CHOICE WITH DELAYED AND PROBABILISTIC REINFORCERS: EFFECTS OF PREREINFORCER AND POSTREINFORCER STIMULI

JAMES E. MAZUR

SOUTHERN CONNECTICUT STATE UNIVERSITY

In Experiment 1, pigeons' pecks on a green key led to a 5-s delay with green houselights, and then food was delivered on 20% (or, in other conditions, 50%) of the trials. Pecks on a red key led to an adjusting delay with red houselights, and then food was delivered on every trial. The adjusting delay was used to estimate indifference points: delays at which the two alternatives were chosen about equally often. Varying the presence or absence of green houselights during the delays that preceded possible food deliveries had large effects on choice. In contrast, varying the presence of the green or red houselights in the intertrial intervals had no effects on choice. In Experiment 2, pecks on the green key led to delays of either 5 s or 30 s with green houselights, and then food was delivered on 20% of the trials. Varying the duration of the green houselights on nonreinforced trials had no effect on choice. The results suggest that the green houselights served as a conditioned reinforcer at some times but not at others, depending on whether or not there was a possibility that a primary reinforcer might be delivered. Given this interpretation of what constitutes a conditioned reinforcer, most of the results were consistent with the view that the strength of a conditioned reinforcer is inversely related to its duration.

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

---

Two factors that are known to affect behavior in choice situations are delay of reinforcement and probability of reinforcement. Experiments with both humans and nonhumans have shown that as the delay between a response and reinforcement increases, preference for that response decreases (e.g., Ainslie, 1974; Green, Fisher, Perlow, & Sherman, 1981; Green, Fry, & Myerson, 1994). Similarly, as the probability of reinforcement for a particular response decreases, preference for that response also decreases (Battalio, Kagel, & McDonald, 1985; Rachlin, Logue, Gibbon, & Frankel, 1986; Rachlin, Raineri, & Cross, 1991; Waddington, Allen, & Heinrich, 1981). A series of experiments with pigeons (e.g., Mazur, 1984, 1987, 1989, 1991, 1995) showed that the effects of both delay and probability of reinforcement could be well described by the following equation:

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

V is the value of a reinforcement schedule composed of *n* different possible delays to reinforcement, where value refers to the schedule's ability to sustain choice responses. P_i is the probability that a delay of D_i seconds will occur on any given trial. *K* is a free parameter that determines how rapidly *V* declines with increasing values of D_i. Because the parenthetical expression in Equation 1 describes a hyperbola, this equation has been called the *hyperbolic decay model.*

The application of Equation 1 to cases involving two delayed reinforcers is fairly straightforward. For example, suppose one response leads to food after a variable delay of either 2 s or 18 s (with each delay occurring on a random half of the trials). If *K* is set equal to 1 (a value that typically provides fairly accurate predictions for pigeons), solving the equation provides a value of *V* = .193. Now suppose we want to predict the duration of a single fixed delay that would be equally preferred to this variable delay. If only one delay is involved, Equation 1 reduces to

$$V = \frac{1}{1 + KD}. \quad (2)$$

Setting *V* equal to .193 and solving for *D* yields a value of 4.2 s, so these equations predict that a reinforcer delivered after variable delay of 2 or 18 s should be equally preferred

to a reinforcer delivered after a fixed delay of 4.2 s. On a qualitative level, this prediction is consistent with the results of studies that have found preference for variable over fixed delays (Cicerone, 1976; Rider, 1983). On a quantitative level, Mazur (1984) used an adjusting-delay procedure to test these equations more rigorously. In this procedure, pigeons chose between a standard alternative (a schedule of delays to reinforcement that remained the same throughout an experimental condition) and an adjusting alternative (a delay that increased or decreased many times per session, depending on the subject's choices). The purpose of these adjustments was to measure indifference points: delays at which the standard and adjusting alternatives were chosen about equally often. The predictions of the hyperbolic decay model proved successful for a wide range of conditions involving choices between fixed and variable delays to reinforcement.

The application of the hyperbolic decay model to cases involving probabilistic reinforcers (reinforcers that are delivered on only a certain percentage of the trials) is a bit more complex, and Mazur (1989, 1991, 1995) found that successful predictions depend on D_i being measured in a specific way. Fairly accurate predictions were obtained if D_i was interpreted as the total duration of the conditioned reinforcers that were presented between an initial choice response and the eventual delivery of the primary reinforcer, food. For example, in one experiment with pigeons (Mazur, 1989), each choice of a red key (the standard alternative) led to a 5-s delay with red houselights, and then food was presented on 20% of the trials. Each choice of a green key (the adjusting alternative) led to an adjusting delay with green houselights, and then food was presented on 100% of the trials. Mazur assumed that the red and green keylights and houselights were conditioned reinforcers because they preceded the delivery of food.

Equation 1 was used to predict the indifference points in this experiment as follows. The red keylight was typically lit for about 1 s per trial, so on about 20% of the trials, D_i was equal to 6 s, because the red stimuli were present for a total of 6 s before food was presented (1 s with a red keylight and 5 s with red houselights). In other cases (on about 16% of the trials), food was presented after the second choice of the red key, and in these cases, D_i was equal to 12 s (two trials with a 1-s red keylight followed by 5 s with red houselights). Similarly, for cases with three, four, or more trials between food reinforcers, the appropriate values of D_i and P_i were used in Equation 1, and a value of V was then calculated in the same way as it was done for variable delays to reinforcement. Mazur (1989) found that this method of calculating the value of a probabilistic reinforcer provided fairly accurate predictions for the results of several experiments on choice with delayed and probabilistic reinforcers.

