CONDITIONED REINFORCEMENT AND CHOICE WITH DELAYED AND UNCERTAIN PRIMARY REINFORCERS

JAMES E. MAZUR

SOUTHERN CONNECTICUT STATE UNIVERSITY

In an adjusting-delay choice procedure, pigeons could peck on either a red key or a green key. A peck on the red key always led to a delay associated with red houselights and then food. The delay was adjusted over trials to estimate an indifference point—a delay at which the two keys were chosen about equally often. In some conditions, a peck on the green key led to food on all trials after delays of either 10 s or 30 s, and green houselights were lit during the delays. In other conditions, food was presented on only half of the green-key trials. If the green houselights continued to occur on both reinforcement and nonreinforcement trials, preference for the green key always decreased. Preference for the green key also decreased if half of the trials had 30-s houselights followed by food and the other half had no green houselights and no food. However, preference for the green key actually increased if half of the trials had 10-s green houselights followed by food and the other half had no green houselights followed by no food. The latter condition therefore demonstrated a case in which preference for an alternative increased when food was removed from half of the trials. The results suggest that the red and green houselights served as conditioned reinforcers. A hyperbolic decay model (Mazur, 1989) provided good predictions for all conditions by assuming that the strength of a conditioned reinforcer is inversely related to the total time spent in its presence before food is delivered.

Key words: choice, probability of reinforcement, delay of reinforcement, conditioned reinforcement, key peck, pigeons

A reinforcer's effectiveness decreases systematically as the delay between response and reinforcer increases. In a series of experiments with animal subjects in choice situations, my colleagues and I have tried to determine, in quantitative terms, the relation between reinforcer delay and reinforcer effectiveness (e.g., Mazur, 1984, 1987; Mazur, Snyderman, & Coe, 1985). The results of these experiments can be predicted quite accurately by assuming that the following equation describes the relation between a reinforcer's delay, D , and the reinforcer's value, V (its effectiveness in sustaining choice responses):

$$V = \frac{1}{1 + KD}, \quad (1)$$

where K is a free parameter that determines how rapidly V declines with increasing values of D . Because Equation 1 describes a hyper-

bola, this account of the effects of reinforcer delay has been called the *hyperbolic decay model*.

For cases in which the delay to a reinforcer varies from trial to trial, the hyperbolic decay model assumes that the overall value of that alternative is an average of the values associated with the different possible delays:

$$V = \sum_{i=1}^n P_i \left(\frac{1}{1 + KD_i} \right). \quad (2)$$

V is now the value of a reinforcer delivered after a variable delay, where there are n different possible delays to reinforcement. P_i is the probability that a delay of D_i seconds will occur on any given trial. Mazur (1984) obtained fairly good predictions if K was set equal to 1, and this value of K will be used throughout this paper.

To see how this model can be used to make specific predictions, suppose that in a choice situation, one alternative is a mixed-time (MT) 10-s 30-s schedule, in which a choice response is followed by a delay of either 10 s or 30 s, and then food is delivered. If either delay is equally likely to occur, then P_i equals .5 for each, and solving Equation 2 yields a value of $V = .062$. Equation 1 can now be used to predict the duration of a fixed-time (FT)

This research was supported by Grant MH 38357 from the National Institute of Mental Health. I thank Jeffrey Colafrancesco, Peter Kelly, and Margaret Nygren for their help in various phases of the research. Correspondence concerning this article should be sent to James E. Mazur, Psychology Department, Southern Connecticut State University, New Haven, Connecticut 06515.

schedule that has the same value as MT 10 s 30 s. If $V = .062$, solving Equation 1 yields $D = 15.1$ s, which implies that FT 15.1 s has the same value as MT 10 s 30 s. These two schedules should therefore be equally preferred in a choice situation.

To test such predictions, Mazur (1984) used an adjusting-delay procedure, in which subjects chose between a standard alternative and an adjusting alternative. The standard alternative delivered food on a schedule (e.g., an MT schedule) that remained constant throughout an experimental condition. The adjusting alternative delivered food after a delay that was adjusted over trials so as to estimate an indifference point—a delay at which both alternatives were chosen about equally often. Using the adjusting-delay procedure, Mazur (1984) obtained indifference points for a variety of different MT and FT schedules, and the results coincided closely with the predictions of the hyperbolic decay model. This model has also been successfully applied to the results from a number of other studies with both animal and human subjects (e.g., Grossbard & Mazur, 1986; Mazur, 1986; Rachlin, Raineri, & Cross, 1991; Rodriguez & Logue, 1988).

In my earlier research on this topic, I assumed that V represents the value of the delayed primary reinforcer (food), and that conditioned reinforcers played a negligible role in the adjusting-delay procedure.

