# OPTIMIZATION AND THE MATCHING LAW AS ACCOUNTS OF INSTRUMENTAL BEHAVIOR

## WILLIAM M. BAUM

UNIVERSITY OF NEW HAMPSHIRE

The interaction between instrumental behavior and environment can be conveniently described at a molar level as a feedback system. Two different possible theories, the matching law and optimization, differ primarily in the reference criterion they suggest for the system. Both offer accounts of most of the known phenomena of performance on concurrent and single variable-interval and variable-ratio schedules. The matching law appears stronger in describing concurrent performances, whereas optimization appears stronger in describing performance on single schedules.

Key words: matching law, optimization, cost-gain analysis, concurrent schedules, single schedules

Recent years have seen the rise of two molar theoretical approaches to understanding instrumental behavior: the matching law and optimization. Although they differ in a number of ways, the two approaches share at least two fundamental assumptions.

First, both assume that all behavior constitutes choice, in the sense that, no matter how limited the situation, an organism always has more than one activity in which it can engage. Besides a measured response, an organism may groom, explore, or rest. A strong form of this assumption further supposes that the total behavior that occurs in any given interval of time is constant (e.g., Herrnstein, 1974; Rachlin, 1978; Staddon, 1979). Restated, the notion means that, if one activity increases in frequency, others must decrease in compensation.

The second assumption both approaches share is that the interaction between environment and behavior is a process of two-way adjustment, because behavior and environment act to change one another. Restated, this means that the interaction can, at least in principle, be described as a feedback system.

The two assumptions together imply that behavior and environment constitute a feedback system that at equilibrium balances a mix of

activities (i.e., choice) with a mix of consequences (i.e., changes in the environment). This balance is described in a straightforward way by the matching law: the frequency of an activity relative to all others matches its reward value relative to all others (Baum, 1973, 1974b; Herrnstein, 1970).

Optimization, on the other hand, assumes that organisms maximize satisfaction. Several recent theoretical accounts are based on this as an initial assumption (Rachlin, 1978; Rachlin, Green, Kagel, & Battalio, 1976; Staddon & Motheral, 1978; Baum, Note 1; Staddon, Note 2). In present terms, we can state it thus: choice tends to produce an optimal mix of consequences. This implies a weighing of cost against benefit. The analysis then has to deal with two problems: what variables determine cost and benefit and by what mathematical relations?

## THE ORGANISM-ENVIRONMENT FEEDBACK SYSTEM

Considering the organism and its environment as a feedback system, we can diagram the system as shown in Figure 1. The organism constitutes the process controlled. Its outputs are the frequencies of its various activities, designated $B_i$; only two appear in Figure 1, although in principle there might be any number. By virtue of the experimental procedure or the nature of the organism and the rest of the world, two sorts of variables depend

on each $B_i$: those beneficial to the organism or the species, denoted as $r_i$ (for "reward"), and those costly to the organism or the species, denoted as $q_i$. The variables $r_i$ and $q_i$ represent the varied punishing and rewarding consequences of changes in the behavioral variable $B_i$. They constitute the feedback that returns to control allocation of behavior. The two combine to produce a single variable, $r_i - q_i$, which could be called the *net gain* returned by the activity $B_i$. This, together with $B_i$, allows computation of an indicator $s_i$, by the function $f(r_i - q_i, B_i)$, yet to be specified. The reference criterion, yet to be specified, tests whether the $s_i$, $i = 1$ to $n$, are all equal. If not, the reference criterion generates error greater than zero. A simple possible expression of the reference criterion and error appears in the figure:

$$\sum_{i=1}^{n} (s_i - \bar{s})^2$$

where $\bar{s}$ is the average of $s_i$, $i = 1$ to $n$. When this quantity, the variance, is zero, the reference criterion is met. A variance greater than zero would constitute error. When this occurs, the variables $B_i$ change so as to diminish the error.

Since practical and theoretical questions such as whether the measures should be rela-



Fig. 1. The organism and environment diagrammed as a feedback system. See text for explanation.

tive or absolute, whether $r$ and $q$ require a common scale, and whether or how to derive a common scale of behavior tend to depend on the indicator function $f$ that enters into the reference criterion, theoretical debates often center explicitly or implicitly around the problem of choosing the indicator. In what follows, this problem is treated as primary.

What would the indicator be for the matching law? A simple possibility emerges if we represent matching across $n$ alternatives as

$$\frac{r_1}{B_1} = \frac{r_2}{B_2} = \ldots = \frac{r_n}{B_n} \tag{1}$$

This is algebraically equivalent to the generalized form of the matching law:

$$\frac{B_i}{\sum_{j=1}^{n} B_j} = \frac{r_i}{\sum_{j=1}^{n} r_j} \tag{2}$$

for $i = 1, 2, \ldots n$. If we expand Equations 1 and 2, as Farley (1980) and de Villiers (1980) suggested, to include punishment (i.e., cost, scaled appropriately) $q$, we have:

$$\frac{r_1 - q_1}{B_1} = \frac{r_2 - q_2}{B_2} = \ldots = \frac{r_n - q_n}{B_n} \tag{3}$$

and:

$$\frac{B_i}{\sum_{j=1}^{n} B_j} = \frac{r_i - q_i}{\sum_{j=1}^{n} r_j - \sum_{j=1}^{n} q_j} \tag{4}$$

for all $i = 1$ to $n$. For the matching law, then, the indicator is:

$$s_i = \frac{r_i - q_i}{B_i}$$

Optimizing translates into maximizing of reinforcement or, when cost is significant, net gain. If we again suppose $n$ reinforcement loops (and, if necessary, $n$ cost loops), then optimizing consists in equating all partial derivatives of net gain with respect to behavioral allocation $B$:

$$\frac{\partial(r_1 - q_1)}{\partial B_1} = \frac{\partial(r_2 - q_2)}{\partial B_2} = \ldots = \frac{\partial(r_n - q_n)}{\partial B_n} \tag{5}$$

If all the $B$'s were mutually independent, then these derivatives would all equal zero. Such a supposition, however, would be unrealistic. It would mean that the performances on two concurrent schedules would match the perfor-
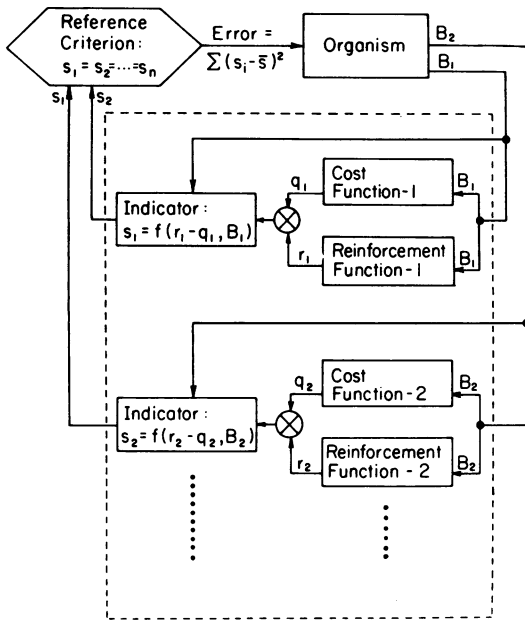
mances on the two schedules presented singly. Performance on concurrent schedules differs from performance on single schedules precisely because the $B$'s are interdependent. Since the alternatives are mutually exclusive, behavior given to one tends to be taken away from other alternatives. The strongest expression of this constraint, already mentioned, assumes that the total of behavior remains constant in the face of changing net gain. This constraint will generally be assumed to apply in what follows. In treating concurrent schedules, however, optimizing needs to assume only that the total behavior allocated to the schedules varies independently of the division of behavior between the schedules. Whatever constraint one assumes, the resulting interdependence of the $B$'s implies that the derivatives in Equation 5 generally will differ from zero. Equation 5 makes clear, however, that for optimization the indicator in Figure 1 is:

$$s_i = \frac{\partial(r_i - q_i)}{\partial B_i}$$

Regardless of whether the reference criterion is taken to be Equation 3 or Equation 5, corrective action when error differs from zero (Figure 1) will be the same. For either equation, $s_i$ varies inversely with $B_i$. When $s_i$ is too small (less than $\bar{s}$), $B_i$ should decrease; when $s_i$ is too large (greater than $s$), $B_i$ should increase.

In comparing the matching law with optimization, the first question that arises is whether optimizing theory can describe the same basic phenomena as the matching law. When and to what extent do their predictions diverge? To provide a partial answer, we consider now some common laboratory situations: concurrent variable-ratio (VR) schedules, concurrent variable-interval (VI) schedules, and concurrent VR VI schedules in which both alternatives deliver the same reinforcer.