Because Equation 1 is only accurate if D_i measures the time spent in the presence of the colored keylights and houselights, it may provide a method for calculating the strength of a conditioned reinforcer (Mazur, 1991, 1993). In essence, the equation states that the strength of a conditioned reinforcer is inversely related to the total time spent in its presence before a primary reinforcer is delivered (even if this time is distributed over several trials). Additional studies found further support for the view that the duration of the conditioned reinforcers is a critical factor. One strong piece of evidence came from studies which showed that preference for the probabilistic reinforcer could be increased simply by decreasing the time spent in the presence of the conditioned reinforcers. For example, in the third experiment of Mazur (1989), the procedure was the same as described above, except that the 5-s red houselights were presented only on the 20% of the trials that were followed by food; on nonreinforced trials, a choice of the standard key led simply to the white houselights of the intertrial interval (ITI). The mean adjusting delay was about 7 s in these conditions, compared to 17 s in conditions that had the 5-s red houselights on both reinforced and nonreinforced trials. In this procedure, shorter adjusting delays indicate greater preference for the standard alternative, so the pigeons showed a much stronger preference for the standard alternative when the total durations of the red houselights were reduced, as predicted by Equation 1 (even though the timing and probability of the primary reinforcer, food, were unchanged).

Later studies confirmed the importance of

conditioned reinforcers in other choice situations involving delayed and probabilistic reinforcers, and they provided further support for the predictions of the hyperbolic decay model (Mazur, 1991, 1995). The model's predictions are also generally consistent, on a qualitative level, with results of similar studies that used concurrent-chains procedures to measure choice (e.g., Dunn & Spetch, 1990; Spetch, Belke, Barnet, Dunn, & Pierce, 1990; Spetch, Mondloch, Belke, & Dunn, 1994). However, the evidence for this approach to conditioned reinforcement has been less convincing when the durations of the putative conditioned reinforcers were increased rather than decreased. For instance, Mazur (1991) used the same general procedures as those just described, but included a condition in which the red houselights were lit for 60 s (rather than 5 s) on nonreinforced trials. According to Equation 1, this large increase in the duration of the red houselights should have decreased preference for the standard alternative, resulting in longer adjusting delays. The adjusting delay did increase for 2 of 4 pigeons, but remained roughly unchanged for the other 2. Using a different procedure, Mazur and Romano (1992) also obtained mixed results when they increased the durations of the colored houselights on nonreinforced trials: There was some decrease in preference for the probabilistic alternative, but not as much as predicted by the hyperbolic decay model, and there was considerable variability among subjects.

In summary, Equation 1 seems to make fairly accurate predictions when distinctive delay-interval stimuli are present on both reinforced and nonreinforced trials and when the distinctive stimuli are omitted on nonreinforced trials, but not when the durations of the distinctive stimuli are increased on nonreinforced trials. Given these results, Mazur (1991, 1995) suggested that some subjects may learn a temporal discrimination in cases in which the distinctive stimuli are extended beyond the time when a primary reinforcer might be delivered. For instance, when the red houselights were present for 5 s on reinforced trials but for 60 s on nonreinforced trials, the first 5 s of red houselights (which were sometimes followed by food) may have become distinct from longer durations of houselights (which were never followed by food). For birds that learned this discrimination, perhaps the first 5 s of red houselights served as a conditioned reinforcer but longer durations of red houselights did not. These birds might show no change in preference from conditions with shorter red-houselight periods. In contrast, if time since the onset of the red houselights exerted no control over the behavior of other subjects, Equation 1 suggests that these subjects would show a decreased preference for the probabilistic reinforcer.

Regardless of whether or not this analysis of previous results has merit, it should be clear that the effects of stimuli presented during and after delays to primary reinforcers are not well understood. The present experiments were designed to investigate further the roles of distinctive stimuli presented at various times during reinforced and nonreinforced trials. In some conditions, colored houselights were present before food but not after, and in other conditions they were present both before food and during the subsequent ITIs. In some conditions, the colored houselights were present at times when no food was ever delivered, and in other conditions there was always a possibility that the colored houselights might be followed by food. The experiments were designed to test the predictions of the hyperbolic decay model, and, more generally, to determine when the presence of a distinctive stimulus will and will not affect preference for a delayed probabilistic reinforcer.

EXPERIMENT 1

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, including adjusting-delay procedures similar to the one used in the present experiment.

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 approxi-

mately 0.15 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. Six 2-W lights (two white, two red, two green) 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

Throughout the experiment, each session lasted for 64 trials or for 60 min, whichever came first. Each block of four consecutive trials consisted of two forced trials followed by two choice trials. At the start of each trial, the center key was illuminated with white light, and the white houselights remained on. A single peck on the center key, positioning the subject's head roughly equidistant from the two side keys, was required to begin the choice period. On choice trials, a peck on the center key darkened this key and illuminated the two side keys, the left key green and the right key red.

