Optimization and Reinforcement

A thesis presented

by

William Vaughan, Jr.

to

The Department of Psychology

and Social Relations

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Experimental Psychology

Harvard University

Cambridge, Massachusetts

May, 1976

Acknowledgements

The theoretical position reported herein was largely developed during a long series of conversations with H. L. Miller, Jr. In addition, his participation in our "early morning sessions" (6 - 8:30 A.M.) allowed us to explore in a variety of directions. I have also. in a number of different ways, been substantially assisted or influenced by W. M. Baum, J. Cerella, M. Cohen, and R. J. Herrnstein.

Both Mrs. Antoinette Papp and Mrs. Virginia Upham helped in the running and caring for animals. In addition, I received valuable technical assistance from S. O. Bradner and D. L. Loveland.

Contents

Abstract

A model accounting for steady-state behavior on a variety of schedules of reinforcement, including interval and ratio schedules, is proposed. Organisms are assumed to behave in a manner which maximizes a variable referred to as "value." In addition to accounting for behavior on single-key schedules, concurrent variable interval and variable ratio schedules are dealt with. In order to evaluate predictions from the model, a family of reinforcement schedules is devised which includes interval and ratio schedules as special cases. A dynamic model which leads to the steady-state model is also suggested. Several experiments are reported, the results of which are consistent with the model.

Introduction

Recently a number of authors (e.g., Jenkins, 1970; Mackintosh, 1975) have questioned the value of investigating behavior maintained by schedules of reinforcement. They suggest, for example, that the behavior maintained by schedules is too complex to be susceptible to analysis at this time. On the other hand, Herrnstein (1970) takes an optimistic view with regard to both what has been acomplished and what can be acomplished in terms of behavior maintained by schedules of reinforcement. The present work falls within the latter domain, and by making a number of simplyfying assumptions attempts to integrate various schedule phenomena under one conceptual framework.

Rosen (1967) points out that, within a fairly constant environment, natural selection can be expected to result in organisms that are optimally suited to that environment. He states: "It is now possible to make the fundamental hypothesis that biological structures, which are optimal in the context of natural selection, are also optimal in the sense that they minimize some cost functional deriving from the engineering characteristics of the situation. This most natural assumption has been called the *Principle of Optimal Design*" (p. 7). By a cost functional is meant some scale along which all possibilities may be evaluated. Those animals lower on the scale can achieve some end incurring less cost than others that are higher.

Applied to behavior, this framework would suggest that animals act in a way that optimizes certain behavioral variables. For example, Herrnstein (1970) states: "Animals do not just repeat the first successful act; they are likely to improve upon it until they find something like the optimal performance. In Thorndike's puzzle box, in the maze, or in Skinner's operant conditioning chamber, animals tend toward faster, easier, and more congenial movements, unless the performances are virtually optimal to begin with" (p. 243). An attempt will be made here to view behavior maintained by schedules of reinforcement as exemplifying optimization principles. The basic point of departure will be differences in behavior under interval and ratio schedules. This is an old problem, relatively speaking: Skinner deals with it in The Behavior of Organisms (1938); the differences in behavior are large; and a number of suggestions have been put forward to account for the differences. After reviewing certain of the data and theories, some principles of optimization will be introduced and applied to this problem.

Behavior on interval and ratio schedules

At extreme values, interval and ratio schedules are identical: at very short intervals, all responses will be reinforced, as they will with a ratio requirement of one. Similarly, both an infinite time requirement and an infinite ratio requirement correspond to extinction. At intermediate values it is generally found that ratio schedules maintain higher response rates than interval schedules. Skinner (1938) ran an experiment that compared interval and ratio schedules. Rats were put on FI 5′ for one session, and then put on FR with number of responses per reinforcement approximately the same as before. All four rats showed an increase in response rate; for three of them this took several sessions to develop. Skinner pointed out that the rats must have exhibited variability in their response rates, and, furthermore, must have detected the correlation between their rate of responding and the obtained rate of reinforcement.

Either one of two factors could be responsible for the increase in response rate. The correlation between response rate and reinforcement rate had changed from close to zero on the interval schedule to close to one on the ratio schedule; such a change in correlation could possibly affect behavior. Secondly, if reinforcement rate rose slightly due to behavioral variability, it might in turn increase rate of responding, thus providing positive feedback. In this case it would not be necessary for the organism to be sensitive to correlations between response rate and reinforcement rate. Ferster and Skinner (1957) ran an experiment designed to separate these two factors. One bird was run on VI 5′, while a second bird was yoked to it so that when the first received reinforcement, it set up for the second. Their weights were adjusted so that their response rates were approximately equal. The first bird was then put on a VR schedule with the number of responses per reinforcement approximately what it had been on VI; the yoked bird continued to have reinforcement set up by the first bird. By means of this design, both birds received reinforcement at approximately the same rate, but only the VR bird had a high correlation between rate of responding and rate of reinforcement (Ferster and Skinner refer to this as "the differential reinforcement of rates or groups of responses" (p. 400)). After 29 sessions, the VR bird was responding at a substantially higher overall rate than the yoked bird, whose rate had changed little. In a second pair of birds, the bird switched to VR failed to maintain responding, possibly because it had few sessions under VI prior to being switched.

While this experiment separated correlation from reinforcement rate, the fact that the effects were shown in different birds leaves open the possibility that the characteristics of the individual birds contributed to the effect. Probably the most persuasive experiment would consist in running birds on both a family of VI and a family of VR schedules, but this does not appear to have been done. However, three kinds of experiments examining both VI and VR in a single organism have been reported: concurrent chains, with VI and VR in the terminal links; concurrent VI VR; and multiple VI VR.

Herrnstein (1964) ran birds on a concurrent chains schedule. Concurrent VI 1′ VI 1′ primed entries into the second links; during six of the nine conditions, VI was paired with VR in the second links. In general, rate of responding during the VR second link was about twice that of the VI second link, even though there was substantial overlap between obtained rates of reinforcement. In spite of these differences in response rates, it appeared that choice during the initial link was governed solely by relative rate of reinforcement. Thus the higher response rate did not detract from what could be called the value of entering the two terminal links.

A concurrent VI VR was reported by Herrnstein (1970). In this experiment, if the relative rate of responding for the VI side was above about .5, birds tended to show exclusive preference for the VI. Below .5, however, the relative number of responses to the VI tended to match the relative number of reinforcements obtained on that side. Herrnstein reported that response rates for VR were about twice those for VI. This implies that while response ratios matched reinforcement ratios, time ratios showed that twice as much time was spent on VI as on VR. In other words, under the assumption that the birds were matching time ratios to value ratios (cf., Baum & Rachlin, 1969), time spent responding for VI was twice as valuable as time spent responding for VR.

A multiple VI VR experiment has been reported by Zuriff (1970), who used component durations of one minute. When equal rates of reinforcement were obtained in the two components, response rate was nearly twice as high for VR as for VI. In a plot of response rate as a function of relative rate of reinforcement, the slope for the VR function was in general 2.5 to 3 times that for the VI component.

A number of suggestions have been put forward to account for the difference in response rates under interval and ratio schedules. Skinner (1938) discussed three possibilities. If an animal is switched to a ratio schedule from an interval schedule, with a requirement that is smaller than the number of responses per reinforcement emitted on the interval schedule, an increase in reinforcement rate will insue. If this leads to an increase in response rate, both rates will climb until some limiting factor is reached. This may be termed a positive feedback model. Skinner also dealt with the case of a larger requirement on the ratio schedule, which generally leads to extinction. This model has more recently been discussed by Herrnstein (1970).

Skinner's second account is of a more local nature. If an animal is switched to a ratio schedule with a requirement the same as the number of responses per reinforcement emitted on the preceding interval schedule, rate of responding should eventually increase because a ratio schedule tends to reinforce responses following short intervals. In contrast, an interval schedule favors responses following long intervals. It has been pointed out (e.g., Morse, 1966) that on a ratio schedule the probability of reinforcement does not change in the absence of responding; nevertheless, interval and ratio schedules still differ in the direction pointed out by Skinner. An explanation in terms of reinforced interresponse times (IRTs) has been subsequently used by Anger (1956) to account for behavior on a VI schedule, and by Shimp (1967) and Reynolds (1968) to account for behavior on interval and ratio schedules.

Finally, Skinner pointed out that while the time preceding the reinforced response was important, the time preceding previous responses could also have an effect, though a diminished one. Anger (1956) also conceded that such effects might be present. A similar suggestion was put forth by Dews (1962), who posited a delay of reinforcement gradient. He suggested that if responding on an FR schedule were rapid, all responses would be closer to reinforcement, and hence be strengthened more, than if responding were slower. Catania (1971) reported an experiment designed to demonstrate the operation of such a process.

In addition to the three possibilities already mentioned, a fourth, based on Skinner's (1948) superstition experiment has been suggested. Both Blough (1966) and Killeen (1969) point out that on interval schedules, as opposed to ratio schedules, behavior other than pecking may tend to be superstitiously reinforced, since the probability of reinforcement grows with periods of not responding.

Several conclusions may be drawn from these experiments and the related theories. Given a variable interreinforcement requirement, schedules may be decomposed into the nearly disjoint subsets of variable interval and variable ratio schedules. They produce corresponding distinct effects on behavior. The heterogeneity of plausible and mutually compatible theories may either lead one to conclude that the behavioral phenomena are adequately explained or, on the other hand, that no definitive answer has been put forth. Unfortunately, few conclusions with any greater strength or generality have been drawn.

Concepts from Optimization

The study of optimization is introduced by Luenberger (1969) as follows:

> It is perhaps natural that the concept of best or optimal decisions should emerge as the fundamental approach for formulating decision problems. In this approach a single real quantity, summarizing the performance or value of a decision, is isolated and optimized (i.e., either maximized or minimized depending on the situation) by proper selection among available alternatives. The resulting optimal decision is taken as the solution to the decision problem. (p. 1)

There exist a number of approaches to optimization theory (cf., Vagners, 1974). The particular approach taken here is generally known as nonlinear programming (Luenberger, 1973), but we may, though somewhat inaccurately, simply refer to it as optimization theory. This approach is very geometrical in character (cf., Wilde, 1964), which will allow us to avoid a great reliance on mathematical notation.

Typically, we begin with a space of some dimension, to every point of which is associated some real number. To be concrete, suppose we have a two-dimensional plane, with a number assigned to every point representing the temperature at that point. We think of there being a function, referred to as the objective function, which assigns to every point in the space the real number associated with that point. In the case of temperature on the plane, the objective function will have the form $z = f(x,y)$, where x and y are the coordinates of a point and z is the temperature there.

For purposes of representation we may think of z as a third dimension. If, as we move in any direction, temperature changes in a continuous manner, then the function f may be represented as a surface in 3-space, as in Figure 1a. In what follows we will generally treat the objective function as being well-behaved in this sense, though if need be the assumption can always be dropped. This surface will be referred to as the response surface. The height of the surface above each point represents the temperature of the plane at that point, to continue the example. It will be convenient to draw what are termed level curves, curves in the surface where f takes on a constant value. In Figure 1a, for example, curves are drawn at points where f

---

Figure 1 about here

---

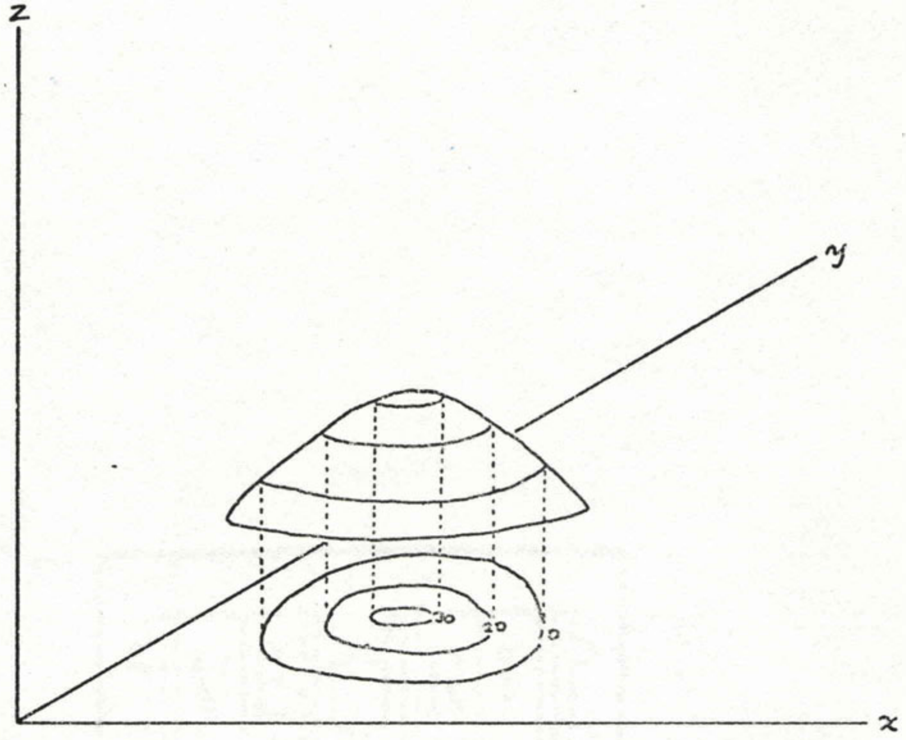
takes on the values 10, 20, and 30. If we then project these curves onto the plane, as in the figure, we may represent the three-dimensional surface in two dimensions. These projections will be referred to as contour lines, or simply contours. A well-known example is the case of curves of constant elevation in topographic maps.

We need to distinguish three classes of contour lines. If a straight line drawn between any two points on a contour does not cross the line except at those points, the contour will be said to be _convex_, otherwise concave. If such a line is neither tangent nor crosses except at the two specified points, the contour will be said to be _strictly_ _convex_. These possibilities are illustrated in Figure 2.

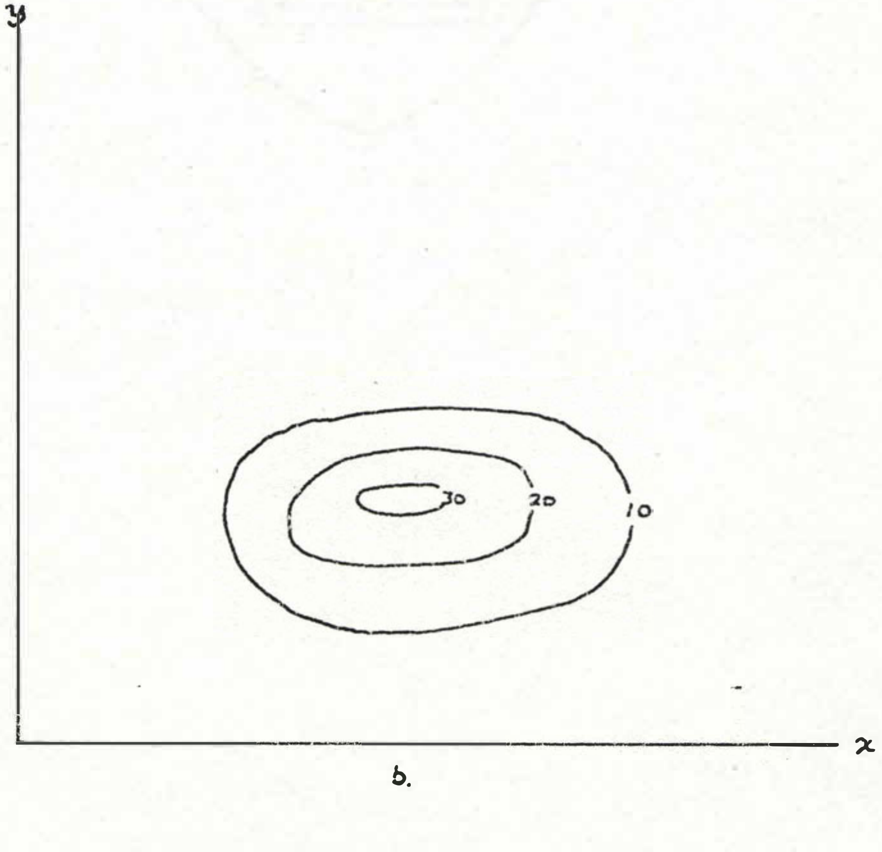
Figure 2 about here

In an optimization problem, one is given some space, on which is defined an objective function f. The solution to the problem consists in finding that point where f is minimized or maximized. To continue with

Figure 1. (a): Representation of a function $z = f(x,y)$ in 3-space. The projections of curves of constant height onto the x-y plane are also indicated. (b): By means of projections onto the x-y plane, the function f may be represented within two dimensions, as it is here.

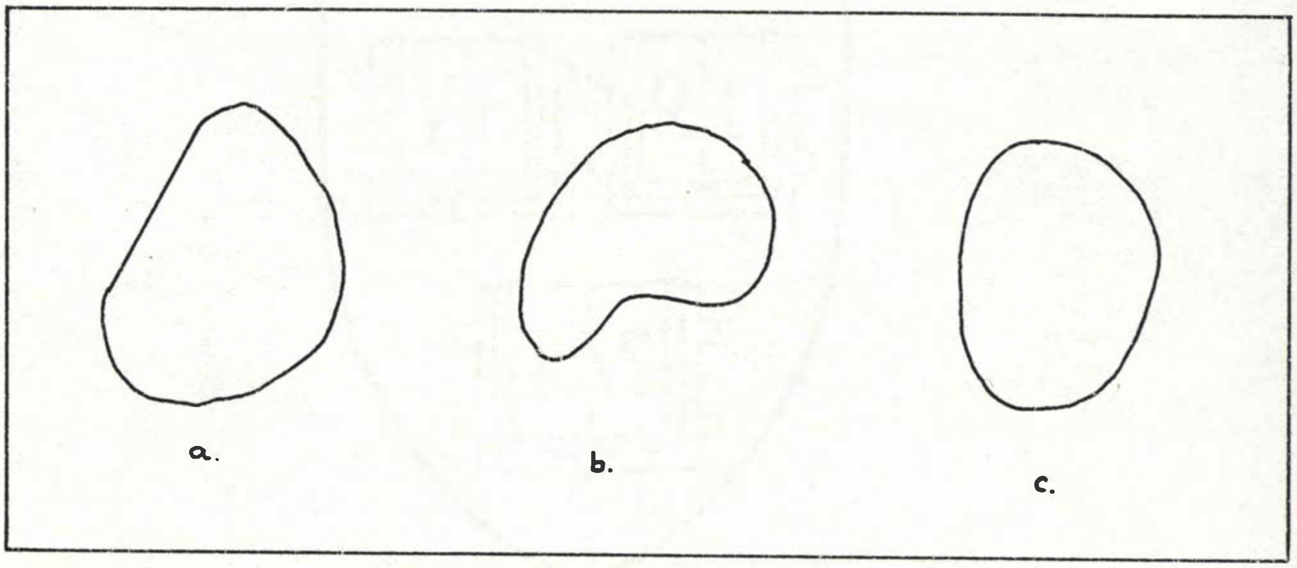


a.



b.

Figure 2. (a): Projection of a convex level curve, (b) concave level curve, and (c) strictly convex level curve.


a.
b.
c.

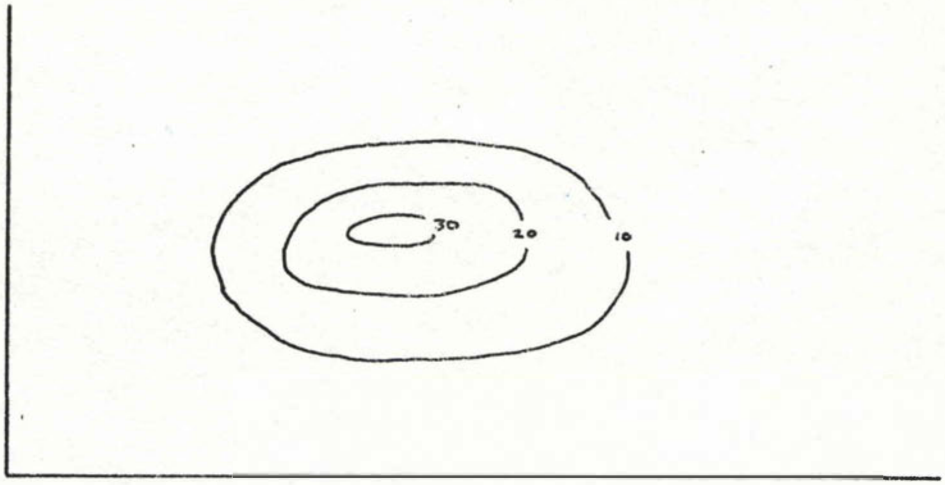
our example of temperature, it is known that for most animals some optimal temperature range exists, the average value of which may be referred to as the preferred temperature (Fraenkel and Gunn, 1940). Instead of plotting temperature of a plane as a surface in 3-space, we may plot closeness to the preferred temperature for some particular species. It would then be expected that each member of the species introduced onto the plane would end up at a local maximum of the surface, a point or area such that all neighboring points are lower in terms of preference. If the surface were unimodal, we might find all members to move to the global maximum, that point or area which is highest.

If movement anywhere within the space is possible, as in the case just considered, we say the problem is one of unconstrained optimization, illustrated in Figure 3a. In general we will be concerned with constrained optimization, of which there are two major categories. If we have an inequality constraint, it is necessary to remain on or to one side of some boundary. Two inequality constraints that must be satisfied

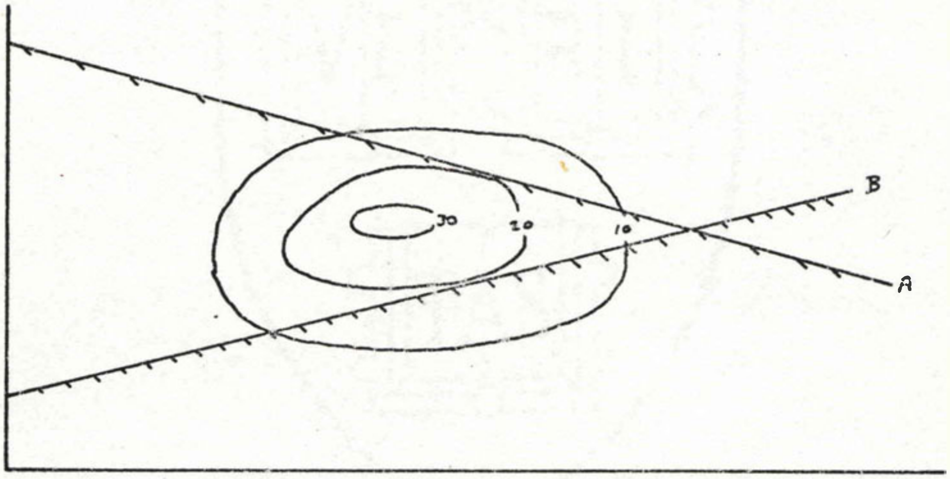
---

Figure 3 about here

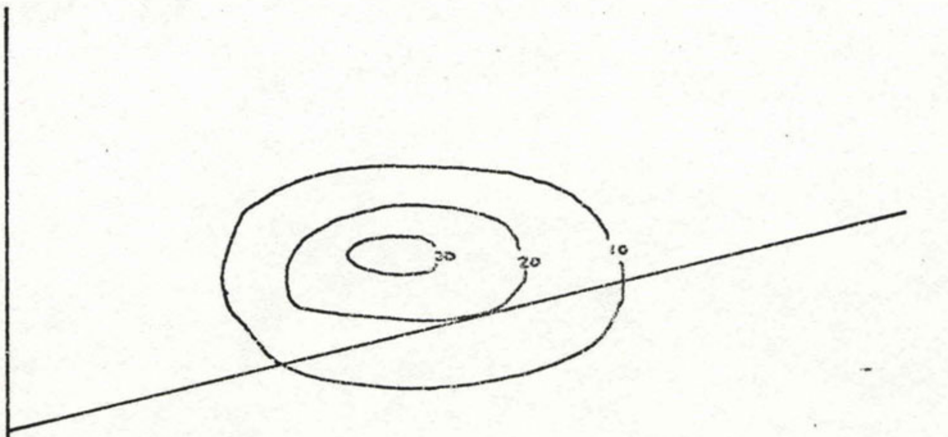
Figure 3. (a): Representation of an unconstrained optimization problem. (b): A problem involving two inequality constraints. The paralled lines indicate on which side of each boundary a solution must be sought. Constraint A is inactive, B is active. (c): An equality constraint. A solution must be sought while remaining on the boundary.



a.



b.



c.

simultaneously are shown in Figure 3b. In this figure the parallel lines are on the side of the boundary on which it is necessary to remain. Given such constraints the optimal point falls on constraint B. In such a case we say constraint B is _active_, while constraint A is _inactive_: the optimal point is an interior point in terms of the set delimited by constraint A. Finally, in the case of an _equality constraint_, one must remain on a border between two regions (Figure 3c). An equality constraint is always active. The region within which one must remain, with or without constraints, is termed the _feasible region_.

Given differentiable constraints (geometrically, continuous smooth curves) and convex contour lines, the maximum of an equality constraint or an active inequality constraint will fall at a point where the constraint is tangent to a contour, unless the constraint passes through a local extremum of the surface. If, in addition, the contours are strictly convex and the constraints linear, the constrained maximum will consist of a single point; otherwise it may consist of an interval.

We now have enough machinery to turn to the problem of free operant behavior. Although behavior is sometimes said to have _strength_ (Herrnstein, 1970), the same results can be described by saying that behavior has, or has gained, _value_. This usage allows us to consider both behavior and reinforcers as being mapped onto a common dimension. (The position of Premack (1971) is similar, but some important differences will be pointed out later.) If both responding and reinforcement may be said to have value, it would appear reasonable to be able to assign a value to the conjunction of any particular rate of responding and rate of reinforcement.

These latter two variables have been used by Baum (1973) to represent a family of VI schedules and the resulting behavior; Figure 4 shows a similar plot. Given this information, suppose we assume the bird to

----

Figure 4 about here

----

be maximizing value. If we take into account a result of Findley (1958), who found that a bird tended to stay with a richer VI schedule when a schedule only advanced when responses occurred on the key associated with it, we may hypothesize a surface with level curves similar to those in Figure 5. The numbers next to each curve

Figure 4. A family of VI schedules represented as functions relating reinforcement rate to response rate. Hypothetical date points are also shown.

