JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1985, 43, 341-351 NUMBER 3 (MAY)

PROBABILITY AND DELAY OF REINFORCEMENT AS FACTORS IN DISCRETE-TRIAL CHOICE

JAMES E. MAZUR

HARVARD UNIVERSITY

Pigeons chose between two alternatives that differed in the probability of reinforcement and the delay to reinforcement. A peck on the red key always produced a delay of 5 s and then a possible reinforcer. The probability of reinforcement for responding on this key varied from .05 to 1.0 in different conditions. A response on the green key produced a delay of adjustable duration and then a possible reinforcer, with the probability of reinforcement ranging from .25 to 1.0 in different conditions. The green-key delay was increased or decreased many times per session, depending on a subject's previous choices. The purpose of these adjustments was to estimate an indifference point, or a delay that resulted in a subject's choosing each alternative about equally often. In conditions where the probability of reinforcement was five times higher on the green key, the green-key delay averaged about 12 s at the indifference point. In conditions where the probability of reinforcement was twice as high on the green key, the green-key delay at the indifference point was about 8 s with high probabilities and about 6 s with low probabilities. An analysis based on these results and those from studies on delay of reinforcement suggests that pigeons' choices are relatively insensitive to variations in the probability of reinforcement between .2 and 1.0, but quite sensitive to variations in probability between .2 and 0.

Key words: probability of reinforcement, delay of reinforcement, adjusting schedule, key peck, pigeons

What is the relationship between a reinforcer's probability of occurrence and its effectiveness or value? The present experiment was designed to gather some information on this question. *Value* might be loosely defined as the control of responses by the reinforcer in a choice situation. One simple possibility is that the value of a reinforcer is proportional to its probability, as depicted in the left panel of Figure 1. This would mean that changing the probability of occurrence of the reinforcer by a certain percentage would change its value by the same percentage. Of course, this is only one of an infinite number of possible relationships; two others are shown in Figure 1. The center panel illustrates a case where behavior shows a large degree of sensitivity to differences among high probability reinforcers but less sensitivity to differences among reinforcers with low probability. The right panel shows the opposite case: Behavior is relatively insensitive to variations among high probability reinforcers but very sensitive to variations among low probability reinforcers.

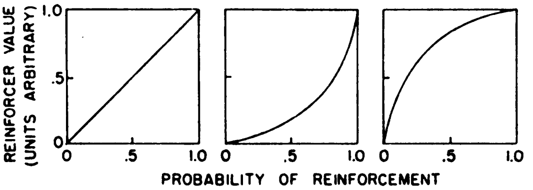

Fig. 1. Three possible relationships between reinforcer probability and value.

Not much information on this question can be extracted from previous experiments in which reinforcers were delivered on a probabilistic basis. One class of experiments has made use of a "percentage reinforcement" procedure with single reinforcement schedules. For example, Staddon and Innis (1969) used fixed-interval (FI) 2-min schedules in which 25% of the reinforcers were replaced with blackouts. This and some other studies on percentage reinforcement (e.g., McMillan, 1971; Zeiler,

---

Preparation of this article was supported in part by Grant No. 1 R01 MH38357-01A1 from the National Institute of Mental Health. I thank David Coe for running the animals in this experiment. Requests for reprints should be sent to James E. Mazur, 768 William James Hall, Harvard University, Cambridge, Massachusetts 02138.

1972) obtained the surprising result that decreases in the probability of reinforcement produced increased response rates. If these response rates were taken to reflect the values of the different reinforcement conditions, the results would of course be inconsistent with all patterns illustrated in Figure 1. It has been suggested, however, that such increased response rates under percentage reinforcement schedules may result from several confounding variables, such as "frustration" effects (Neuringer & Chung, 1967) or reinforcement aftereffects (Staddon, 1970; Staddon & Innis, 1969). The possible influence of these or other factors suggests that measuring response rates on single schedules may not provide unbiased estimates of the values of probabilistic reinforcers.

Kendall (1974) obtained a similar result in a choice situation. Pigeons preferred a reinforcer delivered with a probability of .5 over a reinforcer delivered with a probability of 1.0. However, Kendall's result is an exception to the general pattern of findings from choice situations, and a study by Fantino, Dunn, and Meck (1979) suggested that his result may have been produced by a potentially confusing arrangement of keylight onsets and offsets. Using Kendall's arrangement of stimuli, Fantino et al. replicated his result; but when they used distinctly different stimuli in the choice period, they obtained repeated and consistent preferences for whichever key had the higher probability of reinforcement.

