*MOLAR OPTIMIZATION VERSUS DELAYED REINFORCEMENT
AS EXPLANATIONS OF CHOICE BETWEEN FIXED-RATIO
AND PROGRESSIVE-RATIO SCHEDULES*

JAMES E. MAZUR AND WILLIAM VAUGHAN, JR.

HARVARD UNIVERSITY

In a discrete-trials procedure, pigeons chose between a fixed-ratio 81 schedule and a progressive-ratio schedule by making a single peck at the key correlated with one or the other of these schedules. The response requirement on the progressive-ratio schedule began at 1 and increased by 10 each time the progressive-ratio schedule was chosen. Each time the fixed-ratio schedule was chosen, the requirement on the progressive-ratio schedule was reset to 1 response. In conditions where there was no intertrial interval, subjects chose the progressive-ratio schedule for an average of about five consecutive trials (during which the response requirement increased to 41), and then chose the fixed-ratio schedule. This ratio was larger than that predicted by an optimality analysis that assumes that subjects respond in a pattern that minimizes the response-reinforcer ratio or one that assumes that subjects respond in a pattern that maximizes the overall rate of reinforcement. In conditions with a 25-s or 50-s intertrial interval, subjects chose the progressive-ratio schedule for an average of about eight consecutive trials before choosing the fixed-ratio schedule. This change in performance with the addition of an intertrial interval was also not predicted by an optimality analysis. On the other hand, the results were consistent with the theory that choice is determined by the delays to the reinforcers delivered on the present trial and on subsequent trials.

Key words: fixed ratio, progressive ratio, optimization, delayed reinforcement, key peck, pigeons

In studies with primates, Hodos and Trumbule (1967) and Himeline and Sodetz (1987) studied choice between fixed-ratio (FR) and progressive-ratio (PR) schedules. In both studies, while the number of responses required for reinforcement was constant on the FR schedule, the response requirement on the PR schedule was increased by 20 each time the PR schedule was chosen. For instance, in one condition of the Himeline and Sodetz experiment, rhesus monkeys chose between FR 161 and a PR schedule that required 1 response the first time it was chosen following the FR schedule, 21 responses if chosen immediately again, 41 if chosen again without an intervening FR choice, and so on. The FR and PR schedules were each correlated with a specific response lever and discriminative stimulus (a red or yellow light above the lever). At the start of every trial, both schedules were available, and both stimulus lights were on. The

first response on either lever caused the stimulus light for the other schedule to go off, and the subject was then required to complete the ratio of the chosen schedule. After completion of the ratio and reinforcer delivery, the next trial began, with both schedules available once again.

In "no-reset" conditions, nothing could be done to shorten the PR schedule, which simply increased each time it was chosen throughout a session. The monkeys' choices in these conditions followed a simple pattern: They usually chose the PR schedule until its ratio was equal to that of the FR schedule, and then chose the FR schedule for the rest of the session. This pattern of choice minimized both the number of responses required on the current trial and the overall response-reinforcer ratio across the session.

Himeline and Sodetz observed a different pattern of choices in "reset" conditions, in which the ratio on the PR schedule was reset to 1 each time the FR schedule was chosen. In these conditions, the monkeys chose the FR schedule well before the response requirement on the PR schedule had increased to that of the FR schedule. For example, in the reset condition with FR 161, the monkeys made an

This research was supported by Grant MH 38357 from the National Institute of Mental Health. We thank David Coe for his help in conducting this experiment, and Elizabeth Gray for drawing the figures. Requests for reprints should be sent to James E. Mazur, 768 William James Hall, Harvard University, Cambridge, Massachusetts 02138.