## Concurrent VR VR

Choice between two concurrent ratio schedules usually produces exclusive preference for the smaller ratio (Herrnstein & Loveland, 1975). An account of this as optimizing is straightforward.

Let $B_1$ and $B_2$ be the response rates at Alternatives 1 and 2, and $r_1$ and $r_2$ the rates of reinforcement produced by $B_1$ and $B_2$. We designate the overall response rate as:

$$B \equiv B_1 + B_2,$$

and the proportion of responding at Alternative 1 as:

$$p \equiv \frac{B_1}{B_1 + B_2} \, .$$

Since a ratio schedule provides that the rate of reinforcement is directly proportional to the rate of responding, we write:

$$r_1 = c_1 B_1 \text{ and } r_2 = c_2 B_2 \qquad (6)$$

understanding that $c_1$ and $c_2$ are the reciprocals of the two ratios. From Equation 6 and our definition:

$$r_1 + r_2 = c_1 pB + c_2(1 - p)B$$

Rearranging, we obtain:

$$r_1 + r_2 = B\{c_2 + (c_1 - c_2)p\} \qquad (7)$$

Neglecting for now any differences in cost that might influence choice, we can consider how the gain and loss of overall reinforcement represented in Equation 7 would influence the distribution of responding $p$. Without loss in generality, we can assume that Alternative 1 is the richer alternative; that $c_1$ exceeds $c_2$. The relationship between overall reinforcement and $p$ is linear, with slope $B(c_1 - c_2)$ and intercept $Bc_2$. In the range of possible values of $p$, this relation has its maximum where $p$ equals 1.0—exclusive preference for the richer alternative. If behavior were guided by optimization, $p$ would have to go to 1.0.

As the two ratios differ by less and less, the slope of the relation in Equation 7 gets flatter and flatter. When $c_1$ equals $c_2$, $r_1 + r_2$ becomes independent of $p$. One would expect then that minor differences between the alternatives, omitted from Equation 7, would determine $p$. One could expect also that preference ($p$) would become idiosyncratic from subject to subject. Herrnstein and Loveland (1975) reported just such findings.

A more rigorous account, based on Figure 1 and Equation 5, is straightforward, because $\frac{\partial r_1}{\partial B_1}$ equals $c_1$ and $\frac{\partial r_2}{\partial B_2}$ equals $c_2$. If the ratio schedules are the same, $c_1$ equals $c_2$, and the reference criterion is satisfied regardless of behavioral allocation. If the schedules are unequal, the reference criterion cannot be met. If $c_1$ exceeds $c_2$, then $s_1$ is too large ($s_1 > s$) and $s_2$ is too small ($s_2 < \bar{s}$). Corrective

action increases $B_1$ and decreases $B_2$ until $B_2$ goes to zero and $B_1$ comes under the control characteristic of a single ratio schedule (see later).

Although Herrnstein and Loveland (1975), and others since, supposed that the matching law can predict only exclusive preference for one VR, but not which VR will be preferred, in fact the matching law offers an account of the same set of phenomena. Using Equations 1 and 2, we can write the prediction made in two ways:

$$\frac{B_1}{B_2} = \frac{r_1}{r_2} \tag{8}$$

$$\frac{r_1}{B_1} = \frac{r_2}{B_2} \tag{9}$$

Since the left of Equation 9 is $c_1$ and the right side is $c_2$, the equation shows that if $c_1$ equals $c_2$, matching prevails regardless of how responses might be distributed, whereas if $c_1$ differs from $c_2$, matching cannot prevail as long as $B_1$ and $B_2$ both exceed zero. Equation 8, on the other hand, shows that matching can be satisfied if $B_1$ and $r_1$ or $B_2$ and $r_2$ equal zero. If $c_1$ exceeds $c_2$, and we assume that the organism initially samples both alternatives, then the left side of Equation 9 exceeds the right side. Since Equation 3 provides that the difference between the two will be interpreted as error (see Figure 1), the system will act to correct the discrepancy. Since the left side of Equation 9 could be decreased only by increasing $B_1$, and the right side could be increased only by decreasing $B_2$, the system can act only by increasing $B_1$, decreasing $B_2$, or both—in other words by increasing $p$. Although the corrective action continues, the error never vanishes, until $B_2$ and $r_2$ reach zero and their ratio becomes indeterminate. The system goes to the extreme and stays there. At this point, since both sides of Equation 8 are infinite, the matching law is satisfied.

### Concurrent VI VI

Since this is the situation in which the matching relation was first observed, the account based on the matching law is straightforward. Equations 8 and 9 apply. The system acts to equate rate of payoff across alternatives as in Equation 1 (Rachlin, 1973; Revusky, 1963).

Staddon and Motheral (1978) suggested an account based on optimizing. The following is similar to theirs, but presents the problem more generally and relies on a simpler proof.

Any optimizing account begins with the relation by which a VI schedule makes rate of reinforcement depend on response rate:

$$r = \frac{1}{t + E} \tag{10}$$

where $t$ is the programmed average interreinforcement interval and $E$ represents the time beyond this, from scheduling of a reinforcer to its delivery, added by the behavioral requirement of the schedule. It is a function of $B$, the response rate. When $B$ is high, then $E$ should be small, and $r$ should approach the programmed rate of reinforcement, $1/t$. As $B$ declines, $E$ must grow; as $B$ approaches zero, $E$ must approach infinity. A simple function with these properties is: $E = 1/B$. Figure 2 shows this relation and the curve it defines for Equation 10.

With concurrent VI schedules, overall rate of reinforcement would conform to:

$$r_1 + r_2 = \frac{1}{t_1 + \dfrac{1}{B_1}} + \frac{1}{t_2 + \dfrac{1}{B_2}}$$

Using the same definitions of $p$ and $B$ as we used to write Equation 7, we obtain:

$$r_1 + r_2 = \frac{1}{t_1 + \dfrac{1}{pB}} + \frac{1}{t_2 + \dfrac{1}{(1-p)B}} \tag{11}$$

This relation is shown graphically in Figure 3. In this figure, and in the ones like it to follow (Figures 4 and 6), $B$ was assumed to be constant and equal to 100, so that the relations could be illustrated in two dimensions. Although the curve representing Equation 11 appears almost flat, clearly it has a maximum; at either extreme of preference, rewards come from only one alternative, whereas intermediate preferences produce rewards from both, and therefore represent more nearly optimal performances. Figure 4 shows another method of depicting the interaction between overall reinforcement and distribution of behavior. Instead of the proportion of behavior, the index of distribution is the logarithm of the ratio of the two frequencies. This method makes the existence of an optimal distribution more ob-
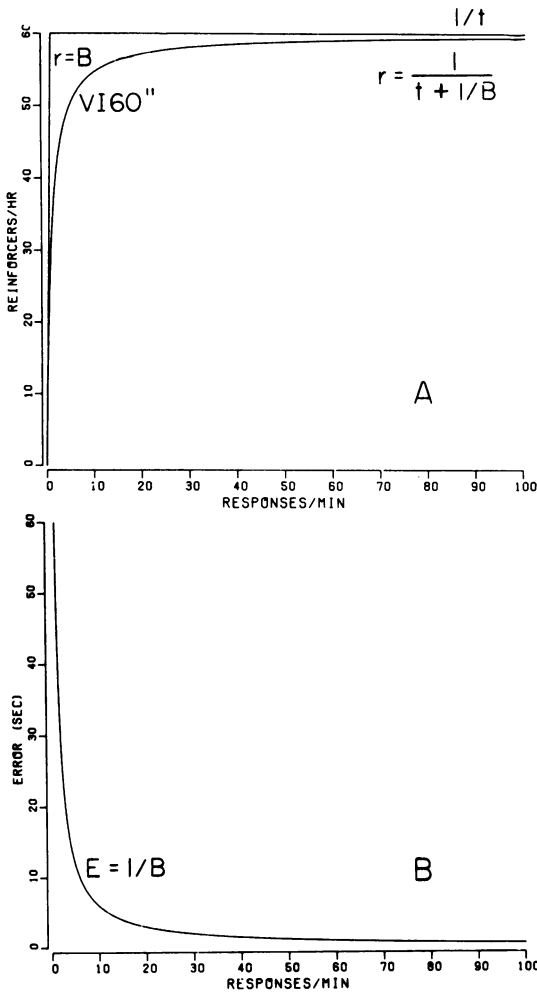
Fig. 2. A possible feedback function for variable-interval schedules, based on the assumption that responses are random in time. A: The feedback curve and its limits for low and high response rates. B: The curve relating the error (time between set-up of reinforcement and its delivery) introduced by the schedule's response requirement and rate of responding.

vious. For the sake of algebraic simplicity, however, we can rely on Equation 11.