Other experiments in which animals chose between two reinforcers of differing probabilities have produced either approximate probability matching (in which response proportions match the reinforcement probabilities) or near-exclusive preference for the alternative with the higher probability of reinforcement (e.g., Behrend & Bitterman, 1961; Bitterman, Wodinsky, & Candland, 1958; Graf, Bullock, & Bitterman, 1964; Wilson, 1960). Shimp (1966, 1967) proposed that pigeons' responses move toward exclusive preference if the training continues long enough. In any case, these experiments show quite clearly that in a choice situation animals prefer the alternative with the higher probability of reinforcement, but they provide little more than ordinal information about the relation between probability of reinforcement and value.

A different class of experiments includes those in which subjects chose between alternatives differing in both probability and amount of reinforcement. For instance, in Young's (1981) experiment with pigeons, pecking at one key produced, with equal probability, either 0 or 10 food pellets. Pecking the other key produced a constant number of pellets on each trial within a condition, with the number varying between conditions. Most of the pigeons preferred 10 pellets with a probability of .5 over 5 pellets with a probability of 1.0, although both alternatives would deliver roughly the same number of pellets in the long run. Assuming that the value of a reinforcer is approximately proportional to the number of pellets delivered, Young's results suggest that the relationship between probability and value was most like the function in the right panel of Figure 1. Of course, it is also possible that 10 pellets had more than twice the value of 5 pellets, but this possibility seems unlikely. Indeed, the possibility of satiation with a large number of pellets and the economic principle of diminishing marginal value suggest that, if anything, 10 pellets might have less than twice the value of 5 pellets.

Staddon and Innis (1966) and Essock and Reese (1974) also conducted experiments with pigeons in which a variable reinforcer amount was contrasted with a fixed reinforcer amount. Unlike Young's experiment, in which responding on the key arranging probabilistic reinforcement resulted in either 10 pellets or none at all, the probabilistic alternative in these experiments always arranged a reinforcer, but it was either small or large. Staddon and Innis found no preference between fixed and variable reinforcer amounts, whereas Essock and Reese found a modest preference for the variable alternative. In these two studies, reinforcer amounts were varied by controlling grain-hopper durations. This made it less likely that a preference for the variable reinforcer would be observed, inasmuch as Epstein (1981) has shown that pigeons typically obtain less than twice as much grain when hopper durations are doubled. On the whole, then, the results of these studies are not conclusive, but they suggest that pigeons may prefer a probabilistic or variable reinforcer over a constant reinforcer. Two other studies, in which the subjects were rats, produced inconsistent results. Leventhal, Morrell, Morgan, and Perkins (1959) found preference for a larger, probabilistic reinforcer in two groups of rats, but no preference in two other groups. In contrast, Logan (1965) found preference for a constant reinforcer of five pellets over a variable reinforcer of either one or nine pellets. The reasons for these different results are not clear.

Like the above studies, the present experiment employed two independent variables, but in this case they were probability and delay of reinforcement. The experiment used a discrete-trial choice situation and an adjusting procedure. If a pigeon pecked what will be called the *standard key*, there was always a 5-s delay, which was followed by reinforcement on a certain percentage of the trials (which varied from condition to condition). The consequences of pecking the other key (the *adjusting key*) were similar, except that the probability of reinforcement was usually different, and the delay was not fixed at 5 s. Instead, the delay was increased or decreased in steps of 1 s many times each session, depending on the subject's previous choices. The purpose of the adjusting delay was to estimate the indifference point, or the delay that equated the values of the two alternatives. Previous experiments with this adjusting procedure (Mazur, 1984, in press) showed that it yields useful estimates of indifference points.

To understand the rationale behind the present experiment, let us begin with the assumption that the value of a reinforcer is some function, f, of the probability of reinforcement. Likewise, we will assume that there is a function, g, that relates a reinforcer's delay to its value. For experiments in which probability and delay are the two variables of interest, we will assume that

$$V_i = C \cdot f(p_i) \cdot g(D_i), \tag{1}$$

where V_i is the value of alternative i, p_i is its probability of reinforcement, and C is a constant that subsumes all other variables that can affect the value of an alternative. Because the units of V_i are arbitrary, C can be fixed at any value. For notational convenience, we will define $p'_i = f(p_i)$, and p'_i will be called the *effective probability* of alternative i. We will assume that at an indifference point between the standard and adjusting alternatives,

$$V_s = V_a, \tag{2}$$

where the subscripts refer to the standard and adjusting alternatives, respectively. It follows from Equations 1 and 2 that at the indifference point,

$$p'_s \cdot g(D_s) = p'_a \cdot g(D_a). \tag{3}$$

The estimation of indifference points was designed to provide information about the relation between probability and value in two ways, one of which makes no assumptions about the quantitative relation between delay and value, and a second which does make some such assumptions. First, this study sought to determine whether the indifference point would shift systematically as the probabilities of reinforcement for both alternatives were changed by the same percentage. For example, suppose that in one condition a subject reached an indifference point when the two alternatives were (1) a reinforcer with a probability of .5 and a delay of 5 s, and (2) a reinforcer with a probability of 1.0 and a delay of 10 s. Based on Equation 3, this would imply that