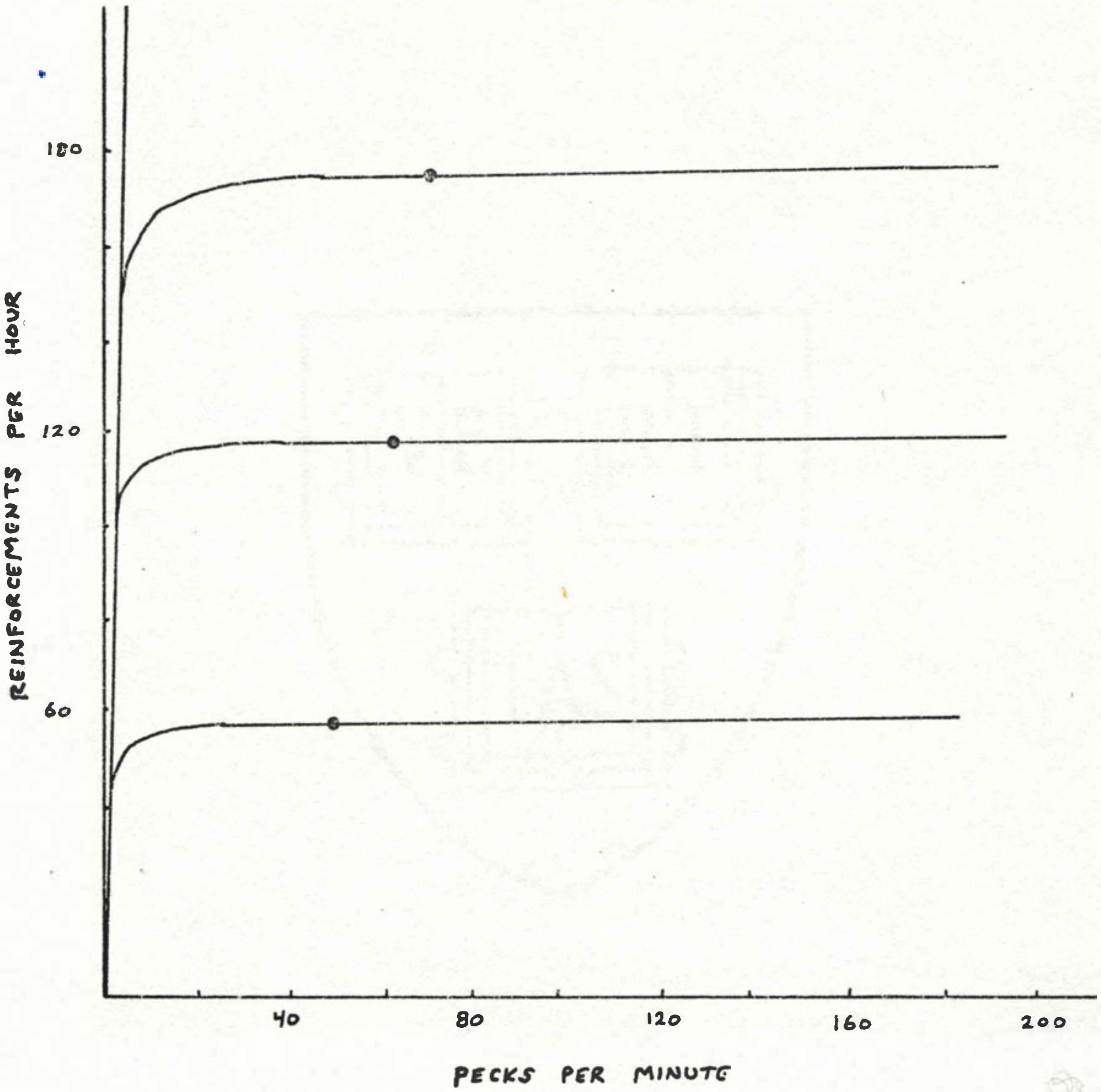

180
120
60
REINFORCEMENTS PER HOUR
40
80
120
160
200
PECKS PER MINUTE

Figure 5 about here

---

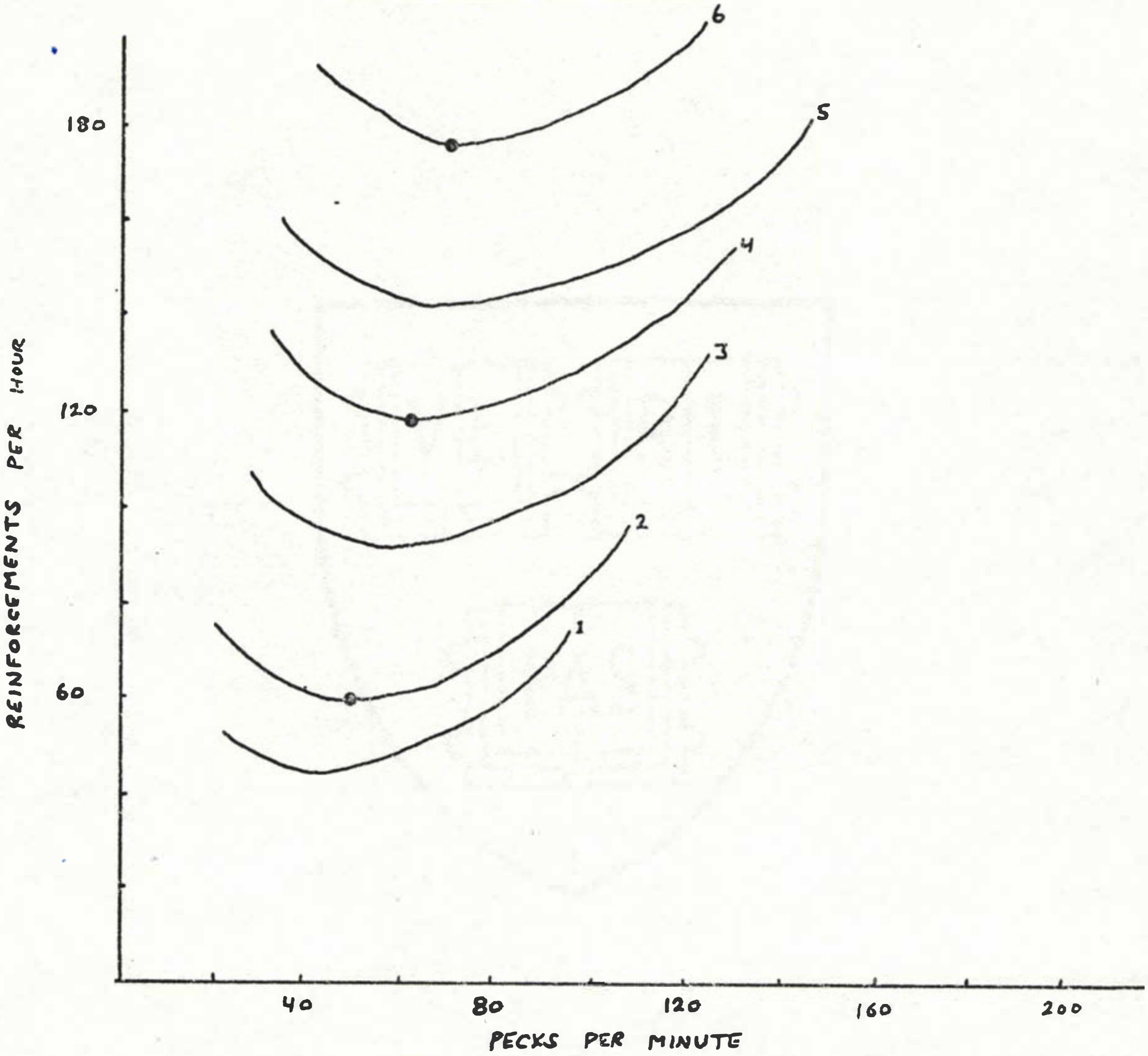
are only meant to indicate the ordinal height of the surface. Curve 3, for example, is projected from a set of points, all of equal height and all higher than those points projected onto curve 2.

To abstract somewhat, it is being suggested that every point with coordinates (P,R), where P is response rate and R is reinforcement rate, has associated with it a certain value: $V = f(P,R)$. Although the function f is unknown, we may hypothesize various properties it may have and attempt to draw conclusions that may be tested.

From the present point of view, a schedule is being viewed as an inequality constraint. Since the organism is free to respond, but does not have to consume reinforcement, its behavior can fall anywhere on or below the function. In practice, given a food-deprived bird, eating in the presence of grain has such a high probability that the inequality constraint is active, in which case it is equivalent to an equality constraint.

Given such a surface, if we consider the case of ratio schedules it can be seen that higher rates of

Figure 5. A family of contour lines that would give rise to the data points in Figure 4. Numbers next to each curve indicate the height from which it was projected.


6
180
5
4
3
120
2
1
60
REINFORCEMENTS PER HOUR
40
80
120
160
200
PECKS PER MINUTE

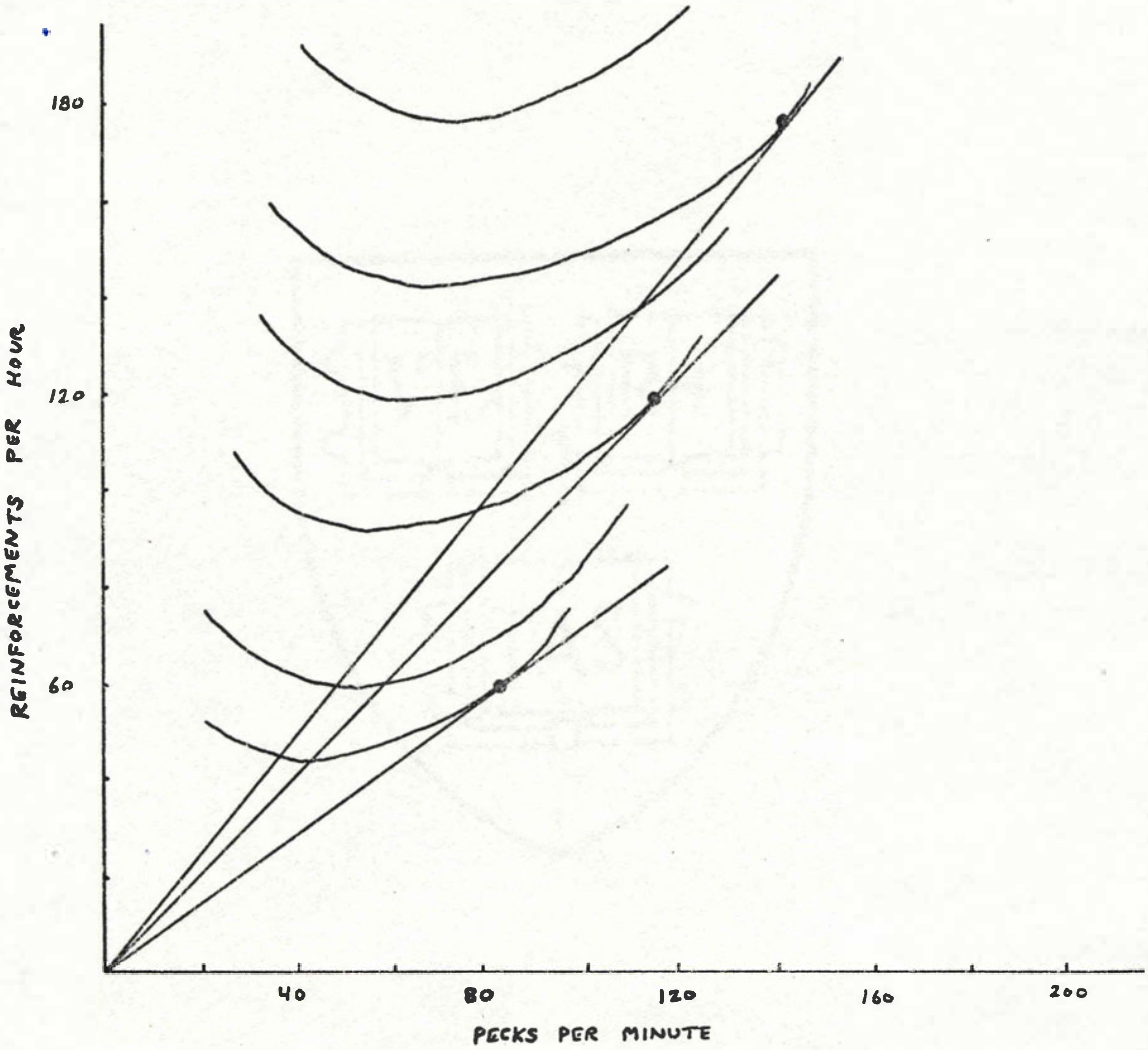
responding for the same rate of reinforcement obtained on an interval schedule can be deduced from the same principle: maximization of value (Figure 6).

Figure 6 about here

This approach is not inconsistent with previous accounts of responding on interval and ratio schedules. Rather, we may view such accounts as suggestions of how different points in the plane acquire value. Thus, it is not being denied that reinforcing a particular band of IRTs has a strong effect on behavior. Rather, it is being suggested that such a schedule confers a high value, relatively speaking, on a particular range of response rates.

We can thus see that by assuming a fairly general response surface, it is possible to deduce how an organism will behave on any VI or VR schedule. Looking at a schedule as a function relating rate of responding and rate of reinforcement, we see that knowing the specific shape of such a response surface for a particular animal is sufficient to account for behavior on any schedule that can be represented as such a function: local maxima on the schedule correspond to points of stability. It is not necessarily possible to

Figure 6. When represented as a function between response rate and reinforcement rate, VR schedules appear as straight lines with positive slope intersecting at the origin. In this figure it can be seen how the contour lines from Figure 5 give rise to higher rates of responding on VR than on VI schedules, for a given obtained rate of reinforcement.


REINFORCEMENTS PER HOUR
180
120
60
40
80
120
160
200
PECKS PER MINUTE

predict how fast an animal will respond, since some schedules may contain multiple points of local maximum value. Although a knowledge of initial conditions could aid with prediction, behavioral variability could introduce a probabilistic element.

One qualification must be mentioned. In the case of interval and ratio schedules, if we record response rate and reinforcement rate over some period of time on the order of a few hours, the point determined by these values will fall close to the function, in spite of behavioral variability. This may be contrasted with what we might expect in the case of a schedule such as DRL (differential reinforcement of low rate). If reinforcement is dependent only on the duration of the immediately preceding IRT, the same average rate of responding can give rise to very different rates of reinforcement, depending on behavioral variability. We may refer to this first case as a global contingency, and the second as a local contingency. In the absence of knowledge of behavioral variability, such a surface could only be applied to global contingencies.

When we say that value is some function of rate of responding and rate of reinforcement, we must distinguish between two cases. In using this approach in regard to data derived from averaging over a number of sessions, it is those average rates that become the arguments of the objective function f. However, in considering an organism in the actual process of responding and being reinforced, we cannot simply plug an instantaneous rate of responding and reinforcement into f, assuming this function to be known, and calculate V. A reasonable approach to this problem would be to view the organism as if it were in some sense calculating instantaneous rates of these two quantities by means of weighting functions, such that more recent events are given greater weight than more distant events. Such a weighting function is in effect a short term memory. Both Catania (1971) and Hawkes and Shimp (1975) have reported results consistent with this view.

In the present discussion the organism is being viewed as mapping a two-component vector (P,R) onto the real line (that is, V takes on real values). Theoretically, any number of components are possible; for example, duration of reinforcement is obviously of some importance. This model is being put forth not as a complete account of behavior, but more as an indication of the direction in which a complete account could lie.

Although we have been referring to value, the dimension under discussion may be more appropriately referred to as stability. Given certain conventions with regard to what is meant by response rate and reinforcement rate, what we may be said to observe in a typical single-key experiment before a steady state is reached is a change in the distribution of time spent responding at different rates. Under such conditions, we may say that one response rate, with its associated rate of reinforcement, is more stable than a second if the ratio if time spent at the first rate to that at the second increases over time. Viewed in this way, contour lines are curves of isostability: they are composed of points of equal stability.

In order to infer the shape of a response surface, the general approach would be to run an animal on a schedule until its behavior is stable. Under the assumption that the animal is maximizing value, it is possible to say that a contour line is in general tangent to the function for the schedule at the point of stability; furthermore, points on the function that are not stable can be assumed to result in less value than points in their immediate neighborhood which are stable. It is not clear whether there is some way of saying how much less value one point may have than another. However, if there is no way of inferring such a quantity, then specific values are probably immaterial. Rather, the ordinal relation between points may be all it is possible or necessary to know. This point will be discussed in a later section.

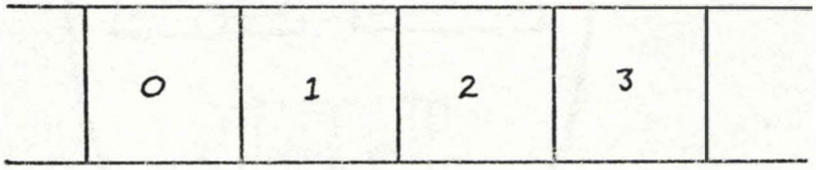
Consider the case of a surface for which every contour is strictly convex, and such that contours with smaller perimeters are projected from greater heights than those with larger perimeters. This implies that any linear constraint has one maximum, which falls at a point, rather than comprising an interval. If the response surface of a particular animal were of this nature, one could theoretically obtain a complete picture of the surface if it were possible to run the animal on schedules corresponding to arbitrary linear functions. Every point in the plane except maxima, minima, and saddle points, corresponds to the maximum of a unique linear function. As every function were studied, the local slope of a contour would be determined. It can be shown mathematically that in the limit, as the number of schedules studied becomes very large, one can infer the shape of contours. In practice, of course, some more approximate method of inferring contours would be used.

The usual way to program an interval schedule is to read a tape loop at constant speed past a microswitch which can detect holes punched in the tape. When a hole is read, reinforcement is set up and the tape reader is programmed to stop until that reinforcement is collected, at which point it again starts. It is possible to program schedules corresponding to linear functions as follows. Interval schedules as usually programmed are nonlinear (for convenience we may treat the schedule and its function as identical) due to the tape reader stopping when reinforcement sets up. We may, however, allow the reader to run continuously, except during reinforcement, and store uncollected reinforcers in a bidirectional stepper, as in Figure 7.

Figure 7 about here

The stepper starts in state 0. Any time a hole is read, the stepper moves one state to the right. If a response occurs when the stepper is in state 1 or greater, reinforcement is delivered and the stepper is decremented one step. Thus, as long as responses occur at least as often as reinforcers set up, the obtained rate of reinforcement will approach the programmed rate as time increases, and the function is linear with a zero slope and positive intercept.

Figure 7. Schematic of bidirectional stepper.


0
1
2
3

In graphing linear schedules it is convenient to ignore the fact that, at low response rates, the schedules becomes CRF (continuous reinforcement). For, while reinforcers are <u>collected</u> according to CRF, they <u>set up</u> according to the linear schedule. In any case, behavior seldom drops so low as to make significant contact with this part of the schedule. In speaking of positive and negative intercepts it must be remembered these are extrapolations from the schedules.

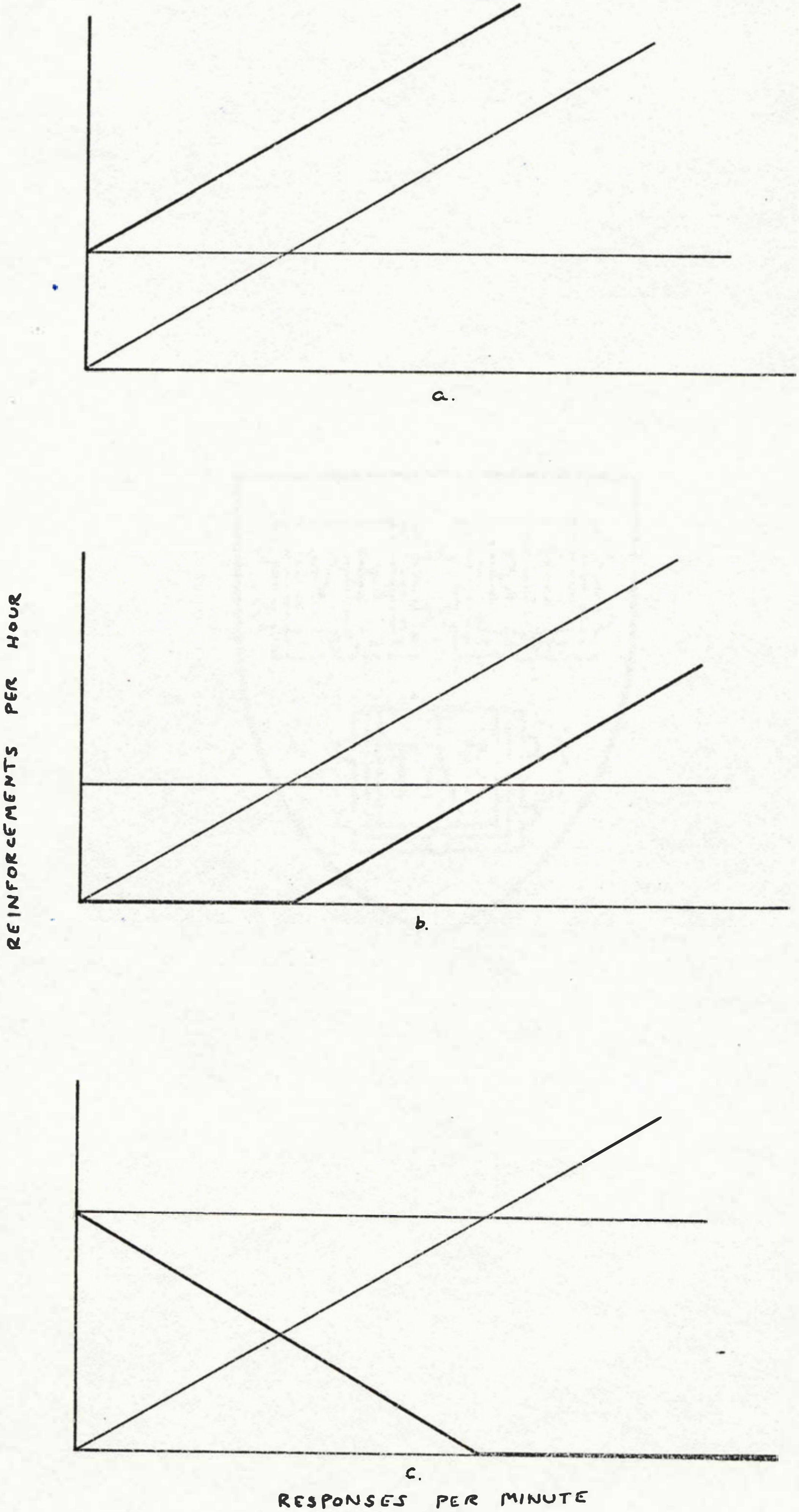
In the case of a ratio schedule, a tape reader is moved a specified distance past a microswitch for every response. When a hole is read, reinforcement is immediately delivered. Such schedules, which provide for a direct proportionality between response rate and reinforcement rate, are linear functions as they stand; they have positive slope and intercept at the origin. If we cause both a ratio reader and an interval reader to feed into a bidirectional stepper in parallel, the resulting function will be the sum of the two functions. It will thus have positive intercept and positive slope.

Two more possibilities exist. If we subtract an interval schedule from a ratio schedule, the resulting function will have positive slope and negative intercept. In order to do this it is necessay to expand the bidirectional stepper to move in the negative direction, below zero. When the interval reader reads a hole, one step in the negative direction is made; when the ratio reader reads a hole, one step in the positive direction is made. The same contingency for delivery of reinforcement as before is present. By subtracting a ratio schedule from an interval schedule, functions with positive intercept and negative slope may be generated. These functions are illustrated in Figure 8.

Figure 8 about here

In the following experiments, when two schedules were to be added, each had a variable interreinforcement requirement. If either had had a fixed requirement, this would have resulted in an unwanted periodicity in received reinforcement rate. When one schedule was subtracted from another, the former had a fixed and the latter a variable requirement. In this case it was assumed that the fixed requirement would not give rise to a detectible periodicity, while subtracting a schedule with a variable requirement might lead to a large variance in the distribution of interreinforcement times.

Figure 8. (a): Sum of interval and ratio schedules. (b): Ratio minus interval schedule. (c): Interval minus ratio schedule.


a.
REINFORCEMENTS PER HOUR
b.
c.
RESPONSES PER MINUTE

Experiment 1: Functions with Negative Slope

As previously mentioned, simple schedules of reinforcement may be fairly readily divided into two classes, providing what may be termed _local_ or _global_ contingencies. To date, the only schedule that has provided for a negative correlation between response rate and reinforcement rate has been the DRL schedule. Since reinforcement on such a schedule depends only on the duration of the immediately preceding IRT, the contingency is a local one. Thus a linear schedule with negative slope fills a gap by providing a global negative correlation schedule. Under such a schedule, any IRT greater than zero may be reinforced. Within the present optimization framework, such schedules must be used in order to examine those areas in which contours have negative slope, if such exist.

There is some evidence that for pecking in pigeons, as response rate decreases the slopes of contours pass from having positive values through zero and become negative. In other words, for a given rate of reinforcement, below a certain rate of responding, value may fall. Staddon and Simmelhag (1971) found that on interval schedules a peck-reinforcement contingency had little effect on rate of responding, though it did cause responding to be directed to the key. Staddon (1972, p. 226) refers to this rate as the "natural" rate of responding. From the present perspective we may think of "natural" rate as that occurring on a function with zero slope, and recognize that this does not exclude the possibility that some tradeoff between a lower response rate and higher reinforcement rate is possible. The present experiment was designed to look for such a tradeoff.

Method

Subjects

Six White Carneaux pigeons and three homing pigeons were used: two Carneaux and one homing pigeon in each of three conditions. They all had prior experimental histories, and were maintained at about 80% of free feeding weight.

Apparatus

A standard pigeon chamber was used, with a single key about 8.5 in. (21 cm) above the floor, centered on an 11 in. (28 cm) wide wall, above a standard feeding magazine. A force of about 14 g (.14 N) was required to operate the key, which was transilluminated with two 7-w red bulbs. An auditory feedback click was provided for each response. The chamber was illuminated with two 7-w white bulbs, except during reinforcement, when only the magazine was illuminated. White noise and a fan masked extraneous noise. Electromechanical equipment was used, in conjunction with a PDP-8/e computer, which provided timed pulses to step a tape reader in order to set up reinforcements.

Procedure

Birds were first run on three negative slopes, produced by subtracting a fixed ratio from a variable interval schedule, the latter composed of 16 intervals generated from Fleshler and Hoffman (1962). The interval schedules used, in this order, were: VI 30" (120 rf/hr), VI 45" (80 rf/hr), and VI 90" (40 rf/hr). The interval tape reader did not stop, except during reinforcement; a bidirectional stepper stored uncollected reinforcers, from the interval reader, or deficits, from the ratio reader. For one group of birds, FR 20 was subtracted; for the second group, FR 40; for the third, FR 60. A minimum of 5 sec had to elapse between the end of reinforcement and the availability of the next, in order to prevent a number of reinforcements from being collected in close succession. Reinforcement consisted of 3 sec access to mixed grain. Each session terminated after 40 minutes or 40 presentations of grain, whichever occurred first. Birds were run seven days a week. After running on negative slopes with three different intercepts, they were run on three flat slopes with reinforcement rate equated to the average of the last five days on each of the negative slopes. In this case only the interval reader was used, pulsed by computer with different parameters for each bird. Each bird was run on a condition for a minimum of about 20 sessions and until its day to day behavior appeared stable. The first condition was run for a minimum of about 60 sessions, in case behavior under these schedules changed in an unusually slow manner. The conditions and the number of sessions each bird was run are summarized in Table 1.