Figures 3 and 4 show that if we assume the division of behavior to be independent of the overall response rate $B$, then optimization becomes equivalent to finding that division of behavior ($p$ here) that maximizes the overall rate of reinforcement (Equation 11). The proof that follows here is formally equivalent to applying Equation 5; it is only more direct conceptually.

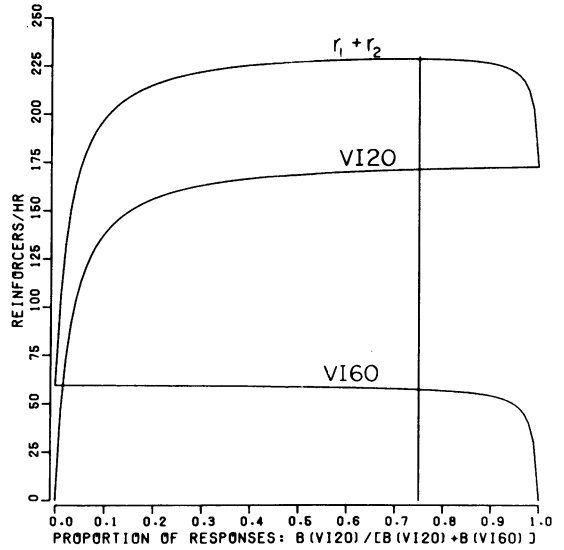To find the optimal value of $p$, we take the partial derivative of Equation 11 with respect



Fig. 3. Rate of reinforcement on conc VI 20″ VI 60″ as a function of proportion of behavior allocated to the VI 20. Lowest curve depicts variation in rate of reinforcement obtained from the VI 60. Middle curve shows variation in rate of reinforcement obtained from the VI 20. Upper curve shows overall rate of reinforcement, the sum of the other two. The vertical line marks the abscicca corresponding to maximal overall rate of reinforcement.

to $p$. (Adding cost in no way changes these statements, because, as we shall see (Figure 11), $\dfrac{\partial q}{\partial B}$ is approximately constant over most of the range of $B$.) This produces:
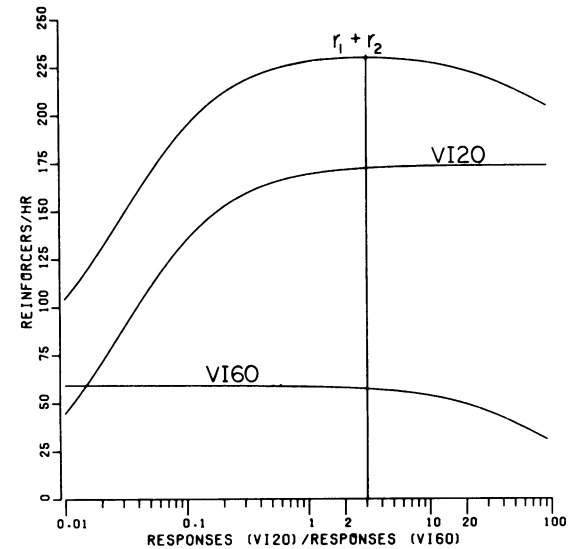


Fig. 4. Rate of reinforcement on conc VI 20″ VI 60″ as a function of ratio of responses at the alternatives. Three curves are as in Figure 3. Note logarithmic x-axis.

$$\frac{\partial(r_1 + r_2)}{\partial p} = \frac{r_1{}^2}{p^2 B} - \frac{r_2{}^2}{(1 - p)^2 B}$$

At the maximum, this derivative equals zero, with the result that:

$$\frac{r_1{}^2}{p^2 B} = \frac{r_2{}^2}{(1 - p)^2 B}$$

This simplifies to:

$$\frac{p}{1 - p} = \frac{r_1}{r_2}$$

From the definition of $p$, we conclude:

$$\frac{B_1}{B_2} = \frac{r_1}{r_2}$$

The optimal value of $p$ is the value that conforms to the matching relation. Optimizing and matching predict the same performance.

A question remains: how critical were the assumptions represented in Figure 2? Assuming that the portion of interreinforcement time dependent on behavior, $E$, equals the average interresponse interval, $1/B$, amounts to assuming that responding occurs randomly through time—that responses are generated by a Poisson process. Although this might be reasonable as a first approximation, doubtless it is inaccurate. For example, if responding tended to occur at a constant tempo, $k$, while the organism was engaged in responding then as response rate increased $E$ would gradually approach $\frac{1}{2}k$. The more accurate VI feedback function presented by Nevin and Baum (1980) captures exactly these properties. Whatever the model, however, the function relating $E$ to $B$ must have a form similar to that shown in Figure 2. Since this determines the shape of the VI reinforcement function, which in turn determines the optimal value of $p$ (Figures 3 and 4), almost any concave-upward function for $E$ should predict at least an approximation to matching. This is true, for example, of the model proposed by Heyman and Luce (1979).

Beginning with a different set of assumptions, Heyman and Luce concluded that optimization predicts performance diverging from matching. The difference between their conclusion and the present one is more apparent than real, however, because they omitted any calculation of relative reinforcement obtained, assuming instead that it remained close to the relative reinforcement programmed. When preference is strong, however, response

rate on the nonpreferred alternative can fall low enough to decrease rate of reinforcement at that alternative substantially. Since the matching relation always refers to reinforcement obtained, only a comparison of optimal choice with relative reinforcement obtained can reveal whether optimization diverges from matching.

Figure 5 shows a replotting of Heyman and Luce's Figure 3, with the addition of the omitted information about obtained reinforcement. It shows, for conc VI 1 VI 3, overall rate of reinforcement as a function of response ratio, for five different rates of changeover (Heyman and Luce's parameter $I$, which is half the average interchangeover interval). The relations are shown in the manner of Figure 4, whereas Heyman and Luce used the coordinates of Figure 3. The symbols in Figure 5 show optimal performances. The diamonds represent the maximum rates of reinforcement (ordinates) and the corresponding response ratios (abscissae). The up-arrows represent, with the same ordinates, the reinforcement ratios obtained by the response ratios (abscissae). The horizontal distance between each pair of symbols indicates the degree of deviation from the matching relation. The deviation remains small even when preference is unrealistically
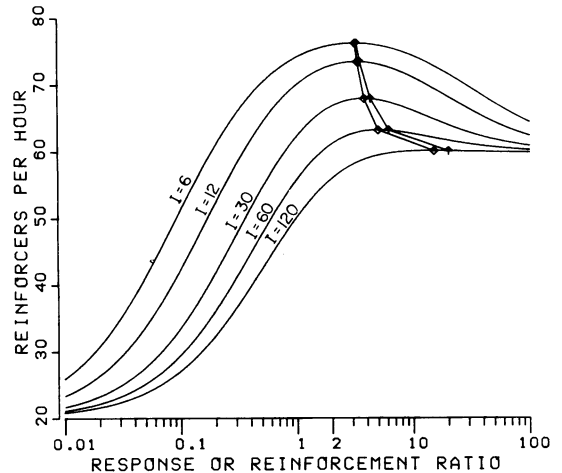


Fig. 5. Overall rate of reinforcement obtained from conc VI 1 VI 3 as a function of response ratio, according to the model of Heyman and Luce (1979). Parameter $I$ equals half the average interchangeover interval in seconds. Symbols indicate optimal performance. Diamonds show optimal response ratios (abscissae). Up-arrows show optimal reinforcement ratios (abscissae). Symbols' ordinates are the maximal reinforcement rates. Note logarithmic x-axis.

great (13 to 1) and rate of changeover unrealistically low ($I = 120$; one changeover every 4 min). For realistic rates of changeover ($I \le 30$; rate of changeover $\ge 1$ per min), the deviations of optimal performance from matching are negligible. Had Heyman and Luce included reinforcement obtained in their figure as Rachlin (1979) suggested, they would have come to a conclusion opposite to the one they published.

*Concurrent VR VI*

Once again, the account in terms of the matching law is straightforward. If the VR reinforcement function is $r_1 = cB_1$, then the system would act to equate the rates of payoff, as in Equation 9. In this case, since $r_1/B_1$ must equal $c$, the only adjustment can be to set $r_2/B_2$ equal to $c$ as well. In two separate experiments, Herrnstein found close approximations to matching (Baum, 1974a; Herrnstein, 1970).

Herrnstein and Heyman (1979) and Staddon and Motheral (1978), using different approaches, have argued that matching is incompatible with optimizing in this situation. The following account suggests that there is a relation between the two predictions and that under some conditions they may converge.