average of three consecutive choices of the PR schedule, and then chose the FR 161 lever although the current response requirement on the PR lever was only 61 responses. Obviously, the monkeys were not simply choosing whichever schedule had the smaller response requirement on a given trial (as might be predicted, for example, by a theory such as molecular maximization; e.g., Hinson & Staddon, 1983). Hineline and Sodetz showed, however, that the monkeys' patterns of choices were advantageous in the long run. In the reset condition with FR 161, for instance, repeatedly following the pattern of three PR choices and one FR 161 choice yielded an overall response-reinforcer ratio of 56. In comparison, choosing the PR schedule for one, two, four, or five consecutive trials produced overall response-reinforcer ratios of 81, 61, 57, and 61, respectively. Hineline and Sodetz found that with all of the FR sizes they presented, the number of consecutive choices of the PR schedule was close to that which minimized the overall response-reinforcer ratio. They concluded that their results provided support for optimal foraging theory (e.g., Krebs, 1978), which predicts that animals will either maximize energy intake or minimize effort or time spent foraging. Lea (1981) has come to the same conclusion about the performance of chimpanzees in the similar study of Hodos and Trumbule (1967). A similar conclusion about performance on PR schedules has been reached by Bhatt and Wasserman (1987). Optimality analyses come in many different forms, but for now we will restrict the discussion to the specific theory presented by Hineline and Sodetz. These writers proposed that the controlling variable in their procedure was the overall response-reinforcer ratio, and that their subjects learned to produce the pattern of choices that minimized this ratio.

An entirely different approach to these results was proposed by Shull and Spear (1987). They suggested that in this type of situation a subject chooses between two series of delayed reinforcers, including the reinforcer delivered on the current trial and those on subsequent trials. Based on the earlier work of McDiarmid and Rilling (1965), they suggested that the influences of several delayed reinforcers are additive, but are inversely related to the delay between a choice response and each reinforcer (cf. Ainslie, 1975). The following

equation is the summation rule proposed by McDiarmid and Rilling:

$$V = \sum_{i=1}^n \frac{1}{D_i}, \quad (1)$$

where V is the value of an alternative, D_i is the delay between a choice response and the delivery of reinforcer i , and n is the number of reinforcers in the series. In a choice situation, it is assumed that the subject will choose whichever alternative has the higher value of V .

To illustrate how this model can be applied to the Hineline-Sodetz procedure, let us make a few simplifying assumptions. First, to translate ratio requirements into delays, assume that the subject responds at a constant rate of one response per second on all ratio schedules. Second, let $n = 4$; that is, assume that only the next four reinforcers influence a choice response. (The following analysis would be qualitatively the same if larger values of n were used.)

Suppose that in one condition of the Hineline-Sodetz procedure, the subject must choose between FR 161 and a PR schedule with a current response requirement of 101. If the subject chooses the PR schedule now and resets the ratio by choosing the FR schedule on the next trial, the next four response requirements would be 101, 161, 1, and 21, respectively. Another option would be to choose the FR schedule now and the PR schedule on the next three trials. This would involve response requirements of 161, 1, 21, and 41. The top two panels of Figure 1 illustrate the predictions of Equation 1. Each vertical line represents the value of one reinforcer, a value which is the reciprocal of its delay. In this example, the next four reinforcers have a higher total value (V in the figure) if the FR schedule is chosen on the current trial. Thus, like the optimality analysis, the delay analysis predicts that a subject will choose the FR schedule when the current PR requirement is still much smaller.

Hineline and Sodetz (1987) considered the delay-based interpretation, but favored an interpretation based on optimizing overall responses per reinforcer (p. 150). The obtained switching points were closer to values predicted by the overall optimization account than to values predicted by the delay approach. However, the predictions of these two approaches are often qualitatively similar, and it

is therefore fairly difficult to distinguish between them on the basis of small differences in predicted reset points. Furthermore, it is possible that the delay approach is essentially correct but that some equation other than Equation 1 provides a better description of the relation between delay and value (and better predictions for this choice situation). Instead of searching for small quantitative differences between the two approaches, the present experiment introduced a manipulation for which the two approaches make unambiguous and qualitatively different predictions. This was accomplished simply by inserting an intertrial interval (ITI) after each reinforcer.