Table 1 about here

Results

The schedules, and the points on those schedules where behavior was stable, are presented in Figure 9, which shows average rates from the last five days.

Figure 9 about here

In Table 2 the overall response rates, reinforcement rates, and session times are shown for the last five

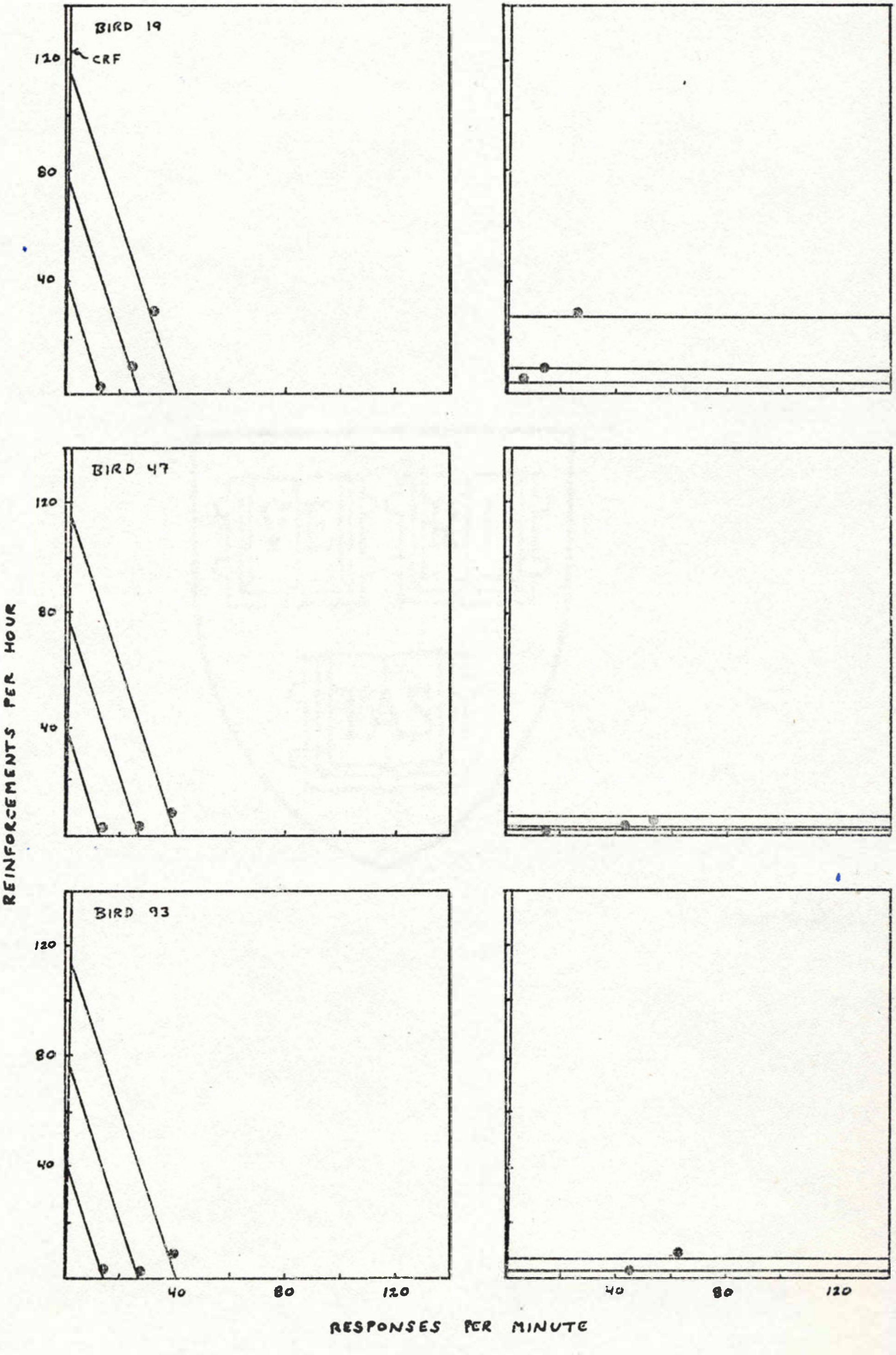
Table 1. Conditions in Experiment 1 and number of sessions each bird run on each condition. In each case, ´flat´ indicates a flat slope with reinforcement rate equated to that received on the corresponding negative slope. All negative slopes were run in the order indicated but before all flat slopes.

Schedules	Subjects		
	Bird 19	Bird 47	Bird 93
a VI30"-FR20	65	67	69
b Flat	29	29	30
c VI45"-FR20	23	23	23
d Flat	25	24	43
e VI60"-FR20	24	24	36
f Flat	23	21	-
	Bird 140	Bird 150	Bird 94
g VI30"-FR40	71	63	70
h Flat	28	34	24
i VI45"-FR40	23	23	23
j Flat	24	23	25
k VI60"-FR40	41	36	39
l Flat	23	21	23

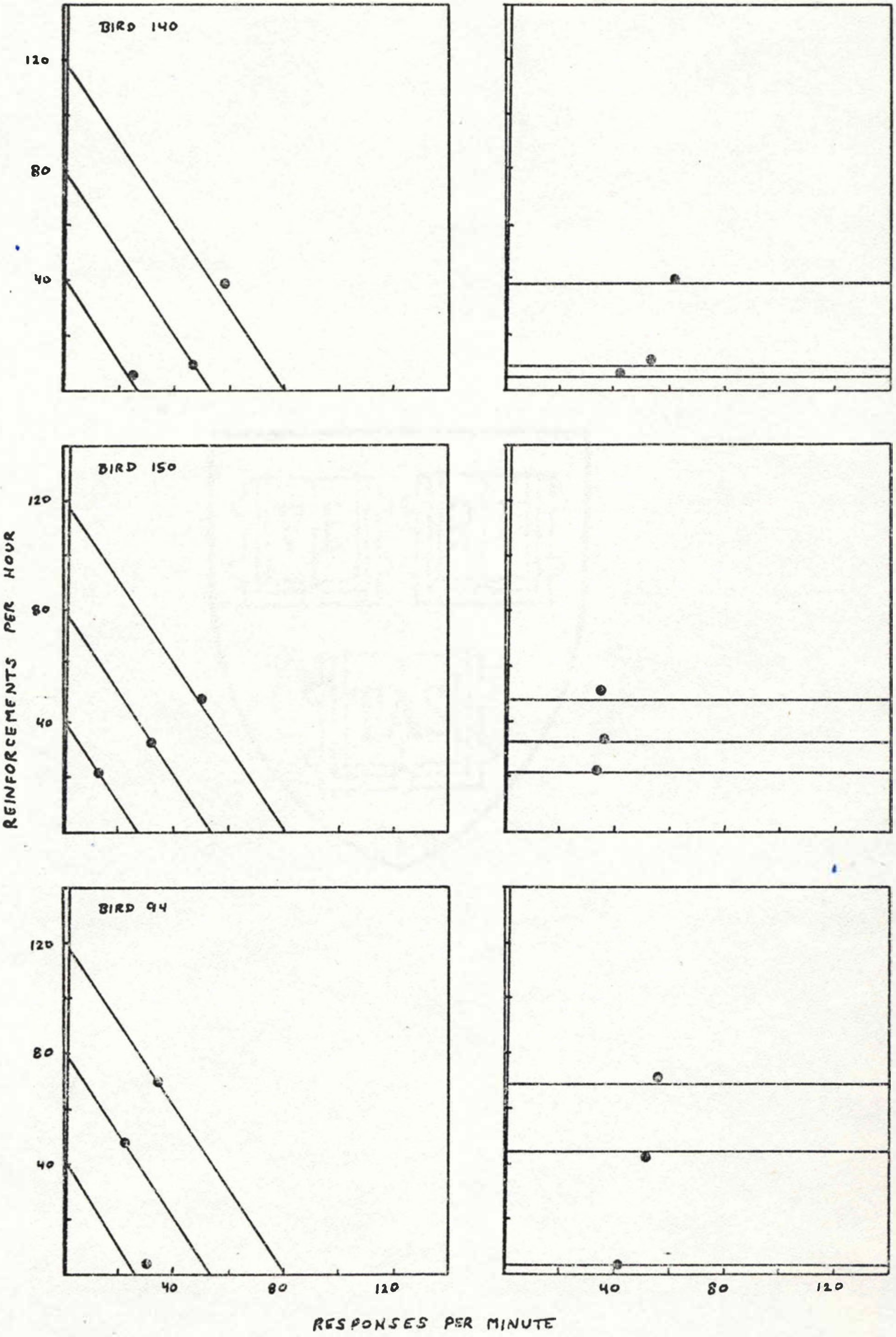
Table 1 (cont.)

	Bird 52	Bird 279	Bird 95
m VI30"-FR60	71	66	70
n Flat	29	34	34
o VI45"-FR60	23	23	28
p Flat	27	22	21
q VI60"-FR60	36	36	32
r Flat	21	22	23

Figure 9. Left columns: behavior under linear schedules with negative slopes. Right columns: Behavior under linear schedules with flat slopes, with reinforcement rate equated to that received under negative slopes.



BIRD 19
CRF
120
80
40
BIRD 47
120
80
40
BIRD 93
120
80
40
40
80
120
40
80
120
REINFORCEMENTS PER HOUR
RESPONSES PER MINUTE


BIRD 140
120
80
40
BIRD 150
120
80
40
BIRD 94
120
80
40
40
80
120
40
80
120
REINFORCEMENTS PER HOUR
RESPONSES PER MINUTE

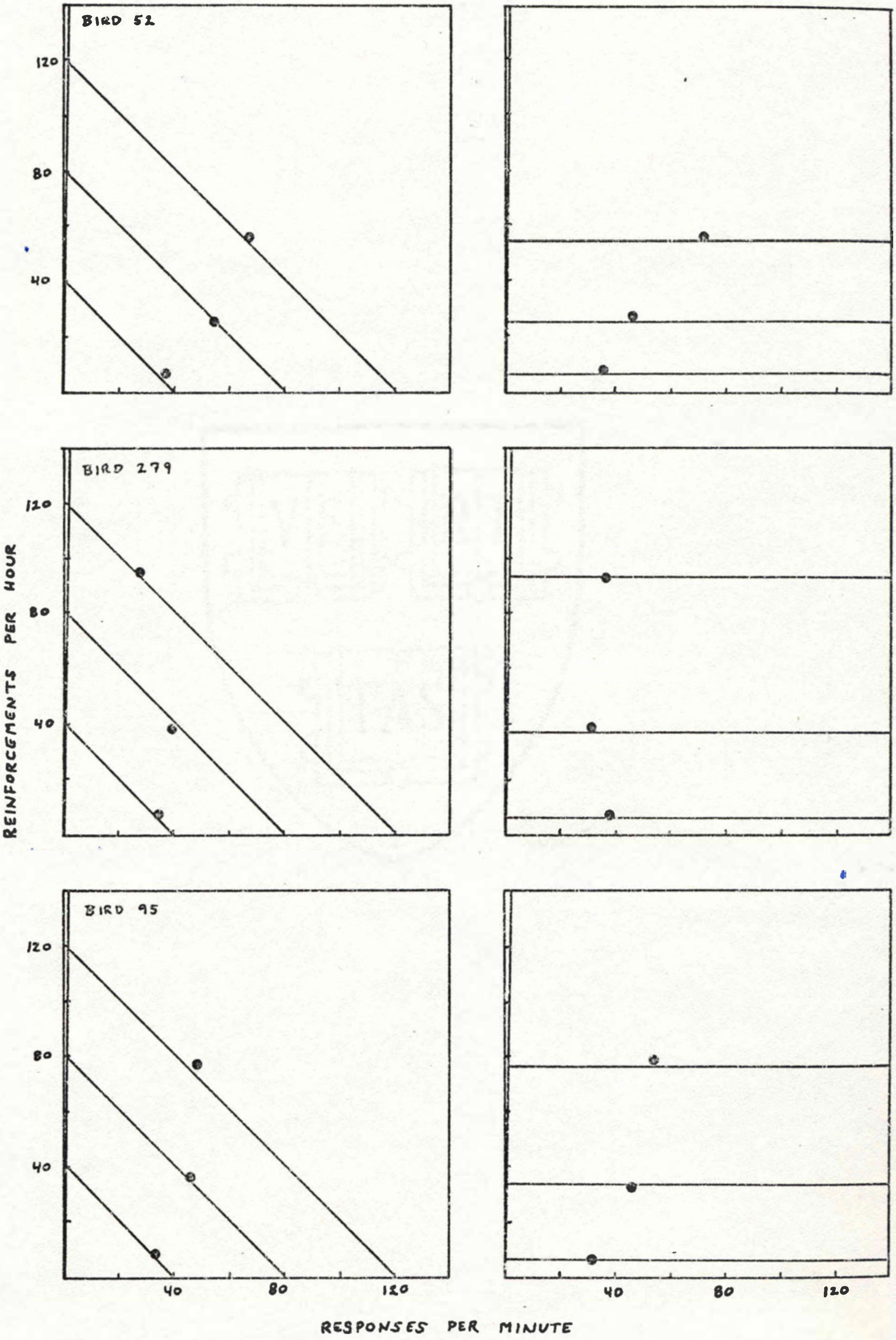

BIRD 52
BIRD 279
BIRD 95
120
80
40
40
80
120
REINFORCEMENTS PER HOUR
RESPONSES PER MINUTE

Table 2 about here

---

days of each bird on each condition. In Figure 10 the average change in response rate and reinforcement rate is shown for each bird, as a function of intercept of

---

Figure 10 about here

---

the schedule. Applying a Wilcoxon matched-pairs signed-ranks test (Siegel, 1956) to each group of birds with the same slope, only those birds with an FR 40 subtracted show a significant change, an increase in response rate with the slope going from negative to zero (n=9; 1 tail, $p<.025$; 2 tail, $p<.05$). With all birds lumped together, by the same test there is a significant increase in response rate going from negative to zero slopes (n=26; 1 tail, $p<.025$; 2 tail, $p<.05$). Thus we can reject the possibility that a significant decrease occurred, which is the only result that would be damaging to the present position. Had such a significant decrease occurred, it would imply an increase in value as response rate decreased and reinforcement rate remained constant. However, a number of experiments suggest that with response rate

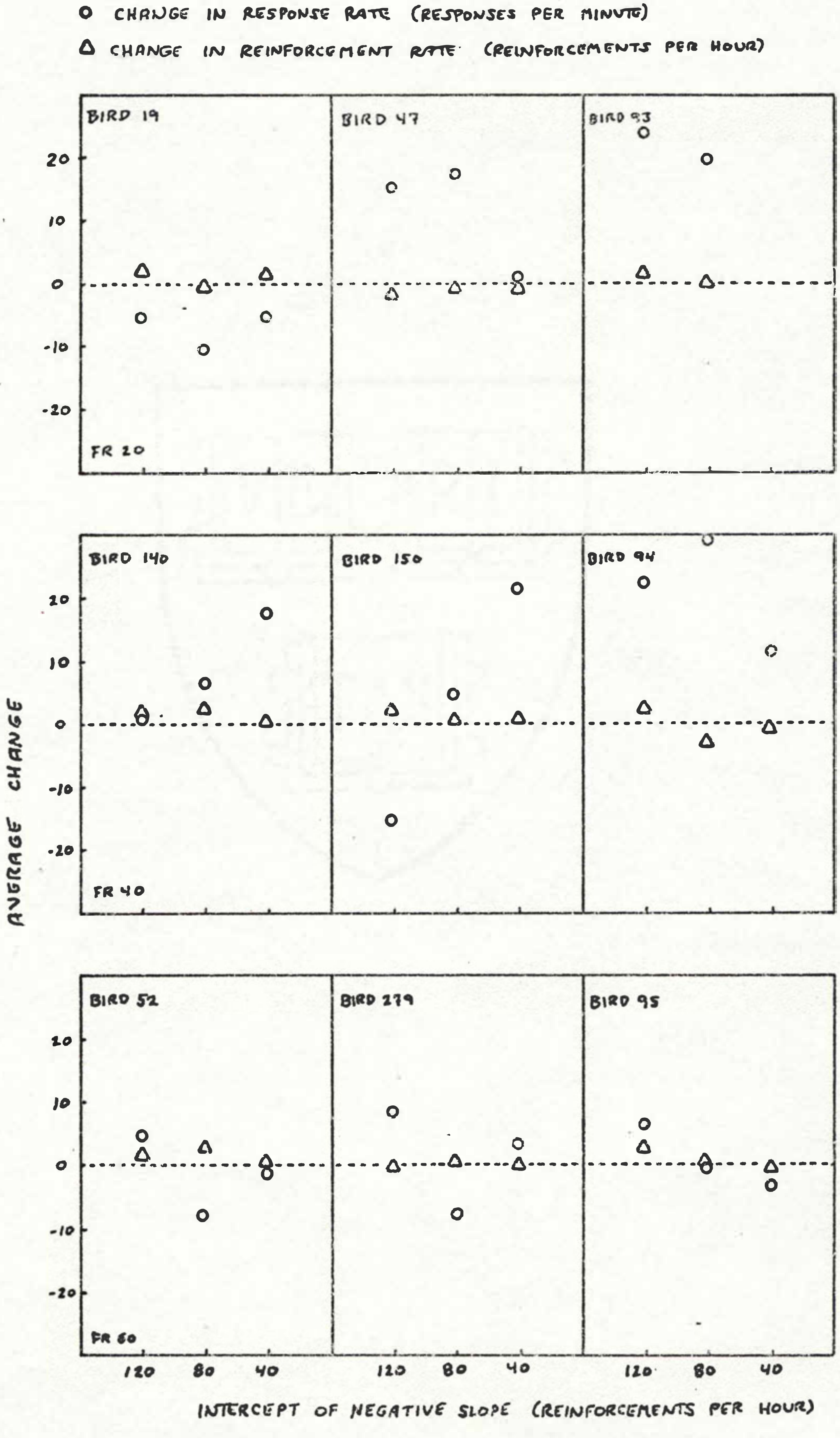
Table 2. Response rates, reinforcement rates, and session times from last five days of each condition in Experiment 1. P/M: Pecks per minute. R/H: Reinforcements per hour. M: Session time in minutes.

	P/M	R/H	M	P/M	R/H	M	P/M	R/H	M
		Bird 19			Bird 47			Bird 93	
a	31.9	27.2	200.95	39.3	7.1	203.06	39.4	7.4	202.70
b	26.6	29.2	199.07	54.2	5.4	200.55	63.4	9.0	200.26
c	24.2	8.8	203.52	26.9	2.9	203.55	26.5	2.6	203.99
d	14.0	8.1	200.34	44.0	2.1	201.00	45.8	2.7	200.81
e	12.8	3.6	202.73	14.4	2.4	203.14	13.9	3.9	201.67
f	7.5	4.8	201.49	15.4	1.5	200.94	-	-	-
		Bird 140			Bird 150			Bird 94	
g	58.8	38.5	200.89	51.6	48.0	200.10	34.9	69.2	173.45
h	59.4	39.8	198.79	36.2	50.2	198.19	57.0	71.5	167.87
i	47.9	9.1	204.01	32.0	32.6	202.24	22.8	44.9	201.95
j	54.3	11.4	200.23	36.4	33.1	199.12	52.5	42.2	198.79
k	25.3	4.7	202.57	12.9	21.4	201.36	30.8	3.6	202.65
l	42.8	5.0	201.59	34.2	21.9	200.18	42.0	2.7	201.28

Table 2 (cont.)

	P/M	R/H	M	P/M	R/H	M	P/M	R/H	M
	Bird 52			Bird 279			Bird 95		
m	67.6	54.2	200.30	28.0	93.0	129.00	48.3	76.8	156.29
n	72.4	55.7	198.14	36.9	92.8	129.23	54.8	79.5	150.99
o	55.4	25.4	203.15	40.6	37.7	202.28	47.8	34.0	201.47
p	47.4	28.0	199.53	32.7	38.3	198.94	46.9	33.7	199.53
q	38.0	6.2	202.09	35.3	6.8	201.89	35.4	6.8	202.09
r	36.1	6.6	201.13	39.0	6.8	201.20	32.2	6.3	201.00

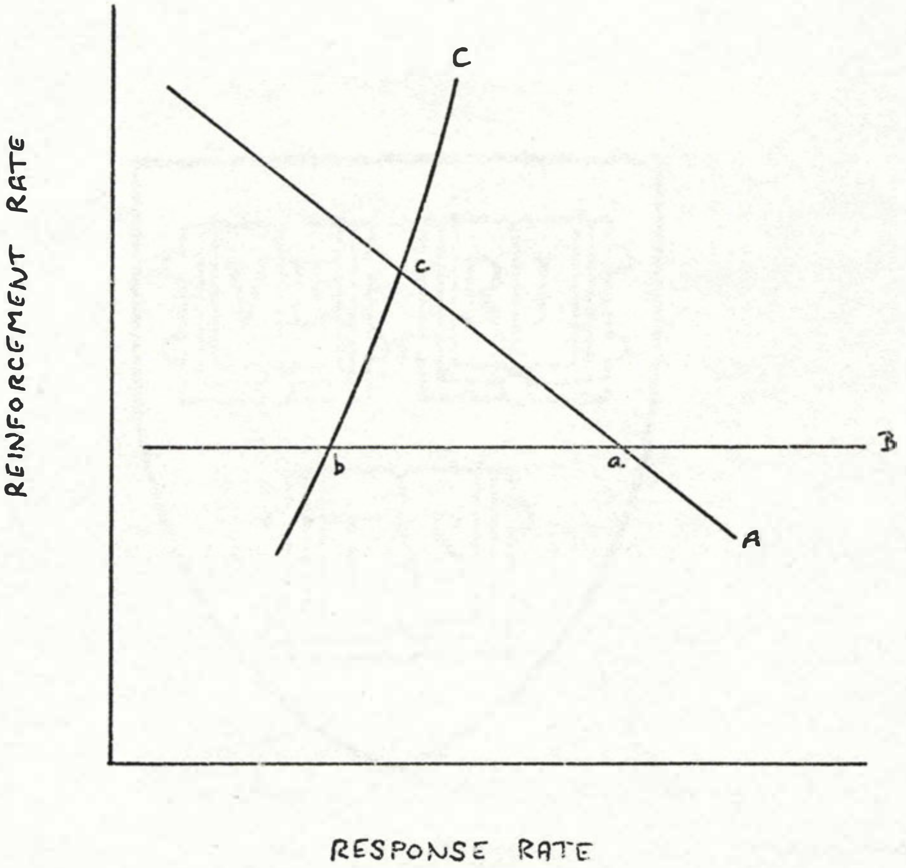
Figure 10. Average change in rate of responding and rate of reinforcement going from negative slopes to flat slopes.


○ CHANGE IN RESPONSE RATE (RESPONSES PER MINUTE)
△ CHANGE IN REINFORCEMENT RATE (REINFORCEMENTS PER HOUR)
BIRD 19
BIRD 47
BIRD 93
FR 20
BIRD 140
BIRD 150
BIRD 94
FR 40
BIRD 52
BIRD 279
BIRD 95
FR 60
20
10
0
-10
-20
AVERAGE CHANGE
120
80
40
INTERCEPT OF NEGATIVE SLOPE (REINFORCEMENTS PER HOUR)

constant or allowed to increase, an increase in reinforcement rate produces an increase in value. This in turn implies that on the negative slope, a point of higher value existed than the one where behavior was stable (see Figure 11). While this is not impossible, it would imply a somewhat irregular response surface which in turn would be a somewhat <u>ad hoc</u> explanation of the result. A larger effect might have been found, were it not for the following procedural difficulty. At the beginning of each session with negative slopes, the bidirectional stepper started in state 0, so that the next hole read in the VI tape reader could produce reinforcement. During most of the session the stepper was in a lower state. Presumably because of the higher probability of reinforcement at the beginning of the session, response rates at that time tended to be higher than later in the session (cf., Catania & Reynolds, 1968). If the stepper could have been programmed to start where it had ended the previous day, this transient higher rate would probably have been eliminated.

Figure 11 about here

Figure 11. Suppose that on schedule A behavior is stable at point a, and on B is stable at point b. Under the assumption that points on C that are higher in terms of reinforcement rate are also higher in terms of value, it can be concluded that some point on A (point c) other than that at which behavior was stable has greater value.


REINFORCEMENT RATE
RESPONSE RATE
C
c
B
b
a
A

Discussion

In all cases, birds did not behave in a way which maximized the rate at which grain was received. The results from this experiment are thus in rough agreement with those from studies on DRL. The results appear to be consistent with an interpretation in terms of a response surface, with maximization of value determining stability. From that perspective, what these results suggest is that, if reinforcers are being delivered according to a flat function, the slope of the local contour line is zero. However, if response rate decreases, the slope of the local contour becomes negative rather quickly, which is equivalent to saying that a large increase in reinforcement rate is necessary in order to significantly reduce response rate. This implies that Staddon's (1972) statement about a "natural" rate of pecking should be modified: for each tradeoff between response rate and reinforcement rate (slope of reinforcement function), and each intercept of the reinforcement function, some rate of responding is "natural" in the sense that it maximizes value. Presumably a different surface would be obtained if another response, such as treadle pressing, were used (cf., Hemmes, 1975).

Experiment 2: Functions with Positive Slope

Before interval and ratio schedules can be meaningfully thought of as elements from a larger space of schedules, it should be shown that schedules from the larger space, other than interval and ratio schedules, can produce behavior that could not be produced by interval or ratio schedules. In other words, if interval and ratio schedules can give us a complete picture of an organism's behavior, from the point of view here under consideration, there is no need to expand the number of schedules under which an organism is to be studied. The results of Experiment 1 suggest that, for a given rate of reinforcement, a function with negative slope can maintain a lower rate of responding than one with zero slope. In this experiment, a zero slope is compared with two positive slopes.

Method

Subjects

Four White Carneaux and two homing pigeons were used; all six birds ran in the same conditions. They had had prior experimental histories, and were maintained at about 80% of free feeding weight.