To discover what optimizing would predict, we follow the same line of reasoning as before. The overall rate of reinforcement is approximated by:

$$r_1 + r_2 = cpB + \cfrac{1}{t + \cfrac{1}{(1 - p)B}} \qquad (12)$$

Figure 6 shows this relation for one pair of schedules. Although there is an optimal value of $p$, it strongly favors the VR schedule. This could be compatible with matching if the tendency to favor the VR schedule affected only a bias in preference (Baum, 1974a). Such a bias can be seen in the data of several experiments, for which Equation 8 must be modified to include a constant $w$:

$$\frac{B_1}{B_2} = w \frac{r_1}{r_2} \qquad (13)$$

where a value of $w$ different from 1.0 represents an asymmetry between alternatives in factors such as unmeasured reinforcement (or cost) or manner of responding (Baum, 1974a).
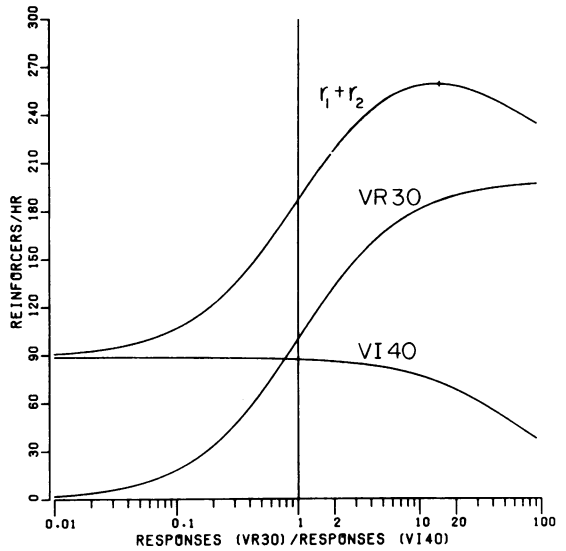


Fig. 6. Rate of reinforcement on conc VR 30 VI 40" as a function of ratio of responses at the alternatives. Curves show rates of reinforcement obtained from the two schedules and overall reinforcement, the sum of the two lower curves. Vertical line indicates the abscissa corresponding to indifference between the alternatives. Note logarithmic x-axis.

Such a result, plotted in logarithmic coordinates, parallels the matching relation, differing from it only in having an intercept different from zero (log $w$).

Taking the partial derivative of Equation 12 with respect to $p$, we obtain:

$$\frac{\partial(r_1 + r_2)}{\partial p} = cB - \frac{r_2^2}{(1 - p)^2 B}$$

Setting this equation to zero, eliminating $B$, and rearranging terms, we find:

$$\frac{p^2}{(1 - p)^2} = \frac{r_1^2}{cr_2^2}$$

From the definition of $p$, this is equivalent to:

$$\frac{B_1^2}{B_2^2} = \frac{1}{c} \cdot \frac{r_1^2}{r_2^2} \qquad (14)$$

Taking the square root of both sides of Equation 14, we see that optimizing predicts biased matching (Equation 13), with bias $w$ equal to $1/\sqrt{c}$, provided that $c$ remains constant. This means that if one varied the VI while holding the VR fixed, optimal performance would parallel matching, with a bias equal to the square root of the VR.

If the VR (and hence c) were varied, other results should ensue, deviating from matching. Suppose one kept the VI schedule constant, varying the VR from condition to condition. We can rewrite Equation 14 as:

$$\frac{B_1^2}{B_2^2} = \frac{B_1}{r_2} \cdot \frac{r_1}{r_2} \qquad (15)$$

As long as $B_2$, the response rate at the VI schedule, remains above some minimal level, the rate of reinforcement obtained from that alternative, $r_2$, will tend to remain invariant. If, moreover, $B_2$ tends to be the lower, as well as the more flexible response rate, then most of the variation in preference would arise from variation in $B_2$, and $B_1$ would tend to remain high and relatively fixed. From Equation 15, we see that these approximate invariances should lead to optimal performances conforming approximately to:

$$\frac{B_1}{B_2} = c' \sqrt{\frac{r_1}{r_2}}$$

where $c'$ equals $\sqrt{B_1/r_2}$.

Figure 7 shows optimal performances, derived by the graphical method depicted in Figure 6. Both the relations suggested in Equations 14 and 16 are verified in Figure 7, which shows optimal performances arrived at

by inspecting, for each VR VI pair, a graph similar to Figure 6. When the VR is held fixed, a line parallel to the matching relation results. This is biased matching, the bias increasing with increases in the fixed VR. When the VI is held fixed, Equation 16 appears in these logarithmic coordinates as a straight line with a slope of .5. The more sparse the VI schedule, the more bias shifts toward the VR, as one would expect from the corresponding decreases in $r_2$.

How well do observed performances conform to these optimal performances? Figure 8 shows the optimal performances from the five conditions of one of Herrnstein's experiments (reported in Baum, 1974a). In three of the conditions, the pigeons' pecks at one key were reinforced on a VR 30, whereas their pecks at the other were reinforced according to a VI that varied from condition to condition (VI 40-sec, VI 30-sec, and VI 15-sec). In accordance with Equation 13, the optimal performances for these three lie on a line parallel to the matching line. Three of the conditions contained the same VI 40-sec at one alternative while pairing it with three different VR schedules (VR 30, VR 45, and VR 60). The optimal performances for these three lie on a line with a slope of .5 in accordance with Equation 16. Given the usual sort of variation in data, such an arrangement of points might appear to conform to biased matching. This
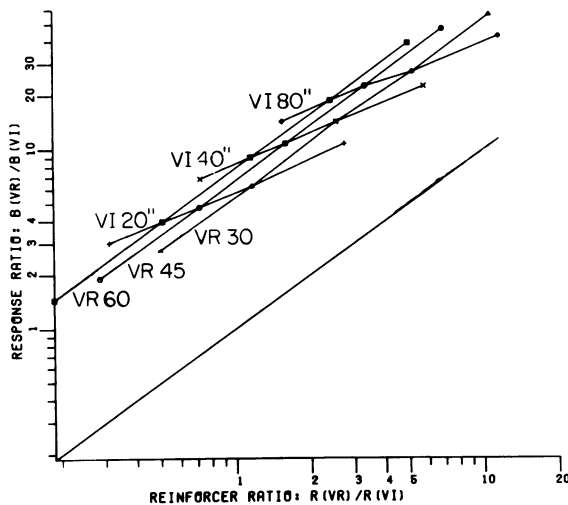


Fig. 7. Optimal performances on various conc VR VI schedules. Ratio of responses (VR to VI) appears as a function of ratio of obtained reinforcement (VR to VI). Lines with symbols show predictions for experiments in which one schedule is held constant while the other is changed. Plain diagonal line shows the matching relation. Note logarithmic axes.
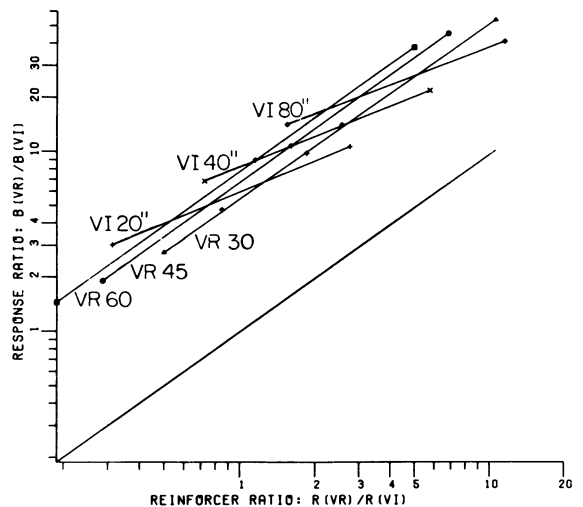


Fig. 8. Optimal performances on conc VR VI, as in Figure 7, showing predictions for the conditions of Herrnstein's experiment (points).

arises mainly from the use of a common VR schedule in three of the conditions.

Figure 9 shows the actual data from Herrnstein's experiment (Baum, 1974a). The points parallel the matching relation closely, more closely than Figure 8 would suggest. Moreover, the bias (1.4) falls far short of the value of 5 or 6 (i.e., $\sqrt{30}$) that one would expect from Figure 8 and Equation 13. The discrepancy in bias alone cannot establish that the performances in Figure 9 were nonoptimal, however, because other factors that affect bias could have reduced it to the level observed. In the absence of additional data, one can only agree with Herrnstein and Heyman (1979) that the observed matching appears incompatible with optimization. The crucial experiment remains to be done: does matching hold when the VI is fixed while the VR is varied? Although Herrnstein's data suggest that it does, the proposition remains to be tested systematically.

