*DISCRETE-TRIAL CHOICE IN PIGEONS:
EFFECTS OF REINFORCER
MAGNITUDE*

JOHN S. YOUNG

BROWN UNIVERSITY

The preference of pigeons for large reinforcers which occasionally followed a response versus small reinforcers which invariably followed a response was studied in a discrete-trial situation. Two differently colored keys were associated with the two reinforcement alternatives, and preference was measured as the proportion of choice trials on which the key associated with uncertain reinforcement was pecked. A combination of choice and guidance trials insured that received distributions of reinforcement equalled the scheduled distributions. For five of six subjects, preference for the uncertain reinforcer appeared to be a linear function of the magnitude of the certain reinforcer. In addition, there was greater preference for the response alternative associated with uncertain reinforcement than would be expected on the basis of net reinforcer value.

Key words: choice, discrete trials, reinforcer magnitude, maximizing, matching law, key peck, pigeon

Recent discussions of choice behavior in animals have centered on two alternative explanatory models. The maximization model, which predicts that subjects will choose the response alternative associated with the greatest net reinforcement, was developed (for animal subjects) in the context of discrete-trial probability learning, and reflects the tendency of subjects in this paradigm to respond exclusively on the alternative associated with the higher probability of reinforcement (e.g. Shimp, 1966, 1969). Recently, a more generalized maximization model has been extended to a variety of other contexts (Rachlin, Green, Kagel, & Battalio, 1976). The alternative model, Herrnstein's matching law (e.g., 1970, 1974), is based on data from free-operant choice studies (concurrent schedules), and holds that the absolute rate of a response is proportional to its relative rate of reinforcement.

Although both models are derived from data on the effects of rate or probability of rein-

forcement on choice, they have been extended to incorporate the effects of other reinforcement parameters such as delay or magnitude. The research described below represents an attempt to provide additional data relevant to the assessment of these models of choice behavior, by examining the effects of reinforcer magnitude on pigeons' choices between certain and uncertain reinforcers in a discrete trial procedure. Several studies (Brownstein, 1971; Catania, 1963; Fantino, Squires, Delbrück, & Peterson, 1972; Schneider, 1973; Todorov, 1973) have assessed the effects of reinforcer magnitude on free-operant choice in pigeons, but the data of these studies can be accommodated by either model. By contrast, the two models make clearly different predictions for discrete-trial procedures employing correction trials, the maximization model predicting exclusive preference for one alternative while the matching law predicts a graded choice function.

An advantage of employing uncertain outcomes in studying the effects of reinforcer magnitude on choice, rather than a simple comparison of different magnitudes, is that the probability differences which offset differences in reinforcer magnitude can be used as an index of the strength of preference induced by the magnitude differences. The present experiment measured preference for a response

---

This research was supported by NIMH grant MH-02456 to Donald Blough. John S. Young was supported by a NSF predoctoral fellowship. The author wishes to thank Donald and Patricia Blough for their many helpful suggestions regarding this manuscript. Reprints may be obtained from John S. Young, The Johns Hopkins University School of Hygiene and Public Health, Department of Environmental Health Sciences, Division of Toxicology, 615 North Wolfe Street, Baltimore, Maryland 21205.

alternative associated with a reinforcer of 10 food pellets presented with a probability of .5 as the magnitude of a certain reinforcer associated with an alternative response varied between 1 and 10 pellets.

METHOD

Subjects

The subjects were six male White Carneaux pigeons (*Columba livia*). Birds 448, 3779, and 3780 had had prior experience in steady-state wavelength generalization tasks, and were approximately three years old at the beginning of the present study. Birds 398, 2554, and 2582 had previously been used in an undergraduate learning laboratory; they were between one and two years old at the beginning of training in this study. All birds were maintained at 80% of their free-feeding weights, receiving the bulk of their food in the experimental sessions. Water and grit were continuously available in their home cages.

Apparatus

The study employed three standard Lehigh Valley Electronics three-key pigeon chambers, with several modifications to the intelligence panels. The center key of each panel was replaced with a metal plate. The two side keys, the centers of which were 25 cm above the floor of the chamber and 6.5 cm to the left or right of the center of the panel, required forces of approximately .15 N to operate. They could be transilluminated with either red or green light. A standard houselight mounted in the center of the panel 3 cm below the ceiling of the chamber provided indirect general illumination at all times except during feeder-light illumination.