However, the results of these experiments are equally consistent with the view that V represents the value of whatever conditioned reinforcers precede the delivery of food. In all my studies with the adjusting-delay procedure, colored keylights and houselights were associated with the standard and adjusting alternatives. For instance, a green key might represent the standard alternative, and a peck on the green key would lead to a delay during which the houselights were green, and at the end of the delay, food would be presented. (The houselights were always white before a choice response was made and during the intertrial intervals.) In a similar way, a red keylight and red houselights might precede the delivery of food for the adjusting alternative. The successful predictions of the hyperbolic decay model could therefore be interpreted as showing that the value of a conditioned reinforcer decreases according to a hyperbolic

function as the duration of that conditioned reinforcer increases.

My more recent work on probabilistic reinforcers (reinforcers delivered with a probability of less than 1) strongly suggests that V in the hyperbolic decay model is best interpreted as the value of the conditioned reinforcers that precede and predict the primary reinforcer, rather than the value of the primary reinforcer itself. A brief description of this research will show why. Extrapolating from the suggestion of Rachlin, Logue, Gibbon, and Frankel (1986) that probabilistic reinforcers are analogous to delayed reinforcers, Mazur (1989) proposed that probabilistic reinforcers can be treated as reinforcers delivered after variable delays. For example, suppose one alternative in a choice situation delivers food with a probability of .2 after a 10-s delay. Basic rules of probability show that such a reinforcer will occur after one 10-s delay with a probability of .2, after two 10-s delays with a probability of .16, after three 10-s delays with a probability of .128, and so on. These probabilities can be used in Equation 2 to calculate the overall value of a reinforcer that is sometimes delivered after 10 s of delay, sometimes after 20 s of delay, sometimes after 30 s of delay, and so on.

Mazur (1989) showed that this approach provided reasonable predictions for choices involving delayed probabilistic reinforcers, but only if D_i in Equation 2 was calculated in a specific way. The predictions were fairly accurate if D_i included the time spent in the presence of the colored keylights and colored houselights associated with the standard and adjusting alternatives, but excluded the durations of the intertrial intervals (ITIs). If ITI durations were included in the calculation of D_i , the predictions were grossly inaccurate. This was the first piece of evidence that conditioned reinforcers played an important role: The time spent in the presence of these colored stimuli was a good predictor of the pigeons' choice behavior, whereas the total time between a choice response and the delivery of the primary reinforcer turned out to be a very poor predictor. A second piece of evidence was that varying the ITI across conditions had no measurable effect on choice. This also suggested that ITI duration should not be included in the equation.

Probably the most compelling evidence for

the role of conditioned reinforcers has come from situations in which the durations of the colored houselights were directly manipulated and the subjects' indifference points changed accordingly. For example, in one condition of Mazur (1989), the standard alternative was food delivered with a probability of .2 after a 5-s delay. Red houselights were present during the 5-s delays on both reinforcement and nonreinforcement trials, and a mean indifference point of 16.6 s was obtained (suggesting that this probabilistic alternative was equally preferred to a reinforcer delivered with certainty after 16.6 s). In another condition, however, the red houselights were omitted on nonreinforcement trials. That is, a peck on the red key led either to red houselights for 5 s and then food or to white houselights and no food. In this condition, the mean indifference point was 7.2 s, indicating a substantially greater preference for the probabilistic alternative when the red houselights were omitted on nonreinforcement trials. (In the adjusting-delay procedure, smaller indifference points reflect greater preference for the standard alternative.) This change in preference was predicted by Equation 2, because removing the red houselights substantially reduced the durations of D_i . Mazur and Romano (1992) used a somewhat different procedure to show the opposite effect: When the durations of the colored houselights associated with a probabilistic reinforcer were increased, preference for that alternative decreased. Similar results have been obtained in studies using concurrent-chains or discrete-trial procedures (Dunn & Spetch, 1990; Spetch, Belke, Barnett, Dunn, & Pierce, 1990; Spetch, Mondloch, Belke, & Dunn, 1994). These studies found that choice responding for an alternative that delivered 50% reinforcement was greater if that alternative had different delay stimuli on reinforcement and nonreinforcement trials (i.e., if the stimulus that preceded food was not present on nonreinforcement trials).