### Single VI and VR Schedules

Herrnstein's (1970) equation, based on the matching law, for describing performance on single schedules is:

$$B = \frac{kr}{r + r_o} \qquad (17)$$

where $B$, $k$, and $r$ are as we have defined them already: response rate, tempo (maximum possible response rate), and rate of reinforcement, respectively. The parameter $r_o$ represents reinforcement from unprogrammed sources, brought into the situation by the nature of the organism itself, such as grooming, exploring, stretching, and possibly adjunctive behavior (Cohen, 1975; Staddon & Simmelhag, 1971). Regardless of whether it is correct or not, Equation 17 embodies an indispensable principle: no matter how carefully controlled the situation, the organism itself brings in behavioral tendencies that compete with scheduled reinforcement. Without the assumption of some sort of intrinsic modulating factors, there can be no explanation of why response rate ever varies as the schedule changes. Hindsight shows us that earlier explanatory propositions, such as the law of least effort and differential reinforcement of interresponse times, reflected this need implicitly. In reply to the question why a VI 30-sec maintains a higher

response rate than a VI 5-min, the law of least effort appeals to the organism's intrinsic sensitivity to the cost of the reinforced behavior, in comparison with the gain it achieves. Differential reinforcement of interresponse times refers to the organism's capacity to be rewarded for pausing—that is, for engaging in activities other than the one reinforced by the schedule. Herrnstein may only have been the first to recognize explicitly the part played by the organism's intrinsic tendencies.

The feedback model in Figure 1 provides a means of describing modulating factors and making the need for them explicit. By itself, a monotonically increasing reinforcement function provides only for positive feedback. With no additional processes, we would have no reason to expect responding to occur at any frequency but the maximum possible. For behavior to stabilize at intermediate frequencies, the system must include negative feedback. This is provided in Figure 1 in two ways: by cost functions and by competing sources of reinforcement.

Costly relations include punishment and effort, the direct deterrents to responding. Competing sources of reinforcement provide negative feedback in the form of lost opportunities for positive reinforcement from alternative sources. The more time is devoted to one activity, the less is available for other, possibly rewarding, activities. Indeed, according to
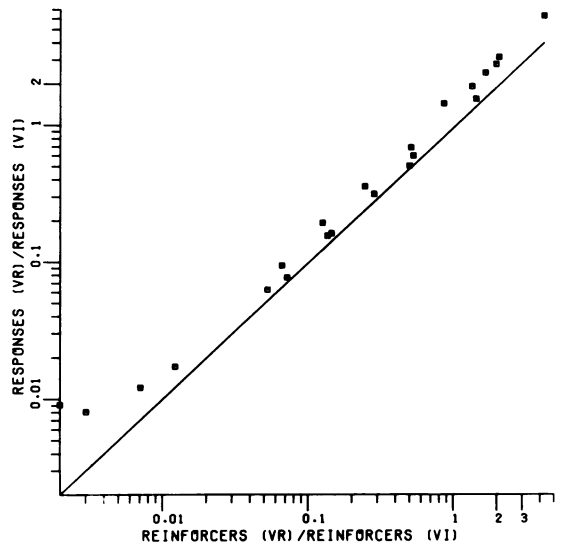


Fig. 9. Herrnstein's results with conc VR VI schedules. Data are from Baum (1974a). Axes as in Figures 7 and 8. Line shows the matching relation.

Herrnstein's (1974) general formulation, $r_o$, which explicitly refers to alternative opportunities for reinforcement, ought to vary under some circumstances. Pear (1975) suggested that it might decrease if programmed reinforcement engaged a large enough portion of behavior. In discussing concurrent schedules, we assumed implicitly that competition between the two programmed sources of reinforcement so dominated the situation as to render other beneficial and costly relations negligible. In considering single schedules, for which programmed reinforcement provides only positive feedback, inclusion of unprogrammed costly relations becomes inescapable.

Two differences between performances on ratio and interval schedules are well documented: (1) ratio schedules engender higher response rates, and (2) interval schedules maintain responding no matter how sparse reinforcement may be, whereas ratio schedules maintain responding only at relatively high rates of reinforcement. To avoid the complexities of pausing in fixed-interval and fixed-ratio performances, we consider only the differences between VI and VR performances (e.g., Ferster & Skinner, 1957; Catania, Matthews, Silverman, & Yohalem, 1977; Zuriff, 1970).
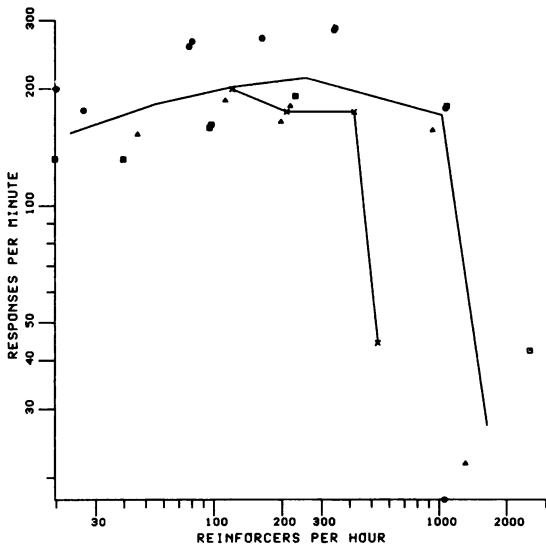


Fig. 10. Performance on variable-ratio schedules: response rate as a function of rate of reinforcement. Unconnected points show data from Brandauer's (1958) three pigeons. Line without symbols connects the average performances. Points marked x show average data from Lieberman's (1972) monkeys. Note logarithmic axes.

Although several sets of data document the general shape of the relation relating response rate to rate of reinforcement as VI schedule varies (e.g., Catania & Reynolds, 1968), few comparable sets exist for VR schedules. Brandauer's (1958) data on random-ratio schedules appear to be the only archival parametric data on single VR schedules. He exposed pigeons to schedules ranging from continuous reinforcement (CRF) to VR 600, which maintained responding in only one of three birds. Figure 10 shows his data, along with some gathered by Lieberman (1972). Lieberman's, from a study of observing responses in monkeys, show the same pattern as Brandauer's.

The two features of VR performance already mentioned appear clearly in Figure 10: response rates were generally high, and no responding could be sustained for rates of reinforcement below 20 per hour (equivalent to a VI 3-min).

A third feature of VR performance appears in the data of both studies. As rate of reinforcement increased, response rate, at first remaining approximately constant, suddenly dropped by five to ten fold. In Brandauer's experiment, this shift occurred between a random-ratio 10 and CRF. In Lieberman's, it occurred between VR 25 and VR 5.

One might expect some decrease in response rate at high rates of reinforcement, simply because a small obligatory pause must follow each reinforcer. It may be no more than the time required to move from the feeder to the key—a fraction of a second. If it included the time required to finish eating or to consume a pellet, the pause would be longer. If one counted postprandial activities like grooming as reflexive, and hence obligatory, the pause might be longer still. At low rates of reinforcement, a pause of even a few seconds can be negligible, but at high rates, even a pause of a fraction of a second can become significant. Presumably only a portion of the postreinforcement pauses that actually occur should be viewed as obligatory in this way. The best estimate of the obligatory pause would be the shortest pause that actually occurs. In Brandauer's experiment, this was .9 sec. If, for every reinforcement, .9 sec is subtracted from the time base, the response rate on CRF rises from 27 responses per minute only to 46—far less than the 150 to 200 required. Clearly we must look beyond mere artifact to account for the

change in performance. Whether a similar shift occurs with interval schedules as they approach CRF remains to be documented. Any complete account of single schedules will at least have to explain the shift for ratio schedules.

*Single Schedules and the Matching Law*

Equation 17 follows from the feedback model in Figure 1, when Equation 3 serves as reference criterion. It is Equation 4 (algebraically equivalent to Equation 3) with cost assumed negligible and only two categories of behavior, $B$ and $B_o$:

$$\frac{B}{B + B_o} = \frac{r}{r + r_o}$$

Since $B + B_o$ equals $k$ by definition, Equation 17 follows directly.

Were cost to be included, Equation 17 would become:

$$B = k \frac{r - q}{r - q + r_o - q_o} \qquad (18)$$

In practice, however, one finds that the two parameters $k$ and $r_o$ suffice for curve-fitting (de Villiers, 1977).