Because an ITI has no effect on the response requirements of either schedule, it should be apparent that the Hineline and Sodetz (1987) analysis predicts exactly the same reset points with or without an ITI. For instance, if choosing the FR schedule on every fourth trial results in a response–reinforcer ratio of 56 when there is no ITI, the response–reinforcer ratio will still be 56 if a long ITI follows each reinforcer. However, in other applications of optimality theory, time has been treated as a critical variable, so it is important to consider how the addition of an ITI affects the predictions of theories that propose that animals in choice situations tend to maximize the overall rate of reinforcement (e.g., Charnov, 1976; Rachlin, Battalio, Kagel, & Green, 1981). It is fairly easy to show that these theories also predict that an ITI will have no effect on the reset point. As a concrete example, suppose that when there is no ITI, resetting the PR on every fourth trial produces an average of one reinforcer every 30 s, whereas resetting on every third trial or on every fifth trial produces one reinforcer every 40 s. A theory of reinforcement rate optimization would predict, of course, that the subject will reset the PR on every fourth trial. If a 100-s ITI is now added, resetting the PR on every third, fourth, or fifth trial will yield one reinforcer every 140 s, 130 s, or 140 s, respectively, so the optimal strategy remains the same. Thus regardless of whether the controlling variable is assumed to be the overall response–reinforcer ratio or the overall rate of reinforcement, an optimality analysis predicts that the addition of an ITI will have no effect on the reset point.

The predictions of the delay analysis are illustrated in the bottom two panels of Figure

Fig. 1. An illustration of the predictions of Equation 1 for two "reset" conditions involving choice between FR 161 and a PR schedule with a current response requirement of 101. As discussed in the text, this illustration assumes that the four reinforcers that follow a choice response contribute to V . It is also assumed that the subject will either choose the FR 161 schedule on the current trial or on the next trial. The curve in each panel plots the function described by Equation 1, and the vertical lines show the values and temporal locations of the four reinforcers. In the top two panels, there is no ITI, and V is greater for the choice of the FR 161 schedule. In the bottom two panels, there is a 100-s ITI, and V is greater for the choice of the PR schedule.

1. The situation depicted here is the same as in the top two panels except that a 100-s ITI occurs after each reinforcer. In this situation, the reinforcers delivered on subsequent trials have less influence on the choice response because the ITIs increase the delays to these reinforcers. The benefits of choosing the FR schedule and resetting the PR ratio are therefore not as great, and according to Equation 1, V is higher for the choice of the PR schedule

and its current response requirement of 101. In general, the delay analysis predicts that reset points should increase when an ITI is inserted.

In addition to using pigeons as subjects and including conditions with and without an ITI, the present experiment differed in one other respect from that of Hineline and Sodetz (1987). Without some discriminative stimulus, subjects might have greater difficulty remembering the current PR response requirement when an ITI was included, and this could affect their performance. To avoid this problem, in many conditions of the present experiment the current PR requirement was signaled by the number of lights above the PR response key that were illuminated on each trial.

METHOD

Subjects

The subjects were 3 White Carneau pigeons (numbers 268, 377, and 391) and 1 racing pigeon (number 331) maintained at 80% of their free-feeding weights. All had previous experience with a variety of experimental procedures.

Apparatus

Two experimental chambers were used, the first for Conditions 1 and 2 and the second for the remaining conditions. The first chamber was 33 cm long, 29 cm wide, and 30 cm high. There were two response keys, each 2 cm in diameter, 15 cm apart. A standard hopper was located below the keys. The second chamber was 29 cm long, 30 cm wide, and 30 cm high. Two response keys, each 2 cm in diameter, were mounted on the front wall of the chamber, 15.5 cm apart (center to center) and 24 cm above the floor.