The hopper feeder in each box was replaced with a pellet feeder designed to deliver 20 mg food pellets (P. J. Noyes formula C pigeon pellets). A tube conveyed the pellets to a hemispherical food dish, which projected from the intelligence panel 6 cm into the experimental chamber. This rubber dish had a diameter of 5.5 cm, and was mounted in a 6 cm by 6.5 cm plywood frame. Its top was 5.2 cm above the chamber floor. The food cup could be illuminated by a light mounted behind the intelligence panel.

White noise to mask extraneous sounds was provided by a Grason-Stadler model 901 noise

generator. Ventilation fans provided a constant supply of fresh air and additional masking sound.

All experimental contingencies and data recording were controlled by a LINC computer.

Procedure

Subjects were first familiarized with the pellets and rubber food dish in their home cages. When they had begun to consume the pellets readily, they were trained in several daily sessions to approach and eat pellets delivered to the illuminated food cup in the experimental chamber. They were then retrained by hand to peck red and green keys on both sides of the intelligence panel.

After pecking was established, the pigeons were given several days of training in a single stimulus procedure. Each trial began with the illumination of a single key. A single peck at a lighted key produced a reinforcer of two pellets (separated by 400 msec) and extinguished the keylight. The food cup light was illuminated for a 5-sec eating period, which was followed by a 5-sec intertrial interval (ITI). The four possible side-color combinations were each presented 10 times, in random order, in a 40-trial block; a daily session consisted of two or three blocks of trials.

Training of the first three subjects on the initial choice procedure began immediately following single-stimulus training. Under this procedure each key color was correlated with a particular distribution of reinforcers. For the first three subjects (448, 3779, 3780) responses to the red key, regardless of side, were always followed by food pellets. Pecks to the green key were followed by 10 pellets on 50% of the trials and no pellets on the other 50% of the trials. Two types of trials were presented. On a choice trial, both left and right keys were illuminated, one with red light and the other with green. The particular side-color assignment was determined by a random process, with either side-color association being equally probable. On a guidance trial, only a single key was illuminated. The outcome of each choice trial was determined by random processes and the pigeon's behavior. Guidance trials were used to control, independently of each subject's pattern of choice, the distribution of reinforcers it received.

Each trial started simultaneously for all three subjects with the illumination of one or

both key lights. A single peck to a lighted key turned out the keylight and houselight, and initiated reinforcer delivery. Pellets were delivered at a rate of one every 400 msec into the food cup. Four sec after the initiation of pellet delivery, the feeder light was illuminated; it remained on for 7.5 sec, after which the houselight came on and the ITI began. The ITI lasted a minimum of 5 sec from the termination of feeder light illumination for the last bird to respond. Pecks by any bird during the ITI delayed the onset of the next trial by 1 sec (timed from the last ITI peck).

For the first three subjects, each session initially consisted of four series of 40 trials. The first trial in each series was a choice trial, and the remaining 39 trials were guidance trials. Sessions were run in five-day blocks, one per week, with the reinforcement distribution constant within a block.

The large proportion of guidance trials reflects an original plan to investigate the effects of variations in the probability distribution of the uncertain reinforcer, as well as those of variations in the magnitude of the certain and uncertain reinforcers. The probability of a large reinforcer following a response to the green key was to vary from .1 to .9, in steps of .1. Presenting the proper distribution of reinforcers and equating exposure to the red key, while counterbalancing for side, required 40 trials. Thus 39 guidance trials followed each choice, to insure the proper reinforcement history on the next choice trial.

Table 1 lists the experimental conditions in the order in which they were presented. The initial value of the certain reinforcer for Subjects 448, 3779, and 3780 was five pellets. Thus the expected values (reinforcer magnitude $\times$ probability) of the reinforcers associated with the two keys were equal. The pigeons were trained under this condition for 16 blocks of sessions (with some sessions missing); each received more than 12,000 trials. The value of the certain reinforcer was increased when performance (percent choice of the variable reinforcer) seemed stable (visual inspection of the data). The third column of Table 1 shows the number of blocks at each value of the certain reinforcer. Performance had generally stabilized by the second block.

After the pigeons had been presented with certain reinforcers of five pellets through 10 pellets, a new procedure began. The size of a

series was changed from 40 to 16 trials; this enabled the collection of data from 10 choices per session, 50 per block. A second group of subjects (398, 2554, and 2582), after pretraining similar to that of the first group, was started on the new procedure at this point. The color-reinforcement assignment for this group was reversed relative to that of the first group.