Apparatus

A standard pigeon chamber was used. The front wall was 12 in. (30 cm) wide, with a key about 10 in. (25 cm) above the floor, 5.5 in. (14 cm) from the right wall; directly below it was a standard grain magazine. A force of about 14 g (.14 N) was required to operate the key, which was transilluminated with two 7-w green bulbs. An auditory feedback click was provided for each response. The chamber was illuminated with two 7-w white bulbs, except during reinforcement, when only the magazine was illuminated. White noise masked extraneous sounds. A PDP-8/e computer controlled the experiment and collected data.

Procedure

Birds were first run on a schedule with a flat slope which provided 60 reinforcements per hour, with variable intervals from Fleshler and Hoffman (1962). After their behavior stabilized, they were run on a VR 100, with either a VI schedule added, or an FI schedule subtracted, so as to maintain reinforcement rate at approximately what had been obtained in the first condition. Parameters were adjusted approximately every two to three days when deviations from the appropriate reinforcement rate occurred. Finally, they were run on VR 200 with the same contingencies in effect. Reinforcement consisted of 3 sec presentation of mixed grain. Following reinforcement, a minimum of

5 sec had to elapse before reinforcement could again be obtained. Sessions ran until 40 reinforcements were presented or 40 min elapsed, whichever occurred first. Sessions were run seven days a week. Each bird was run until all birds had run at least 30 sessions and there were five consequetive days that appeared stable. The conditions and the number of sessions for each bird are shown in Table 3.

Table 3 about here

Results

Looking at the last ten days of all birds for the two conditions which involved changes of parameters (conditions with a VR component), there were a total of eight changes of parameters out of these 12 cases. This is evidence that by the end of these conditions, day to day reinforcement rates were essentially stable. In Figure 12 are shown the schedules and the points showing average response rates and reinforcement rates for the last five days.

Figure 12 about here

In Table 4 are shown

Table 3. Conditions in Experiment 2 and number of sessions each bird was run on each condition. Both VR 100 and VR 200 conditions had, in addition, interval schedules added or subtracted.

Schedules	Subjects		
	Bird 9	Bird 53	Bird 96
a VI 60"	45	45	45
b VR 200	44	46	48
c VR 100	54	51	51
	Bird 55	Bird 83	Bird 97
a VI 60"	45	45	45
b VR 200	48	48	46
c VR 100	51	52	49

Figure 12. Rates of responding and reinforcement under schedules with (a) flat slope, (b) slope determined by VR 200, and (c) slope determined by VR 100.

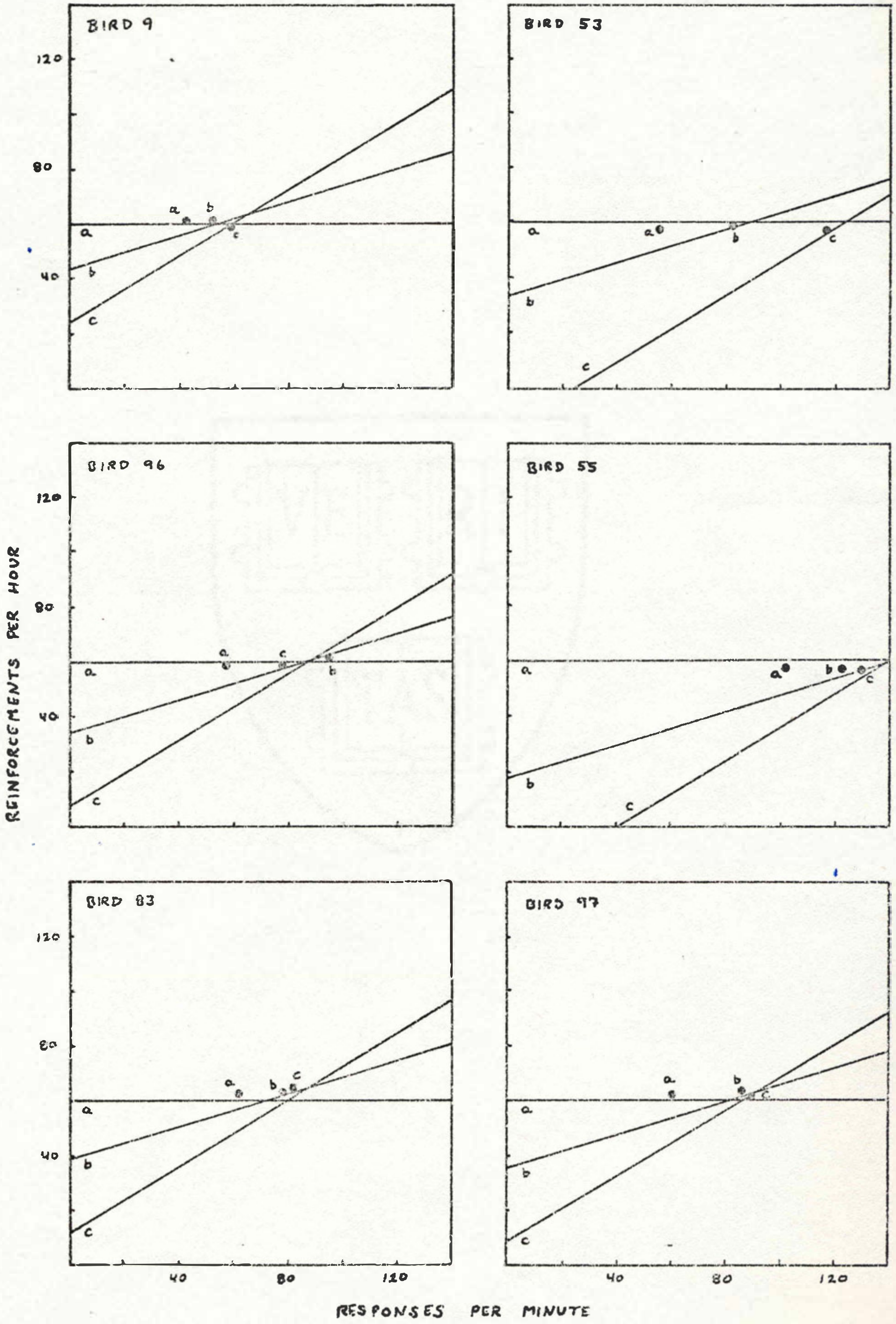

BIRD 9
BIRD 53
BIRD 96
BIRD 55
BIRD 83
BIRD 97
a
b
c
120
80
40
40
80
120
REINFORCEMENTS PER HOUR
RESPONSES PER MINUTE

Table 4 about here

----

the overall response rates, reinforcement rates, and session times for the last five days for each bird. These data are reploted in Figure 13 to emphasize rate information. It can be seen that in five cases out of six there is a monotonic increasing relation between slope of the reinforcement function and rate of responding. A Wilcoxon matched-pairs signed-ranks test was done on the differences in response rates between VI and VR 200, and VR 200 and VR 100 conditions. Overall, these differences showed a significant increase (n=12; 1 tail, p<.005; 2 tail, p<.01).

Figure 13 about here

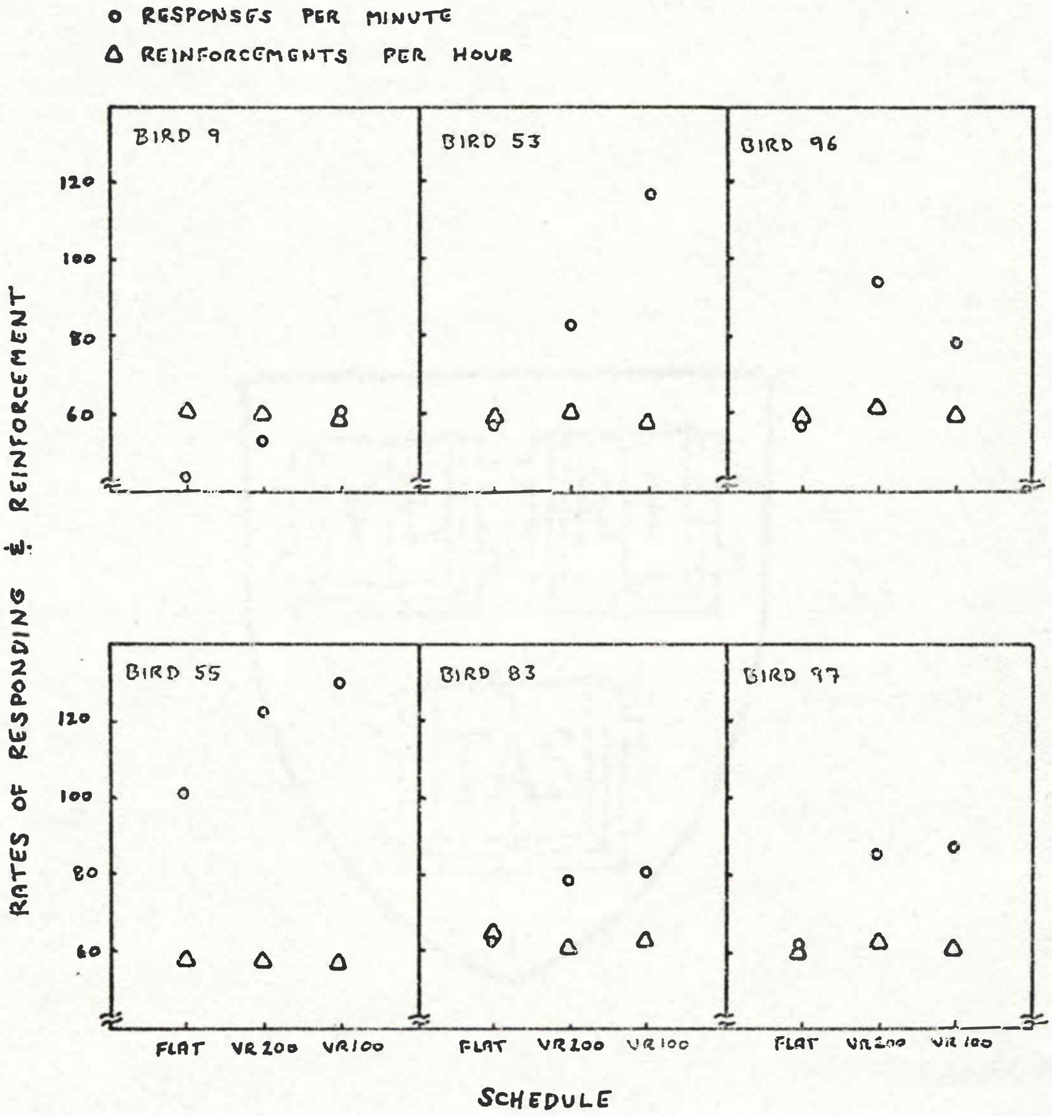
Discussion

These results are consistent with the concept of a response surface with contours that, moving to the right, pass through zero and increase. The results appear inconsistent with the view that interval and ratio schedules are two disparate cases; rather, it appears possible to go continuously from one to the other and produce a continuous change in behavior.

Table 4. Response rates, reinforcement rates, and session times from last five days of each condition in Experiment 2. P/M: Pecks per minute. R/H: Reinforcements per hour. M: Session time in minutes.

	P/M	R/H	M	P/M	R/H	M	P/M	R/H	M
	Bird 9			Bird 53			Bird 96		
a	43.1	60.2	194.2	56.5	58.1	199.2	57.7	58.9	195.6
b	52.4	60.0	193.9	83.6	59.1	197.9	94.3	61.3	191.8
c	60.0	59.6	195.2	117.0	57.4	199.7	78.2	59.1	196.8
	Bird 55			Bird 83			Bird 97		
a	101.9	57.2	197.0	62.3	62.6	189.6	60.5	61.9	189.9
b	122.9	57.7	196.6	78.7	62.1	187.4	85.7	61.4	191.6
c	130.0	56.7	199.9	80.3	63.1	189.2	86.9	60.4	188.6

Figure 13. Rates of responding and reinforcement as a function of schedules with different slopes.


o RESPONSES PER MINUTE
Δ REINFORCEMENTS PER HOUR
BIRD 9
BIRD 53
BIRD 96
BIRD 55
BIRD 83
BIRD 97
120
100
80
60
RATES OF RESPONDING & REINFORCEMENT
FLAT
VR 200
VR 100
SCHEDULE

It could be argued that behavior on the sum of interval and ratio schedules simply reflects a conjunction of the effects of these two schedules. For example, the interval component might reinforce long IRTs, and the ratio component short IRTs. However, the present approach is not intended to disprove such a possibility. Rather, it is designed to provide a coherent summary of those effects that are found.

Concurrent Schedules

On concurrent schedules, two basic distributions of behavior have been reported. On concurrent VI schedules, a matching relation between relative response rate and relative reinforcement rate is often found (Herrnstein, 1961, 1970). On the other hand, on concurrent VI schedules with only one tape reader running (Findley, 1958), on concurrent VR schedules (Herrnstein & Loveland, 1975), and on the similar discrete trials probabilistic reinforcement where every response has a nonzero probability of reinforcement (Shimp, 1966), nearly exclusive preference for the side providing the higher rate of reinforcement (given that they are different) is found. Later the question of whether this is also matching will be brought up.

There exist a number of explanations of these results. Herrnstein (1970) argues that the behavior of matching is the expression of a matching law which is an inherent part of an organism. This view may to some extent be characterized as a denial that matching is the outcome of other processes. For example, Shimp (1966, 1969) argues that matching is the result of an organism's emitting that response which has the higher probability of reinforcement at the time; this process is referred to as <u>momentary maximizing</u>. According to

Shimp, such a process in turn implies global maximizing: the overall rate of reinforcement is greatest if momentary maximizing is followed. Shimp holds that this process can account for both probabilistically reinforced choices and concurrent VI schedules.

A third possibility, which is perhaps closer to the position of Baum (1973) than anyone else, would view matching as maximization of global reinforcement rate, but in a way that was not as tightly constrained on a molecular level as Shimp suggests is the case. For example, behavioral variability could lead to oscillations around matching; deviations from matching, resulting in a lowered overall rate of reinforcement, could move the distribution of responses in the other direction. This position may be called _global maximization_.

A fourth explanation is based on the fact that animals tend to match both time and response ratios (Catania, 1966). This in turn implies that the local rates of reinforcement and responding are equal for the different alternatives. Rachlin (1973) imputes to Killeen (1972) the view that responding in the presence of a lower rate of reinforcement becomes aversive, causing a bird to switch to a key with a higher rate of reinforcement. This behavioral mechanism, coupled with concurrent VI schedules, in which the longer one spends on a side the lower the local rate of reinforcement, is sufficient to predict time and response matching. We may refer to this position as <u>local maximization</u>. Rachlin then points out that data of Nevin (1969) showing a decreased probability of changing over as more responses are emitted on one side goes against Killeen´s position; Rachlin himself then adopts the position that matching on a local level, so that the relative rates of responding and reinforcement are .5, is basic. This position does not substantially differ from that of Herrnstein; only a different time base is involved.

Let us see how these different positions relate to the data. An account in terms of matching can account well for behavior on concurrent VI schedules. However, there is a limitation. If a bird put all of its responses on one side, thereby earning all reinforcements from that side, matching would still be preserved. Thus matching does not exclude exclusive preference on concurrent VI schedules. In the absence of knowledge in regard to how matching comes about in a situation in which an organism is not originally matching, one might expect matching by means of exclusive preference to be as common as matching with behavior distributed between the two alternatives. On concurrent VR schedules matching predicts exclusive preference if the schedules differ, but does not say which of the two schedules will be preferred. Herrnstein and Loveland (1975) appear to conjoin matching with a global maximization process to account for the nearly exclusive preference for the better of two VR schedules which they find. The same considerations apply to Findley's (1958) finding of exclusive preference for the better of two VI schedules when only one tape runs at a time. Thus, although matching is consistent with these data, it is also consistent with possibilities which do not obtain.

Global maximization, in the sense of maximizing rate of grain presentation, is inconsistent with the results of Experiment 1 with negative slopes. In that experiment, when run on negative slopes, response rates were sufficiently high to keep reinforcement rates much lower than they could have been (see Figure 9). Thus this principle cannot constitute a general account of behavior. In the case of ratio schedules, or more generally linear schedules with positive slope, both momentary maximizing and global maximization can avoid the prediction of infinite response rates only on an ad hoc basis. We will return to these positions when discussing data from some concurrent experiments below.

A local maximization approach may be modified in the following direction. We may think of a bird as distributing its time between two alternatives in some fashion. If the time spent on one alternative results in a higher rate of reinforcement while on that side, that should result in an increase in time spent on that side, simply by virtue of a higher rate of reinforcement being obtained on that side. As Herrnstein and Loveland (1975) state: "It is axiomatic that, given incompatible responses that differ in reinforcement and given that the difference has been detected, a subject will choose the more highly reinforced alternative" (pp. 113-114). However, as Rachlin (1973) points out, by putting more time on the locally better side on concurrent VI schedules, the local rate of reinforcement on that side decreases and increases on the other.

The only mechanism it is necessary to postulate is one that will shunt behavior in a continuous manner from a locally poorer alternative to a locally richer alternative. On concurrent VI schedules, this mechanism would damp out deviations from matching because they result in an increase in disparity between local reinforcement rates, in a direction tending to drive the distribution of behavior back toward matching. On concurrent VR schedules, and concurrent VI with one tape reader moving, the same mechanism will tend to drive behavior toward exclusive preference for the better side.

Previously, it was suggested that if while responding on one key, if one rate of responding has greater value than another, relatively more time than before will be spent responding at the rate with the greater value. The schedule may or may not change as more time is spent at a particular rate of responding. On concurrent schedules, we may consider value, rather than simply reinforcement rate, to affect distribution of behavior. If, while responding on one side, greater value is produced per unit time than is obtained on the other side, the same mechanism that adjusts response rate on a single key will distribute more time to the better side. This suggests a hierarchical organization of behavior. On one level (e.g., while responding on a key) choice, or rate or responding, is governed by value deriving from such responding. On a higher level (e.g., choosing between two keys) choice is governed by a comparison of the values resulting from responding on each of the keys. However, another possibility also exists. Suppose a bird responds on a single key at two rates in alternation, each giving rise to a different value. More time should subsequently be spent at the rate with higher value. It appears plausible that it may make little difference whether the two rates occur on a single key or whether they are on separate keys. In either case, the identical operation appears to be in effect.

Experiment 3: Concurrent Flat Slopes

On a single VI schedule as usually programmed, reinforcement rate is affected to some extent by response rate, since after a reinforcer is made available, but before it is collected, the tape reader is stopped. On concurrent VI schedules we may expect such an effect to be amplified, since reinforcement may set up while responding is occurring on the other key; in addition, a COD may lengthen the time during which a tape reader does not operate. With a long COD, for example, Shull and Pliskoff (1967) found the obtained relative rates of reinforcement to differ substantially from the programmed relative rates. Stubbs and Pliskoff (1969) used a procedure that controlled relative rates of reinforcement, but absolute rates could still be affected by behavior. Since, whenever one tape reader sets up both readers stop, it is plausible that overall reinforcement rate will be more affected by behavior than if the tapes moved independently.

If a linear schedule with flat slope is used, rather than a normal VI schedule, good control is obtained over absolute reinforcement rate, as long as responses occur at least as often as reinforcement sets up. Used in a concurrent situation, this means that both absolute and relative reinforcement rates are under greater experimental control than is produced by other procedures. An implication of this is that, given at least some minimum response rate on each side, practically any distribution of responses will earn the programmed rate of reinforcement. This procedure thus excludes the possibility that only at matching is the global rate of reinforcement maximized. In this experiment, run with H. L. Miller Jr. and D. H. Loveland, concurrent flat slopes were studied.

Method

Subjects

Four White Carneaux pigeons with prior experimental histories were used. They were maintained at about 80% of free feeding weight.

Apparatus

A standard pigeon chamber with two keys was used. The two keys were at the same level, 5 in. (13 cm) apart, centered on a wall 11.5 in. (29 cm) wide, about 9 in. (23 cm) from the floor. A standard feeder was centered on the wall below the keys. A force of about 14 g (.14 N) was required to operate the keys. The left key was transilluminated with two 7-w green bulbs, the right key with red bulbs. The chamber was illuminated, except during reinforcement, with two 7-w white bulbs; during reinforcement only the hopper was illuminated. White noise and a fan masked extraneous noise. A PDP-8/e computer controlled the experiment, in conjunction with external tape readers.

Procedure

Five conditions were run, in each of which a total of 90 reinforcements per hour were programmed. Two VI tape readers with 12 intervals generated from Fleshler and Hoffman (1962) ran continuously during the session, except during reinforcement. Uncollected reinforcers were stored by computer in a simulated bidirectional stepper. The VI values used, on left and right keys respectively, were, in this order: VI 120" and VI 60"; VI 60" and VI 120"; VI 48" and VI 240"; VI 80" and VI 80"; VI 240" and VI 48". A 3 sec COD was in effect; in addition, a minimum of 3 sec had to elapse before a second reinforcement (or more) could be collected from one side. Sessions terminated with the presentation of 40 reinforcements, and were run seven days a week. In addition to response and reinforcement count, distributions of the number of responses to a side before changing over were collected. Sessions were run until all birds appeared stable. In Table 5 are shown the conditions and the number of sessions each bird

---

Table 5

---

was run on each condition.

Results

Figure 14 shows, for each bird, logged ratios of responses as a function of logged ratios of reinforcers received, along with a least-squares linear fit. These ratios were calculated from the average response and reinforcement rates for the last five days. In general, response ratios exhibited some undermatching. The 95% confidence intervals for the slopes are also shown. For two out of four birds this range includes a slope of one. In Table 6 are shown overall response rates to the two keys, reinforcement rates, and session times from the last five days for each bird. Figure 15 shows the data from all birds, with each bird´s data

---

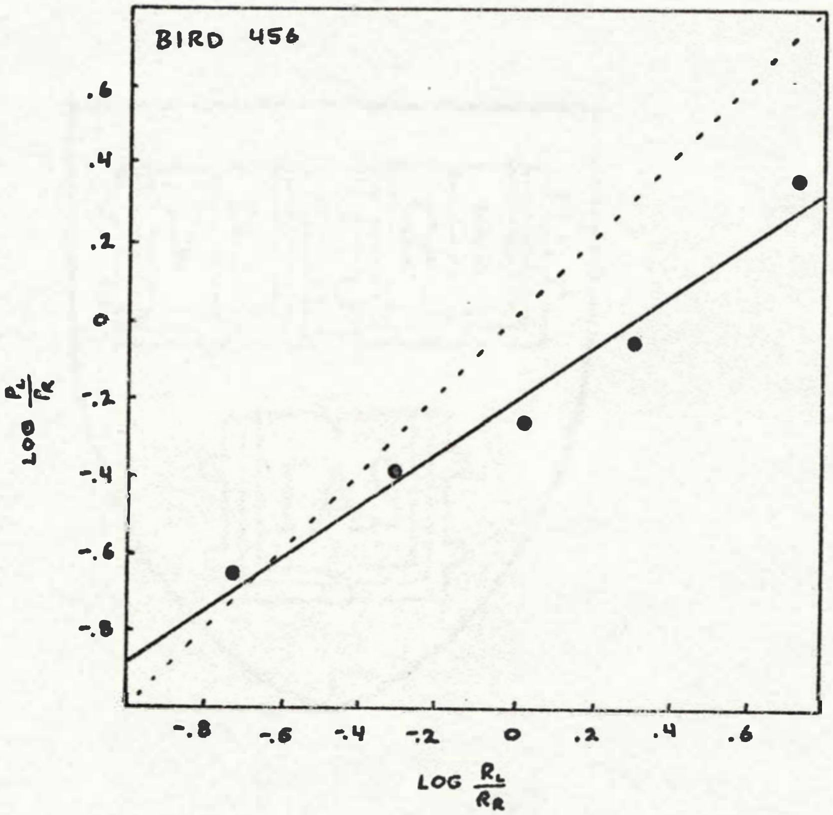
Figure 14 about here

Table 6 about here

Table 5. Conditions in Experiment 3 and number of sessions each bird was run on each condition.

VI values	Bird 456	Bird 475	Bird 468	Bird 300
a 120" 60"	25	25	25	25
b 60" 120"	37	37	36	36
c 48" 240"	35	35	34	32
d 80" 80"	22	25	24	20
e 240" 48"	24	28	31	35

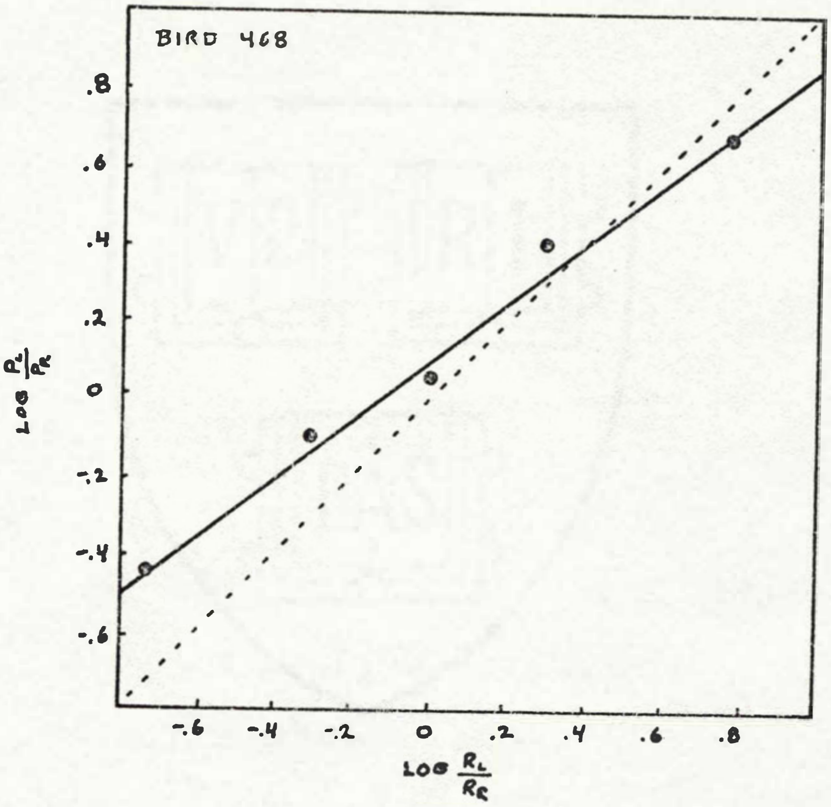
Figure 14. Logged ratio of responses to two keys as a function of logged ratio of reinforcers received on those keys. P_L, P_R: total pecks left and right. R_L, R_R: total reinforcers left and right. The heavy line indicates the least-squares linear fit, whose equation appears below the graph. The dotted line indicates matching. The percentage of variance accounted for by the linear equation, and the 95% confidence interval of the slope are indicated.