These two parameters, if allowed to vary, can account for most of the features of VR and VI performance. Herrnstein and Loveland (1974) and Pear (1975) suggested that some of the unprogrammed reinforcement $r_o$ is ratio-like, and therefore competes with concurrent programmed reinforcement differently, depending on whether the programmed reinforcement flows from a ratio or an interval schedule. More generally, we might suppose that $r_o$, like any reinforcement, is governed by a reinforcement function—that there is a positive relation between the magnitude of $r_o$ and the time devoted to the activities $(B_o)$ that produce it. The time available for producing $r_o$, of course, is the time left over from the time devoted to the programmed reinforcement. For low rates of responding, which leave a lot of time over, $B_o$ is high, and we expect $r_o$ to be near a maximum. Although over a wide range of rates $r_o$ may remain approximately constant, when responding at the programmed source occurs at a high enough rate, it must decrease $B_o$ enough to decrease $r_o$. A

ratio schedule, which maintains a much higher response rate for any given value of $r$ than an interval schedule, should force a decrease in $r_o$, which accords with a higher rate in Equation 17. For example, if $k$ equals 200 and $r$ equals 40, then $r_o$ equal to 10 predicts a rate of 160 responses per minute, whereas $r_o$ close to zero predicts close to 200 responses per minute.

The example makes clear, however, that change in $r_o$ alone cannot account for the difference in rates. In two types of experiments—one in which pairs of aniamls receive equal rates of reinforcement, one animal on a VI schedule and the other on a VR (the "yoked-box" experiment; e.g., Catania et al., 1977), and one in which a single animal receives reinforcement on a VI schedule and a VR schedule presented alternately in the presence of distinguishing stimuli (a multiple schedule; e.g., Zuriff, 1970)—the ratio of VR response rate to VI response rate, for the same rate of reinforcement, generally equals two to one or more. The typical range of variation of $r_o$ (10 to 20 reinforcers per hour) cannot accommodate differences this large, because changes in $r_o$ leave the predicted maximum rate $k$ unchanged. Typical values of $k$ fitted to VI performance rarely exceed 100 responses per minute (de Villiers, 1977), whereas pigeons often peck a key associated with a VR schedule at rates exceeding 200 responses per minute (e.g., Figure 10). The difference must reflect a change in $k$.

The higher value of $k$ for ratio schedules could arise from a difference in the manner of the pigeon's interaction with the response key. Laboratory lore and high-speed photography (Smith, 1974) suggest this is almost certainly the case. Whereas on interval schedules pigeons tend to make single discrete pecks, on ratio schedules they tend to vibrate or swipe at the key, producing several operations for a single cycle of extension and retraction.

To account for the dip in responding at high rates of reinforcement (Figure 10), Equation 17 again requires that $k$ change value. As reinforcement becomes dense, and a single response or only a few responses are required for each reinforcer, the matching law requires that the manner of responding become less efficient—i.e., that $k$ decrease. Although an increase in $r_o$ could theoretically account for the drop, Brandauer's data would require an im-

plausibly large value of $r_o$ in the CRF condition, on the order of 8,000 reinforcers per hour. The longer postreinforcement pause, therefore, would have to reflect, at least in part, a more lackadaisical manner of responding.

Pear (1975) pointed out that the VR reinforcement function (Equation 6) together with Equation 17 predict the cut-off in responding for large ratios. Substituting $r = cB$ in Equation 17, we find:

$$B = k - \frac{r_o}{c} \qquad (19)$$

which states that whenever the product of the VR times $r_o$ equals or exceeds $k$, no responding will occur.

One feature of VR performance may remain intractable for the matching law: the absence of responding maintained at low rates of reinforcement. Nothing about Equation 19 suggests that the function relating response rate to rate of reinforcement (Equation 17) should end abruptly at what, for interval schedules, would be a moderately high rate of reinforcement. Yet Brandauer's data (Figure 10) and all accounts of "ratio strain" (e.g., Ferster & Skinner, 1957) indicate that such is the case. Equations 17 and 18 are continuous for rates of reinforcement all the way down to zero. For $k$ equal to 210 and $r_o$ equal to 1.0, Equations 17 and 18 predict that a VR 200 will maintain 10 responses per minute at 3 reinforcers per hour. Such performances fail to occur. Some additional principle would need to be added to the matching law to account for the discontinuity.

Many details remain to be worked out if the matching law is to be considered an adequate theory of performance on single schedules. At the least, a better understanding of variation in $k$ would be required, the manner in which $r_o$ varies with the time available to it would need to be specified, the effects of direct deterrents to responding, punishment and other forms of cost, would need to be incorporated (Equation 18).

## Optimization and Single Schedules

Optimization offers an account of all the phenomena in VR and VI performance. It requires that the costly relations of Figure 1 be specified. Earlier accounts that pointed to the likelihood that interval schedules differentially reinforce longer interresponse times presupposed, usually implicitly, either the law of least effort or the maximization of frequency of reinforcement (e.g., Anger, 1956; Dews, 1962; Morse, 1966). The longer the pause, the more likely the next response will be reinforced. But if a 5-sec pause is better than a 1-sec pause, so too is a 1-hr pause still better than either. Why don't organisms respond so slowly that each response or interresponse time is certain of reinforcement? Why, on the other hand, do organisms respond at high rates on ratio schedules when all interresponse times are reinforced with equal probability? The answers invariably imply sensitivity to average frequency of reinforcement. Why does response rate ever fall short of the maximum possible? The answer always implies that the performance tends to minimize effort.

The optimizing account requires both unprogrammed reinforcement $r_o$ and cost, but it emphasizes cost. Figure 1 shows that every situation must include at least two sources of reinforcement or cost. (Cost is included here, because performance could consist in balancing two sources of cost, such as concurrent avoidance schedules.) For a single schedule of reinforcement, the second source consists in $r_o$. We assume, however, that in optimal performance the net gain from $B_o$, $r_o - q_o$, will always be at its maximum. From Equation 5:

$$\frac{\partial(r_o - q_o)}{\partial B_o} = \frac{\partial(r - q)}{\partial B} = 0$$

This means that optimizing for a single schedule consists in maximizing the net gain $r - q$ as a function of $B$. Hence we emphasize the modulating effects of the cost function.

Assuming that Equation 10 and Figure 2 describe the VI reinforcement function, that Equation 17 describes stable performance on VI schedules, and that stable performance maximizes the difference between gain and cost, one can derive a cost function analytically. Its exact form may be of little interest, because it depends on so many assumptions. As an alternative, one can begin with a few simple assumptions, and derive a similar function.

## A Cost Function

We may suppose that there are two sources of cost: effort and discomfort from neglecting activities that satisfy basic bodily needs (e.g., grooming, stretching, scratching). Assuming

the sum total of behavior to be constant, we define:

$$B + B_o = k$$

where $B$ is response rate of the activity under study, $B_o$ is the behavior, in units of $B$, allocated to all other activities, and $k$ has the same meaning as earlier.

For effort, we can assume most simply a direct proportionality to response rate $B$:

$$q_1 = mB \qquad (20)$$

The discomfort due to neglecting the body might be directly proportional to the expected delay to service of a need. If we assume that this is zero when the organism is not responding, then it equals the probability that the organism is responding times the average delay to service $d$:

$$q_2 = b\,\frac{Bd}{k} \qquad (21)$$

If we assume further that switching between the activity $B$ and other activities $B_o$ is entirely random, then $d$ is inversely proportional to the probability of the other activities: $B_o/k$. Substituting for $d$, we rewrite Equation 21:

$$q_2 = b\,\frac{B}{B_o} = b\,\frac{B}{k - B} \qquad (22)$$

Adding Equations 20 and 22, we write:

$$q = mB + b\,\frac{B}{k - B} \qquad (23)$$

The cost function in Figure 11 represents Equation 23 with $m$ equal to .03, $b$ equal to .95, and $k$ equal to 100.

Although the assumptions leading to Equation 23 might vary, the cost function it specifies has two important properties: it is positively accelerated and its slope at the origin exceeds zero. Taking the derivative:

$$\frac{dq}{dB} = m + b\,\frac{k}{(k - B)^2}$$

Since $m$ and $b$ are positive, this derivative is positive and grows as $B$ approaches $k$. When $B$ equals zero, the derivative equals $m + b/k$, a number greater than zero.