$$\frac{p'_{1.0}}{p'_{.5}} = \frac{g(D_5)}{g(D_{10})}, \tag{4}$$

where the subscripts refer to the probabilities and delays for the two alternatives. Now suppose the two probabilities were reduced by half, to .25 and .5, respectively. If the delay for the larger probability reinforcer at the indifference point remained at 10 s, this would imply that

$$\frac{p'_{.5}}{p'_{.25}} = \frac{g(D_5)}{g(D_{10})}. \tag{5}$$

It follows from Equations 4 and 5 that

$$\frac{p'_{1.0}}{p'_{.5}} = \frac{p'_{.5}}{p'_{.25}}. \tag{6}$$

This proportionality is consistent with the relation between probability and value depicted in the left panel of Figure 1.

As a second possibility, suppose that the delay for the second reinforcer at the indifference point decreased from 10 s to 8 s when the probabilities were decreased to .25 and .5. This would imply that

$$\frac{p'_{1.0}}{p'_{.5}} > \frac{p'_{.5}}{p'_{.25}}, \tag{7}$$

a result that is consistent with the center panel in Figure 1.

As a third possibility, suppose that the delay for the second reinforcer increased from 10 s to 15 s. This would imply that

$$\frac{p'_{1.0}}{p'_{.5}} < \frac{p'_{.5}}{p'_{.25}}, \tag{8}$$

a result that is consistent with the right panel in Figure 1. These are only three simple possibilities, however, and the actual results from many different conditions presented a more complex picture, as will be shown.

The second method used to analyze the data depended on the use of a specific equation for g, the function relating delay of reinforcement and value. If this function were known, the indifference points obtained in this experiment would provide much more direct information about reinforcement probability. For example, suppose we knew that a reinforcer delayed 10 s had only 60% of the value of a reinforcer delayed 5 s. If an indifference point was reached between a reinforcer with a .5 probability and a 5-s delay and one with a 1.0 probability and a 10-s delay, Equation 3 would indicate that $p'_{.5}$ is 60% of $p'_{1.0}$. Although the function g is not known with certainty, in two different experiments on discrete-trial choice (one involving choice between different delay-amount combinations, and the other involving choice between fixed and variable delays before reinforcement), Mazur (1984, in press) found evidence that supported the following equation:

$$g(D_i) = \frac{1}{1+KD_i}. \tag{9}$$

The parameter K determines how rapidly value decreases with increasing delay. Estimates of K varied from subject to subject in Mazur's experiments, but the average value of K was about 1.0. Combining Equations 9 and 1, we obtain

$$V_i = C \cdot p'_i \cdot \left(\frac{1}{1+KD_i}\right). \tag{10}$$

Based on the assumption that Equation 9 provides at least rough estimates of the values of reinforcers with different delays, Equation 10 was applied to the results of this experiment to estimate the shape of the function that relates probability of reinforcement and value.

METHOD

Subjects

Four White Carneaux pigeons were maintained at 80% of their free-feeding weights. All subjects had served in several experiments involving the adjusting procedure, delay, and probability of reinforcement (Mazur, in press, Experiments 2 and 3, and some unpublished studies).

Apparatus

The experimental chamber was 30 cm long, 30 cm wide, and 33 cm high. Three response keys, each 2.5 cm in diameter, were mounted in the front wall of the chamber, 20.5 cm above the floor. A force of approximately 0.10 N was required to operate each key, and each effective response produced a feedback click. A hopper below the center key provided controlled access to mixed grain, and when grain was available, the hopper was illuminated with two 6-W white lights. Six 6-W lights (two white, two blue, and two orange) were mounted above the wire-mesh ceiling of the chamber. The chamber was enclosed in a sound-attenuating box that contained an air blower for ventilation and a speaker producing continuous white noise to mask extraneous sounds. A PDP-8 computer in another room was programmed in SUPERSKED to control the stimuli and record responses.

Procedure

The experiment consisted of 13 conditions that differed from one another only in the probabilities of reinforcement produced by responses on either of the two keys. A peck on the standard key was always followed by a delay of 5 s, and the procedure was designed to determine what delay on the adjusting key would result in indifference between the two choices. The conditions followed each other continuously except that an interval of approximately 8 months separated Conditions 10 and 11.