Both groups were trained for four blocks in the new procedure, with the value of the certain reinforcer at 10 pellets. Subsequently, all six birds were exposed to two blocks (1,600 trials) at each value of the certain reinforcer between 10 pellets and 1 pellet, in descending order. The original three subjects were also given two sessions with the value of the certain reinforcer set at zero, but they performed very poorly and failed to complete the second session for that condition. Finally, these three subjects were given two more blocks at each value of the certain reinforcer between two and five pellets, in ascending order.

RESULTS

Figure 1 shows the percentage of choice trials on which the stimulus associated with uncertain reinforcement was pecked, across different magnitudes of the certain reinforcer, for each subject. Data are for the last 50 choices (one block) in the 16-trial series conditions, and for the last 40 choices (two blocks) in the 40-trial series conditions. These data are also presented in Table 1, along with the standard deviations of the daily percentages. These standard deviations should be interpreted cautiously. A difference in one choice between two days corresponds to a difference of 25% in the 40-trial series conditions, and 10% in the 16-trial series conditions. For the first three subjects, the difference between replications, as seen in Figure 1, is probably a better measure of variability.

The most striking aspect of the data is that choice was clearly a graded function of the value of the certain reinforcer. Preference for either response alternative did not approach exclusivity except at extreme values of the independent variable. Regressions employing a variety of models (power, logarithmic, linear, exponential, and quadratic) indicated that percent choice of the uncertain reinforcer is best characterized as a linear func-

Table 1

Percent choice of the uncertain reinforcer across experimental conditions (means and standard deviations of daily values).

First groups of subjects^a

Condition Number	Value of Certain S^R	Number of Blocks	Subjects					
			448		3779		3780	
			X	δ	X	δ	X	δ
1	5	16	60.0	29.3	77.5	21.9	87.5	17.7
2	6	3	62.5	27.0	67.5	23.7	77.5	14.2
3	7	5	45.0	19.7	45.0	10.5	87.5	17.7
4	8	5	47.5	18.4	10.0	17.5	67.5	16.9
5	9	2	45.0	35.0	7.5	12.1	70.0	15.8
6	10	4	22.5	24.9	5.0	10.5	35.0	24.2
7	10	4	28.0	13.0	4.0	5.5	50.0	12.2
8	9	2	48.0	14.8	8.0	8.4	44.0	8.9
9	8	2	42.0	14.8	16.0	13.4	46.0	20.7
10	7	2	48.0	17.9	12.0	11.0	48.0	11.0
11	6	2	64.0	15.2	40.0	10.0	66.0	11.4
12	5	2	78.0	11.0	44.0	13.4	62.0	11.0
13	4	2	76.0	11.4	60.0	14.1	78.0	17.9
14	3	2	96.0	8.9	68.0	17.9	92.0	8.4
15	2	2	88.0	8.4	88.0	11.0	84.0	11.4
16	1	2	98.0	4.5	96.0	8.9	98.0	4.5
17	2	2 ^c	96.0	5.5	94.0	5.5	90.0	10.0
18	3	3 ^d	90.0	7.1	82.0	13.0	92.0	11.0
19	4	2	86.0	16.7	66.0	25.1	98.0	4.5
20	5	2	98.0	4.5	36.0	8.9	92.0	13.0

Second group of subjects^b

Condition Number	Value of Certain S^R	Number of Blocks	Subjects					
			398		2554		2582	
			X	δ	X	δ	X	δ
1	10	4	40.0	15.8	32.0	21.7	8.0	8.4
2	9	2	50.0	10.0	38.0	19.2	8.0	8.4
3	8	2	56.0	5.5	52.0	23.9	30.0	10.0
4	7	2	46.0	15.2	64.0	16.7	46.0	11.4
5	6	2	64.0	22.9	92.0	8.4	62.0	13.0
6	5	2	42.0	8.4	84.0	15.2	48.0	16.4
7	4	2	52.0	8.4	92.0	8.4	50.0	17.3
8	3	2	36.0	13.4	88.0	11.0	60.0	0.0
9	2	2	60.0	18.7	100.0	0.0	56.0	18.2
10	1	3 ^d	66.0	11.4	96.0	5.5	82.0	19.2

a. Data for Conditions 1 to 6 are from last two 5-day blocks of 4-choice sessions; data for Conditions 7 to 20 are from final 5-day block of 10-choice sessions.

b. All data are from final 5-day block of 10-choice daily sessions.

c. Data missing from two sessions.

d. Includes two partial blocks of sessions.

tion of the magnitude of the certain reinforcer, although frequently a minor improvement in fit could be obtained by adding a small quadratic term to the model. The data from Bird 398 are something of an exception; none of the models provided a clearly better fit to that pigeon's pattern of relative indifference at all values of the certain reinforcer. Figure 2, which plots the mean of the choice percentages for all subjects except Bird 398, illustrates the linear trend of the data.