To summarize, this application of the hyperbolic decay model assumes that the strength of a conditioned reinforcer is inversely related to the total amount of time spent in its presence before the primary reinforcer is delivered. Notice, however, that the calculation of V is complicated by at least two factors. First, multiple values of D_i are needed in the case of probabilistic reinforcers, because the primary rein-

forcer might be delivered on the first trial or only after one or more nonreinforcement trials. Second, for those times when the primary reinforcer is not delivered on the first trial, D_i does not represent the duration of one continuous presentation of the conditioned reinforcer, but is instead the total duration of two or more presentations of this stimulus, each separated by an ITI (and sometimes by one or more trials on which the other alternative was chosen). Neither the ITIs nor the trials with the other alternative are included in D_i . A similar method for calculating the total "duration" of a stimulus that is only sometimes followed by reinforcement has been used in Gibbon's scalar expectancy theory (Gibbon & Balsam, 1981). This method was adopted by the hyperbolic decay model because it provided accurate predictions for the results of several experiments, whereas other approaches to the problem did not (see Mazur, 1989, 1991, 1993).

The present experiment was designed to make further tests of this model of conditioned reinforcement by combining two features used separately in previous studies with the adjusting-delay procedure—MT schedules and probabilistic reinforcers. Some conditions of this experiment involved a choice between MT 10 s 30 s (with green houselights during the delays) and an adjusting delay (with red houselights). As explained previously, with $K = 1$, Equation 2 predicts an indifference point of approximately 15 s. In other conditions, only half of the trials with the MT schedule delivered food. For example, in one condition food was delivered after the 30-s delay but not after the 10-s delay. In another condition, food was delivered after the 10-s delay but not after the 30-s delay. The hyperbolic decay model predicts that removing half of the reinforcers will have a large effect on choice in the first condition but only a minor effect in the second. More interesting cases were those in which both the food and the green houselights were removed for either the 10-s or the 30-s delay components. The model predicts that preference should decrease if the 10-s green houselights and subsequent food are omitted (and replaced with simply the white houselights of the ITI). However, if instead the 30-s green houselights and subsequent food are omitted, the model predicts that subjects will show greater preference for this alternative than for the original MT schedule that delivered a re-

inforcer on every trial. Thus the model makes the counterintuitive prediction that removing half of the reinforcers (and the stimuli that precede them) can sometimes lead to an increase in preference. The model makes this prediction because removing the 30-s green houselights produces substantial reductions in D_i that more than offset the absence of food on half of the trials. The experiment also included a few other combinations of stimuli and 50% reinforcement, and for each condition the actual indifference points were compared to the specific, quantitative predictions of the model.

METHOD

Subjects

Four White Carneau pigeons were maintained at about 80% of their free-feeding weights. All had previous experience with a variety of experimental procedures.

Apparatus

The experimental chamber was 30 cm long, 30 cm wide, and 33 cm high. Three response keys, each 1.8 cm in diameter, were mounted in the front wall of the chamber, 20.5 cm above the floor. A force of approximately 0.1 N was required to operate each key, and each effective response produced a feedback click. Each key could be transilluminated with lights of different colors. A hopper below the center key provided controlled access to grain, and when grain was available, the hopper was illuminated with a 2-W white light. Eight 2-W lights (two white, two red, two green, and two orange) were mounted above the wire-mesh ceiling of the chamber. The chamber was enclosed in a sound-attenuating box containing a ventilation fan. All stimuli were controlled and responses recorded by an IBM®-compatible personal computer using the Medstate® programming language.

Procedure

The experiment consisted of 10 conditions that differed only in the reinforcement schedule and stimuli used as the standard alternative. The procedure for Conditions 1 and 5 (identical conditions) will be described in detail, and then the variations made in subsequent conditions can be described more briefly.

Conditions 1 and 5. Each session lasted for 64 trials or for 60 min, whichever came first.

Each block of four trials consisted of two forced trials followed by two choice trials. Figure 1 illustrates the possible sequences of events that could occur on a choice trial in these conditions, which featured an MT 10-s 30-s schedule as the standard alternative. After an ITI during which only the white houselights were lit, the center key was illuminated with white light to start a trial. A single peck on the center key was required to begin the choice period. The purpose of this center peck was to make it more likely that the subject's head was equidistant from the two side keys when the choice period began. On choice trials, a peck on the center key darkened this key and illuminated both the left green key and the right red key. The green key was always the standard-delay key, and the red key was always the adjusting-delay key.

A peck on the green key extinguished both side keys and initiated a delay of either 10 s or 30 s, during which the green houselights were lit. The delays were chosen on a random basis, with the constraint that every 16 trials with the standard alternative included eight instances of each delay. At the end of the standard delay, the green houselights were extinguished, and grain was presented for 3 s.

A peck on the red key during the choice period extinguished both side keys and initiated the adjusting delay, during which the red houselights were lit. The duration of the adjusting delay was changed over trials, as explained below. At the end of the adjusting delay, the red houselights were extinguished, and grain was presented for 3 s. For both the standard and adjusting alternatives, only the white lights above the grain were lit during reinforcement periods. After each reinforcement period, there was an ITI during which only the white houselights were lit. The duration of the ITI was adjusted so that the total time from the choice response to the end of the trial was 50 s in all conditions, for both the standard and adjusting alternatives.