LOG $\frac{P_L}{P_R}$ = .67 · LOG $\frac{R_L}{R_R}$ −.21

$r^2 = .97$

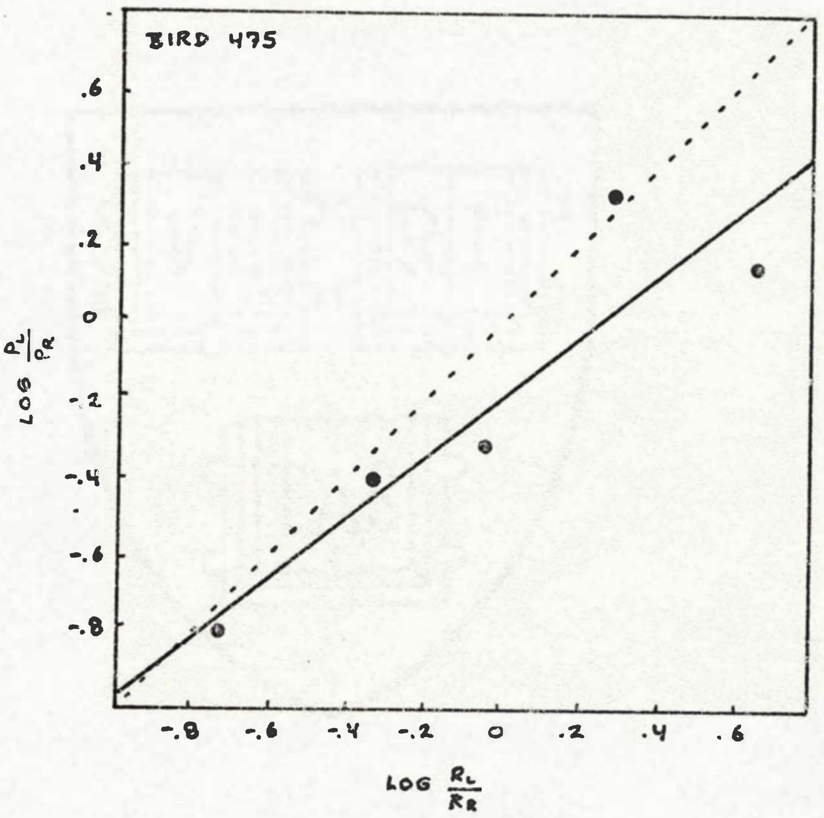
95% CONFIDENCE INTERVAL: .46 - .88



$$\log \frac{P_L}{P_R} = .76 \cdot \log \frac{R_L}{R_R} + .11$$

$r^2 = .99$

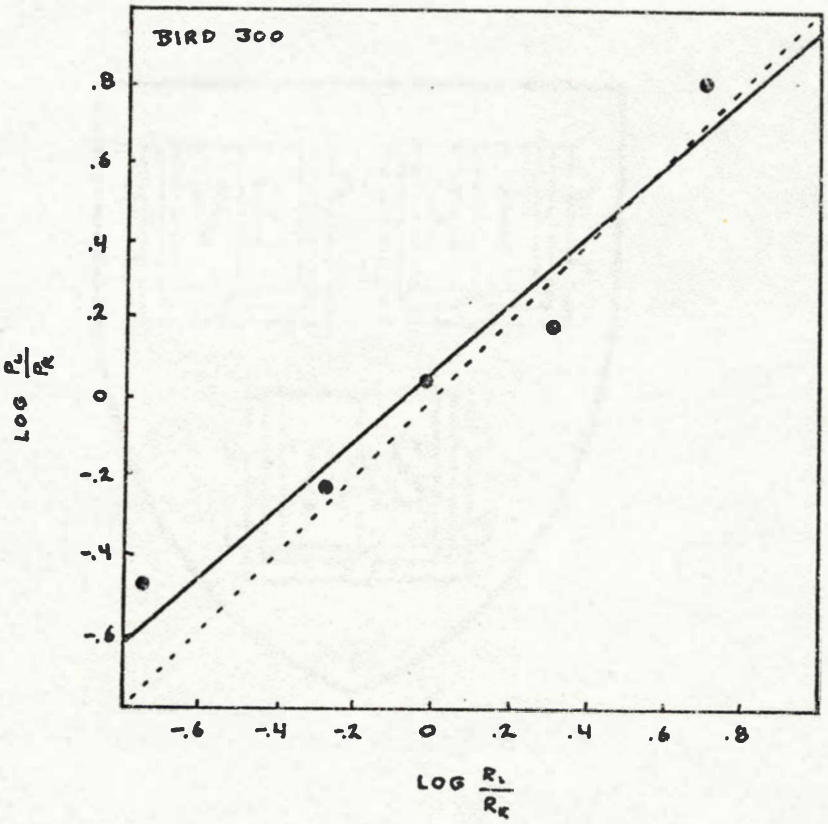
95% CONFIDENCE INTERVAL: .60 - .91



$$\text{LOG}\ \frac{P_L}{P_R} = .77 \cdot \text{LOG}\ \frac{R_L}{R_R} - .19$$

$r^2 = .84$

95% CONFIDENCE INTERVAL: .16 - 1.38



$$\text{LOG } \frac{P_L}{P_R} = .88 \cdot \text{LOG } \frac{R_L}{R_R} + .07$$

$r^2 = .94$

95% CONFIDENCE INTERVAL: .49- 1.27

Table 6. Overall response rates and reinforcement rates to each key, and session times from last five days of each condition in Experiment 3. P/M: Pecks per minute. R/H: Reinforcements per hour. M: Session time in minutes. L,R: Left, right keys.

	P/M		R/H		M
	L	R	L	R	
			Bird 456		
a	17.3	42.4	31.1	61.7	129.40
b	21.7	24.8	62.6	30.8	128.34
c	30.0	13.1	80.6	14.8	125.81
d	19.8	36.4	48.0	45.2	128.81
e	9.0	40.8	15.0	78.6	128.29
			Bird 475		
a	14.6	36.4	28.8	61.3	133.18
b	40.8	19.2	60.9	30.7	131.02
c	16.0	11.5	76.8	16.6	129.67
d	10.8	21.8	44.7	48.4	128.95
e	4.9	31.2	14.6	76.9	131.05

Table 6 (cont.)

	P/M		R/H		M
	L	R	L	R	
Bird 468					
a	37.5	46.4	30.0	60.8	132.21
b	42.6	17.0	60.4	31.1	131.19
c	52.0	11.1	79.9	13.5	128.47
d	45.0	38.9	44.8	44.3	129.90
e	19.5	54.9	15.1	79.3	127.08
Bird 300					
a	36.3	61.1	32.3	58.7	131.82
b	51.0	32.6	60.5	28.9	134.81
c	57.7	8.7	74.7	14.8	134.12
d	41.3	36.6	44.7	45.6	132.96
e	18.0	56.0	14.0	76.3	132.89

Figure 15 about here

---

adjusted so that its linear regression function passes through the origin. The 95% confidence interval does not include a slope of one. In Figure 16 are plotted changeovers per opportunity for Bird 300, whose

---

Figure 16 about here

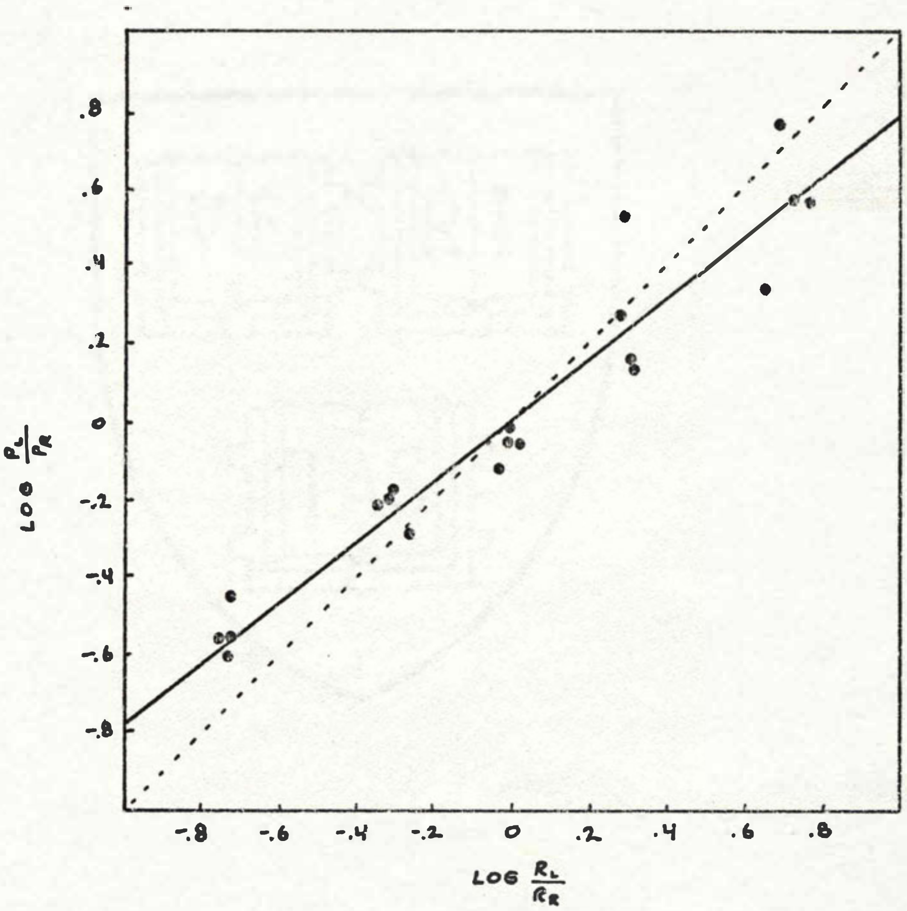
---

behavior was closest to matching. These points were calculated from the distribution of number of responses before changing over in a manner analogous to that used by Anger (1956) to calculate IRT per opportunity. These graphs thus give the probabilitty of changing over as response count on a side increases.

Discussion

Although there are deviations from matching, the closeness to matching that is found argues against an explanation of matching in terms of maximization of global reinforcement rate. If such an explanation were true, for any given ratio of reinforcements we should expect to find a ratio of responses that was independent of the reinforcement ratio; these results suggest a strong functional relation. A t-test to

Figure 15. Logged ratio of responses for all four birds as a function of logged ratios of reinforcers received. The intercepts for each bird have been adjusted to pass through the origin.

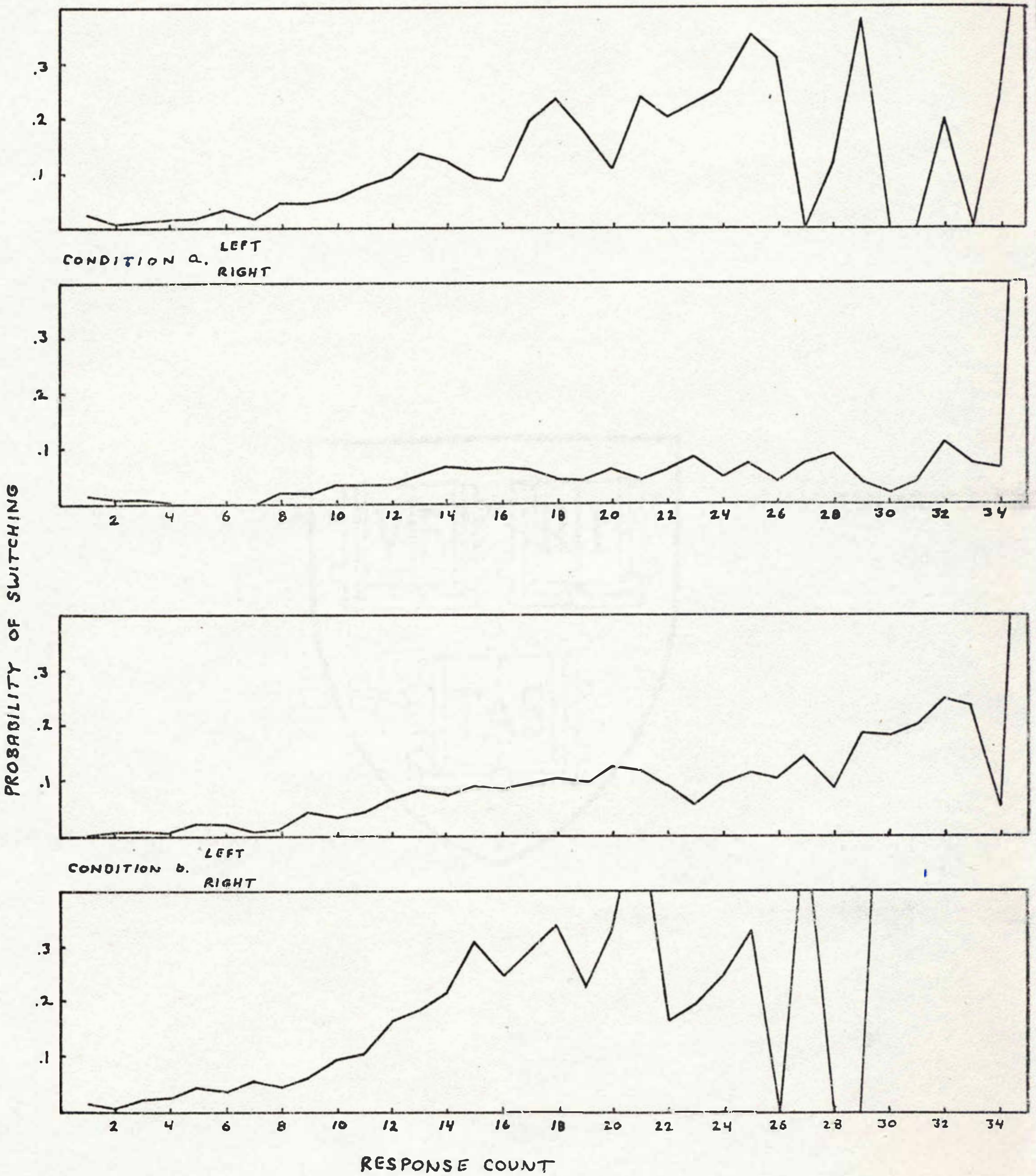


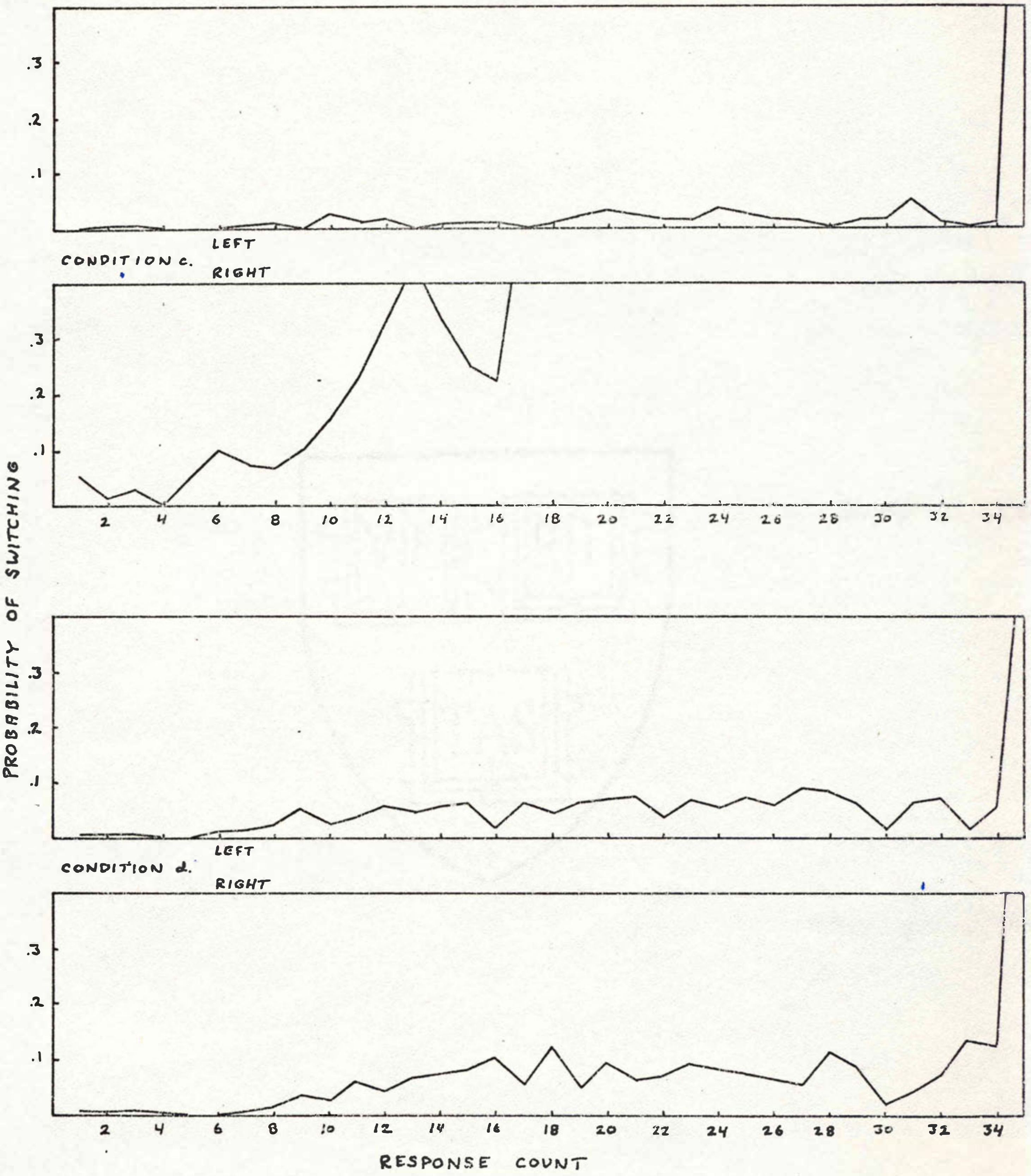
$$\text{LOG } \frac{P_L}{P_R} = .77 \cdot \text{LOG } \frac{R_L}{R_R}$$

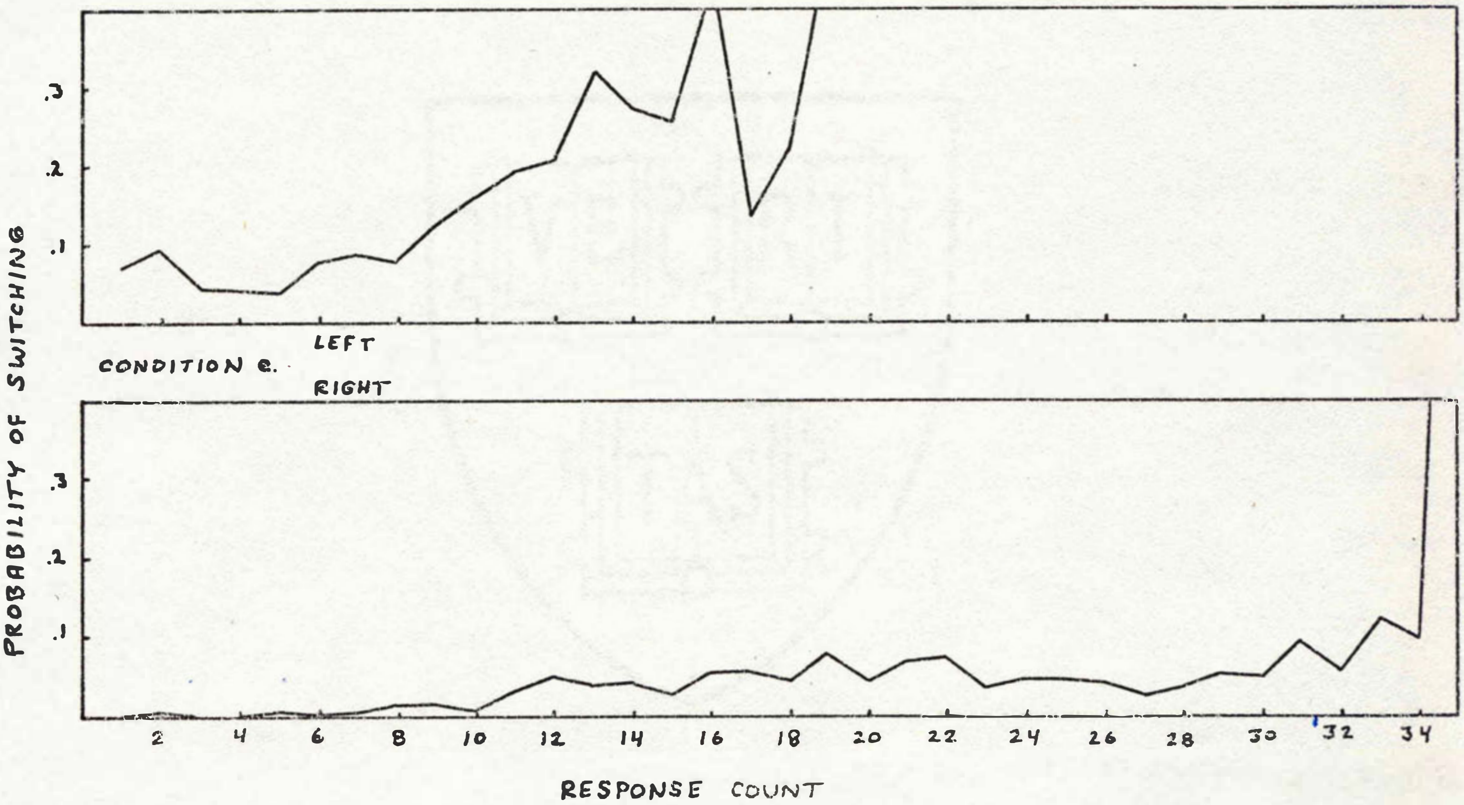
$r^2 = .92$

95% CONFIDENCE INTERVAL: .66- .88

Figure 16. Changeovers per opportunity for Bird 300. These graphs show, for each consecutive response to a side the probability that the next response will be directed at the other key.


PROBABILITY OF SWITCHING
.3
.2
.1
CONDITION a.
LEFT
RIGHT
.3
.2
.1
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
.3
.2
.1
CONDITION b.
LEFT
RIGHT
.3
.2
.1
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
RESPONSE COUNT


.3
.2
.1
LEFT
CONDITION c.
RIGHT
.3
.2
.1
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
PROBABILITY OF SWITCHING
.3
.2
.1
LEFT
CONDITION d.
RIGHT
.3
.2
.1
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
RESPONSE COUNT


PROBABILITY OF SWITCHING
.3
.2
.1
LEFT
CONDITION e.
RIGHT
.3
.2
.1
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
RESPONSE COUNT

determine whether the slopes were significantly different from zero was done for each bird, since lack of a functional relation should tend to produce a zero slope. The levels of significance, using a 2-tail test, were: .001, .05, .002, .01. The results thus differ significantly from what would be predicted if overall rate of reinforcement were governing the distribution of responses between the two alternatives. Thus, a distribution of responses that is close to matching occurs when overall reinforcement rate is independent of the distribution of responses. The undermatching that occurs may result in some way from the inhomogeneity of reinforcement rate on each key. If several stored reinforcers were collected within a short period of time, the effect might be different from what would have occurred had the tape readers stopped upon setting up reinforcement. What appears as an increase in variance to the right in the group data (Figure 16) probably results from deviations from linearity of Birds 475 and 300.

It appears likely that Shimp would predict a well-defined step in the plot of changeover per opportunity, since in a discrete trial situation momentary maximizing predicts a specific number of responses to each side before changing over (Shimp, 1966). It seems plausible that in this continuous procedure, momentary maximizing would also predict a certain number of responses to each side before changing. In two cases out of 40 (two keys, four birds, five conditions) there does appear to be a sharp increase in changeover probability, so such a result is not impossible. The general finding, however, is similar to the pattern exhibited by Bird 300. This may be characterized as a near zero probability soon after changing to a key, followed by a fairly rapid increase to a constant probability of changing, that is, to a Poisson process.

The initial low probability of changing is likely to be due to the COD, at the end of which any reinforcement set up while responding on the other side or during the COD will be collected. (The step seen at the right in some cases results from the method of collecting data: in the last bin, the probability of changing must be one.) This pattern of changing differs from the results of Nevin (1969), who found a decrese (in general), starting from a high probability at the beginning. This difference may be accounted for by the fact that Nevin was using a discrete trials procedure, with no penality for changing over (cf., Herrnstein, 1961). In summary, although this pattern is not predicted from a local maximization process, neither is it excluded by such a process. It is, on the other hand, excluded by a momentary maximization process.

Experiment 4: Concurrent Negative Slopes

Concurrent flat slopes neither reinforce nor punish matching, in that most deviations from matching do not affect reinforcement frequency. On concurrent negative slopes the situation can arise where, if a bird is responding at some overall rate and matching, a redistribution of responses will both increase reinforcement rate and produce a deviation from matching. Given that responding is sufficiently rapid, such a procedure can thus pit global maximization against matching.

Method

Subjects

Six White Carneaux pigeons were used, three in each of two conditions. They did not have long experimental histories, and were run at about 80% of free feeding weight.

Apparatus

Two standard two-key pigeon chambers were used, one for each condition. The first chamber was the same as that used in Experiment 3; the two experiments were run at different times of the day. In the second chamber two keys 6 in. (15 cm) apart at the same level were centered on a 12 in. (30.5 cm) wide wall; they were 9 in. (23 cm) from the floor. A standard food hopper was centered below the keys. The left key was transilluminated with white light from two 28-v D.C. (2-w) bulbs; the right key with green light. The chamber was continuously illuminated by two 7-w white bulbs. A force of about 14 g (.14 N) was required to operate the keys. Responses produced an auditory feedback click. During reinforcement, the key lights were extinguished, the house lights remained on and the hopper was illuminated. Both chambers were controlled by a PDP-9T computer.

Procedure

In the first chamber, concurrent negative slopes with the same intercepts were used. Two independent VI 45 sec tapes were simulated by the computer, with 20 intervals from Fleshler and Hoffman (1962). These added into two simulated bidirectional steppers, from which fixed ratios were subtracted. The sum of the two fixed ratios was 80. Five conditions were run. The fixed ratios used, for left and right respectively, were, in this order: 60 and 20; 30 and 50; 50 and 30; 20 and 60; 40 and 40. In the second chamber the same contingencies were in effect, except that the VI for the left key was 60 sec, and for the right key, 30 sec. At the end of each session the positions of the VI schedule, FR schedule, and stepper were printed out, and entered at the beginning of that bird's session the next day. The program was thus in the same state when it stopped one day and started the next. A 3 sec COD was in effect, and a minimum of 3 sec had to elapse between one reinforcement and the next. Reinforcement consisted of 3 sec access to mixed grain. Sessions terminated after 40 presentations of grain or 35 min, whichever occurred first. Sessions were run seven days a week, until behavior appeared stable. In Table 7 are shown the various conditions and the number of sessions each bird was run in each condition.