A single peck on an illuminated side key constituted a choice response. The green key always served as the standard key and the red key as the adjusting key. If the red key was pecked during the choice period, both keylights were extinguished and the adjusting delay began, during which the red houselights were lit. The adjusting delay was always followed by 3-s access to grain (with all houselights off) and then an ITI with white houselights (except in Conditions 2, 3, and 11, in which red houselights were present during the ITI, as explained below). Throughout the experiment, the duration of the ITI was adjusted so that the total time from a choice response to the start of the next trial (including the reinforcer duration) was 40 s.

A peck on the green key during the choice period served as a choice of the standard alternative, and the consequences of this choice varied across conditions. As with choices of the adjusting alternative, the total time from a standard choice response to the start of the next trial was 40 s in all conditions. The experiment included five different stimulus arrangements that were used in one, two, or three conditions each, for a total of 12 conditions. To make the explanation of these different stimulus arrangements easier, the procedures used in the green-white conditions will be described in detail, and then the remaining procedures can be described more briefly.

Green-white conditions (Conditions 1, 4, and 10). Each choice of the green key led to a 5-s delay with green houselights, and then food was presented for 3 s on either 20% of the trials (Conditions 1 and 10) or 50% of the trials (Condition 4). A 32-s ITI with white houselights followed each food presentation. On no-food trials, the 5-s period with green houselights was immediately followed by a 35-s ITI with white houselights. The top panel in Figure 1 shows the sequences of events for both food and no-food trials in these conditions.

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

After every four-trial block, the delay for the adjusting key was increased by 1 s (up to a maximum possible duration of 35 s) if the adjusting key was chosen on both choice trials in the preceding block, decreased by 1 s (down to a minimum possible duration of 0 s) if the standard key was chosen on both choice trials, and remained unchanged if each key was chosen once. In all three cases, this adjusting delay remained in effect for the next block of four trials. For the first session of each condition, the adjusting delay began at 0 s. At the start of all subsequent sessions, the adjusting delay was determined by the above rules as if it were a continuation of the preceding session.

No-food-white condition (Condition 9). As shown in the second panel of Figure 1, this condition was identical to the 20% green-white conditions, except that 5-s green houselights were omitted on no-food trials. This
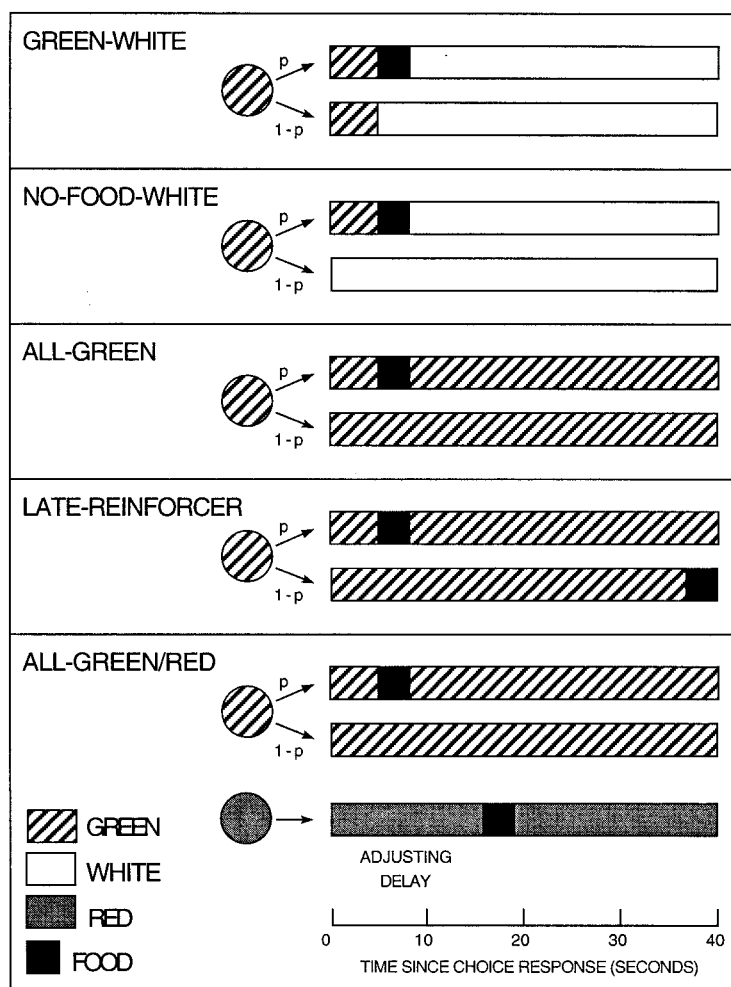
Fig. 1. For each of the five types of conditions in Experiment 1, the consequences of a peck on the green (standard) key are shown. p represents the proportion of trials on which food was delivered after a 5-s period with green houselights. For the all-green/red conditions, the consequences of a peck on the red (adjusting) key are also shown.

condition was similar to procedures used in several previous experiments (Mazur, 1989, 1991), and based on the previous results, the omission of the green houselights on no-food trials should produce a large increase in preference for the standard alternative.

All-green conditions (Conditions 5, 6, and 12). As shown in the third panel of Figure 1, these conditions were identical to the green-white conditions, except that green houselights were present instead of white houselights in the ITIs for the standard choice. Thus, the green houselights were present throughout each trial with the standard alternative, except during reinforcer deliveries. If the value

of the standard alternative is inversely related to the durations of the green stimuli, these conditions should produce a dramatic decrease in preference for the standard alternative compared to the green-white conditions (unless subjects acquired a temporal discrimination, as described in the introduction). Food was delivered on 20% of the trials with the standard alternative in Conditions 5 and 12 and on 50% of the standard trials in Condition 6.