For both chambers, each key could be trans-illuminated with either red or green light produced by two 2-W bulbs mounted behind the keys. A force of approximately 0.10 N was required to operate each key. When either key was illuminated, an effective peck on that key produced a feedback click; no feedback was delivered for a peck at a dark key. A standard hopper provided controlled access to mixed grain. Two 2-W white lights above the hopper were lit whenever grain was available. Two 2-W white houselights were mounted just be-

low the ceiling in the rear of the chamber. Each chamber was enclosed in a sound-attenuating box that contained an air blower for ventilation. A PDP-8® computer in another room used a SuperSKED® program to control the stimuli and record responses for all conditions.

In the second chamber only, a row of 20 stimulus lights, 25 cm in length, extended across the top of the front wall of the chamber, 1 cm from the ceiling. Each 1.1-W light had either a red or green translucent cover that was 0.8 cm in diameter. A strip of clear Plexiglas in front of the lights prevented a subject from touching the lights or dislodging their covers.

Procedure

Condition 1. On each trial, a subject first chose between an FR 81 schedule on the right (red) key and a PR schedule on the left (green) key. After the choice response, the subject was required to complete the ratio on the chosen key. The ratio requirement on the PR schedule was 1 response at the start of the session, but this requirement was increased by 10 responses each time the PR schedule was chosen without an intervening choice of the FR schedule. One choice of the FR 81 schedule always reset the requirement on the PR schedule to 1 response.

At the start of each trial, both keys were lit. After a single peck on either key, the other keylight was turned off, and the subject was required to complete the ratio schedule on the chosen key. When the ratio was completed, the keylight was turned off, and grain was presented for 3 s. After reinforcement the next trial began. A session ended after 60 trials or 100 min, whichever came first.

Condition 2. Unlike Condition 1, after reinforcement there was an ITI of 50 s during which only the houselights were on, and then the next trial began.

Condition 3. This and subsequent conditions were run in the second chamber described above. The 12 stimulus lights farthest to the left had green covers, and the other eight lights had red covers. The white houselights were lit throughout each session except during reinforcement.

The only difference between this condition and Condition 2 was that at the start of each trial, both keys were lit, along with some of the lights above the keys. Because the schedule on the right (red) key was FR 81, the eight

red lights above this key were lit (one light for every 10 responses required after the choice response). Similarly, one green light above the left (green) key was lit for every 10 postchoice responses that were required on the PR schedule on that trial. On the first trial of the session, no green lights were lit, because the reinforcer was delivered immediately after the choice response if the PR key were chosen. On a trial after the PR schedule had been chosen once, one green light (the one farthest to the left) was lit, because 10 postchoice responses were required if the PR schedule were chosen again. Similarly, the two green lights farthest to the left were lit when 20 postchoice responses were required, and so on. If the ratio requirement on the PR key were 120 or higher, all 12 green lights were lit.

After a single peck on either key the other keylight was turned off, as were the stimulus lights associated with the other schedule, and the subject was required to complete the ratio schedule on the chosen key. When the ratio was completed the keylight and stimulus lights were turned off, and grain was presented for 3 s. After reinforcement, there was an ITI of 50 s during which only the houselights were on, and then the next trial began.

Conditions 4 through 9. These conditions were identical to Condition 3, with the following exceptions. In Conditions 4, 6, and 8, there was no ITI—a trial began immediately after reinforcement for the previous trial. In Condition 5, the ITI was 25 s. Condition 9 had a 50-s ITI, but regardless of the current PR requirement, at each choice point the eight green lights farthest to the left were lit, along with the eight red lights associated with the FR schedule. This condition was included to determine how subjects would perform under a long ITI when there were no discriminative stimuli indicating the current response requirement on the PR schedule. Condition 7 was included to test for color and position biases. The PR schedule was available on the right (red) key, and the FR schedule on the left (green) key. The eight lights farthest to the left had green covers and the other 12 had red covers. As in the other conditions, one light was lit for every 10 postchoice responses required by each schedule. Table 1 summarizes the procedures for each condition.

Conditions 1 and 3 through 5 lasted for a minimum of 20 sessions, and subsequent con-

Table 1

Order of experimental conditions and number of sessions per condition.