Throughout the experiment, sessions lasted for 64 trials or for 50 min, whichever came first. Each block of four consecutive trials consisted of two forced-choice trials followed by two free-choice trials. Figure 2 diagrams the sequence of events on a free-choice trial. Each trial was preceded by a 30-s intertrial interval (ITI), during which the white houselights were illuminated. At the start of the trial, the center key was illuminated with white light, and a single peck on the center key was required to begin the choice period. The purpose of this response on the center key was to ensure that the subject's head was equally distant from the two side keys when the choice period began. A peck on the center key darkened this key and illuminated the two side keys, one green and one red. The locations of these colors (left key or right key) varied randomly from trial to trial to control for position preferences.

The red key was always the standard key, and the green key was the adjusting key. A peck on the red key extinguished both side keys and initiated a 5-s delay, during which the blue houselights were lit instead of the white houselights. At the end of the standard delay, the blue houselights were extinguished and either (1) grain was presented for 2 s, followed by the next ITI, or (2) there was no food presentation, and the next ITI began immediately. During the choice period, a peck on the green key extinguished both side keys and led to the adjusting delay, during which the orange houselights were lit. As with the standard delay, the adjusting delay was followed either by 2-s access to grain or by the next ITI. For all cases where the probability of reinforcement was less than 1.0, reinforcer delivery was determined by a pseudorandom sequence that ensured that the actual probability of reinforcement at least approximately equaled the nominal probability in each session. The scheduled probabilities of reinforcement on the two keys are presented in Table 1 for each condition.

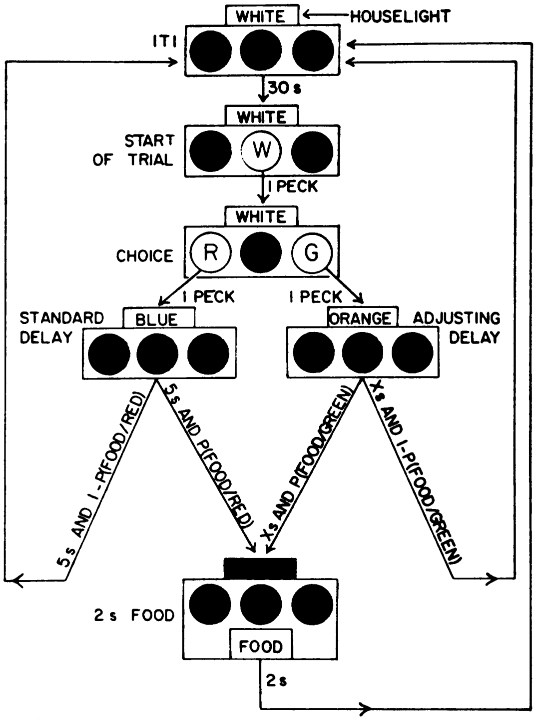

Fig. 2. A diagram of the sequence of events on a free-choice trial. The three circles represent the three response keys, and the small rectangle above them represents the houselight. The terms "p(food/red)" and "p(food/green)" represent the probabilities of reinforcement after a red and a green peck, respectively. The notation "X s" indicates that the delay after a peck at the green key was adjustable, as described in the text.

The procedure on forced-choice trials was the same as on free-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-choice trials, one involved the red key and the other the green key. The temporal order of these two types of trials varied randomly.

After every two free-choice trials, the delay for the adjusting key might be changed. If a subject had chosen the adjusting key on both free-choice trials, the adjusting delay was increased by 1 s. If the subject had chosen the standard key on both trials, the adjusting de-

Table 1

Order of Experimental Conditions and Number of Sessions per Condition

Condition	*Probability of Reinforcement* *Adjusting Key*	 *Standard Key*	 *Adjusting: Standard*	*Number of Sessions* *Subj. 1*	 *Subj. 2*	 *Subj. 3*	 *Subj. 4*
1	1.0	0.5	2:1	26	27	27	30
2	1.0	0.2	5:1	24	16	26	21
3	0.25	0.125	2:1	18	13	20	18
4	0.5	0.1	5:1	14	15	21	18
5	0.5	0.25	2:1	12	14	13	11
6	1.0	0.2	5:1	16	16	21	19
7	1.0	0.5	2:1	15	28	15	11
8	0.5	0.1	5:1	13	14	24	10
9	0.5	0.25	2:1	16	16	13	12
10	0.25	0.125	2:1	15	10	27	11
11	0.25	0.05	5:1	12	20	25	16
12	1.0	1.0	1:1	11	19	20	10
13	0.25	0.05	5:1	15	11	11	15

lay was decreased by 1 s unless it was already zero. If the subject had chosen 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 the first session of the experiment, the adjusting delays were set at four different durations between 1 and 20 s for the four subjects. At the start of every other session, the adjusting delay was determined by the above rules as if it were a continuation of the preceding session.