Late-reinforcer conditions (Conditions 8 and 9). The fourth panel in Figure 1 shows that these conditions were similar to the all-green conditions, except that on trials when no food

was delivered after a 5-s delay, food was delivered at the end of the trial (i.e., after a 37-s delay with green houselights). These conditions were included to address the possibility that in the all-green conditions, subjects might learn a discrimination between the first 5 s of green houselights (which were sometimes followed by food) and longer houselight durations (which were never followed by food). Although such a discrimination might be learned in the all-green conditions, it should not be learned in the late-reinforcer conditions, because food was often delivered after a longer period of green houselights. Food was delivered after a 5-s delay on 20% of the trials with the standard alternative in Condition 8 and on 50% of the standard trials in Condition 9.

All-green/red conditions (Conditions 2, 3, and 11). As shown in Figure 1, the choices of the standard alternative were identical to those of the all-green conditions (with green houselights present throughout a trial except during reinforcement periods). However, these conditions were different from all other conditions of the experiment because the red houselights were presented both before and after the food on choices of the adjusting alternative. That is, the red houselights were present throughout the trial, except during reinforcement periods. If the value of the adjusting alternative is inversely related to the durations of the red houselights, these conditions might show a decrease in preference for the adjusting alternative compared to the all-green conditions. Food was delivered on 20% of the trials with the standard alternative in Conditions 2 and 11 and on 50% of the standard trials in Condition 3.

Stability criteria. All 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 a condition was terminated when the following three criteria were met, using the data from all subsequent sessions: (a) Neither the highest nor the lowest single-block mean of a condition could occur in the last six blocks of the condition. (b)

Table 1

Order of conditions, mean adjusting delays (in seconds), and sessions to meet stability criteria (in parentheses) for each subject in Experiment 1.

Order	Condition	Mean adjusting delay (sessions to stability)			
		Bird 1	Bird 2	Bird 3	Bird 4
1	20% green-white	22.6 (15)	16.2 (16)	12.5 (19)	16.7 (14)
2	20% all-green/red	24.2 (37)	16.3 (19)	17.5 (18)	23.3 (16)
3	50% all-green/red	7.1 (15)	7.9 (19)	7.7 (25)	10.2 (15)
4	50% green-white	13.7 (19)	11.9 (15)	6.5 (18)	11.2 (16)
5	20% all green	19.5 (15)	19.0 (16)	11.5 (18)	17.8 (14)
6	50% all green	7.1 (16)	11.1 (24)	6.7 (16)	10.4 (19)
7	50% late reinforcer	8.7 (15)	8.5 (14)	5.5 (18)	11.2 (17)
8	20% late reinforcer	15.8 (17)	15.6 (14)	22.3 (14)	17.2 (15)
9	20% no-food white	11.5 (17)	5.9 (16)	7.7 (28)	9.9 (19)
10	20% green-white	23.6 (15)	13.7 (20)	13.5 (16)	17.8 (19)
11	20% all-green/red	15.5 (20)	13.0 (21)	16.6 (16)	17.0 (16)
12	20% all green	16.8 (16)	11.3 (17)	11.6 (19)	17.9 (14)

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).

RESULTS AND DISCUSSION

For all conditions, the mean adjusting delay from the six half-session blocks that satisfied the stability criteria was used as a measure of the indifference point. The mean adjusting delays and the number of trials needed to satisfy the stability criteria are presented in Table 1 for all subjects and all conditions.

Figure 2 also shows the indifference points from the different conditions, except that the means of the two replications are shown for conditions that occurred twice. The error bars show one standard deviation above and below each mean, calculated using the means from the six half-session blocks that met the
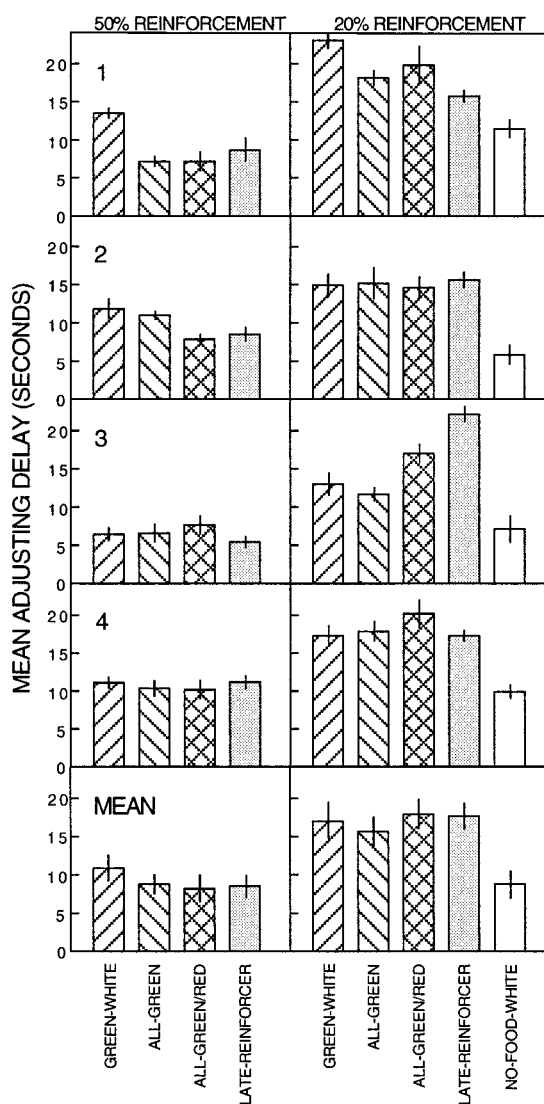
Fig. 2. Mean adjusting delays are shown for each subject and for the group in each type of condition in Experiment 1. For conditions that were repeated twice, the means of the two replications are shown. The error bars represent one standard deviation above and below each mean. The results from the 50% reinforcement conditions are presented on the left, and those from the 20% conditions are on the right.