Table 7 about here

Results

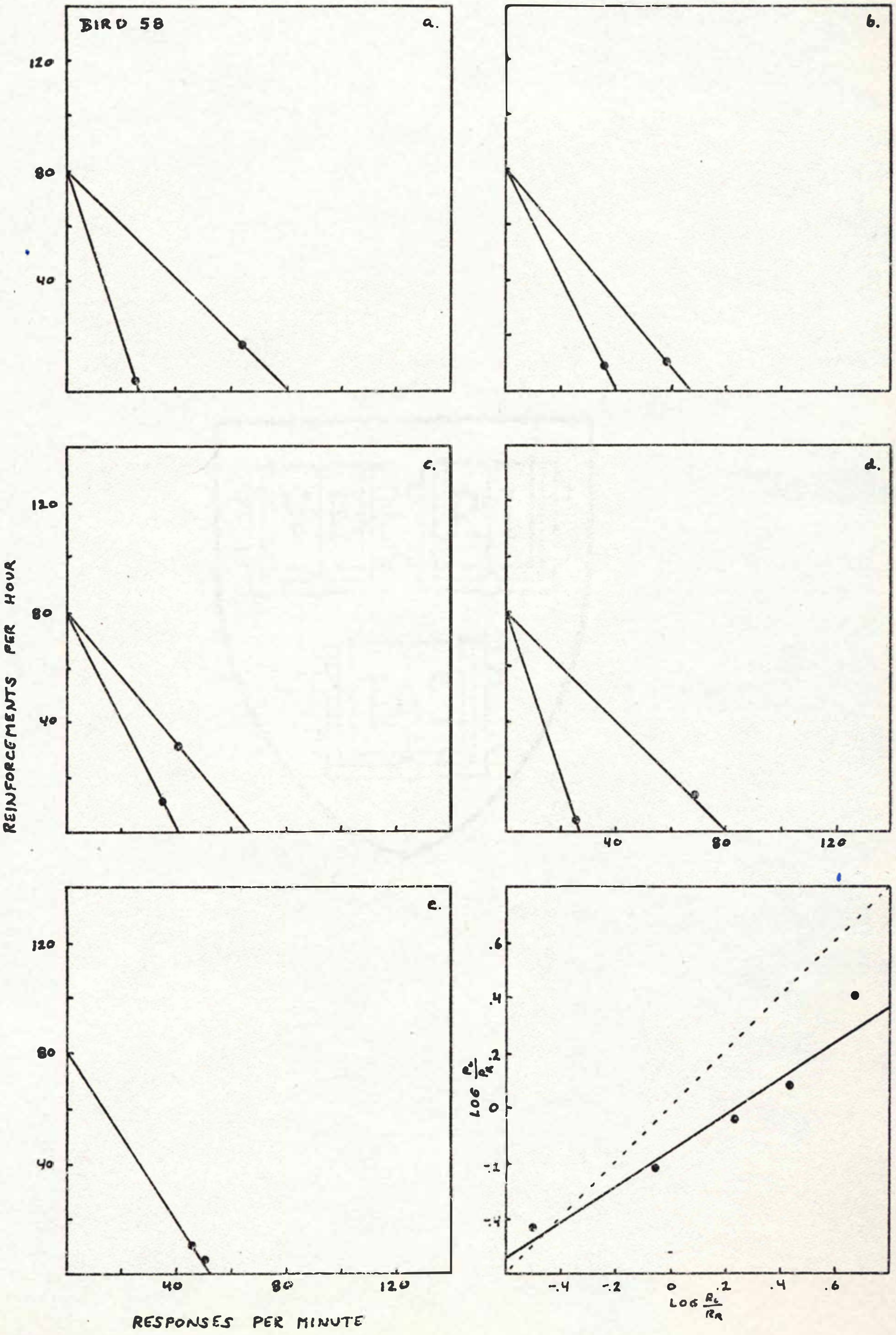
In Figure 17 are shown the schedules on which birds were run, and the logged ratios of responses as a function of logged ratios of reinforcers received; these data are averages of the last five days on a condition. In addition, the best fitting linear regression is shown. The 95% confidence interval for

Figure 17 about here

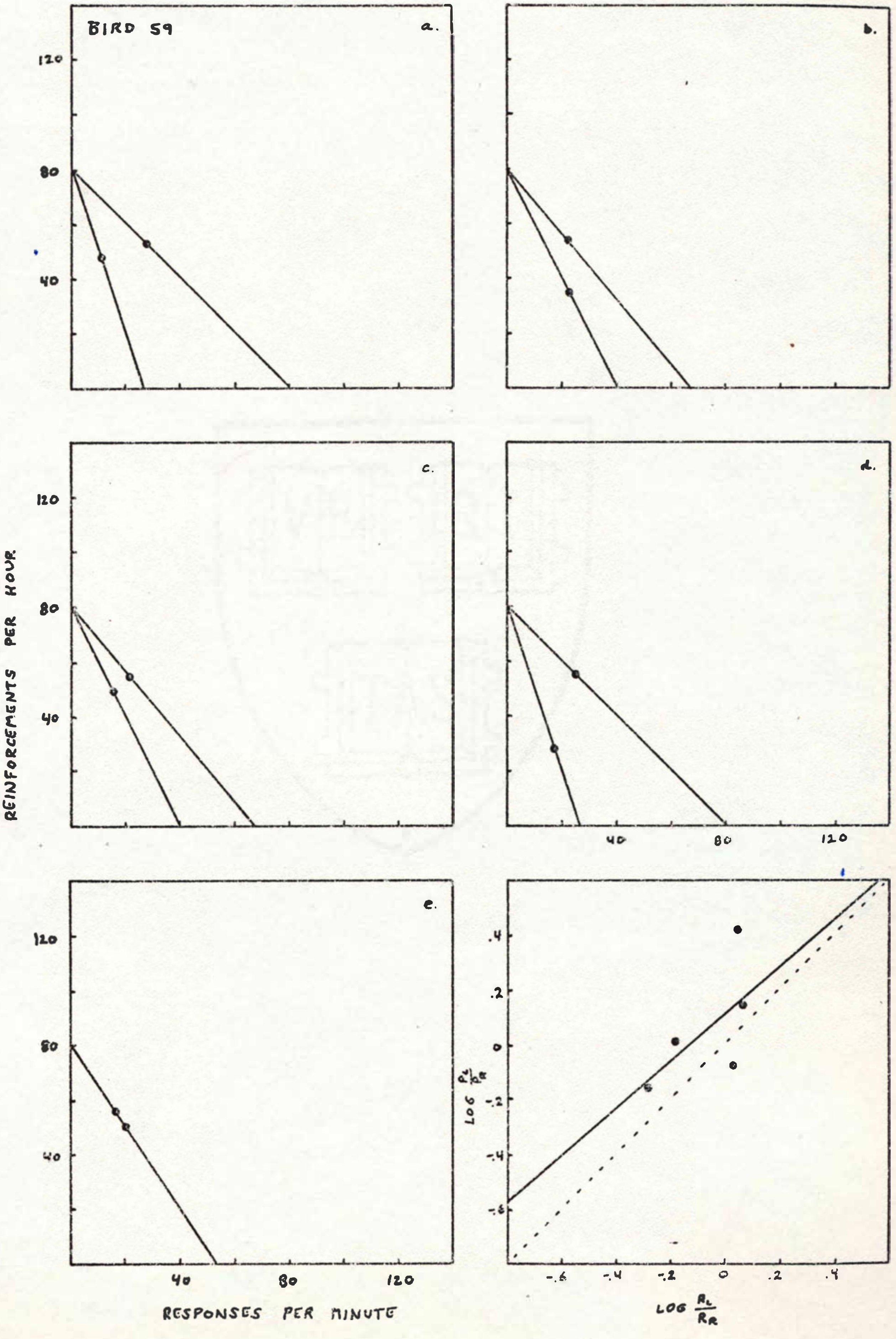
Table 7. Conditions in Experiment 4 and number of sessions each bird was run on each condition.

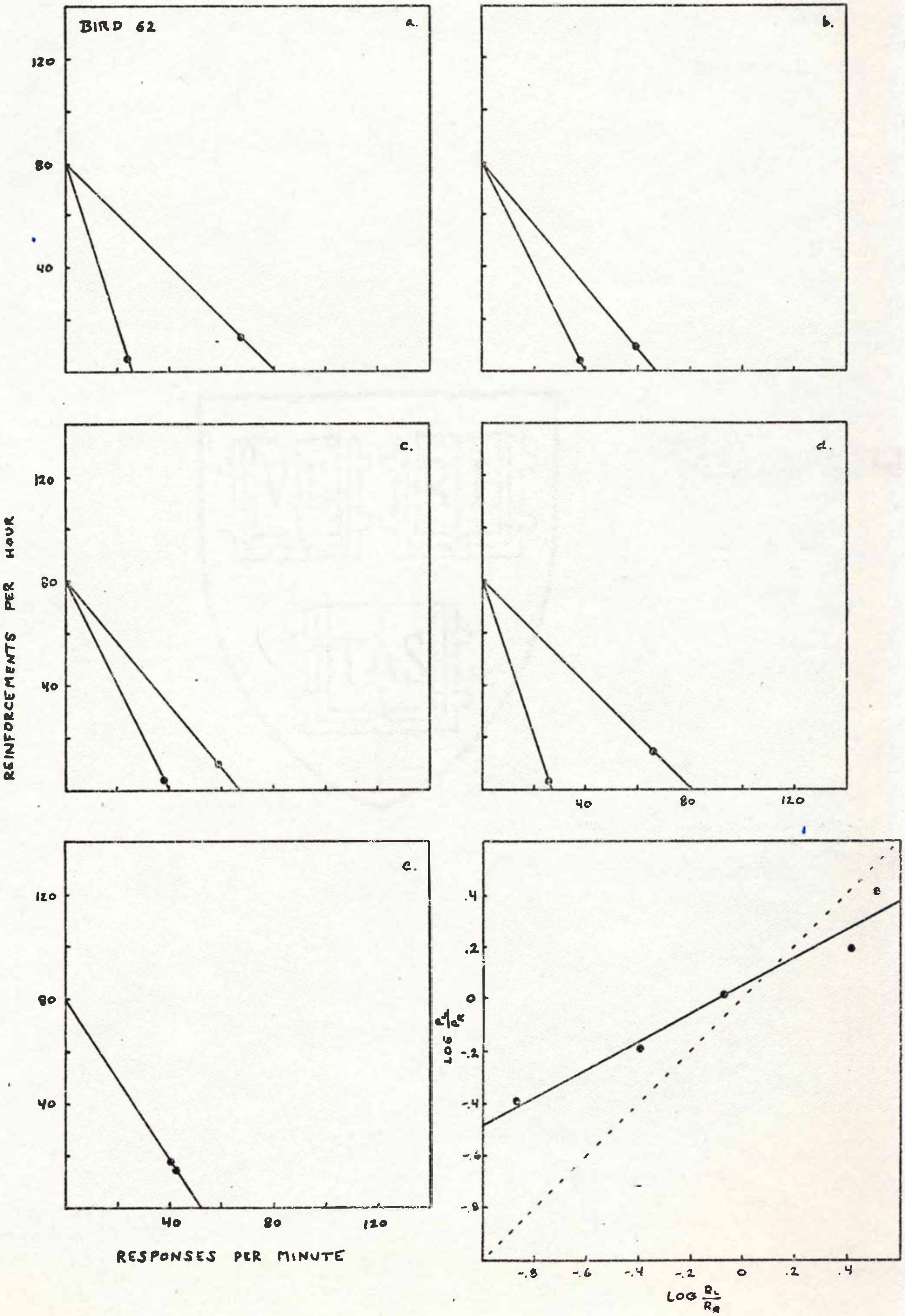
Schedules	Subjects		
VI 45" 45"	Bird 58	Bird 59	Bird 62
a 60,20	20	21	20
b 30,50	24	31	21
c 50,30	24	25	28
d 20,60	23	23	27
e 40,40	29	24	28
VI 60" 30"	Bird 175	Bird 176	Bird 111
f 60,20	23	24	24
g 30,50	21	20	22
h 50,30	28	28	27
i 20,60	28	27	25
j 40,40	28	30	28

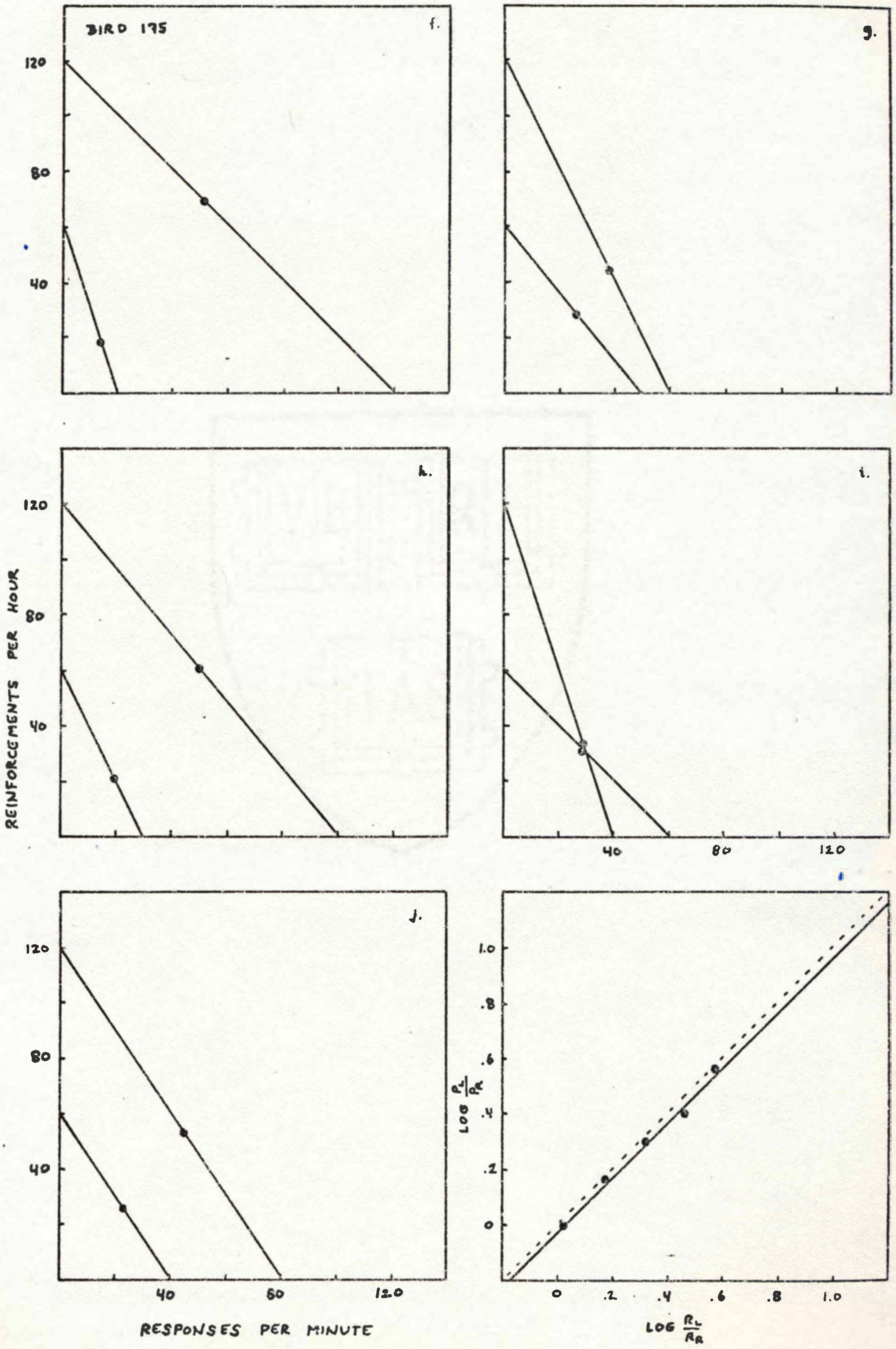
Figure 17. For all six birds, the schedules under which it was run are shown. In addition, logged ratios of responses as a function of logged ratios of reinforcers are shown, along with a least-squares linear fit.

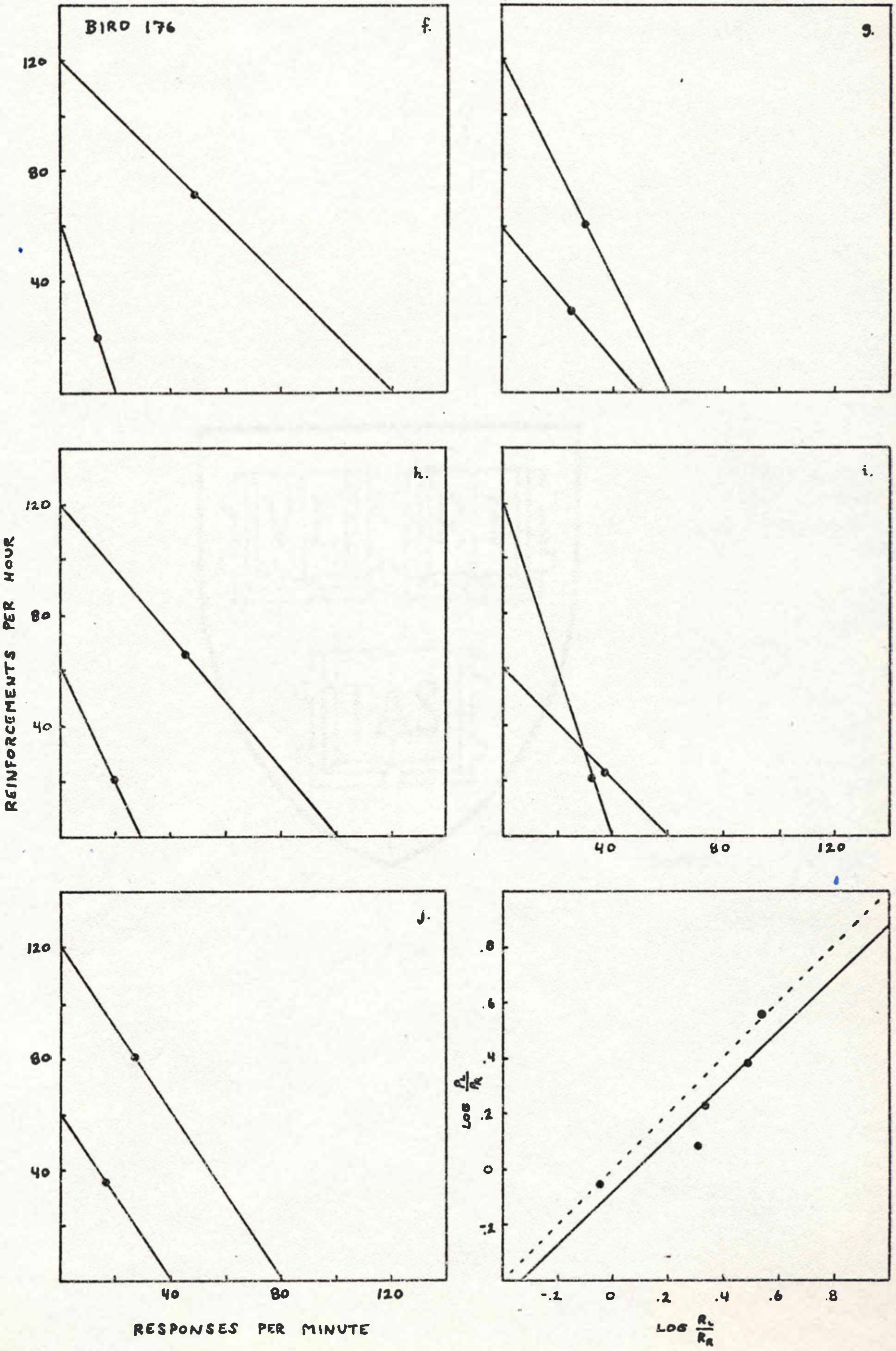


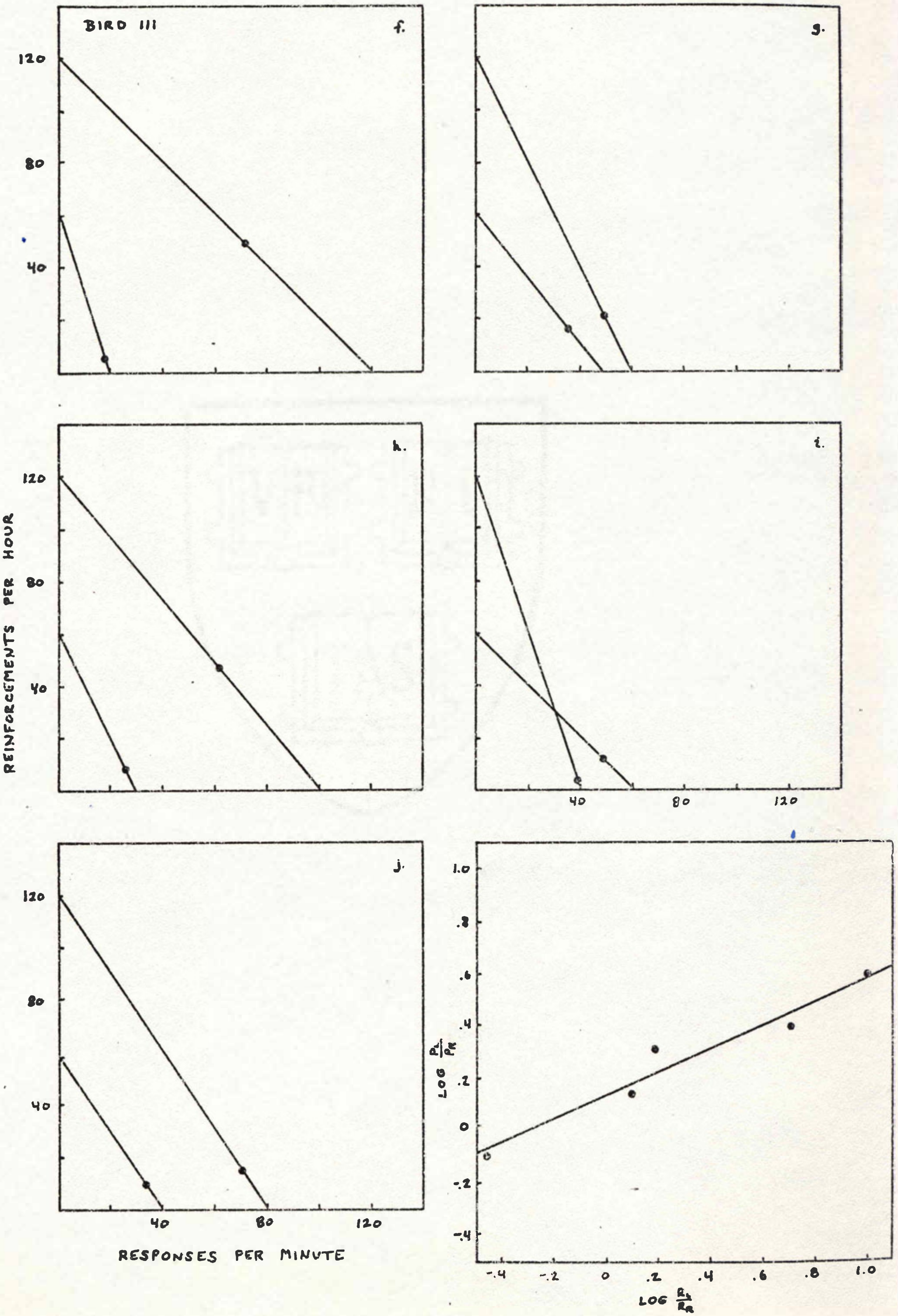
BIRD 58
a.
b.
c.
d.
e.
120
80
40
REINFORCEMENTS PER HOUR
40
80
120
RESPONSES PER MINUTE
.6
.4
.2
0
-.2
-.4
LOG $\frac{R_L}{R_R}$
-.4
-.2
0
.2
.4
.6
LOG $\frac{R_L}{R_R}$


BIRD 59
a.
b.
c.
d.
e.
120
80
40
40
80
120
REINFORCEMENTS PER HOUR
RESPONSES PER MINUTE
.4
.2
0
-.2
-.4
-.6
-.6
-.4
-.2
0
.2
.4
LOG $\frac{P_L}{P_R}$
LOG $\frac{R_L}{R_R}$


BIRD 62
a.
b.
c.
d.
e.
120
80
40
REINFORCEMENTS PER HOUR
40
80
120
RESPONSES PER MINUTE
.4
.2
0
-.2
-.4
-.6
-.8
LOG r_L/r_R
-.8
-.6
-.4
-.2
0
.2
.4
LOG R_L/R_R


BIRD 175
f.
g.
h.
i.
j.
120
80
40
REINFORCEMENTS PER HOUR
40
80
120
RESPONSES PER MINUTE
1.0
.8
.6
.4
.2
0
LOG $\frac{R_L}{R_R}$
0
.2
.4
.6
.8
1.0
LOG $\frac{R_L}{R_R}$


BIRD 176
f.
g.
h.
i.
j.
REINFORCEMENTS PER HOUR
120
80
40
RESPONSES PER MINUTE
40
80
120
LOG $\frac{R_L}{R_R}$
LOG $\frac{r_L}{r_R}$
-.2
0
.2
.4
.6
.8



BIRD 111
f.
g.
h.
i.
j.
120
80
40
REINFORCEMENTS PER HOUR
40
80
120
RESPONSES PER MINUTE
1.0
.8
.6
.4
.2
0
-.2
-.4
LOG $\frac{P_L}{P_R}$
-.4
-.2
0
.2
.4
.6
.8
1.0
LOG $\frac{R_L}{R_R}$

the slope is also indicated. In the case of equal intercepts there is a tendency for undermatching; for two out of three birds in the case of unequal intercepts the slope of the least-squares linear fit is close to one. In Figure 18 the response ratios as a function of reinforcement ratios are plotted separately for the two groups of birds. The overall response rates, reinforcement rates, and session times are shown in Table 8.