Table 1 shows that the first 10 conditions consisted of two presentations of each of five different pairs of probabilities. In all 10 conditions the adjusting key had the higher probability of reinforcement. Three pairs involved a 2:1 ratio of probabilities (1.0 to .5, .5 to .25, and .25 to .125) and the other two pairs involved a 5:1 ratio of probabilities (1.0 to .2, and .5 to .1). Conditions 11 and 13 examined a third pair of probabilities with a 5:1 ratio (.25 to .05). In Condition 12 the probability of reinforcement was 1.0 on both keys.

The first condition of the experiment lasted for a minimum of 24 sessions, and all other conditions lasted for a minimum of 10 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 were not used, and a condition was terminated when the following three criteria were met, using the data from all subsequent sessions:

(1) Neither the highest nor the lowest single-block mean of a condition could occur in the last eight blocks of the condition.

(2) The mean adjusting delay across the last eight blocks could not be the highest nor the lowest eight-block mean of the condition.

(3) The mean delay of the last eight blocks could not differ from the mean of the preceding eight blocks by more than 10% or by more than 1 s (whichever was larger).

RESULTS

The right portion of Table 1 lists the number of sessions required to satisfy the stability criteria by each subject in each condition. All analyses were based on the results from the eight half-session blocks that met the stability criteria. The mean adjusting key delay of these eight half-session blocks was used as an estimate of the indifference point—the delay that equated the values of the two alternatives.

Although the actual probabilities of reinforcement usually varied somewhat from the scheduled probabilities listed in Table 1, these variations were small. For instance, in the 24 cases (4 subjects times 6 conditions) where the scheduled probability of reinforcement was .5,

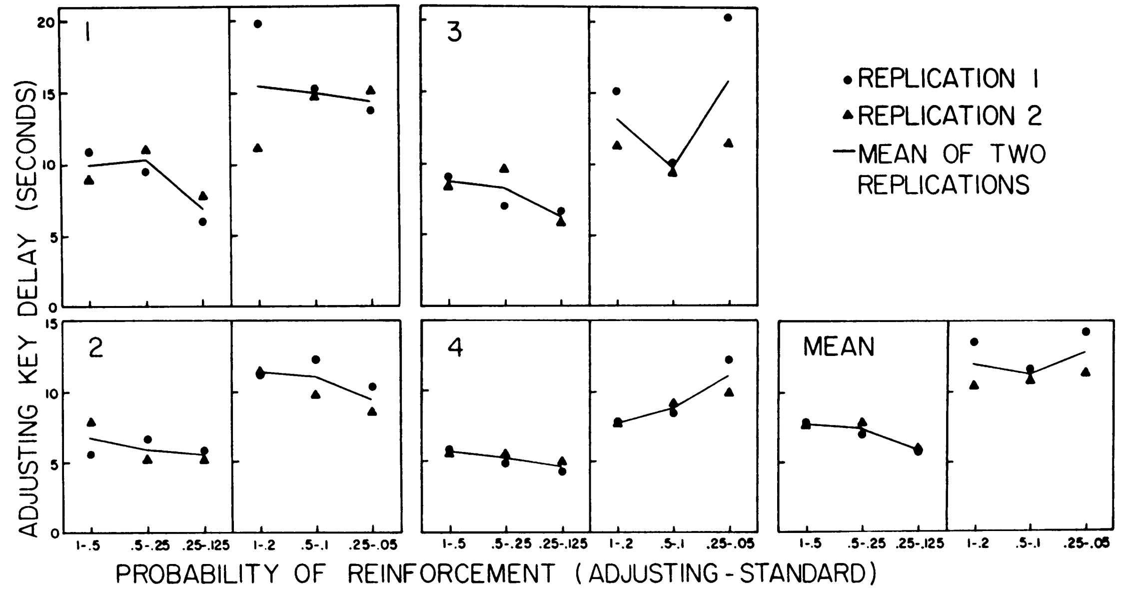

Fig. 3. For each subject and for the group mean, indifference points from conditions in which the probability of reinforcement was twice as high on the adjusting key (left panels) or five times as high (right panels). The lines in each panel mark the means of two replications.

the mean deviation from this probability (absolute value, scheduled minus obtained probability) was .018 in the eight half-session blocks used to calculate the indifference points. With lower scheduled reinforcement probabilities, the deviations were smaller. When the scheduled probabilities were .25, .2, .125, .1, and .05, the mean deviations from these probabilities were, respectively, .016, .016, .014, .010, and .006. Furthermore, these are conservative measures of the accuracy of the obtained reinforcement probabilities because they are based on only the last eight half-session blocks of a condition, whereas subjects received a minimum of 20 half-session blocks in each condition. The deviations between scheduled and obtained probabilities were smaller if more sessions are considered. For example, in the eight cases with the largest deviations from the scheduled probabilities, the mean deviation was .035. If the last 16 half-session blocks are considered, however, the mean deviation decreased to .022. Not surprisingly, variations of the indifference points across subjects, across replications, and across conditions were not systematically related to these small deviations from the scheduled reinforcement probabilities.