stability criteria in each condition. The most obvious finding was that indifference points were consistently longer in the 20% reinforcement conditions. A series of planned comparisons found significant differences between all 20% reinforcement conditions and their corresponding 50% conditions: For green-white conditions, $F(1, 33) = 14.40$; for

all-green conditions, $F(1, 33) = 17.51$; for all-green/red conditions, $F(1, 33) = 31.13$; and for late-reinforcer conditions, $F(1, 33) = 23.63$; $p < .001$ in all cases. These results show that regardless of the arrangement of the colored houselights, subjects' choices were sensitive to the differing reinforcement percentages.

As predicted by the hyperbolic decay model when D_i was measured as the total duration of the colored keylights and houselights between food presentations, the indifference points in the 20% no-food-white conditions were much shorter ($M = 8.8$ s) than in the 20% green-white conditions ($M = 17.0$ s), and a planned comparison showed that this difference was significant, $F(1, 33) = 25.50$, $p < .001$. This result replicates the findings of Mazur (1989, 1991, 1995) that preference for a delayed probabilistic alternative increases if the delay-interval stimuli are not presented on nonreinforced trials. However, a planned comparison found no significant differences between the 20% green-white conditions and the other three 20% reinforcement conditions, $F(1, 33) = 0.01$. Similarly, a planned comparison found no significant differences between the 50% green-white conditions and the other three 50% reinforcement conditions, $F(1, 33) = 2.21$.

These results were compared to the predictions generated by the hyperbolic decay model when D_i was measured as the total duration of the colored keylights and houselights between food presentations. Using this interpretation of D_i, the predictions of the model are fairly accurate for some conditions of this experiment but not for others. (To obtain the predictions discussed below, K was set equal to 1, and it was assumed that the green and red response latencies were both 1 s. Actual response latencies were typically about 1 s in duration. However, the model's predictions would be qualitatively similar if other values of K and other response latencies were used.)

For the 50% and 20% green-white conditions, the model predicts indifference points of 7.9 s and 14.6 s, respectively, which are somewhat shorter than the group means of 10.8 s and 17.0 s. For the 20% no-food-white condition, the model predicts an indifference point of 7.8 s, slightly shorter than the group mean of 8.8 s. (These quantitative differences between predictions and results could be due

to a position or color bias, an inappropriate value of K, or other factors.)

For the remaining conditions, all of which included longer houselight durations, the predictions based on this method of measuring D_i were not supported. For the 50% and 20% late-reinforcer conditions, the model predicts indifference points of 17.2 s and 21.5 s, respectively, but Figure 2 shows that the actual indifference points did not differ systematically from those of the green-white conditions. For both the 50% and 20% all-green conditions, in which the green houselights were present for a minimum of 37 s between food presentations, the model predicts indifference points so large that the adjusting delay should have risen to its maximum possible value of 35 s. However, none of the subjects showed such increases, and the actual indifference points were about the same as in the green-white conditions. Predictions for the all-green/red conditions are ambiguous, because the red houselights were lit for a total of 37 s per food reinforcer regardless of the size of the adjusting delay. However, if preference for the adjusting alternative decreased when the duration of the red houselights increased, the indifference points should have decreased compared to the all-green conditions. Contrary to this prediction, Figure 2 shows no systematic differences between the all-green and all-green/red conditions.

As explained in the introduction, one possible explanation for the inaccurate predictions for the conditions with longer houselight durations is that subjects may have learned to discriminate the first 5 s of green houselights (which were sometimes followed by food) from longer green houselight presentations (which were never followed by food). If so, then better predictions might result if D_i included only those portions of the houselight periods that had some possibility of being followed by food.

A second set of predictions was based on this assumption. For the green-white and no-food-white conditions, the predictions are the same as before, because every presentation of the green houselights had some possibility of being followed by food. For the all-green and all-green/red conditions, food was presented only after 5-s green houselights for the standard alternative, so only the first 5 s of green houselights on nonreinforced trials were in-

cluded in calculating D_i. Similarly, in the all-green/red conditions, the red houselights that were presented in the ITI after food were never followed by another food delivery, so these time periods were not included in D_i for the adjusting alternative. With these changes, the model now correctly predicts that the indifference points should be the same in these two conditions as in the green-white conditions.