The procedure on forced trials was the same as on choice trials, except that only one side key was lit, red or green, and a peck on this key led to the appropriate delay. A peck on the opposite key, which was dark, had no effect. Of every two forced trials, one involved the red key and the other involved the green key. The temporal order of these two types of trials varied randomly.

After every two choice trials, the delay for the adjusting key might be changed. If a subject chose the adjusting key on both choice trials, the adjusting delay was increased by 1 s. If the subject chose the standard key on both trials, the adjusting delay was decreased by 1 s. If the subject chose each key on one trial, no change was made in the adjusting delay. In all three cases, this adjusting delay remained in effect for the next block of four trials. At the start of each session, the adjusting delay was determined by the above rules as if it were a continuation of the preceding session.

To facilitate presentation and discussion of the results, each condition will be given an abbreviation that indicates, for the standard alternative, the houselight colors, the delay durations, and whether food was delivered (+) or not delivered (-). Conditions 1 and 5 will therefore be referred to as G10+,G30+ because both the 10-s and 30-s green houselights were followed by food.

Other conditions. In all conditions, the number and organization of trials were the same as in Condition 1, as was the adjusting alternative. The only differences involved the stimuli and schedules of reinforcers for the standard alternative. For each condition, Figure 2 diagrams the different possible sequences of events that could follow a peck on the green key. Each of the sequences depicted for a given condition occurred with equal probability.

Condition G20+ (Condition 8) was included to measure any bias that might be due to key color, position, or a preference for or against the adjusting alternative. The standard alternative was fixed-time (FT) 20 s: a 20-s delay with green houselights followed by food. If there were no bias, the mean adjusting delay should also be about 20 s.

In each of the remaining seven conditions, the overall probability of reinforcement for the standard alternative was .5 (i.e., half of the standard trials ended in reinforcement, and half did not). In Condition G10+,G30- (Condition 3), only 10-s delays (signaled by the green houselights) were followed by food. The other half of the standard trials had 30-s delays with green houselights, but no food was presented. In Condition G10+,O30- (Condition 10), green houselights were present during the 10-s (reinforcement) delays, but orange houselights were present during the 30-s (nonreinforcement) delays. In Condition G10+,

Fig. 1. The sequences of events that could occur on a choice trial in Conditions 1 and 5, depending on whether the red or green key was pecked during the choice period.

W50- (used in both Conditions 4 and 7), green houselights were again present during the 10-s (reinforcement) delays, but only white houselights were present on nonreinforcement trials. That is, on nonreinforcement trials, a peck on the green key led immediately to the white houselights of the ITI, which lasted for the full 50 s in this case. This set of conditions was designed to examine the effects of presenting different stimuli on nonreinforcement trials.

In Condition G30+,G10- (Condition 2), only 30-s delays (signaled by the green houselights) were followed by food. The other half of the standard trials had a 10-s delay with green houselights, but no food was presented. In Condition G30+,W50- (Condition 6), green houselights were again present during

Fig. 2. For each condition, the different possible sequences of events that could occur after the standard (green) key was pecked. For each condition, the different sequences shown occurred with equal probability.

the 30-s (reinforcement) delays, but only white houselights were present on nonreinforcement trials.

Finally, Condition G10+/-, G30+/- (Condition 9) included delays of 10 s and 30 s, both signaled by the green houselights, but in this case each delay duration had a reinforcement probability of .5. That is, each of the four sequences shown in Figure 2 was

scheduled to occur on one fourth of the standard trials.

Criteria for terminating conditions. Condition 1 lasted for a minimum of 20 sessions, and all other conditions lasted for a minimum of 14 sessions. After the minimum number of sessions, a condition was terminated for each subject individually when several stability criteria were met. To assess stability, each session was divided into two 32-trial blocks, and for each block the mean delay on the adjusting key was calculated. The results from the first two sessions of a condition were not used, and the condition was terminated when all of the following criteria were met, using the data from all subsequent sessions: (a) Neither the highest nor the lowest single-block mean of a phase could occur in the last six blocks of a condition. (b) The mean adjusting delay across the last six blocks could not be the highest or the lowest six-block mean of the condition. (c) The mean delay of the last six blocks could not differ from the mean of the preceding six blocks by more than 10% or by more than 1 s (whichever was larger). (d) Visual inspection showed no large and systematic trends over the last six sessions.

RESULTS

The number of sessions needed to satisfy the stability criteria ranged from 14 to 24 (median = 17). For each subject and condition, the mean adjusting delay from the six half-session blocks that satisfied the stability criteria was used as an estimate of the indifference point. Table 1 presents the indifference points for each subject and condition. A repeated measures analysis of variance found a significant effect of condition, $F(9, 27) = 28.92, p < .001$. Planned comparisons were then used to test the predictions of the hyperbolic decay model, as described below.