In Condition 12, where the probability of reinforcement was 1.0 for both alternatives, the mean adjusting key delays were 5.63, 4.30, 8.89, and 4.93 s, respectively, for the 4 subjects. The group mean was 5.94 s, and considering the variability in the results from this and other conditions, there was no evidence that there was a bias either for or against the alternative with the adjusting delay.

The results from the other 12 conditions are shown in Figure 3. For each subject, the left panel contains the results from the six conditions in which the probability of reinforcement was twice as high on the adjusting key, and the right panel contains the results from the six conditions where it was five times as high. The solid line in each panel marks the mean of the two presentations of each probability pair. In a few cases, the results from the two replications were substantially different—there were two instances in which the indifference points in the two replications differed by over 8 s. On the other hand, in most cases the results from the two replications were fairly similar. In 11 of 24 cases in which a condition was replicated, the two indifference points were less than 1 s apart, and the median difference between replications was 1.4 s.

The results in Figure 3 are presented in a format that makes the first analysis described in the introduction easy to visualize. If there were a simple proportionality between reinforcer probability and value (as in the left panel of Figure 1), then the results in each panel of Figure 3 should form a horizontal line. To a first approximation they do, for the most obvious result shown in Figure 3 is the difference between the indifference points of the 2:1 conditions and those of the 5:1 conditions. The mean adjusting-key delay across all 2:1 conditions was 7.0 s, and the mean across all 5:1 conditions was 11.9 s. For all subjects, there was virtually no overlap between these two sets of conditions; the only exception was the second replication of the .5 to.25 condition for Subject 3.

Although a first glance suggests that there was little difference among conditions with the same adjusting:standard probability ratio, a closer examination reveals a small but systematic tendency for the indifference points to decrease in the 2:1 conditions as reinforcement probabilities declined. This trend was evident in the results from all 4 subjects, and an analysis of variance applied to the results of the 2:1 conditions revealed a significant effect of probability [$F(2,6) = 8.00$, $p < .05$].

In the absence of other data, this declining pattern would suggest that the relation between probability and value was similar to the pattern shown in the middle panel of Figure 1 (i.e., a 2:1 ratio of probabilities has a greater effect when these probabilities are large than when they are small). However, the results from the 5:1 conditions did not replicate this pattern. The indifference points from these conditions were more variable. Subject 2 showed some evidence for smaller indifference points with decreasing probabilities, but Subject 4 showed the opposite pattern, and Subjects 1 and 3 showed no discernible trends. An analysis of variance showed no significant effect of reinforcement probability in the 5:1 conditions [$F(2,6) = .65$, ns].

Because the trends in the 2:1 and 5:1 conditions were different, they provided no clear support for any of the three patterns pictured in Figure 1. On the other hand, a subsequent analysis based on Equation 10 revealed a pattern that is consistent with the results of both the 2:1 and 5:1 conditions. It follows from Equations 2 and 10 that

$$Cp'_a \frac{1}{1+KD_a} = Cp'_s \frac{1}{1+KD_s}. \quad (11)$$

Solving for p'_s, we obtain

$$p'_s = p'_a \frac{1+KD_s}{1+KD_a}. \quad (12)$$

Equation 12 was used along with the group means shown in Figure 3 to estimate the relationship between p and p'. To begin, K was set equal to 1.0, based on the results of previous experiments mentioned in the introduction. In all conditions of this experiment, D_s was 5 s. In the first calculation, $p'_{1.0}$ was arbitrarily set at 1.0. Based on a mean adjusting-key delay (D_a) of 7.78 s in the 1.0 to .5 conditions, it follows from Equation 12 that $p'_{.5} = .683$. This value was then used as an estimate of p'_a for the .5 to .25 conditions, and Equation 12 estimated that $p'_{.25} = .486$. A similar procedure was used to estimate p' for the other four probabilities used in this experiment. The results of this analysis are the filled circles in Figure 4, in which effective probability is plotted as a function of objective probability. As can be seen, this function suggests that as p decreases from a value of 1.0, p' decreases more slowly. In short, this analysis suggests that choice responses were less influenced by decreased probability of reinforcement than an objective measure of probability would predict. For instance, it suggests that as the probability of reinforcement was decreased from 1.0 to .2, the effective probability of the reinforcer (and hence its value) decreased by only about 54%, not 80%. On the other hand, the function suggests that effective probabilities decreased precipitously as objective probabilities decreased from about .15 toward 0.