Con- di- tion	Procedure	Subject			
		268	377	391	331
		Number of sessions			
1	0-s ITI, no stimulus lights	25	20	26	20
2	50-s ITI, no stimulus lights	21	21	18	20
3	50-s ITI	20	22	24	23
4	0-s ITI	23	20	25	20
5	25-s ITI	30	20	25	20
6	0-s ITI	15	17	20	15
7	0-s ITI, colors and keys reversed	21	22	20	26
8	0-s ITI	23	17	23	26
9	50-s ITI, unchanging stimulus lights	24	17	19	25

ditions lasted for a minimum of 15 sessions. Except in Condition 2, each condition was terminated for each subject individually when the following criteria were met. For each session, the average ratio requirement a subject completed on the PR schedule before switching to the FR 81 schedule was calculated. For convenience this number will be called the *mean reset ratio*. The condition was terminated when (a) none of the mean reset ratios in the last five sessions were the highest or lowest of the condition, excluding the first five sessions of the condition, and (b) the mean reset ratio across the last five sessions was within 10% of the mean reset ratio of the previous five sessions. Because we initially did not plan to use the data from the first two conditions, the stability criteria were not initially applied, and they could be applied retrospectively in Condition 1 but not in Condition 2. Condition 2 was terminated at an arbitrary point, and the data analyses were based on the last five sessions for each subject.

RESULTS

The right side of Table 1 lists the number of sessions each subject received in each condition. All analyses were based on the last five sessions of every condition.

Figure 2 presents the mean reset ratios in all conditions for each subject. The group medians are also shown. The results from Conditions 3 through 6 and 8 are most relevant to

Fig. 2. The mean reset ratios (i.e., average ratio requirement a subject completed on the progressive-ratio schedule before switching to the FR 81 schedule) from each condition are plotted for each subject. The group medians are also shown.

the distinction between the two theoretical positions described in the Introduction, because in these five conditions the stimulus lights were operational and the PR schedule was presented on the left key. The points from Conditions 4, 6, and 8, which had a 0-s ITI, are connected by solid lines. Solid lines also connect the results from Conditions 3 and 5, which had ITIs of 50 s and 25 s, respectively. Figure 2 shows that the mean reset ratios were consistently larger in the two conditions with non-zero ITIs. For every subject, the mean reset ratios were larger in Conditions 3 and 5 than in Conditions 4, 6, and 8.

For 3 subjects, reset ratios were larger in Condition 7 than in the other three conditions

with 0-s ITIs. This may reflect a bias for the right, red key, because the keys and colors for the two schedules were reversed in this condition. In the remaining three conditions (1, 2, and 9), no stimulus lights signaled the current size of the PR schedule. Despite the absence of these discriminative stimuli, the mean reset ratios in Condition 1, which had a 0-s ITI, were comparable to those from the other three conditions with a 0-s ITI and the PR schedule on the left key. Similarly, mean reset times in Conditions 2 and 9, which included a 50-s ITI, did not differ systematically from those of Condition 3, the other condition with a 50-s ITI. Although the mean reset ratios revealed no large differences between the con-

0-s ITI

50-s ITI

50-s ITI, UNCHANGING LIGHTS

Fig. 3. Cumulative records for Subject 268 from single complete sessions in Conditions 4, 3, and 9, respectively. The recording pen was returned to the x -axis after each reinforcer. The filled circles above each record indicate trials on which the FR 81 schedule was chosen. In the 50-s ITI conditions, the cumulative recorder continued to run during the ITI.

Fig. 4. Frequency distributions of reset ratios (the ratio requirement a subject completed on the PR schedule before switching to the FR 81 schedule) from the last five sessions of every condition are shown for each subject. The last bin in each panel includes all reset ratios of 161 or higher.

ditions with the stimulus lights and those without them, these stimuli had substantial effects on trial-to-trial consistency, as will be shown below.