In Condition 8, which used an FT 20-s schedule as the standard alternative, the indifference points should be close to 20 s if there was no response bias. For the 4 subjects, indifference points ranged from 15.9 s to 20.3 s and averaged 18.7 s, possibly indicating a slight bias toward the standard alternative for the group as a whole. This possible bias should be taken into account when interpreting the results from other conditions.

Figure 3 presents the indifference points in

Fig. 3. For each subject, mean adjusting delays (indifference points) are shown for seven different stimulus arrangements used in the experiment. The data are averaged across replications for conditions that occurred twice.

graphical form for each subject, averaged across replications where applicable. This figure shows that the results from the different conditions were usually similar for the 4 subjects. Indifference points from Condition G10+, G30+ were all less than 16 s, although the average delay for the standard alternative was 20 s. This result is an example of the "preference for variability" that has been found in

numerous other experiments (e.g., Davison, 1969; Fantino, 1967; Mazur, 1984; Rider, 1983).

The next three bars in each panel of Figure 3 show the results from the three types of conditions that always included green house-lights for the standard alternative, but only half of the standard trials ended in reinforcement. For all three conditions, the hyperbolic

Table 1
Order of conditions and mean adjusting delays for each subject and the group mean.

Order	Condition	Mean adjusting delay (seconds)				
		Bird 1	Bird 2	Bird 3	Bird 4	Mean
1	G10+,G30+	13.7	13.2	14.5	15.4	14.2
2	G30+,G10-	32.8	40.2	39.1	28.7	35.2
3	G10+,G30-	13.6	21.1	27.7	8.4	17.7
4	G10+,W50-	6.4	10.5	16.4	4.8	9.5
5	G10+,G30+	16.9	14.7	15.7	9.2	14.1
6	G30+,W50-	26.5	25.2	28.5	24.4	26.2
7	G10+,W50-	7.3	6.6	11.8	7.3	8.2
8	G20+	20.2	20.2	18.3	15.9	18.7
9	G10+/-,G30+/-	17.6	22.6	24.6	19.3	21.0
10	G10+,O30-	8.8	6.3	14.3	7.2	9.1

decay model predicts that indifference points should be larger than in the condition with 100% reinforcement (indicating a decreased preference for the standard alternative). However, the model predicts clear differences among these three conditions: Indifference points should be longest in Condition G30+,G10-, intermediate in Condition G10+/-,G30+/-, and shortest in Condition G10+,G30-. Figure 3 shows that this prediction was supported by the results from individual subjects, except for one reversal for Bird 3. A planned comparison found a significant linear trend among these three conditions in the predicted direction, $F(1, 27) = 61.71, p < .01$.

Although the indifference points were similar across subjects in most conditions, this was not the case in Condition G10+,G30-, where they ranged from 8.4 s to 27.7 s. A possible reason for this variability among subjects will be suggested in the Discussion section.

The next two bars in each panel of Figure 3 show the results from the 50% reinforcement conditions in which only white houselights were present on nonreinforcement trials. The hyperbolic decay model predicts that indifference points from these conditions should be smaller than those from the corresponding conditions that included green houselights on nonreinforcement trials (thus showing increased preference for the standard alternative when the green houselights were omitted). Figure 3 shows that this prediction was supported for all 4 subjects in both of the relevant comparisons. That is, indifference points were always smaller in Condition G30+,W50- than in Condition G30+,G10-, and were always smaller in Condition G10+,W50- than in Condition G10+,G30-. More important, the model predicts that indifference points in Condition G10+,W50- should actually be smaller than those in Condition G10+,G30+ (thus showing an increase in preference for the standard alternative when half of the primary reinforcers and the associated stimulus lights were removed). Figure 3 shows that the results from every subject were in the predicted direction, although the size of the difference varied among subjects. A planned comparison found a significant difference when comparing the two replications of Condition G10+,W50- with the two replications of G10+,G30+, $F(1, 27) = 11.33, p < .01$.

The remaining bars in each panel of Figure

3 show the indifference points from Condition G10+,O30-, in which 30-s orange houselights occurred on the nonreinforcement trials. Because these orange houselights were never paired with food, the hyperbolic decay model does not treat them as conditioned reinforcers, and the durations of these stimuli are therefore not included in D_i in Equation 2. The model predicts that indifference points in this condition should be the same as those in Condition G10+,W50-, and therefore that this condition should also produce smaller indifference points than the condition with 100% reinforcement (G10+,G30+). The indifference points were smaller in the G10+,O30- condition for all 4 subjects, although they were only slightly smaller for Subject 3. A planned comparison found a significant difference between the G10+,O30- condition and the two G10+,G30+ conditions, $F(1, 27) = 6.83, p < .05$.