Because the value of K might not have been 1.0 for the subjects in this experiment, Figure 4 shows analyses based on two other values of K. With $K = .2$, the discrepancy between p and p' becomes more pronounced. The third function in Figure 4 shows that as K approaches infinity, the discrepancy between p and p' is only slightly reduced. The conclusion

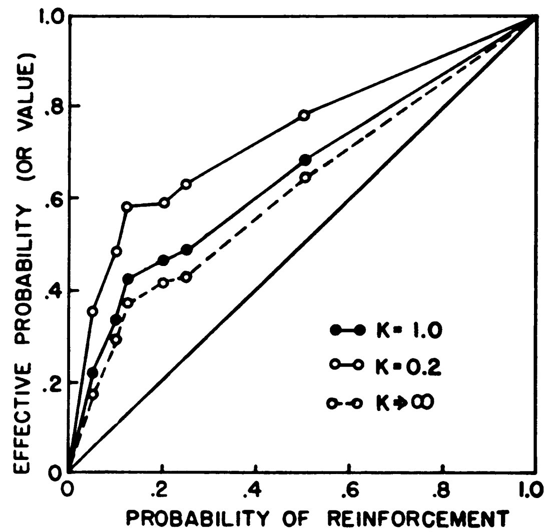

Fig. 4. Three estimates of the relationship between p and p', based on Equation 10 and the mean results shown in Figure 3. The three functions were generated by using three different values of K in Equation 10.

from this analysis can be stated more generally: Given the results of this experiment, if Equation 9 correctly describes the relationship between delay and value, then for all possible values of K, the function relating p and p' will remain above the diagonal line in Figure 4.

DISCUSSION

The results of this experiment demonstrate that it is easy to be misled about the shape of the function relating probability and value if only a restricted range of probability pairs is examined. By themselves, the results from the 2:1 conditions were consistent with the sort of concave-upward function depicted in the middle panel of Figure 1. The results from the 5:1 conditions were not consistent with a function of this shape, however, so it was necessary to search for a function that could accommodate the entire pattern of results. The second analysis described in the Results section, which made use of the results from all conditions, generated the relation depicted in Figure 4. This figure suggests that the relation between reinforcer probability and value is not linear for pigeons, but this relation was estimated using a particular equation relating reinforcer delay and value. It is therefore fair to ask to what extent the conclusions one draws from this experiment depend on one's choice of equations. The answer is that the particular equation is not especially important. We can reach similar conclusions using a simpler analysis that relies only on the present and previous data, not on any specific equation relating delay and value. This analysis is described below.

Mazur (in press) obtained estimates of indifference points between various delay-amount combinations, using an adjusting procedure similar to the one used in the present experiment. The two alternatives always delivered 2 s and 6 s of grain, respectively. For the present analysis, all we need to assume about these two reinforcer amounts is that 6-s access to grain has no more than three times the value of 2-s access. This is a conservative assumption, considering Epstein's (1981) finding that pigeons eat only about 2.15 times as much grain during 6-s access as during 2-s access. Mazur's results suggested that a choice between 2 s of grain delayed 5 s and 6 s of grain delayed about 15 s constitutes an indifference point for pigeons. This result was about the same for all 4 pigeons. It follows from the preceding assumption that with all else equal, a reinforcer delayed 15 s has at least 33% of the value of a reinforcer delayed 5 s. But in the 1.0 to .2 conditions of the present experiment, responding was roughly indifferent between a .2-probability reinforcer delayed 5 s and 1.0-probability reinforcer delayed only about 12 s. Because a reinforcer delayed 12 s must have a higher value than a reinforcer delayed 15 s, it follows that a reinforcer with a probability of .2 has more than 33% of the value of a reinforcer with probability of 1.0. Notice that this analysis is based on cautious assumptions about the relationship between reinforcer amount and value. Nevertheless, the conclusion is that the value of the .2-probability reinforcer is higher relative to a 1.0-probability reinforcer than the actual probabilities would suggest. A less conservative calculation based on Epstein's findings would start with the assumption that 6-s access to grain has about 2.15 times the value of a 2-s reinforcer, and this would lead to the conclusion that a .2-probability reinforcer has at least 46% of the value of a 1.0-probability reinforcer. (The estimate from Equation 10 with $K = 1.0$ was 46.3%.)

In addition, it cannot be said that the assumptions about the relationship between delay and value are based only on results from Mazur's (in press) adjusting procedure. For instance, Rachlin and Green (1972) used their results and those of previous studies to suggest that the effectiveness of a reinforcer is roughly an inverse function of delay. In the notation of this manuscript, they assumed that $g(D_i) = 1/D_i$. This equation is close to Equation 9 with $K = 1.0$. For the present purposes, it does not matter which equation is more accurate. As long as reinforcer value is at least approximately an inverse function of delay, the results of the present experiment imply that the function relating reinforcer probability and value has the general shape depicted in Figure 4.