Figure 3 presents three typical cumulative records for Subject 268 for one complete session from Conditions 4, 3, and 9, respectively. The recording pen was returned to the *x*-axis after each reinforcer. The filled circles above each record indicate trials on which the FR 81 schedule was chosen. In the 50-s ITI conditions, the cumulative recorder continued to run during the ITIs. The top record shows that in the 0-s ITI conditions, responding was rapid and steady, with preratio pauses preceding some of the longer ratios. The size of the maximum PR ratio varied from one series of PR trials to the next, but the subject almost always switched to the FR 81 schedule before the PR requirement reached 81. The middle

record shows that with a 50-s ITI the subject chose the PR schedule until its ratio was about the same as on the FR 81 key, and again the maximum ratio varied from one series to the next. The bottom record shows how this subject's performance became more variable in the condition with the unchanging lights. In this session, the subject switched to the FR 81 schedule after a PR ratio as low as 41 on two occasions, but at the end of the session the subject had increased the PR ratio to 141. This session also included some long pauses at or near the beginning of a new ratio.

Figure 4 presents more systematic information about how reset ratios varied across series of trials within a condition. For each subject, this figure displays the relative frequency of each reset ratio in the last five sessions of every condition. One important comparison is between the distributions from

conditions with 0-s ITIs (top row of panels) and from those of the conditions with 25-s and 50-s ITIs (second row). The stimulus lights were operational in these five conditions. The main difference between the distributions in the top two rows is that with nonzero ITIs, the peaks of the distributions are shifted to higher reset ratios. The overall shapes of these distributions are similar to those in the 0-s ITI conditions, with the exception of Subject 391 in the 25-s ITI condition. This suggests that the larger reset ratios in Conditions 3 and 5 were not due to weaker stimulus control (which might have occurred because an ITI separated one trial from the next).

The third row of panels in Figure 4 presents the frequency distributions from the two conditions that had 50-s ITIs and either no stimulus lights (Condition 2) or unchanging stimulus lights (Condition 9). A comparison of these distributions with those in the second row shows how the absence of these discriminative stimuli affected performance. Six of the eight distributions in the third row are broad and flat with little or no indication of a peak, which shows that there was great variability in the reset ratios from one series of trials to the next. The two exceptions are Subjects 391 and 331 in Condition 9. For these two, there was much less variability in reset ratios in Condition 9 than in Condition 2. This indicates that with a 50-s ITI and without the discriminative stimuli, the current size of the PR schedule exerted more control over the performance of these 2 subjects at the end of the experiment than at the beginning.

The frequency distributions from Condition 1 (bottom row) show that the absence of the stimulus lights had much less of an effect on performance when there was no ITI. The trial-to-trial variability in reset ratios was, on the whole, just slightly greater than in Conditions 4, 6, and 8.

DISCUSSION

Two aspects of the present findings relate to optimization theory. First, performance in the 0-s ITI conditions of this experiment deviated substantially from the predictions of the optimization approach. Whereas a reset ratio of 21 would have minimized the response-reinforcer ratio, the mean reset ratio in Conditions 2, 4, and 6 was 46.2. With this reset

ratio, the response-reinforcer ratio was about 13% larger than it would have been with a reset ratio of 21. The behavior of these pigeons was therefore quite different from that of the rhesus monkeys in the Hineline and Sodetz (1987) study, whose reset ratios were nearly optimal across a wide range of FR requirements. Although there were many procedural differences between the two studies, it is possible that these two species behave differently in this choice situation, at least at a quantitative level.

Second, when conditions with and without an ITI were compared, the reset ratios were larger in the former than in the latter. This result is also inconsistent with the optimality analysis of Hineline and Sodetz (1987), which holds that subjects select a reset ratio minimizing the response-reinforcer ratio. This theory cannot explain the present results because adding an ITI does not affect the response-reinforcer ratio. The results are also inconsistent with the hypothesis that subjects choose the reset ratio that maximizes the overall rate of reinforcement, because, as shown in the Introduction, inserting an ITI does not change the reset ratio that maximizes the rate of reinforcement. In sum, the present results provide no evidence that choice was controlled by the overall response-reinforcer ratio or by the overall rate of reinforcement.