Figure 4 compares the group means to the quantitative predictions of Equation 2. To derive these predictions, K was set equal to 1, so the predictions were made without free parameters. In previous studies of this type (Mazur, 1989, 1991; Mazur & Romano, 1992), time spent in the presence of the green and red keylights (i.e., the time between onset of these keylights and the single peck that extinguishes them) was also treated as part of D_i . This was also done for the present experiment, but for simplicity, the calculations assumed a response latency of 1 s for each key, because the mean response latencies were approximately 1 s. If actual response latencies were used, the differences in the predictions of Equation 2 would be negligible. Similarly, excluding response latencies from D_i altogether had little effect for this set of data, but actually yielded slightly better predictions than those shown in Figure 4.

Figure 4 shows a fairly close correspondence between the predictions and the group means. The predictions of the model accounted for 91.8% of the variance in the group means. Although this percentage represents a good fit to the data, the actual indifference points were slightly smaller than predicted in all seven comparisons. This systematic discrepancy could be due to a slight bias toward the standard alternative. As noted above, the mean indifference point in the G20+ condition was 18.7 s, which is 6.5% shorter than a bias-free indifference point of 20 s. If this difference re-

Fig. 4. Group means of the indifference points are shown for seven different stimulus arrangements used in the experiment. The data are averaged across replications for conditions that occurred twice. The horizontal lines above each bar are the predictions of Equation 2 with $K = 1$.

flects a bias of 6.5% toward the standard alternative, we can compensate for the bias by reducing each predicted indifference point by 6.5%. If this is done, the correspondence between predicted and obtained indifference points improves, and the predictions now account for 97.5% of the variance in the group means.

DISCUSSION

This experiment found that a model of conditioned reinforcer value could provide accurate, parameter-free predictions for choices involving MT schedules and probabilistic reinforcement under a variety of different stimulus and reinforcement arrangements. When every standard trial ended in reinforcement after an average of 20 s (Condition G10+,G30+), subjects exhibited the pre-

dicted preference for variability, with indifference points averaging about 14 s. When food was then omitted on 50% of the standard trials, indifference points increased substantially in some conditions but actually decreased in others, depending on which trials had the food removed and whether or not the green houselights were removed as well. When the green houselights remained illuminated on nonreinforcement trials, there was a large decrease in preference for the standard alternative when only the 30-s houselights were followed by food (G30+,G10) but there was a much smaller decrease in preference when only the 10-s houselights were followed by food (G10+,G30-). The hyperbolic decay model predicted this result because in the first case the green houselights were always present for at least 30 s before food was delivered, but in the second case food was sometimes delivered after only 10 s of green houselights.

In studies with FT schedules, Mazur (1989, 1991) found that preference for a probabilistic reinforcer increased if the stimuli normally presented during the delay (colored houselights) were omitted on nonreinforcement trials. The present experiment extended this result to MT schedules: In both Conditions G30+,W50- and G10+,W50-, indifference points were significantly shorter (indicating greater preference) than those in the corresponding conditions with the green houselights present on nonreinforcement trials. This effect was found for every subject in every relevant comparison. In addition, the use of MT schedules made it possible to test the counterintuitive prediction that in certain cases, subjects should show greater preference for 50% reinforcement than for 100% reinforcement. In Condition G10+,W50-, only half of the standard trials ended with food, but this food was always delivered after relatively short exposure to the green stimuli. In Condition G10+,G30+, food was presented on every standard trial, but sometimes only after a long (30-s) exposure to the green houselights. The hyperbolic decay model predicted greater preference for the 50% alternative, because although the total time between food presentations was longer, the time spent in the presence of the green stimuli was shorter. This prediction was also confirmed, because indifference points were significantly shorter in Condition G10+,W50-.

The hyperbolic decay model predicts the same degree of preference for the 50% alternative when some other stimulus is presented on nonreinforcement trials. This prediction was supported by the results from Condition G10+,O30-, which included 30-s orange houselights on nonreinforcement trials. Indifference points in this condition were also significantly shorter than those in Condition G10+,G30+, thus providing a second example of greater preference for 50% reinforcement than for 100% reinforcement. This result shows that simply including some unique stimulus after a choice response does not have the same effect as presenting the conditioned reinforcer—the stimulus that precedes food on reinforcement trials. In this case, the orange houselights could not be considered conditioned reinforcers because they were never paired with food; thus, their durations were not included as part of D_i in Equation 2. If their durations had been included, an incorrect prediction would result.