A positively accelerated cost function means that at low response rates cost is practically negligible, but that as response rate grows, cost grows disproportionately. The same increment in rate adds more to cost when rate is high than when rate is low. Staddon and Motheral (1978, Appendix B) showed that assuming cost directly proportional to response rate ($q = wB$) leads to:

$$B = a(1/\sqrt{w} - 1) \qquad (24)$$

where $a$ is the programmed rate of reinforcement of a VI schedule. If $w$ is a constant, then Equation 24 makes the erroneous prediction that response rate on VI schedules is directly proportional to programmed rate of reinforcement. Herrnstein (1970) originally proposed Equation 17 precisely because, as VI schedules are improved (that is, as the average interval is decreased), response rate grows nonlinearly: rapidly at first, but ever slower and slower, apparently approaching a limit ($k$). Equation 24 cannot be correct unless $w$ varies. To accommodate the negatively accelerated curve that VI responding follows, $w$ must increase as the response rate $B$ increases. This means a positively accelerated cost function, because $w$ represents the increment in cost for a unit increase in response rate. If $w$ increases as $B$ increases, then cost grows more rapidly for higher rates than for lower rates.

## Cost and Performance

Figure 11 shows how the properties of the cost function can account for the differences between VI and VR performances. The upper part shows reinforcement functions for a VI schedule and a VR schedule along with a cost function. The optimal response rates, at which gain differs maximally from cost, for VI and VR are indicated as $a$ and $b$, respectively. The lower portion of Figure 11 makes this clearer by showing the differences between gain and cost for the two reinforcement functions. At response rate $a$ or $b$, net gain reaches a maximum.

The optimal response rate for the VR schedule ($b$) is greater than that for the VI schedule ($a$). Indeed the geometry of the curves indicates that it would be difficult, if not impossible, to find a pair of VI and VR schedules such that the optimal VR response rate was the lower of the two. The shapes of the reinforcement functions, given a positively accelerated cost function, insure that VI responding remains the lower of the two as rate of reinforcement increases. Extremely short interval schedules must operate identically with the extreme for ratio schedules, CRF; the VI reinforcement
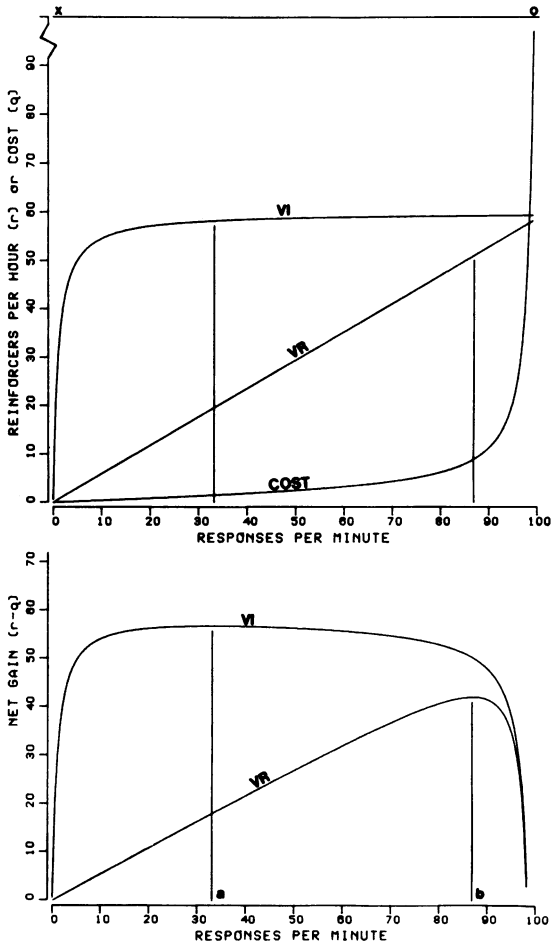
Fig. 11. Cost (q) and benefit (r: rate of reinforcement) as a function of response rate on variable-interval and variable-ratio schedules. Upper graph shows the two reinforcement functions and the cost function. Lower graph shows the two curves describing variations in difference between benefit and cost (net gain). Vertical lines indicate abscissae corresponding to optimal performances: *a* indicates optimal VI performance; *b* indicates optimal VR performance. See text for further explanation.

function would be effectively linear over the range of possible response rates. Apart from convergence at this extreme, optimizing predicts that maintained VR responding will occur at a higher rate than VI responding.

The difference between VI and VR performances at low rates of reinforcement also can be readily predicted from Figure 11. As the rate of reinforcement programmed by the VI schedule decreases, the asymptote of the reinforcement function falls, and the rate of approach to the asymptote increases—that is, as

the curve lowers, it also becomes flatter. Given the positively accelerated cost function, almost no matter how sparse the programmed reinforcement might be, still a low, but optimal, response rate will exist. The negatively accelerated reinforcement function will still diverge from the positively accelerated cost function. For ratio schedules, in contrast, as the schedule grows leaner and the slope of the reinforcement line falls, the line must approach the cost curve. Eventually it must become indistinguishable from the lower portion of the cost curve, or even fall below it. Since the slope of the cost function is greater than zero at the origin, the reinforcement line not only may cease to lie above the cost curve for practical purposes, but even may do so mathematically. For such schedules, there exists no optimal response rate; the organism cannot gain by responding. This means that responding on VR schedules should occur at high rates until the ratio is increased beyond a certain value; at that point, the cut-off point, responding should cease. Optimization, therefore, predicts not only the cut-off in ratio performance, but also what the matching law cannot: inability of ratio schedules to maintain responding at low rates of reinforemcent.

To account for the precipitous drop in response rate at the highest rates of reinforcement (Figure 10), we need one additional, but reasonable, assumption: that there is a practical upper limit on rate of reinforcement. That is, creatures are constructed so that reinforcing activities like eating, drinking, copulating, and so on have upper limits as to how often they can be repeated. The acts involved take time, and the bodily processes they initiate (e.g., digestion, recovery from fatigue) take time as well.

When the feedback function is sufficiently steep that it would predict an optimal response rate that would produce a rate of reinforcement higher than this upper limit, the optimal response rate becomes the one at which the feedback function intersects the upper limit (line *xo* in Figure 11). In Brandauer's data (Figure 10), for example, the upper limit would be around 1,600 reinforcers per hour. For maximal response rates of about 200 responses per min, 1,600 reinforcers per hr is exceeded when the schedule falls below VR 7.5. Beyond this, the optimal response rate begins to fall. For VR 7, it is 187; for VR 5,

it is 133; for VR 3, it is 80; for VR 2, it is 53. As a result, the performance curve (Figure 10) turns an abrupt corner at an abscissa value of 1,600, and response rate falls precipitously.

If this reasoning is correct, the upper limit ought to apply to performance on interval schedules as well. As the programmed inter-reinforcement interval is decreased, an interval schedule ought to function more and more like a ratio schedule and should eventually show the same drop in responding at some high rate of reinforcement. This remains to be documented.
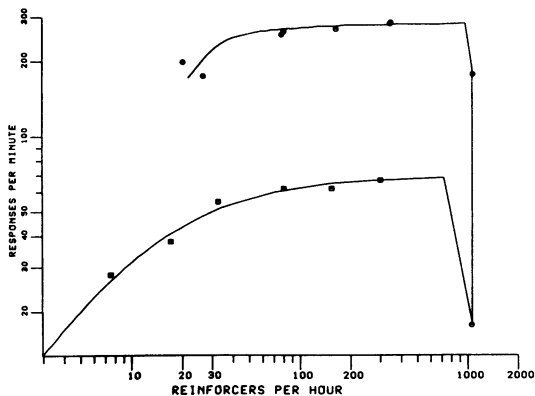


Fig. 12. Optimal performances on VR and VI schedules. The upper set of data (VR) is from Brandauer's (1958) Pigeon 15. The lower set of data (VI) is from Catania and Reynolds's (1968) Pigeon 279. See text for explanation.

Performance curves resulting from this sort of optimizing model appear in Figure 12. The upper set of points represents the data from Brandauer's Pigeon 15 (circles in Figure 10). Although this bird's response rates generally exceeded those of the other birds, the pattern of its points resembles theirs; it was selected because it produced the most data. The lower set of points represents the data from a typical pigeon exposed to a range of VI schedules (Catania & Reynolds, 1968; Pigeon 279). The curves give the loci of points representing optimal performances. The cost function derived earlier was used. For VR performance, the maximum possible response rate, $k$, equaled 290, and the parameters $m$ and $b$ equaled .12 and .2. For VI performance, $k$ equaled 70, and $m$ and $b$ equaled .0001 and .033.

The curves in Figure 12 show all the features expected. Responding on VI schedules is predicted to decrease continuously as the

schedule becomes leaner. Responding on VR schedules is predicted to remain relatively constant, decreasing gradually for large ratios, and then abruptly breaking off at a still moderate rate of reinforcement. Both curves turn an abrupt corner at the maximum possible rate of reinforcement, the value of which (1,065) was chosen to conform to the performance of Brandauer's pigeon on the two richest schedules (CRF and VR 10).