On the other hand, the results of both the zero and nonzero ITI conditions are consistent with an analysis of FR-PR choice proposed by Shull and Spear (1987). In essence, this theory states that choice is determined by the delays to the reinforcers on the current trial and on a number of subsequent trials. According to this theory, the addition of ITIs produced longer reset ratios because the time between trials increased the delays to all subsequent reinforcers, thereby diminishing their impact on a subject's choice behavior.

Advocates of the optimization approach might argue that the addition of ITIs made it more difficult for subjects to remember the current PR response requirement, and that the larger reset ratios were somehow caused by these memory deficits. Stated in different terminology, this argument might go as follows: With nonzero ITIs, the current PR requirement exerted less control over a subject's behavior because the discriminative stimulus signaling the current response requirement (i.e.,

the size of the previous ratio) was further removed in time. In the present experiment the stimulus lights above the response keys appear to have minimized any such loss of stimulus control. The frequency distributions of reset ratios in Figure 4 provide the clearest evidence on this point. If stimulus control were weaker in conditions with nonzero ITIs, this would have produced greater variability in reset ratios within a condition, which would be reflected in broader and flatter frequency distributions. However, a comparison of the results in the top two rows of Figure 4 shows that, with one exception, the distributions were similar in shape, differing primarily in their modes (as well as their means, see Figure 2), not their variability. The individual differences in frequency distributions from the conditions in which the stimulus lights did not signal the current ratio size are not particularly relevant to this theoretical issue, except insofar as they show how a weakening of stimulus control can affect performance in this situation. The results suggest that the stimulus lights exerted considerable control over behavior with a 50-s ITI, but probably less control with a 0-s ITI, where performance with and without the lights was quite similar.

The reset ratio predicted by the delay approach, as expressed in Equation 1, depends on the number of subsequent reinforcers that are included in the calculations. The predicted reset ratio is 51 if two reinforcers are considered (the one on the present trial and on the next trial), 41 if between 3 and 42 reinforcers are considered, and 31 if more than 42 trials are considered. It is not clear how many reinforcers should be included in the analysis, but because all of these predictions are greater than 21, they are closer to the obtained results than those of the optimization approach. A problem with Equation 1, however, is that because its predicted reset ratios are usually larger than that predicted by optimization, it cannot predict the optimal performance that has been observed in some instances of FR-PR choice (e.g., Hineline & Sodetz, 1987). It should be emphasized, however, that the limitations of this specific equation are not necessarily damaging to the general approach proposed by Shull and Spear (1987). A similar equation, but one with greater flexibility in accommodating different types of performance, has been used by Mazur (1984, 1986; Mazur, Snyder-

man, & Coe, 1985) with considerable success:

$$V = \sum_{i=1}^n \frac{1}{1 + KD_i}, \quad (2)$$

where K is a free parameter that specifies how sharply value decreases with increases in delay. The reset ratios predicted by Equation 2 depend on the value of K . With K equal to 1.0, the predictions of Equation 2 are similar to those of Equation 1 for the present experiment. As K approaches zero, however, the predictions of this equation converge on those of the optimization approach. Thus according to Equation 2, optimal performance is a special case that will occur if delay has very little effect on a reinforcer's value.

Although the results from 0-s ITI conditions are inconsistent with the optimization approach and consistent with Equation 2, they are not necessarily predicted by the latter, because there are free parameters. However, the increase in the reset ratio in the nonzero ITI conditions is not only inconsistent with optimization, but both consistent with and predicted by Equation 2.

A somewhat different approach, which is in qualitative agreement with the results presented here, has been suggested by Vaughan (1985). In that model, although the value of an alternative is a function of successive events, the events do not act in parallel, as Equations 1 and 2 imply. Rather, the value of one reinforcer is itself increased if it is followed by another reinforcer. In this approach, then, events concatenate in a serial manner. To our knowledge no explicit test of a parallel versus a serial model of multiple future events has been reported.