For the late-reinforcer conditions, food was delivered after either 5 s or 37 s of green houselights, so both of these durations must be included in calculating D_i. However, the green houselights that followed food deliveries on some trials (see Figure 1) were not included because these periods never ended with food. Given these changes, the model predicts small increases in the indifference points compared to the green-white conditions (from 7.9 s to 9.9 s in the 50% reinforcement conditions and from 14.6 s to 18.4 s in the 20% reinforcement conditions). As already noted, however, there were no systematic differences between the results from the green-white and late-reinforcer conditions.

Although the results of Experiment 1 present a somewhat confusing picture, one conclusion is clear: The presence or absence of distinctive stimuli during times when there was no possibility of a food delivery (such as during ITIs) had no detectable effects of preference. In contrast, the presence or absence of distinctive stimuli during delays that might be followed by food had large effects on preference. This was shown in the comparison between the green-white and no-food-white conditions of this experiment as well as in similar results from previous studies (e.g., Mazur, 1989, 1991, 1995).

If stimuli of the sort used in this experiment are in fact conditioned reinforcers, then it appears that their temporal placement is an important factor. Thus the first 5 s of green houselights may have served as a conditioned reinforcer in the all-green and all-green/red conditions (because food might be presented after 5 s), whereas longer durations of the green houselights were not conditioned reinforcers. The results suggest that such a stimulus serves as a conditioned reinforcer only at times when there is some possibility that it will be followed by the primary reinforcer. Using this assumption to es-

timate the durations of D_i, Equation 1 yielded fairly accurate predictions for all conditions except the late-reinforcer conditions. In these conditions, the addition of a late reinforcer did not lead to the longer indifference points predicted by the model. However, the predicted changes of only 2 to 4 s might have been difficult to detect, given the within-subject variability. Perhaps the effects of lengthening the stimuli on nonreinforced trials might be seen in a situation for which the model predicted a larger difference between conditions. Experiment 2 was designed to arrange such a situation.

EXPERIMENT 2

In this experiment, the standard alternative delivered food on some trials, after a delay with green houselights that lasted either 5 s or 30 s. In normal-green conditions, the 5-s and 30-s houselight presentations each occurred on a random half of the trials, and both houselight durations were followed by food on 20% of the trials. Because each presentation of the green houselights might be followed by food, each was included in calculating the durations of D_i in Equation 1. In long-green conditions, there were no differences on the 20% of the trials that ended with reinforcement: Food was delivered after a 5-s or a 30-s period with green houselights. However, all 80% of the nonreinforced trials with the standard alternative had 30-s periods of green houselights. Because any presentation of green houselights, whether 5 s or 30 s, might be followed by food, all green houselight periods were included in calculating D_i. The main question of this study was whether indifference points would be longer in the long-green conditions than in the normal-green conditions. Assuming that D should include all the time spent in the presence of a color stimulus, the hyperbolic decay model predicts longer indifference points in the long-green conditions because a larger proportion of the trials had 30-s periods of green houselights. Using the same assumptions as before ($K = 1$, and response latencies of 1 s), the hyperbolic decay model predicts indifference points of 30.6 s for the normal-green conditions and 37.5 s for the long-green conditions.

Table 2

Order of conditions and sessions to meet stability criteria for each subject in Experiment 2.

Order	Condition	Sessions to stability			
		Bird 1	Bird 2	Bird 3	Bird 4
1	Normal green	24	16	14	19
2	Long green	14	23	14	15
3	Normal green	14	15	18	18
4	Long green	16	17	14	16
5	Normal green	17	14	14	17
6	Long green	14	19	14	14

METHOD

Subjects and Apparatus

The subjects were the same 4 pigeons as in Experiment 1, and the same apparatus was used.

Procedure

This experiment was conducted several months after the completion of Experiment 1. Between the two experiments, the pigeons participated in several conditions of a different, unpublished experiment.