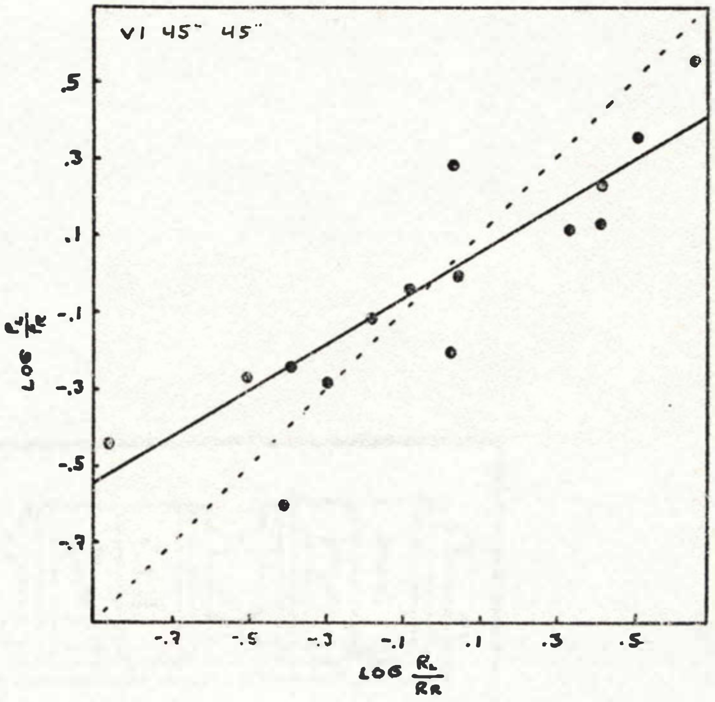
Figure 18 about here

Table 8 about here

Discussion

This experiment extends the results of Experiment 1, and shows that global maximization of reinforcement rate, if pitted against matching, does not control the distribution of responses. For two out of three birds in each of the two conditions the 95% confidence interval includes a slope of one. In some cases the plots of logged response ratios do not fall very close to a straight line. Under concurrent variable interval schedules, reinforcement rates on each side are only

Figure 18. Logged ratio of responses as a function of logged ratio of reinforcers for the two groups of birds. Prior to plotting, each bird's data points were adjusted so that the least-squares linear fit passed through the origin.


VI 45" 45"
.5
.3
.1
-.1
-.3
-.5
-.7
LOG R'L/R'R
-.7
-.5
-.3
-.1
.1
.3
.5
LOG RL/RR

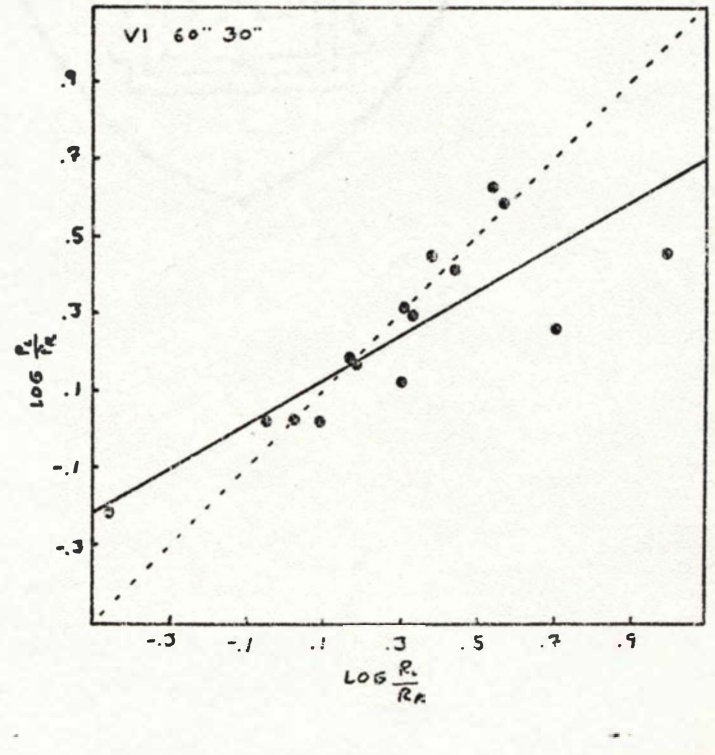

VI 60" 30"
.9
.7
.5
.3
.1
-.1
-.3
LOG R'L/R'R
-.3
-.1
.1
.3
.5
.7
.9
LOG RL/RR

Table 8. Overall response rates and reinforcement rates to each key, and session times from last five days of each condition in Experiment 4. P/M: Pecks per minute. R/H: Reinforcements per hour. M: Session time in minutes. L,R: Left, right keys.

	P/M		R/H		M
	L	R	L	R	
Bird 58					
a	63.9	25.2	16.1	3.4	175.0
b	35.5	58.3	9.2	10.3	175.0
c	41.2	34.3	31.1	11.4	173.8
d	26.0	69.2	4.1	13.0	175.0
e	46.0	50.3	10.6	4.8	175.0
Bird 59					
a	28.0	10.9	52.4	47.4	120.2
b	22.8	22.3	34.4	52.7	137.8
c	21.4	16.2	54.2	48.1	117.3
d	17.2	24.9	28.2	54.8	144.4
e	16.7	20.0	54.7	50.5	114.1
Bird 62					
a	66.6	25.7	13.7	4.1	175.0
b	37.8	58.6	3.8	9.2	175.0
c	58.3	37.8	10.1	3.8	172.4
d	26.6	65.1	2.0	15.1	175.0
e	43.3	41.6	15.3	17.4	172.2

Table 8 (cont.)

	P/M		R/H		M
	L	R	L	R	
			Bird 175		
f	51.0	14.0	69.4	18.4	136.5
g	38.3	26.1	44.6	29.4	161.3
h	49.9	20.1	60.6	20.7	147.5
i	29.0	29.3	33.3	31.1	167.7
j	44.7	22.6	53.2	25.6	152.2
			Bird 176		
f	47.0	13.1	72.1	20.9	129.0
g	30.1	25.1	59.8	29.5	134.4
h	46.0	19.4	64.9	21.0	139.7
i	33.4	37.7	20.2	22.3	175.0
j	27.6	16.4	79.0	36.3	104.0
			Bird 111		
f	71.4	18.4	49.0	4.8	175.0
g	50.2	36.2	21.2	16.8	175.0
h	61.6	25.3	46.3	9.0	173.5
i	39.2	49.3	3.8	10.6	175.0
j	70.2	34.3	14.4	9.2	175.0

slightly under the organism's control. In the present experiment, a change in behavior could have a substantial effect on reinforcement rates; this fact may account for the somewhat noisy results in some cases.

Summary, and a dynamic model

To summarize, there are two classes of theories which relate to the present account. The first class deals with behavior on single-key interval and ratio schedules. In general, these theories appeal to the strengthening effect of reinforcement on the preceding behavior, although a more global positive feedback model has also been proposed. Within the present account, these theories are viewed as quite possibly correct, but essentially incomplete, approaches to single-key responding.

Previous accounts of responding on concurrent schedules, the second class of theory, are viewed somewhat more critically. Experiment 3 strongly suggests that momentary maximizing, emitting that response which has the highest momentary probability of reinforcement, is not a viable theory. Global maximization, although not strongly espoused by anyone, was also shown not to account for certain results. And it may be said that the phenomenon of matching has in a sense been explaned, as resulting from the operation of a particular form of maximization.

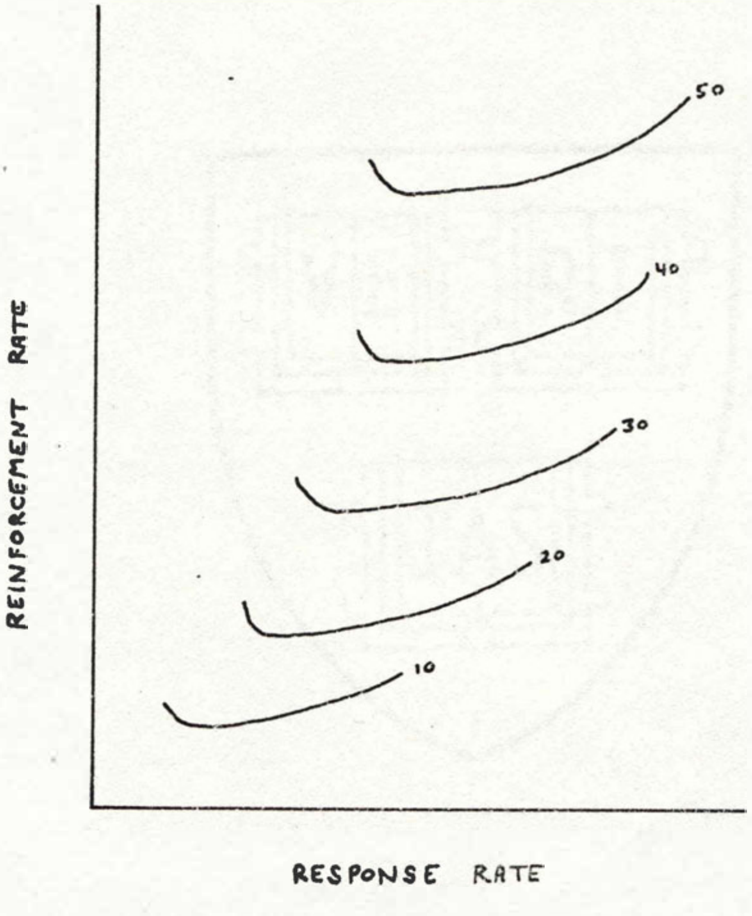
To summarize the present model, on a single manipulandum value is assumed to be some function, possibly unimodal, of response rate and reinforcement rate. Viewing a schedule of reinforcement as a constraint, the organism is assumed to behave in a manner which maximizes value, at least locally, while satisfying whatever constraints are present. The first experiment suggests that, in the case of pigeons pecking a key for grain, the slopes of the contours become negative and quite steep after passing through zero, as response rate decreases. Although limited to only one rate of reinforcement, the second experiment suggests that the slopes of contours increase in a gradual manner from zero, as response rate increases.

Figure 19 shows a response surface that is in qualitative agreement with Experiments 1 and 2.

Figure 19 about here

Looking at the points where contours have a slope of zero, an increase in reinforcement rate gives rise to an increase in response rate, which may asymptote. Data from Catania and Reynolds (1968) derived from a family of VI schedules suggest this characteristic, as do the flat schedules in Experiment 1. Going to the left, contours become steep rather quickly. This is intended to account for the difficulty of decreasing response rate when a negative slope is compared to a

Figure 19. Hypothetical response surface consistent with results of Experiments 1 and 2.


50
40
30
20
10
REINFORCEMENT RATE
RESPONSE RATE

flat slope, with equal reinforcement rates. Going to the right past the zero slopes, the slopes increase rather gradually. This is suggested by the gradual increase in response rate with an increase in slope found in Experiment 2.

Concurrent schedules are approached by assuming that individual keys are evaluated separately, and that if one key gives rise to greater value than a second, more time will tend to be distributed to the first and less to the second than was the case in the past. In Figure 20 we see how local rate of reinforcement on the two keys in the third experiment theoretically varied as a function of relative time on the two keys. When matching obtains, the local rates of reinforcement are equal. This may be expressed as:

Figure 20 about here

$$r_1 / T_1 = r_2 / T_2$$

where r_i is the total number of reinforcers on one side, and T_i is total time on that side.

Consider now concurrent VR VR, assuming equal rates of responding on the two sides. The local

Figure 20. Theoretical local rates of reinforcement in Experiment 2 with Conc VI 2´ VI 1´ in effect, as a function of relative time on the right key. The intersection of the two functions corresponds to matching.

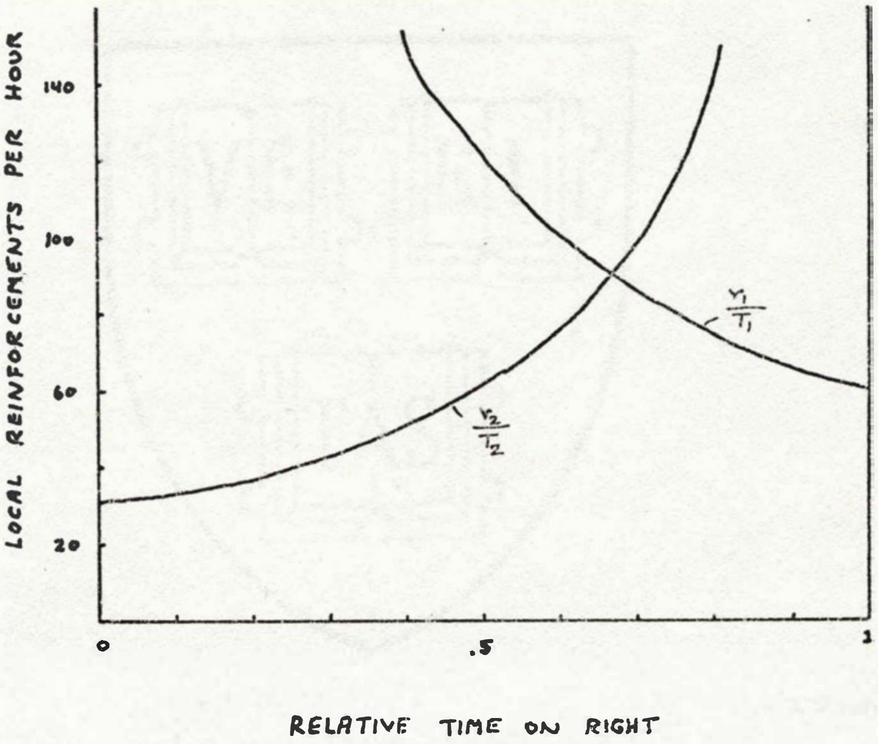

LOCAL REINFORCEMENTS PER HOUR
140
100
60
20
0
.5
1
RELATIVE TIME ON RIGHT
$\frac{r_1}{T_1}$
$\frac{r_2}{T_2}$

Figure 21 about here

---

reinforcement rates are shown in Figure 21. In this case the following holds:

$$r_1/T_1 = C_1 > C_2 = r_2/T_2.$$

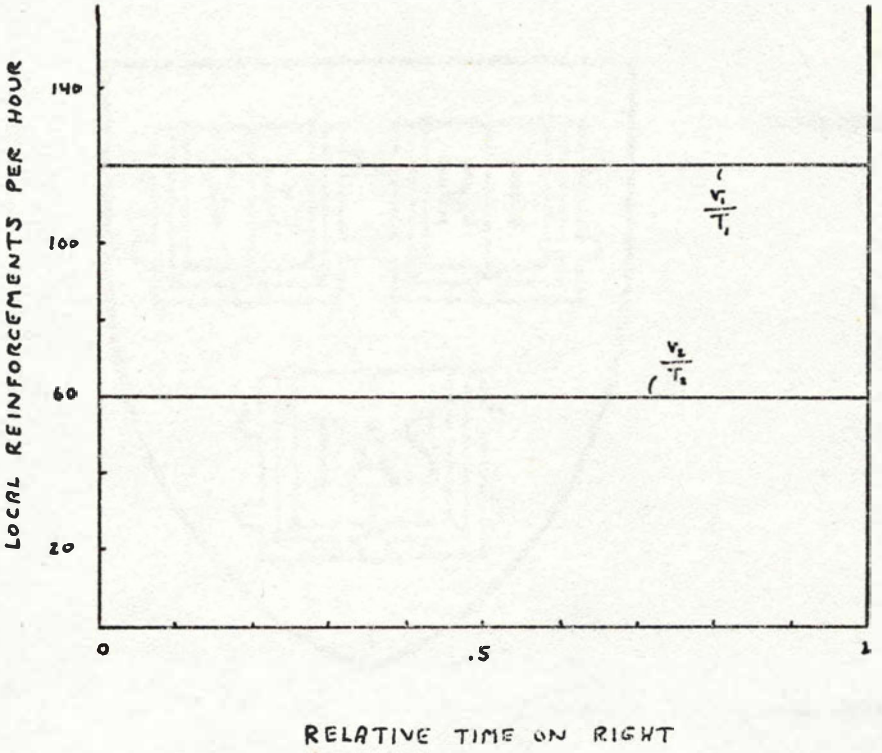
That is, the local reinforcement rates are constants, and differ if the VR schedules are different. As $T_1/(T_1 + T_2)$ approaches one, both r_2 and T_2 approach zero, but their ratio remains constant. At exclusive preference, then:

$$r_1/T_1 > \lim_{T_2 \to 0} r_2/T_2$$

From this point of view it would not be appropriate to say that the local reinforcement rates had become equal. If we take this equality of local reinforcement rates as our definition of matching, then we cannot say that matching obtains in the case of a concurrent VR VR with different ratio requirements.

The issues under discussion can be brought into better focus by setting forth a dynamic model of behavior: one that attempts to describe how behavior changes over time. We know how a dynamic model must behave under constant conditions as time grows large, since it must then lead to the steady-state model already presented. Previously it was said that the

Figure 21. Local rates of reinforcement on concurrent VR VR as a function of relative time on the right key, assuming a constant rate of responding to both keys.


LOCAL REINFORCEMENTS PER HOUR
140
100
60
20
0
.5
1
$\frac{V_1}{T_1}$
$\frac{V_2}{T_2}$
RELATIVE TIME ON RIGHT

only assumption it is necessary to make is that, over time, more time tends to be distributed to a locally better alternative and less time to a locally poorer alternative. This would suggest a model of the form

$$dT_1/dt = g(V_1 - V_2) = -dT_2/dt$$

where V_i is the value in situation i and T_i is time in that situation. We may assume g to be monotonically increasing and symmetrical about the origin (so that $g(0) = 0$). A simple function satisfying these constraints is the identity function, which will now be employed with the proviso that this is a tentative assumption. This gives:

$$dT_1/dt = V_1 - V_2$$

However, unless $\lim V_2 = V_1$ as $T_2 \rightarrow 0$, this function makes the erroneous prediction that time in a situation can continue to increase, even when all time is spent in that situation. We can remedy this by introducing a multiplicative factor $h(T_i)$ which goes to zero as time in the poorer situation i goes to zero; again an identity function is employed:

$$dT_1/dt = T_2(V_1 - V_2) = -dT_2/dt$$

where $V_1 > V_2$. More abstractly, between any two situations i and j, this function should hold:

$$dT_i/dt = - dT_j/dt = T_k(V_i - V_j)$$

where $k = i$ if $V_i < V_j$, and $k = j$ if $V_i > V_j$. If $V_i =$

V_j the value of T_k is immaterial (provided it is finite) since $V_i - V_j = 0$.

At equilibrium, $dT_i/dt = dT_j/dt = 0$; that is to say, the distribution of time is not changing. According to this dynamic model, this may come about in two ways, either or both of which may occur. In the first case, $V_i = V_j$. This obtains in the usual matching found on concurrent VI VI: there local response rates (P_i and P_j) are equal, as are local reinforcement rates (R_i and R_j), and so both $V_i = f(P_i, R_i)$ and $V_j = f(p_j, R_j)$ must be equal. On concurrent VR VR, with the requirement for i less than that for j, $V_i > V_j$, and T_i will increase until $T_j = 0$. In this case equlibrium comes about because of a time limitation and not because of equality of local values. If $\lim V_j = V_i$ as $T_j \rightarrow 0$, and for $T_j > 0$, $V_i > V_j$, then at equlibrium matching would obtain and exclusive preference would be exhibited. Thus at exclusive preference this model says that matching may or may not obtain. Although the steady-state outcomes may differ, a single dynamic process appears sufficient to account for those results.

The level of measurement appropriate to different heights on the response surface is at least an ordinal scale. On a single schedule, for example, points in the vicinity of a point of stability on the reinforcement function are assumed to have less value than the point of stability. In contemporary economics this appears to be the level of measurement assumed to be appropriate for the utility surface, which is quite similar to the present response surface (Mansfield, 1975). A possible means of gaining an interval scale might derive from the use of conjoint measurement, either additive or polynomial (Krantz, Luce, Suppes, & Tversky, 1971). However, in order to do so it would be necessary to assume that value can be represented as an additive or polynomial function of response rate and reinforcement rate. It is not feasible to evaluate this possibility with data presently available.

A second possibility is suggested by Lange (1934) in a discussion of the level of measurement appropriate for the utility surface. If it is just assumed that any pair of commodities may be ordered in terms of preference (allowing indifference), an ordinal scale is determined. If, in addition, a person can order differences between pairs of commodities, the following construction may be employed to give an interval scale of measurement. Given commodities a and b, with b preferred to a, find a third commodity c such that the difference from a to b is the same as that from b to c.

This difference can then be taken as our unit of measurement. Lange points out that while the preference ordering is derived from behavioral data, the ordering of differences must come from introspection.

In the case of the response surface, if a dynamic model of the sort suggested above holds, this opens the theoretical possibility of obtaining behavioral data about the ordering of differences in value between situations. This possibility exists because we are assuming that a difference in value does not lead to an instantaneous change in performance, but a change in performance whose rate is some function of that difference in values. Thus if it were possible to measure that rate, an inference might be made as to the difference in value.

Implications for the concept of reinforcement

Finally, we may consider some of the implications of this approach for the concept of reinforcement; specifically, the present approach allows one to view predictions of Premack from another point of view. The view of reinforcement usually cited, deriving from Thorndike (1911) and Skinner (1938), states that a reinforcer is a stimulus which, if it follows some behavior, will increase the frequency of that behavior. It can immediately be seen that such an effect is one of three possibilities. The second is that a stimulus may produce no change; these are called neutral stimuli. The third possibility is that a stimulus may decrease behavior; such stimuli are called punishers.

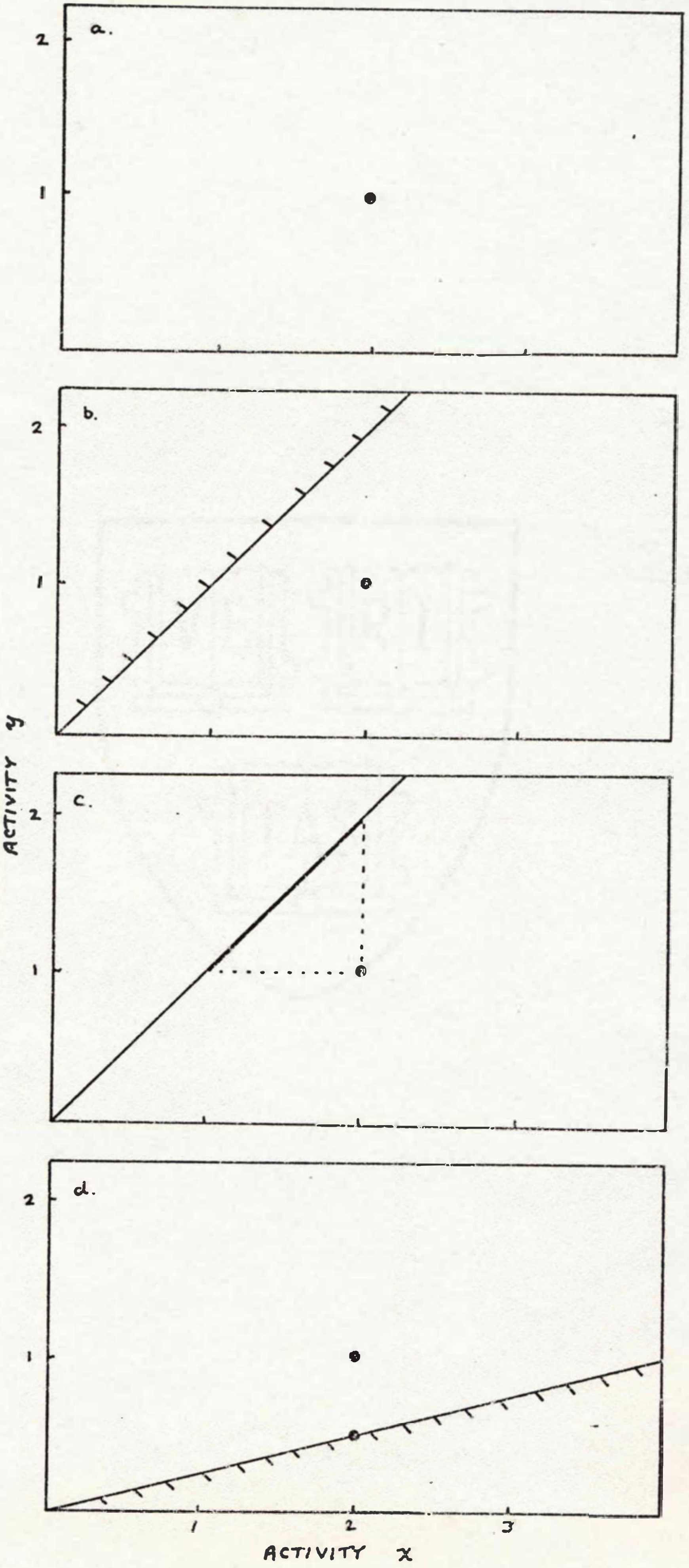
This scheme was shown to have certain limitations by Premack (1959), who found that one response with a higher probability of occurring could reinforce one with a lower probability, but the latter could not reinforce the former. In general, what would earlier be thought of as a stimulus becomes for Premack a response: instead of food, we consider eating. We may speak of events which make contact with the animal, and not lose what is meant in either case. Premack points out that his position eliminates the possibility of dividing events into reinforcers, neutral stimuli, and punishers. Rather, the status of a contingent event depends upon its relation to the instrumental event. Premack's position requires that events be compared on a common scale; otherwise not every two events would be comparable in terms of probability. He first suggested using intact chains of behavior as a common unit (called smallest possible units); later (1965) he suggested that the amount of time an activity is engaged in is a valid measure of its probability. Thus if more time is spent in one activity than another, the first should reinforce the second, but not the other way around. Although it does not follow from his position, he has found that, in order for reinforcement to occur, the contingent event must be occurring less than in baseline. For example, if a rat engages in running and drinking ab lib, one response may occur more than the other. In spite of the fact that on numerous instances the more probable response follows the less probable, no reinforcement (i.e., increase in the less probable response) occurs.

We may illustrate what Premack is suggesting by means of a surface with constraints. In Figure 22a we

---

Figure 22 about here

Figure 22. (a): Hypothetical point representing base rates of activity x and activity y. (b): Inequality constraint enforcing at least as much of y as x. (c): Range of activity as predicted by Premack under this constraint. (d): Point predicted by Premack under inequality constraint enforcing four times as much of x as of y.


a.
b.
c.
d.
2
1
ACTIVITY y
1
2
3
ACTIVITY x

see a point representing the amount of time spent on two activities, x and y; x is two times as probable as y. Suppose we make x contingent on y, one unit of x contingent on one unit of y. This is equivalent to an inequality constraint: the animal may move to any point on or above the diagonal (Figure 22b). Premack says that the animal will move to a point above the line y = 1: this corresponds to the lower probability behavior being reinforced. Assuming the animal stays on the constraint, Premack says behavior will fall somewhere above y = 1 and below y = 2 (Figure 22c).

On the other hand, the situation of y being contingent on x is illustrated in Figure 22d. In this particular case, the animal may go anywhere on or below the function y = x/4. If the animal stays on the function, no change in x puts him at the point x = 2, y = .5.