This conclusion is consistent with Young's (1981) finding that pigeons preferred a .5 probability of 10 food pellets over a 1.0 probability of 5 pellets. It is also consistent with the conclusions of Caraco, Martindale, and Whittam (1980), who suggested that yellow-eyed juncos are "risk prone" when they are deprived of food. By risk prone they meant only that the probability-value function is similar to that depicted in the right panel of Figure 1. However, Caraco et al. also concluded that under conditions of low food deprivation, juncos are risk averse (i.e., their probability-value functions resemble that of the center panel in Figure 1). Caraco et al. proposed an explanation of their results based on the particular feeding habits and survival needs of juncos; other species might behave differently when choosing between probabilistic reinforcers under conditions of high and low food deprivation. It would be interesting to learn whether the results of the present experiment would be altered if the pigeons were maintained closer to their free-feeding weights. Along these same lines, it would be interesting to determine how the results depicted in Figure 4 would compare to those obtained from other species. It is certainly possible that the relationship between reinforcer probability and value might be closer to a strict proportionality for rats or chimpanzees than it is for pigeons. Further research is needed to determine whether the pattern depicted in Figure 4 is limited to the particular species, deprivation conditions, and delay and amount values used in the present experiment, or whether it is more generally applicable as a description of the relationship between reinforcer probability and value.

REFERENCES

Behrend, E. R., & Bitterman, M. E. (1961). Probability-matching in the fish. *American Journal of Psychology,* **74,** 542-551.

Bitterman, M. E., Wodinsky, J., & Candland, D. K. (1958). Some comparative psychology. *American Journal of Psychology,* **71,** 94-110.

Caraco, T., Martindale, S., & Whittam, T. S. (1980). An empirical demonstration of risk-sensitive foraging preferences. *Animal Behaviour,* **28,** 820-830.

Epstein, R. (1981). Amount consumed as a function of magazine-cycle duration. *Behaviour Analysis Letters,* **1,** 63-66.

Essock, S. M., & Reese, E. P. (1974). Preference for and effects of variable- as opposed to fixed-reinforcer duration. *Journal of the Experimental Analysis of Behavior,* **21,** 89-97.

Fantino, E., Dunn, R., & Meck, W. (1979). Percentage reinforcement and choice. *Journal of the Experimental Analysis of Behavior,* **32,** 335-340.

Graf, V., Bullock, D. H., & Bitterman, M. E. (1964). Further experiments on probability-matching in the pigeon. *Journal of the Experimental Analysis of Behavior,* **7,** 151-157.

Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior,* **21,** 463-473.

Leventhal, A. M., Morrell, R. F., Morgan, E. F., Jr., & Perkins, C. C., Jr. (1959). The relation between mean reward and mean reinforcement. *Journal of Experimental Psychology,* **57,** 284-287.

Logan, F. A. (1965). Decision making by rats: Uncertain outcome choices. *Journal of Comparative and Physiological Psychology,* **59,** 246-251.

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. (in press). An adjusting schedule for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. Reinforcer value: The effect of delay and intervening events.* Cambridge, MA: Ballinger.

McMillan, J. C. (1971). Percentage reinforcement of fixed-ratio and variable-interval performances. *Journal of the Experimental Analysis of Behavior,* **15,** 297-302.

Neuringer, A. J., & Chung, S. H. (1967). Quasi-reinforcement: Control of responding by a percentage-reinforcement schedule. *Journal of the Experimental Analysis of Behavior,* **10,** 45-54.

Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior,* **17,** 15-22.

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

Shimp, C. P. (1967). Reinforcement of least-frequent sequences of choices. *Journal of the Experimental Analysis of Behavior,* **10,** 57-65.

Staddon, J. E. R. (1970). Temporal effect on reinforcement: A negative "frustration" effect. *Learning and Motivation,* **1,** 227-247.

Staddon, J. E. R., & Innis, N. K. (1966). Preference for fixed vs. variable amounts of reward. *Psychonomic Science,* **4,** 193-194.

Staddon, J. E. R., & Innis, N. K. (1969). Reinforcement omission on fixed-interval schedules. *Journal of the Experimental Analysis of Behavior,* **12,** 689-700.

Wilson, W. A., Jr. (1960). Supplementary report: Two-choice behavior of monkeys. *Journal of Experimental Pyschology,* **59,** 207-208.

Young, J. S. (1981). Discrete-trial choice in pigeons: Effects of reinforcer magnitude. *Journal of the Experimental Analysis of Behavior,* **35,** 23-29.

Zeiler, M. D. (1972). Fixed-interval behavior: Effects of percentage reinforcement. *Journal of the Experimental Analysis of Behavior,* **17,** 177-189.

Received May 9, 1984
Final acceptance March 12, 1985