The experiment consisted of six conditions that used the same adjusting-delay procedure as in Experiment 1 (Table 2). A peck on the right red key led to an adjusting delay with red houselights, followed by 3-s of food and then an ITI with white houselights. A peck on the left green key led to a delay of either 5 s or 30 s, and then food was delivered on 20% of the trials. In Conditions 1, 3, and 5, the normal-green conditions, the durations of the green houselights were the same on reinforced and nonreinforced trials. That is, the green houselights were presented for either 5 s or 30 s, with equal probability, on both reinforced and nonreinforced trials. In Conditions 2, 4, and 6, the long-green conditions, the durations of the green houselights were different on reinforced and nonreinforced trials. On reinforced trials, the green houselight durations were the same as in the normal-green conditions, but they were always 30 s on nonreinforced trials. White houselights were present during the ITIs that followed either food or the green houselights. On all trials, with both the standard and adjusting alternatives, ITI duration was adjusted so that the total time from a choice response to the start of the next trial
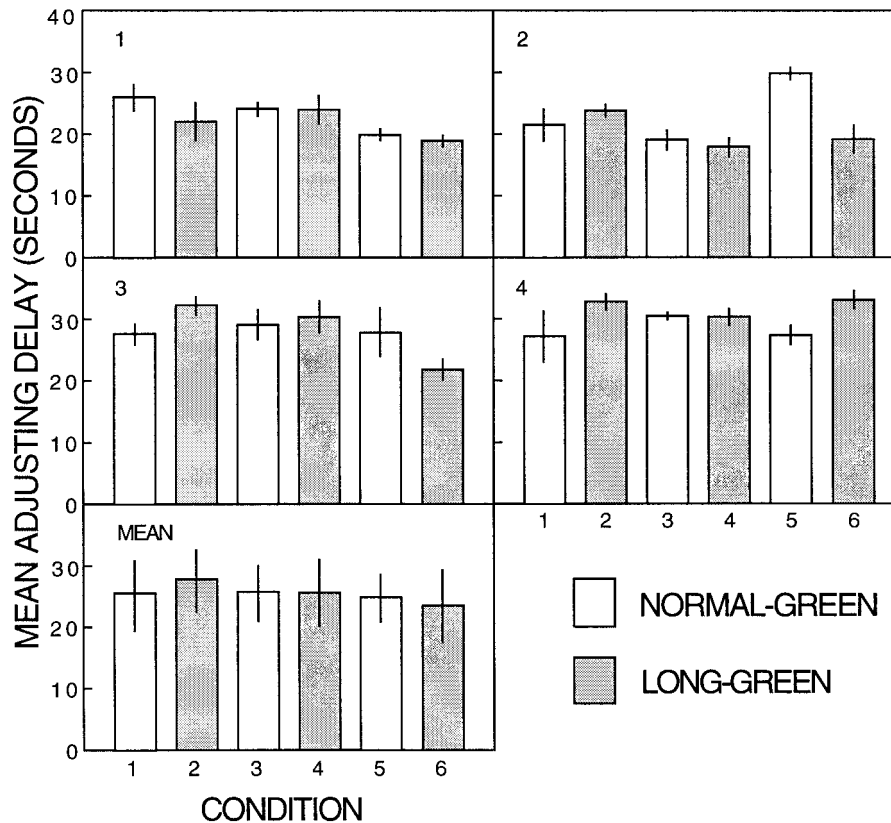
Fig. 3. Mean adjusting delays are shown for each subject in each condition of Experiment 2. The error bars represent one standard deviation above and below each mean.

(including the reinforcer duration) was 50 s. In all other respects, the procedure was the same as in Experiment 1, and the same stability criteria were used to terminate conditions.

RESULTS AND DISCUSSION

For all conditions, the mean adjusting delay from the six half-session blocks that satisfied the stability criteria was used as a measure of the indifference point. Figure 3 shows the mean indifference points and standard deviations from each condition for each subject. The mean indifference points were 25.3 s in the normal-green conditions and 25.4 s in the long-green conditions. The indifference points for normal-green conditions were somewhat shorter than the predicted durations of 30.6 s. More important, whereas the model predicts longer indifference points in the long-green conditions, there were no systematic differences between normal-green

and long-green conditions for any of the 4 subjects.

These results pose a problem for the prediction of the hyperbolic decay model that preference for a delayed probabilistic reinforcer will be inversely related to D_i, which is defined as the total duration of the conditioned reinforcers that occur between food presentations. In some of the conditions of Experiment 1, there was ambiguity about which time periods should and should not be included in D_i, because some green houselight presentations (e.g., those during ITIs) were never followed by food. However, this ambiguity was presumably avoided in Experiment 2, because both the 5-s and 30-s green houselights were at least occasionally followed by food in all conditions. By the usual definition of conditioned reinforcer, all of these green houselight periods should have served as conditioned reinforcers. If the strength or value of a conditioned reinforcer

is inversely related to its duration, as predicted by the hyperbolic decay model, the longer houselight durations in the long-green conditions should have resulted in longer indifference points, but they did not.

The absence of such an effect poses a problem, not just for the hyperbolic decay model, but more generally for attempts to predict the strength of a conditioned reinforcer. Although they differ from the hyperbolic decay model in their details, a number of other theories also propose that the strength of a conditioned reinforcer is inversely related to its duration (e.g., Fantino, 1977; Killeen, 1982; Vaughan, 1985). Some possible reasons why the durations of the green houselights did not affect preference in this experiment will be examined below.

GENERAL DISCUSSION

A recent review of research on the hyperbolic decay model (Mazur, 1997) suggested that one of the more difficult issues in its approach to conditioned reinforcement is the question of when a stimulus will and will not serve as a conditioned reinforcer. This statement was based on a few studies in which the durations of putative conditioned reinforcers were lengthened, and the results were sometimes inconsistent with the predictions of the model and were variable from subject to subject (Mazur, 1991; Mazur & Romano, 1992).

The present experiments were conducted to examine this matter more closely. These experiments, combined with the results of previous studies, help to clarify certain points. First, it is clear that the presence or absence of a conditioned reinforcer in the delay interval that normally precedes the primary reinforcer is a critical factor. This effect can be seen in the comparison of the green-white and no-food-white conditions of Experiment 1 as well as in several previous studies (Mazur, 1989, 1991, 1995). These results support the position of the hyperbolic decay model and of several other theories that the strength of a conditioned reinforcer is inversely related to its duration (Fantino, 1977; Killeen, 1982; Vaughan, 1985).

Second, Experiment 1 showed that when the same stimulus that preceded the primary reinforcer was also presented after the primary reinforcer, this additional presentation

had no detectable effects on preference. This result was obtained both in the all-green conditions, in which the green houselights were again presented during the ITIs that followed food, and in the all-green/red conditions, in which both the red and green houselights were presented in the ITIs that followed food. This finding suggests that all subjects discriminated between the houselights presented before food and houselights presented after food, and that whereas the former served as conditioned reinforcers, the latter did not.