Other optimizing models of performance on single schedules have been proposed. Since Allison's only applies in a straightforward way to ratio schedules, we need not discuss it here (Allison, Miller, & Wozny, 1979). Two others derive so-called "bitonic" performance curves —that is, they predict decreased responding at both low and high rates of reinforcement. Although the curves in Figure 12 could be called "bitonic," they differ significantly from the other curves that have been proposed. Staddon (1979) and Rachlin and Burkhard (1978) fail to predict the cut-off in VR performance at moderately low rates of reinforcement. Instead, they predict VR responding to fall continuously as rate of reinforcement falls. They fail to predict also the abrupt shift in VR responding at high rates of reinforcement, predicting instead a gradual decline. In practice, data must be carefully gathered and treated to reveal an abrupt shift in a dependent variable. Averaging and any kind of error tend to make sharp corners appear round and vertical drops appear sloped. The data Staddon used to test his model, for example in his Figure 7, contain both of these problems. Despite this, however, his plots reveal that the drop in ratio responding is precipitous—decreases of 50 to 1 and more in responding for increases of only 2 to 1 in rate of reinforcement.

Rachlin (1978) is ambiguous as to whether he predicts the cut-off in VR responding for large ratios, because he applies two different models to small and large ratio schedules. He predicts unstable performance for large ratios under some circumstances. His model fails entirely, however, to predict the decline in response rate for small ratio schedules.

Although all of these models were tested against data, the data used were generally inappropriate. Many were averages across groups of animals. Many were from fixed-interval, fixed-ratio, and mult VR VI sched-

ules. For modeling performance of individual organisms on single schedules, it is best to use data of individual organisms on single schedules. Averaging generally distorts. This can be seen even in Figure 10, where the average performance suggests a gradual decline in responding at high rates of reinforcement, whereas examination of the data of individual pigeons suggests an abrupt drop. In mult VR VI, the effects of alternating VR components with VI components remain to be learned. Similarly, the postreinforcement pauses in fixed schedules remain to be understood. Can we simply ignore the inhomogeneity they represent?

Why none of the other models was tested against Brandauer's (1958) data remains unclear. It appears to be the only set of data on single VR schedules. More such data are needed, particularly comparing performances of the same organism on VR and VI schedules.

## CONCLUSION

Many complications remain to be explored, in the light of both matching and optimizing. This paper has ignored the effects of varying amount and quality of reinforcement, for example. Fixed-interval and fixed-ratio schedules produce different results than VI and VR schedules; the role of the postreinforcement pause remains to be assessed. Other questions and extensions are inevitable. The present work represents only a start.

Likening the interaction between instrumental behavior and environment to a feedback system seems to provide a handy framework for comparing different theoretical accounts. In contrast to many earlier approaches, it has the advantage of suggesting experimental tests that can be informative no matter how they turn out. Several general possibilities remain to be explored. Since at present the matching law appears to account better for concurrent performances, whereas optimization appears to account better for performances on single schedules, we need to consider whether at some times one principle might apply and at other times the other. Finally, we need to consider whether the definition of optimal performance in the laboratory might differ from optimization in foraging in nature (Pyke, Pulliam, & Charnov, 1977).

## REFERENCE NOTES

1. Baum, W. M. Optimization and instrumental behavior. Talk given at *Midwestern Association for Behavior Analysis*, Chicago, May, 1978.
2. Staddon, J. E. R. Some implications of optimality and economic analyses of operant behavior. Talk given at *Midwestern Association for Behavior Analysis*, Chicago, May, 1978.

## REFERENCES

Allison, J., Miller, M., & Wozny, M. Conservation in behavior. *Journal of Experimental Psychology: General*, 1979, 108, 4-34.

Anger, D. The dependence of interresponse times upon the relative reinforcement of different interresponse times. *Journal of Experimental Psychology*, 1956, 52, 145-161.

Baum, W. M. The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, 1973, 20, 137-153.

Baum, W. M. On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 1974, 22, 231-242. (a)

Baum, W. M. Choice in free-ranging wild pigeons. *Science*, 1974, 185, 78-79. (b)

Brandauer, C. M. The effects of uniform probabilities of reinforcement upon the response rate of the pigeon. Doctoral dissertation, Columbia University, 1958.

Catania, A. C., & Reynolds, G. S. A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1968, 11, 327-383.

Catania, A. C., Matthews, T. J., Silverman, P. J., & Yohalem, R. Yoked variable-ratio and variable-interval responding in pigeons. *Journal of the Experimental Analysis of Behavior*, 1977, 28, 155-161.

Cohen, I. L. The reinforcement value of schedule-induced drinking. *Journal of the Experimental Analysis of Behavior*, 1975, 23, 37-44.

de Villiers, P. Choice in concurrent schedules and a quantitative formulation of the law of effect. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior*. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

de Villiers, P. A. Toward a quantitative theory of punishment. *Journal of the Experimental Analysis of Behavior*, 1980, 33, 15-25.

Dews, P. B. The effect of multiple $S^\Delta$ periods on responding on a fixed-interval schedule. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 369-374.

Farley, J. Reinforcement and punishment effects in concurrent schedules: A test of two models. *Journal of the Experimental Analysis of Behavior*, 1980, 33, 311-326.

Ferster, C. B., & Skinner, B. F. *Schedules of Reinforcement*. New York: Appleton-Century-Crofts, 1957.

Herrnstein, R. J. On the law of effect. *Journal of the Experimental Analysis of Behavior*, 1970, 13, 243-266.

Herrnstein, R. J. Formal properties of the matching

law. *Journal of the Experimental Analysis of Behavior*, 1974, **21**, 159-164.

Herrnstein, R. J., & Heyman, G. M. Is matching compatible with reinforcement maximization on concurrent variable interval, variable ratio? *Journal of the Experimental Analysis of Behavior*, 1979, **31**, 209-223.

Herrnstein, R. J., & Loveland, D. H. Hunger and contrast in a multiple schedule. *Journal of the Experimental Analysis of Behavior*, 1974, **21**, 511-517.

Herrnstein, R. J., & Loveland, D. H. Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior*, 1975, **24**, 107-116.

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

Lieberman, D. A. Secondary reinforcement and information as determinants of observing behavior in monkeys (*Macaca mulatta*). *Learning and Motivation*, 1972, **3**, 341-358.

Morse, W. H. Intermittent reinforcement. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application*. New York: Appleton-Century-Crofts, 1966.

Nevin, J. A., & Baum, W. M. Feedback functions for variable-interval reinforcement. *Journal of the Experimental Analysis of Behavior*, 1980, **34**, 207-217.

Pear, J. J. Implications of the matching law for ratio responding. *Journal of the Experimental Analysis of Behavior*, 1975, **23**, 139-140.

Pyke, G. H., Pulliam, H. R., & Charnov, E. L. Optimal foraging: A selective review of theory and tests. *Quarterly Review of Biology*, 1977, **52**, 137-154.

Rachlin, H. Contrast and matching. *Psychological Review*, 1973, **80**, 217-234.

Rachlin, H. A molar theory of reinforcement sched-

ules. *Journal of the Experimental Analysis of Behavior*, 1978, **30**, 345-360.

Rachlin, H. Comment on Heyman and Luce: "Operant matching is not a logical consequence of maximizing reinforcement rate." *Animal Learning and Behavior*, 1979, **7**, 267-268.

Rachlin, H., & Burkhard, B. The temporal triangle: Response substitution in instrumental conditioning. *Psychological Review*, 1978, **85**, 22-47.

Rachlin, H., Green, L., Kagel, J. H., & Battalio, R. C. Economic demand theory and psychological studies of choice. *The Psychology of Learning and Motivation*, 1976, **10**, 129-154.

Revusky, S. H. A relationship between responses per reinforcement and preference during concurrent VI VI. *Journal of the Experimental Analysis of Behavior*, 1963, **6**, 518.

Smith, R. F. Topography of the food-reinforced key peck and the source of 30-millisecond interresponse times. *Journal of the Experimental Analysis of Behavior*, 1974, **21**, 541-551.

Staddon, J. E. R. Operant behavior as adaptation to constraint. *Journal of Experimental Psychology: General*, 1979, **108**, 48-67.

Staddon, J. E. R., & Motheral, S. On matching and maximizing in operant choice experiments. *Psychological Review*, 1978, **85**, 436-444.

Staddon, J. E. R., & Simmelhag, V. L. The "superstition" experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review*, 1971, **78**, 3-43.

Zuriff, G. E. A comparison of variable-ratio and variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1970, **13**, 369-374.