By translating Premack´s position into these terms, it appears that while reasonable, what he has said does not make strong predictions. Furthermore, one finds no statement in regard to how greater formal rigor could be incorporated into this position. On the other hand, Premack is certainly responsible for having closely scrutinized the concept of reinforcement.

Premack (1971) makes an attempt to deal with the observation that some behaviors occur seldom but, when they occur, have high probability; other behaviors occur more frequently but with lower probability when they do. This leads him to suggest that it is momentary probability that determines what behavior will reinforce what. Such information can be lost if only average probabilities are considered. This strategy makes prediction less straightforward, for an outcome can only be predicted by knowing future momentary probabilities.

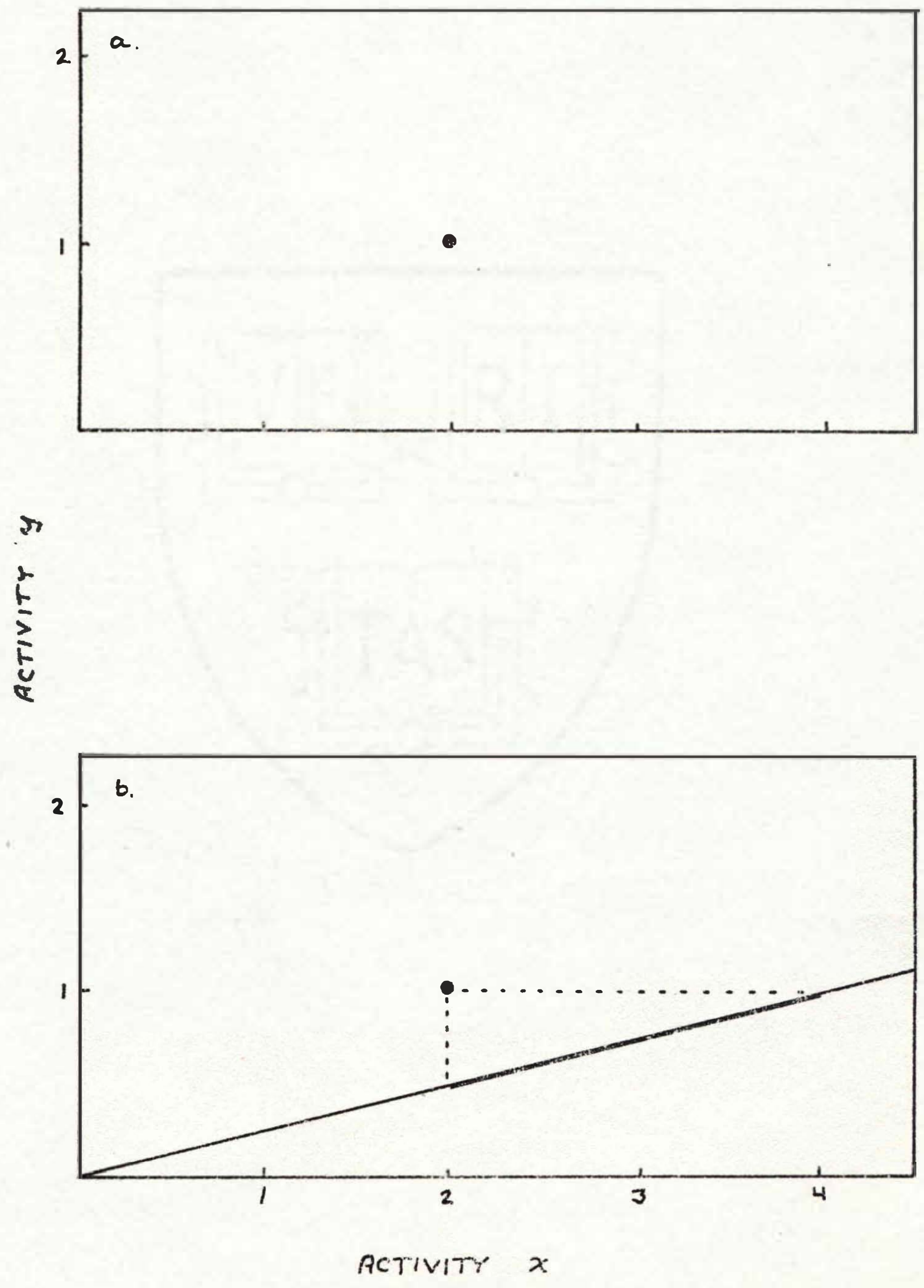
Timberlake and Allison (1974) take a slightly different view, which they term an adaptive model of performance. Given two responses, x and y, one first determines base rates by allowing free access to these events simultaneously. If the animal is then constrained so that by engaging in the base rate of x it can only engage in less than the base rate of y, and if engaging in more than the base rate of x will allow it to engage in more of y than otherwise, then x will be reinforced. It does not matter whether x or y had higher probability to begin with; deprivation of either allows for the possibility of reinforcement of the other. Further, the instrumental event will not increase above the point at which the contingent event is occurring at baseline level, but it may not reach that high.

Suppose, as in Figure 23a, an animal engages in x twice as much as y. If we now constrain the animal so that engaging in one unit of y is contingent on engaging in four units of x, only by engaging in eight units of x can the baseline of two units of y be obtained. As suggested by Figure 23b, they predict the animal will go somewhere along the function, between x = 4 and x = 8. They predict the same outcome as Premack if a higher probability response is contingent on a lower probability response. Timberlake and Allison mention that it is not necessary to measure the two behaviors on a common scale. This reduces the number of possibly arbitrary assumptons that must be made. They, as well as Premack, discuss no systematic way of predicting where on some constraint an animal will go.

Figure 23 about here

We thus have three views of reinforcement that have been discussed. In the first, reinforcement occurs when some behavior is followed by a reinforcing stimulus. In the second, reinforcement occurs when a lower probability response is followed by a higher

Figure 23. (a): Hypothetical amounts of activities x and y. (b): Range predicted by Timberlake and Allison under the inequality constraint shown.



a.
2
1
ACTIVITY y
b.
2
1
1
2
3
4
ACTIVITY x

probability response. In the third, reinforcement occurs if an increase in one response allows a second response, occurring below baseline, to approach its baseline level.

Within the general point of view being developed here, we make no assumptions about what events correspond to the coordinates of our space. Suppose an animal is allowed to engage in two events, x and y, at baseline rates (Figure 24a). We may hypothesize that

---

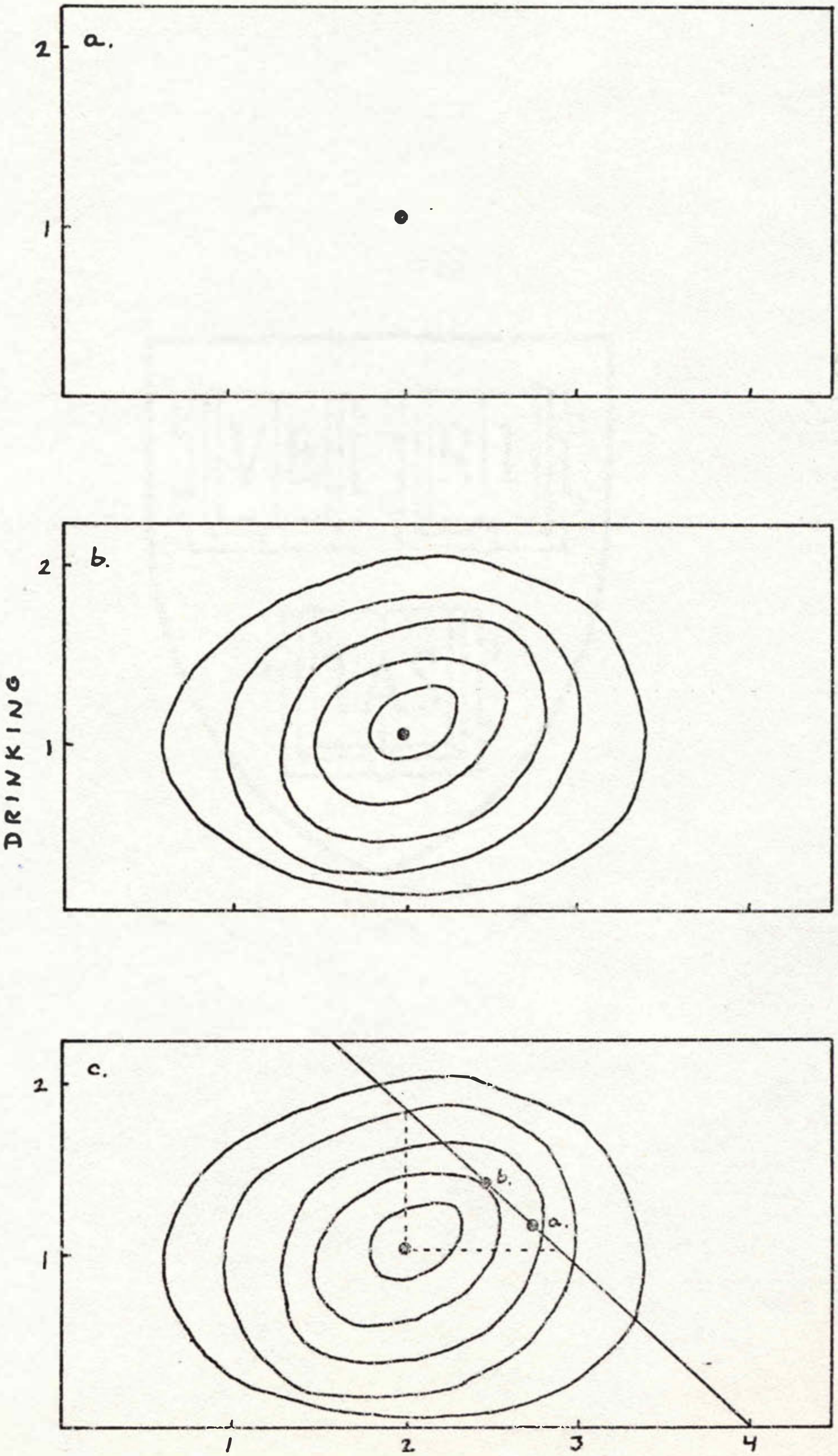
Figure 24 about here

---

by doing so it is thereby maximizing value (at least locally). Thus other points in the vicinity should have less value associated with them, and we may hypothesize the existence of level curves. In Figure 24b we see possible level curves, in this case strictly convex.

Given the existence of such curves, predictions in the presence of constraints become straightforward: movement away from a point only occurs to the extent that value increases. Consider the constraint in Figure 24c. Here, both x and y have increased. Such a situation could arise as follows: baseline rates of running and drinking result in x of running and y of

Figure 24. (a): Hypothetical data point representing base rates of running and drinking. (b): Hypothetical contours that would give rise to behavior observed in (a). (c): Under the constraint shown, behavior should move to point b. Timberlake and Allison do not discriminate between points a and b.


a.
b.
c.
2
1
DRINKING
RUNNING
1
2
3
4
b.
a.

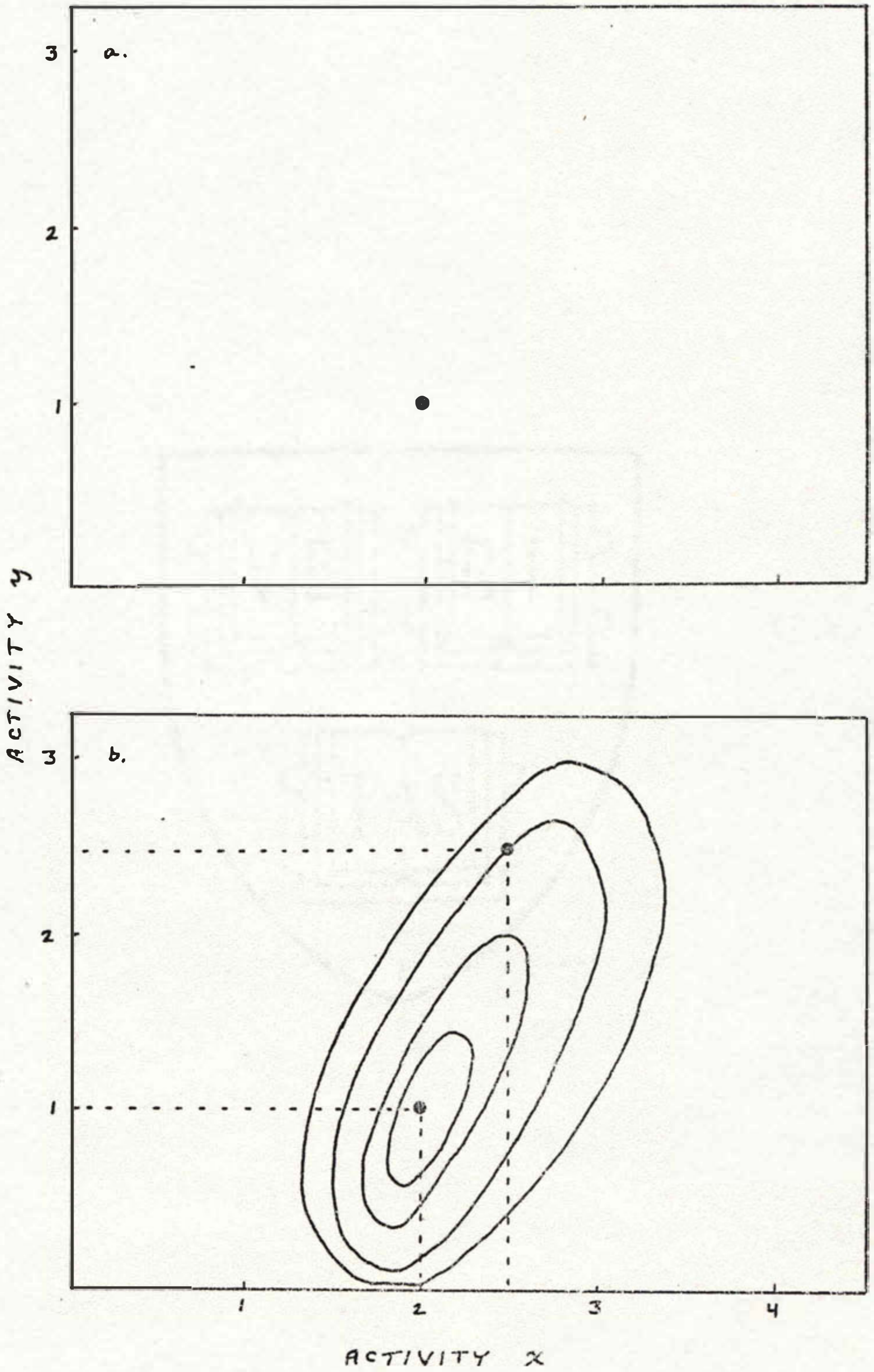
drinking. Now, the running wheel is forced to turn twice as much as before; if drinking occurs it will reduce the amount of running necessary. It is reasonable to think that both drinking and running will end up at higher than baseline rates, since either that must occur or else one of them must be very much above baseline. This goes against predictions of Premack, for, from his point of view, both responses have been reinforced, not just the one with the lower probability in baseline. Suppose the animal starts out at point a. Timberlake and Allison cannot account for a move to b, because no deprivation exists.

Interaction effects could also give rise to interesting possibilities. Suppose a subject responds at a in Figure 25a, and suppose further that y interacts with x in such a way that an increase in y will make the optimal x value greater. For example, x might be drinking water and y licking salt. Contours could be oriented as in Figure 25b, where a constraint is shown that enforces equal amounts of x and y. Here, both x and y could be increased above baselines. Neither Premack nor Timberlake and Allison are prepared

----

Figure 25 about here

Figure 25. (a): Hypothetical data point. (b): Hypothetical contours that might result if activity y potentiated activity x.


3
a.
2
1
3
b.
2
1
ACTIVITY y
1
2
3
4
ACTIVITY x

to deal with such cases.

The problem raised by infrequent behaviors with high probability when they occur may be dealt with as follows. Suppose we observe a person to spend one hour a day eating, and ten hours a day chewing gum; further, that any deviation from that amount of eating is aversive, but little change in value results from large changes in gum chewing. In Figure 26a we see possible relations between value and amount that would

---

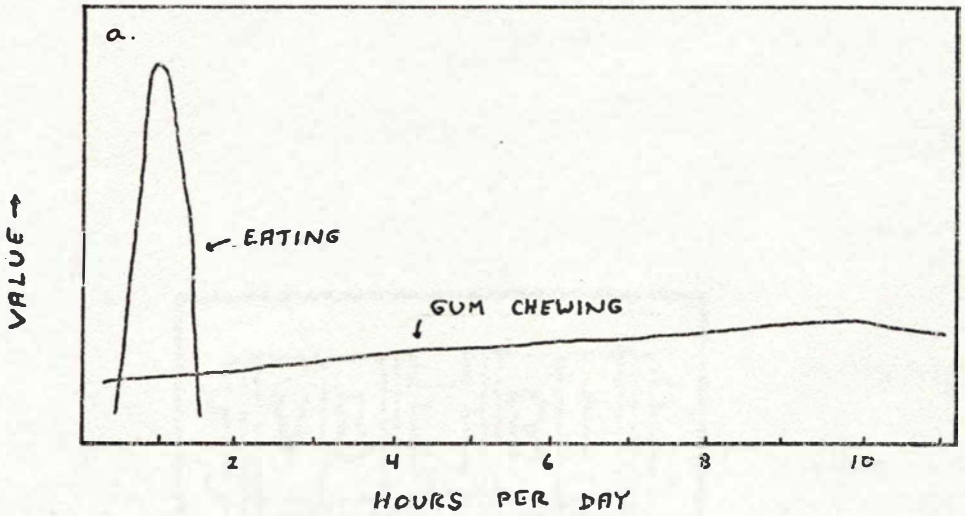
Figure 26 about here

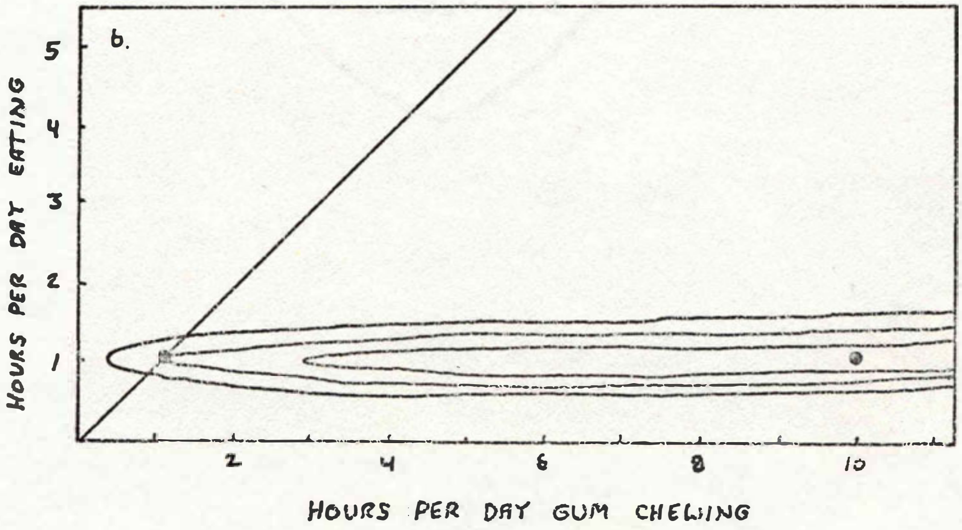
---

correspond to these cases. Assuming no interactions, level curves would be similar to those in Figure 26b. We see that a constraint that forces these two events to occur in equal amounts changes one greatly and the other a small amount. The simplest sort of interaction would probably orient the axes of the contours (viewing them as elipses) at some other angle.

The process of reinforcement is usually said to occur if some response increases above its baseline level, or operant level. If we consider what happens from the present point of view, we see that without constraints the animal will go to some point in the space, and if constrained from doing so it will go to

Figure 26. (a): Hypothetical functions showing value as a function of amount for two activities. (b): Approximate contours that would result from functions in (a), and behavior maintained by a constraint.


a.
VALUE →
EATING
GUM CHEWING
2
4
6
8
10
HOURS PER DAY


b.
HOURS PER DAY EATING
5
4
3
2
1
2
4
6
8
10
HOURS PER DAY GUM CHEWING

some other point. But in either case the same process is involved: movements toward higher value are retained while those toward lower value are not. Reinforcement does not appear unique to the constrained situation, but occurs in any situation.

By identifying value with reinforcement, we see that maximizing value is equivalent to doing what is most reinforcing out of the sampled alternatives. If one behavior is more reinforcing than another, we can only identify it as so because of an increase in the ratio of that behavior to the other. It may appear that the present account suffers from circularity, which would imply that it cannot be disconfirmed (cf, Popper, 1959). However, we need to distinguish between specific hypotheses put forth, such as monotonicity of the response surface, and the more general approach being advocated. The specific hypotheses discussed are certainly susceptible to disconfirmation, so we need deal only with the general approach. That approach is based on the assumption that how an organism distributes its time between two or more alternatives depends in a lawful and orderly manner on the characteristics of those alternatives. While this assumption may never be disconfirmed, it is also likely few people would reject it for this reason.

References

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., & Rachlin, H. C. Choice as time allocation. Journal of the Experimental Analysis of Behavior, 1969, 12, 861-874.

Blough, D. S. The reinforcement of least-frequent interresponse times. Journal of the Experimental Analysis of Behavior, 1966, 9, 581-591.

Catania, A. C. Concurrent operants. In W. K. Honig (Ed.), Operant behavior: areas of research and application. New York: Appleton-Century-Crofts, 1966.

Catania, A. C. Reinforcement schedules: the role of responses preceding the one that produces reinforcement. Journal of the Experimental Analysis of Behavior, 1971, 15, 271-287.

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.

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

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

Fleshler, M., & Hoffman, H. S. A progression for generating variable interval schedules. *Journal of the Experimental Analysis of Behavior*, 1962, *5*, 529-530.

Fraenkel, G. S., & Gunn, D. L. *The orientation of animals: kineses, taxes and compass reactions*. Oxford: Clarendon Press: 1940.

Hawkes, L., & Shimp, C. P. Reinforcement of behavioral patterns: shaping a scallop. *Journal of the Experimental Analysis of Behavior*, 1975, *23*, 3-16.

Hemmes, N. S. Pigeons' performance under differential reinforcement of low rate schedules depends upon the operant. *Learning and Motivation*, 1975, *6*, 344-357.

Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis* of *Behavior*, 1961, *4*, 267-272.

Herrnstein, R. J. Secondary reinforcement and rate of primary reinforcement. *Journal of the Experimental Analysis of Behavior*, 1964, *7*, 27-36.

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

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.

Jenkins, H. M. Sequential organization in schedules of reinforcement. In W. N. Schoenfeld (ed.), *The theory of reinforcement schedules*. New York: Appleton-Century-Crofts, 1970.

Killeen, P. Reinforcement frequency and contingency as factors in fixed-ratio behavior. *Journal of the Experimental Analysis of Behavior*, 1969, *12*, 391-395.

Killeen, P. A yoked-chamber comparison of concurrent and multiple schedules. *Journal of the Experimental*

Analysis of Behavior, 1972, 18, 13-22.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A., Foundations of measurement. Vol. 1. Additive and polynomial representations. New York: Academic Press, 1971.

Lange, O. The determinateness of the utility function. The Review of Economic Studies, 1934, 1, 218-225.

Levins, R. Evolution in changing environments. Princeton: Princeton University Press, 1968.

Luenberger, D. G. Optimization by vector space methods. New York: John Wiley, 1969.

Luenberger, D. G. Introduction to linear and nonlinear programming. Reading, Mass.: Addison-Wesley, 1973.

Mackintosh, N. J. The psychology of animal learning. London: Academic Press, 1974.

Maier, S. F., Seligman, M. E. P., & Solomon, R. L. Pavlovian fear conditioning and learned helplessness: effects on escape and avoidance behavior of (a) the CS-US contingency and (b) the independence of the US and voluntary responding. In B. A. Campbell & R. M. Brown (Eds.), Punishment and aversive behavior. New York: Appleton-Century-Crofts, 1969.

Mansfield, E. Microeconomics: theory and applications (2nd ed.). New York: W. W. Norton, 1975.

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. Interval reinforcement of choice behavior in discrete trials. Journal of the Experimental Analysis of Behavior, 1969, 12, 875-885.

Popper, K. The logic of scientific discovery. New York: Basic Books, 1959.

Premack, D. Toward empirical behavioral laws: I. Positive reinforcement. Psychological Review, 1959, 66, 219-233.

Premack, D. Reinforcement theory. In D. Levine (Ed.), Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1965.

Premack, D. Catching up with common sense or two sides of a generalization: reinforcement and punishment. In R. Glaser (Ed.), The nature of reinforcement. New York: Academic Press, 1971.

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

Rachlin, H., & Baum, W. M. Effects of alternative reinforcement: does the source matter? Journal of the Experimental Analysis of Behavior, 1972, 18, 231-241.

Reynolds, G. S. A primer of operant conditioning. Glenview, Ill: Scott, Foresman and Co., 1968.

Rosen, R. Optimality principles in biology. New York: Plenum Press, 1967.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Shimp, C. P. Probabilistically reinforced choice behavior in pigeons. Journal of the Experimental Analysis of Behavior, 1966, 9, 443-455.

Shimp, C. P. The reinforcement of short interresponse times. Journal of the Experimental Analysis of Behavior, 1967, 10, 425-434.

Shimp, C. P. Optimal behavior in free-operant experiments. Psychological Review, 1969, 76, 97-112.

Shull, R. L., & Pliskoff, S. S. Changeover delay and concurrent performances: some effects on relative performance measures. Journal of the Experimental Analysis of Behavior, 1967, 10, 517-527.

Skinner, B. F. The behavior of organisms. New York: Appleton- Century, 1938.

Skinner, B. F. "Superstition" in the pigeon. Journal of Experimental Psychology, 1948, 38, 168-172.

Staddon, J. E. R. Temporal control and the theory of reinforcement schedules. In R. M. Gilbert & J. R.

Millenson (Eds.), _Reinforcement: behavioral analyses_. New York: Academic Press, 1972.

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.

Stubbs, D. A., & Pliskoff, S. S. Concurrent responding with fixed relative rate of reinforcement. _Journal of the Experimental Analysis of Behavior_, 1969, _12_, 887-895.

Thorndike, E. L. _Animal intelligence_. New York: Macmillan, 1911.

Timberlake, W., & Allison, J. Response deprivation: an empirical approach to instrumental performance. _Psychological Review_, 1974, _81_, 146-164.

Vagners, J. Optimization techniques. In C. E. Pearson (Ed.), _Handbook of applied mathematics_. New York: Van Nostrand Reinhold, 1974.

Wilde, D. J. _Optimum seeking methods_. Englewood Cliffs: Prentice-Hall, 1964.

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.