The possibility that a stimulus can serve as a conditioned reinforcer at some times but not at others may pose problems for researchers, but similar phenomena have been observed in other areas. In classical conditioning, phenomena such as inhibition of delay and temporal conditioning (Pavlov, 1927) show that the passage of time can modulate the conditioned responses that are elicited by a conditioned stimulus. In operant conditioning, the peak procedure (Roberts, 1981) and the time-left procedure (Gibbon & Church, 1981) are two of many examples that show how responding can change systematically in the presence of an unchanging stimulus as time passes. It seems reasonable to suppose that similar temporal discriminations can be learned in situations involving delayed and probabilistic reinforcers.

If it is assumed that a stimulus will serve as a conditioned reinforcer only at those times when there is some possibility that it will be followed by a primary reinforcer, the hyperbolic decay model can account for most of the results of Experiment 1. The only apparent exception came in the late-reinforcer conditions, for which the model predicted slight increases in indifference points, but none were observed. The results of Experiment 2, however, pose a more serious challenge for this model, as well as for other theories that assume an inverse relation between the duration of a conditioned reinforcer and its strength. Because every presentation of the green houselights had some possibility of being followed by food, the longer green houselight presentations in the long-green conditions should have produced longer indifference points. Yet none of the subjects showed any systematic differences between the normal-green and long-green conditions.

Although the reasons for this result are not

clear, at least two possibilities can be considered. One is that the differences between the normal-green and long-green conditions were too subtle to gain differential control over the subjects' choices. However, pigeons in a previous study with mixed delays and probabilistic reinforcement exhibited systematic changes in preference as the conditions changed (Mazur, 1995), and it seems unlikely that the pigeons in the present experiment were unable to make similar discriminations.

A second explanation relates to the idea that subjects may have learned to discriminate between the first 5 s of green houselights and longer presentations. This explanation may be easier to understand by imagining hypothetical conditions similar to those of Experiment 2 in which the green houselights actually changed colors (e.g., from a dark green to lighter shade of green) after the first 5 s. Given such an arrangement, the dark green houselights would be present for 5 s on every trial in all conditions of Experiment 2, and food would follow the dark-green houselights on 10% of the trials. In the normal-green conditions, 25-s periods of light green houselights would occur on 50% of the trials with the standard alternative, and 20% of these presentations would be followed by food. In the long-green conditions, 25-s periods of light-green houselights would occur on 90% of the trials with the standard alternative, but only 11% of these presentations would be followed by food. Given these durations and percentages, Equation 1 was used to calculate separate values for the dark-green and light-green stimuli, which were then summed to obtain an overall value for the standard alternative in both types of conditions. These calculations resulted in predicted indifference points of 20.4 s and 19.7 s for the normal-green and long-green conditions, respectively.

Of course, the green houselights did not actually change colors in Experiment 2. But if the pigeons learned to discriminate the first 5 s of green houselights from the remaining 25 s such that the two periods functioned as two different conditioned reinforcers, the same calculations might apply. These predictions are more consistent with the actual results, because they predict virtually no difference between the normal-green and long-green conditions. (In addition, the actual in-

difference points were slightly longer than the predictions, as was the case in most of the conditions of Experiment 1.) This account is admittedly speculative, but it offers one possible explanation for the results of Experiment 2.

At the very least, the results of the present experiments support the idea that a stimulus can serve as a conditioned reinforcer at some times but not at others, depending on whether there is or is not a possibility that a primary reinforcer will be delivered. An unanswered question is whether, in a similar fashion, a stimulus can serve as a strong conditioned reinforcer at some times and as a weak conditioned reinforcer at other times, depending on whether the delay to primary reinforcement is likely to be short or long. Additional research, perhaps including actual changes in stimulus colors as described hypothetically above, might help to answer this question.

REFERENCES

Ainslie, G. W. (1974). Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior, 21,* 485–489.

Battalio, R. C., Kagel, J. H., & McDonald, D. N. (1985). Animals' choices over uncertain outcomes: Some initial experimental results. *American Economic Review, 75,* 597–613.

Cicerone, R. A. (1976). Preference for mixed versus constant delay of reinforcement. *Journal of the Experimental Analysis of Behavior, 25,* 257–261.

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. (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., & Church, R. M. (1981). Time left: Linear versus logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 87–108.

Green, L., Fisher, E. B., Perlow, S., & Sherman, L. (1981). Preference reversal and self control: Choice as a function of reward amount and delay. *Behaviour Analysis Letters, 1,* 43–51.

Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science, 5,* 33–36.

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. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative anal-*

yses 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. (1995). Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. *Journal of the Experimental Analysis of Behavior, 63,* 139–150.

Mazur, J. E. (1997). Choice, delay, probability, and conditioned reinforcement. *Animal Learning & Behavior, 25,* 131–147.

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.

Pavlov, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, trans.). London: Oxford University Press.

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.

Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 242–268.

Spetch, M. L., Belke, T. W., Barnet, 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.

Vaughan, W. (1985). Choice: A local analysis. *Journal of the Experimental Analysis of Behavior, 43,* 383–405.

Waddington, K. D., Allen, T., & Heinrich, B. (1981). Floral preferences of bumblebees (*Bombus edwardsii*) in relation to intermittent versus continuous rewards. *Animal Behavior, 29,* 779–784.