A second interesting feature of the subjects' patterns of choice is that three of the six birds clearly showed greater preference for the uncertain reinforcer than would be expected on the basis of the objective expected value of the two alternatives. They preferred the uncertain reinforcer when it was equal in expected value to the certain reinforcer, and were indifferent when the expected value of the certain reinforcer was somewhat greater than that of the uncertain reinforcer. Bird 398 was again an

Fig. 1. Percent choice of the uncertain reinforcer as a function of the value of the certain reinforcer. A. Bird 448. Solid line represents decreasing values of the certain reinforcer, dotted line represents initial ascending series, dashed line represents final ascending series. B. Bird 3779; same conventions as Bird 448. C. Bird 3780; same conventions as Bird 448. D. Second set of subjects. Solid line represents Bird 398, dashed line Bird 2554, dotted line Bird 2582.

exception, and the choices of Birds 3779 and 2582 were more nearly balanced, but no subject showed a preference for the certain reinforcer.

The performance of the two groups of subjects did differ to a certain degree. Bird 398 was an extreme example, but Bird 2582 also exhibited an extended range of indifference, and correspondingly less preference for the response alternative associated with uncertain reinforcement. Both these subjects, unlike the other four, exhibited pronounced left-key preferences as well. Whether this side preference underlay the indifference to key color, or

simply reflects this indifference, is impossible to say.

As noted above, the extent of preference for the uncertain reinforcement alternative also varied across replications of the different comparisons in the first group of subjects. These changes seem limited to the elevations of the choice functions (degree of preference for the uncertain reinforcer), rather than the slopes of the functions. It is uncertain whether these changes are systematic, or simply represent random variation. There may be an influence of procedure (40- vs 16-trial series) or prior experience, as two subjects show decreased preference

Fig. 2. Percent choice of the uncertain reinforcer as a function of the value of the certain reinforcer. Mean value of individual subject means.

for the uncertain reinforcement alternative on the descending series of certain reinforcer values following the initial ascending series. No consistent differences, however, are found between the descending series and the final ascending series, and the three replications of the comparison with a certain reinforcer of five pellets show no consistent direction of change. In any event, no subject displayed a preference for the certain reinforcer.

DISCUSSION

The finding that the pigeons' preference for the uncertain reinforcement alternative changed continuously, rather than abruptly, as the magnitude of the certain reinforcer changed is clearly inconsistent with a reinforcement maximization principle. It is also somewhat unexpected, as previous studies employing a similar procedure with rats in runway apparatus (Leventhal, Morrell, Morgan, & Perkins, 1959; Logan, 1965) found that preferences were essentially all-or-none.

The accommodation of the data to Herrnstein's matching law is more problematic. The subjects' choice proportions clearly do not match the objective expected values of the reinforcement alternatives. The uncertain reinforcer is more preferred than would be expected on the basis of expected value, and preference is more extreme than would be expected when both certain and uncertain reinforcers are set at 10 pellets.

This pattern of preference suggests that the averaging of reinforcement events associated with the response alternatives is not linear, but weighted. Alternately stated, there is a nonlinear scaling of reinforcer magnitude or probability. Deviations of this sort from strict matching are not uncommon in the experimental literature, and a variety of modified matching laws, incorporating additional parameters, have been proposed to accommodate them. De Villiers (1977) discusses some of these, which incorporate weighting (scaling) factors for reinforcer magnitude, delay, or probability. Another alternative is a generalized matching law, incorporating two parameters, as proposed by Baum (e.g., 1979).

Such additions of free parameters to the matching law are troublesome in the absence of supporting logic and demonstrations of transsituational generality. In this regard, the model advanced by Baum enjoys the advantage of having a plausible interpretation for its parameters, with testable implications. Interestingly, fitting the present data to Baum's model requires that the slope parameter, a , be noticeably greater than one (approximately 2.0). This is a rare occurrence in the data he discusses, and one to which his analysis of factors influencing the parameter does not extend.

While deviations from strict matching are not unusual, the direction of the deviation seen in the present case is somewhat surprising. It would appear to contradict a fairly general finding from concurrent schedule research (e.g. Schneider, 1973; Todorov, 1973) that small reinforcers delivered at a high rate are preferred to larger reinforcers delivered at a lower rate, even if the net values of the two reinforcement alternatives are objectively equal.