The idea that the strength of a conditioned reinforcer varies inversely with its duration is not new. A number of other theories of choice include similar assumptions about conditioned reinforcers (e.g., Fantino, 1977; Killeen, 1982; Squires & Fantino, 1971; Vaughan, 1985). In fact, Vaughan's (1985) model uses a hyperbolic equation similar to Equation 1 to calculate the value of a conditioned reinforcer. An important difference, however, is that both Vaughan's model and Fantino's delay-reduction theory use only the mean delay to primary reinforcement in calculating the effect of a conditioned reinforcer. In contrast, the hyperbolic decay model calculates separate values for each different possible delay and averages these values to determine the total strength of a conditioned reinforcer. This strategy allows the model to account for preference for variable delays (Mazur, 1984) and for the relatively small impact of nonreinforcement trials with probabilistic reinforcers (Mazur, 1991).

Another difference between this model and delay-reduction theory is that, in the case of probabilistic reinforcers, D_i is calculated as the total duration of a conditioned reinforcer from the first choice response until the primary reinforcer is delivered, and excludes the durations of ITIs and other events that occur when the conditioned reinforcer is not present. This method of calculating D_i is not very different from methods commonly used to analyze the results from other choice procedures. For example, in a changeover-key procedure, one response key (the schedule key) alternates between two different colors, each associated with a different reinforcement schedule. Each response on a second key (the changeover key) changes the color and reinforcement schedule in effect on the schedule key. A subject might change the color of the schedule key several times before a reinforcer is delivered. Yet in analyzing such data, researchers typically sort responses, time, and reinforcers separately for each key color, thus assuming (with good reason) that these two sets of events exert separate control over the subject's behavior. Likewise, the present model assumes that times spent in the presence and absence of conditioned reinforcers have very different effects on a subject's choice behavior. This assumption did not originate from some a priori theoretical position but rather from experimental results that seem to demand it.

One other important difference between de-

lay-reduction theory and the hyperbolic decay model concerns the role of "context." Delay-reduction theory assumes that the strength of a conditioned reinforcer depends on the context in which the stimulus occurs (the average time to reinforcement for all choice alternatives). In its present form, the hyperbolic decay model affords no role to context: ITI duration is excluded from all calculations, and the duration of one alternative has no effect on the value of another.

One reason the hyperbolic decay model has been able to avoid the issue of context is that it has been applied primarily to procedures with very brief "initial links"—choice phases involving a single response. Thus, for the present experiment, D , was calculated as the total duration of the colored keylights (the initial links) and the colored houselights (the terminal links). This strategy was used primarily for convenience, and, as already mentioned, the predictions of the model would not be substantially changed if keylight durations were excluded altogether. Although this strategy was sufficient for the present experiment and for previous studies (Mazur, 1989, 1991; Mazur & Romano, 1992), it will need modification if the model is to be applied to a broader range of choice situations. For example, how would choice behavior be affected if the keylights and houselights were different colors, and how would the hyperbolic decay model take such factors into account? How can the model be applied to situations with longer initial links, such as the variable-interval schedules commonly used in concurrent-chains procedures? As it has been developed so far, the hyperbolic decay model makes no specific predictions for these cases.

Another unresolved issue concerns the evidence from several studies showing that animals may choose a delayed alternative delivering food on 50% of the trials over an equally delayed alternative delivering food on 100% of the trials (Dunn & Spetch, 1990; Kendall, 1974, 1985; Spetch et al., 1994). This preference for the 50% alternative is observed only if different stimuli are presented during the delays that end with food and those that do not. Spetch and her colleagues have proposed a variation of delay-reduction theory that accounts for this effect by considering the different contexts in which the conditioned reinforcers are presented for the 50% and 100% alternatives. However, the hyperbolic decay

model does not predict a preference for the 50% alternative in this situation—it predicts either indifference or a slight preference for the 100% alternative.

A final unresolved issue is whether the same physical stimulus can serve as a conditioned reinforcer at some times but not at other times. For instance, in Condition G10+,G30–, the green houselights were presented for 30 s on nonreinforcement trials. For the first 10 s of green houselights, there was no way to tell whether or not food would be delivered on that trial. But if food was not delivered after 10 s of green houselights, it would not be delivered at all. Subjects could have learned to discriminate between the first 10 s of green houselights (which were sometimes followed by food) and the remaining 20 s (which were never followed by food). Is it possible that the first 10 s of green houselights served as a conditioned reinforcer, whereas the remaining 20 s did not? The large individual differences in this condition suggest the possibility that 2 subjects learned this discrimination but the other 2 did not. The shorter indifference points for Subjects 1 and 4 in this condition might have occurred because the 30-s green houselights did not have their full detrimental effect on the value of the standard alternative. Not much can be concluded from the results of this single condition, but similar intersubject variability has been observed in conditions of previous studies in which a putative conditioned reinforcer was extended beyond the time when there was any possibility that the primary reinforcer would be delivered (Mazur, 1989, 1991).