Two simple versions of optimization theory, which assume that subjects either minimize the response-reinforcer ratio or maximize the overall rate of reinforcement, cannot account for the results of this study. We have focused on these two theories both because previous writers have explicitly advocated them (e.g., Charnov, 1976; Hineline & Sodetz, 1987; Lea, 1981; Rachlin et al., 1981) and because their predictions for the present experiment are unambiguous. We do not mean to suggest, however, that the results are incompatible with all possible theories that assume some sort of optimization. Indeed, the theory of delayed reinforcement we have described can be called

an optimality analysis because it assumes that the subject will choose whichever alternative has the greater value at the moment of choice. Furthermore, it might well be possible to start with a theory like that of Hineline and Sodetz, then add a temporal discounting function and thereby develop a theory that is compatible with our results (cf. Rachlin et al., 1981, p. 387). We are aware of no attempts to develop such a theory that are precise enough to make concrete predictions for FR-PR choice or similar situations. We suspect, however, that if a temporal discounting function is added to an optimality analysis, the theory may be indistinguishable from the general delayed-reinforcement approach we have presented. Whether or not such a theory should still be called an optimality theory is a matter of semantics, but it should be clear that neither the total food per response nor the total food per unit time is optimized under this theory.

REFERENCES

- Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, **82**, 463-496.
- Bhatt, R. S., & Wasserman, E. A. (1987). Choice behavior of pigeons on progressive and multiple schedules: A test of optimal foraging theory. *Journal of Experimental Psychology: Animal Behavior Processes*, **13**, 40-51.
- Charnov, E. L. (1976). Optimal foraging: The marginal value theorem. *Theoretical Population Biology*, **9**, 129-136.
- Hineline, P. N., & Sodetz, F. J. (1987). Appetitive and aversive schedule preferences: Schedule transitions as intervening events. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 141-157). Hillsdale, NJ: Erlbaum.
- Hinson, J. M., & Staddon, J. E. R. (1983). Matching, maximizing, and hill-climbing. *Journal of the Experimental Analysis of Behavior*, **40**, 321-331.
- Hodos, W., & Trumbule, G. H. (1967). Strategies of schedule preference in chimpanzees. *Journal of the Experimental Analysis of Behavior*, **10**, 503-514.
- Krebs, J. R. (1978). Optimal foraging: Decision rules for predators. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (pp. 23-63). Sunderland, MA: Sinauer.
- Lea, S. E. G. (1981). Correlation and contiguity in foraging behaviour. In P. Harzem & M. D. Zeiler (Eds.), *Advances in analysis of behaviour: Predictability, correlation, and contiguity* (pp. 355-406). Chichester, England: Wiley.
- Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. *Journal of Experimental Psychology: Animal Behavior Processes*, **10**, 426-436.
- Mazur, J. E. (1986). Fixed and variable ratios and delays: Further tests of an equivalence rule. *Journal of Experimental Psychology: Animal Behavior Processes*, **12**, 116-124.
- Mazur, J. E., Snyderman, M., & Coe, D. (1985). Influences of delay and rate of reinforcement on discrete-trial choice. *Journal of Experimental Psychology: Animal Behavior Processes*, **11**, 565-575.
- McDiarmid, C. G., & Rilling, M. E. (1965). Reinforcement delay and reinforcement rate as determinants of schedule preference. *Psychonomic Science*, **2**, 195-196.
- Rachlin, H., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology. *Behavioral and Brain Sciences*, **4**, 371-417. (Includes commentary)
- Shull, R. L., & Spear, D. J. (1987). Detention time after reinforcement: Effects due to delay of reinforcement? In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 187-204). Hillsdale, NJ: Erlbaum.
- Vaughan, W., Jr. (1985). Choice: A local analysis. *Journal of the Experimental Analysis of Behavior*, **43**, 383-405.

Received August 20, 1986
Final acceptance May 12, 1987