The explanation of this discrepancy is uncertain. If the results reflect the scaling of reinforcer magnitude, the difference between studies may reflect differences in the range of reinforcer magnitudes employed. Reinforcers in the present study ranged from 0 to .2 g, whereas in Schneider's (1973) study, the reinforcers ranged from .18 g to 1.35 g. Comparison with studies using hopper access duration as the measure of reinforcer magnitude (Brownstein, 1971; Catania, 1963; Fantino et al., 1972; Todorov, 1973) are more difficult, but the range of durations employed in these

studies (1.5 to 8 sec) again probably corresponds to greater weights of food than those used in the present study. A function relating the behavioral value of a reinforcer to its magnitude which was positively accelerated at low magnitudes and negatively accelerated at higher values is plausible, and would predict the divergent patterns of results in the different studies. It would also predict that a study employing concurrent schedules and small reinforcers should find a preference for larger, less frequent reinforcers.

An alternative interpretation is that the data of the present study reflect a preference for uncertain reinforcement per se. It should be noted that this interpretation is distinct from a nonlinear scaling of reinforcer probability. If the subjective probability of response outcome is independent of the value of that outcome and the probabilities of all possible outcomes must sum to 1.0, as suggested by most discussions of the concept, a situation in which each of two events is presented with a probability of .5 should be dealt with objectively.

If the present results do reflect a preference for uncertainty as such, one would expect to see a different pattern of preference when both reinforcement alternatives were uncertain, as in concurrent variable interval schedules, or when both were certain. The two studies of rats mentioned earlier (Leventhal et al., 1959; Logan, 1965) lend credence to this possibility. Both used reinforcers of approximately the same overall magnitude, but in Leventhal et al.'s study one reinforcement alternative was uncertain (.5 g or 0 g food) while in Logan's it was variable (9 45-mg pellets or 1 45-mg pellet). Leventhal et al.'s results were in the same direction as those of the present study, while Logan's showed an opposite pattern of preference. If pigeons do in fact have a preference for uncertainty, a discrete-trial study employing variable, rather than uncertain, reinforcement should obtain results more in accordance with those of concurrent schedule research.

REFERENCES

- Baum, W. M. Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, 1979, **32**, 269-281.
- Brownstein, A. J. Concurrent schedules of response-independent reinforcement: Duration of a reinforcing stimulus. *Journal of the Experimental Analysis of Behavior*, 1971, **15**, 211-214.
- Catania, A. C. Concurrent performances: A baseline for the study of reinforcement magnitude. *Journal of the Experimental Analysis of Behavior*, 1963, **6**, 299-300.
- de Villiers, P. Choice in concurrent schedules and a quantitative formulation of the law of effect. In W. K. Honig & J. E. R. Staddon (Eds.) *Handbook of operant behavior*, Englewood Cliffs, N.J.: Prentice Hall, 1977.
- Fantino, E., Squires, N., Delbrück, N., & Peterson, C. Choice behavior and the accessibility of the reinforcer. *Journal of the Experimental Analysis of Behavior*, 1972, **18**, 35-45.
- Herrnstein, R. J. On the law of effect. *Journal of the Experimental Analysis of Behavior*, 1970, **13**, 243-266.
- Herrnstein, R. J. Formal properties of the matching law. *Journal of the Experimental Analysis of Behavior*, 1974, **21**, 159-164.
- Leventhal, A. M., Morrell, R. F., Morgan, E. F. Jr., & Perkins, C. C. Jr. The relation between mean reward and mean reinforcement. *Journal of Experimental Psychology*, 1959, **57**, 284-287.
- Logan, F. A. Decision making by rats: uncertain outcome choices. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 246-251.
- Rachlin, H., Green, L., Kagel, J. H., & Battalio, R. C. Economic demand theory and psychological studies of choice. In G. Bower (Ed.) *The psychology of learning and motivation* (Vol. 10). New York: Academic Press, 1976.
- Schneider, J. W. Reinforcer effectiveness as a function of reinforcer rate and magnitude: A comparison of concurrent performances. *Journal of the Experimental Analysis of Behavior*, 1973, **20**, 461-471.
- Shimp, C. P. Probabilistically reinforced choice behavior in pigeons. *Journal of the Experimental Analysis of Behavior*, 1966, **9**, 443-455.
- Shimp, C. P. Optimal behavior in free-operant experiments. *Psychological Review*, 1969, **76**, 97-112.
- Todorov, J. C. Interaction of frequency and magnitude of reinforcement on concurrent performances. *Journal of the Experimental Analysis of Behavior*, 1973, **19**, 451-458.

Received February 7, 1980

Final acceptance July 3, 1980