary consequence of maximizing reinforcement rate. Animal Learning and Behavior, 1979, 7, 133–140.


Morse, W., & Kelleher, R. Determinants of reinforcement and punishment. In W. K. Honig,
Sanrclson, P. Royana, R. Rozirn, R. Raprcott, R. Rcrhlin, R. Pulf
I'ress, l955. 
4. Optimality Analyses of Operant Behavior


Staddon, J. E. R. Thirst—a static analysis. *Behavioral and Brain Sciences*, 1979, 2, 120–121. (b)