If it turns out that the same stimulus can serve as a conditioned reinforcer at some times but not at others, this will add additional complexity to the analysis of conditioned reinforcer strength. It will also be difficult to predict the behavior of individual subjects. Fortunately, this ambiguity arises only in cases in which a temporal discrimination can be learned, and this was possible only in Condition G10+,G30– of the present experiment. For all other conditions, the predictions of the hyperbolic decay model were straightforward, and the actual results were more uniform across subjects. Whether this model offers the best approach to the problem of conditioned reinforcer strength remains to be seen, but it does make several assertions that appear to be supported by the data. First, the strength of a

conditioned reinforcer is inversely related to its duration. Second, the strength of a conditioned reinforcer depends not simply on its mean duration but on the trial-to-trial variability in the delay to primary reinforcement that its presence signals. Third, at least in some cases, time spent in the absence of the conditioned reinforcer has little or no effect on choice.

REFERENCES

- Davison, M. C. (1969). Preference for mixed-interval versus fixed-interval schedules. *Journal of the Experimental Analysis of Behavior*, 12, 247-252.
- Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: Conditioned reinforcement effects. *Journal of the Experimental Analysis of Behavior*, 53, 201-218.
- Fantino, E. (1967). Preference for mixed- versus fixed-ratio schedules. *Journal of the Experimental Analysis of Behavior*, 10, 35-43.
- Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313-339). Englewood Cliffs, NJ: Prentice-Hall.
- Gibbon, J., & Balsam, P. (1981). Spreading activation in time. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 219-253). New York: Academic Press.
- Grossbard, C. L., & Mazur, J. E. (1986). A comparison of delays and ratio requirements in self-control choice. *Journal of the Experimental Analysis of Behavior*, 45, 305-315.
- Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, 21, 463-473.
- Kendall, S. B. (1985). A further study of choice and percentage reinforcement. *Behavioural Processes*, 10, 399-413.
- Killeen, P. R. (1982). Incentive theory: II. Models for choice. *Journal of the Experimental Analysis of Behavior*, 38, 217-232.
- 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). Choice between single and multiple delayed reinforcers. *Journal of the Experimental Analysis of Behavior*, 46, 67-77.
- Mazur, J. E. (1987). An adjusting procedure for studying delayed 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. 55-73). Hillsdale, NJ: Erlbaum.
- Mazur, J. E. (1989). Theories of probabilistic reinforcement. *Journal of the Experimental Analysis of Behavior*, 51, 87-99.
- Mazur, J. E. (1991). Choice with probabilistic reinforcement: Effects of delay and conditioned reinforcers. *Journal of the Experimental Analysis of Behavior*, 55, 63-77.
- Mazur, J. E. (1993). Predicting the strength of a conditioned reinforcer: Effects of delay and uncertainty. *Current Directions in Psychological Science*, 2, 70-74.
- Mazur, J. E., & Romano, A. (1992). Choice with delayed and probabilistic reinforcers: Effects of variability, time between trials, and conditioned reinforcers. *Journal of the Experimental Analysis of Behavior*, 58, 513-525.
- 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.
- Rachlin, H., Logue, A. W., Gibbon, J., & Frankel, M. (1986). Cognition and behavior in studies of choice. *Psychological Review*, 93, 33-45.
- Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, 55, 233-244.
- Rider, D. P. (1983). Preference for mixed versus constant delays of reinforcement: Effect of probability of the short, mixed delay. *Journal of the Experimental Analysis of Behavior*, 39, 257-266.
- Rodriguez, M. L., & Logue, A. W. (1988). Adjusting delay to reinforcement: Comparing choice in pigeons and humans. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 105-117.
- Spetch, M. L., Belke, T. W., Barnett, R. C., Dunn, R., & Pierce, W. D. (1990). Suboptimal choice in a percentage-reinforcement procedure: Effects of signal condition and terminal-link length. *Journal of the Experimental Analysis of Behavior*, 53, 219-234.
- Spetch, M. L., Mondloch, M. V., Belke, T. W., & Dunn, R. (1994). Determinants of pigeons' choice between certain and probabilistic outcomes. *Animal Learning & Behavior*, 22, 239-251.
- Squires, N., & Fantino, E. (1971). A model for choice in simple concurrent and concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior*, 15, 27-38.
- Vaughan, W. (1985). Choice: A local analysis. *Journal of the Experimental Analysis of Behavior*, 43, 383-405.

Received June 27, 1994

Final acceptance October 19, 1994