

Choice behavior in transition: Development of preference in a free-operant procedure

JAMES E. MAZUR and THERESA A. RATTI
Southern Connecticut State University, New Haven, Connecticut

Twenty acquisition curves were obtained from each of 8 pigeons in a free-operant choice procedure. Every condition began with a phase in which two response keys had equal probabilities of reinforcement, and, as a result, subjects' responses were divided fairly evenly between the two keys. This was followed by a phase in which one key had a higher probability of reinforcement than the other, and the development of preference was observed. In all but a few cases, response proportions increased for the key with the higher probability of reinforcement. In most conditions, the two probabilities differed by .06, but the actual probabilities varied (from .16 and .10 in one condition to .07 and .01 in another). Development of preference for the key with the higher probability of reinforcement was slower when the ratio of the two reinforcement probabilities was small (.16/.10) than when it was large (.07/.01). This finding is inconsistent with the predictions of several different quantitative models of acquisition, including the kinetic model (Myerson & Miezin, 1980) and the ratio-invariance model (Horner & Staddon, 1987). However, the finding is consistent with a hypothesis based on Weber's law, which states that the two alternatives are more discriminable when the ratio of their reinforcement probabilities is larger, and, as a result, the acquisition of preference is faster.

In recent years, there has been increasing interest in the question of how subjects develop a preference for one alternative or the other in a choice situation. A number of different theoretical approaches to the acquisition of choice behavior have been proposed. Since the 1950s, linear-operator models have been frequently used to describe acquisition curves (e.g., Bush & Mosteller, 1955). Several optimal-foraging models of acquisition have used variations of the linear-operator model (e.g., Harley, 1981; Kacelnik, Krebs, & Ens, 1987; Lester, 1984). Linear-operator models have not been universally accepted, however, and other theories of acquisition have arisen from research on the matching law, including melioration theory (Herrnstein & Vaughan, 1980; Vaughan & Herrnstein, 1987) and the kinetic model (Myerson & Hale, 1988; Myerson & Miezin, 1980). Still another approach is provided by Staddon's ratio-invariance model (Horner & Staddon, 1987; Staddon, 1988; Staddon & Horner, 1989).

Compared with the abundance of theoretical approaches, empirical research on acquisition in choice situations has been quite scarce. Myerson has analyzed the results of studies that supported some predictions of his kinetic model (Catania, 1969; Killeen, 1972; Myerson & Hale, 1988). Kacelnik et al. (1987) reported an experiment that failed to support the predictions of some optimal-foraging

theories based on the linear-operator model, and they developed a more complex model to deal with their results. Staddon and Horner (1989) presented some results that favored their model over a simple linear-operator model, but their analyses were based on long-term aggregates of choice behavior rather than on patterns of acquisition over time. Except for these studies and perhaps a few others, the sizable literature on choice behavior has included only measures of the "steady-state" behavior that is achieved after a subject has experienced the same choice situation for many sessions. Some characteristics of steady-state behavior can be used to test acquisition theories (see Myerson & Miezin, 1980), but only indirectly.

It seems clear that much more research on the acquisition of choice behavior will be necessary before the strengths and weaknesses of the different theoretical approaches can be evaluated. To compare the predictions of several different acquisition theories, Bailey and Mazur (1990) used a procedure in which the development of preference could be examined repeatedly with the same subjects. This procedure allowed them to make within-subjects comparisons of acquisition rates with different pairs of reinforcement schedules. Because the present experiment used a similar procedure, it will be described in some detail.

Bailey and Mazur (1990) obtained one acquisition curve from each pigeon subject in each condition of their experiment. Every condition began with a few *one-probability sessions*, which were followed by one *transition session* and then several *two-probability sessions*. In the one-probability sessions, a single reinforcement schedule assigned reinforcers to either of two response keys, with

This research was supported by Grant MH 38357 from the National Institute of Mental Health. We thank Millie Hemming for her help in various phases of this research. Correspondence should be addressed to James E. Mazur, Psychology Department, Southern Connecticut State University, New Haven, CT 06515.

equal probability. Once a reinforcer was assigned to one key, no additional reinforcers were assigned to either key until that reinforcer was collected. This interdependent scheduling of reinforcers ensured that roughly the same number of reinforcers was delivered by each key (cf. Stubbs & Pliskoff, 1969). As a result, subjects responded about equally often on the two keys during the one-probability sessions. The transition session used the same schedule as the preceding one-probability session for the first 100 trials, but for the remainder of the session, two independent schedules assigned the reinforcers for the two keys. The probability of reinforcement was higher for one key (the left key in some conditions and the right key in others) than for the other key. We will use the notation ϕ_1 to represent the higher probability of reinforcement and ϕ_2 to represent the lower probability of reinforcement. These two independent schedules, with the same values of ϕ_1 and ϕ_2 , remained in effect for all of the two-probability sessions that followed a transition session. In almost every case, the proportion of responses on Key 1 (the key with the higher probability of reinforcement) gradually increased during the two-probability phase. In brief, each condition began with sessions in which response proportions were about equal on the two keys; then the development of preference was observed when the probability of reinforcement became higher on one key.

Bailey and Mazur (1990) varied the values of ϕ_1 and ϕ_2 across conditions to determine how different pairs of reinforcement probabilities would affect the rate of acquisition. In particular, they used four different pairs of reinforcement probabilities in which the difference $\phi_1 - \phi_2$ was the same, but the ratio ϕ_1/ϕ_2 was different. The four pairs of reinforcement probabilities were .40, .30, .20, .10, .15, .05, and .12, .02. Notice that the reinforcement probabilities differ by .10 in each pair, but the ratio of the two probabilities ranges from 4/3 to 6/1. Bailey and Mazur found that the acquisition of preference for the key with the higher probability of reinforcement was significantly faster when the probabilities differed by a large ratio (i.e., 6/1) than when they differed by a small ratio (i.e., 4/3). This result is not consistent with the predictions of several of the theories mentioned above. Bailey and Mazur used computer simulations to determine the predictions of Staddon's (1988) ratio-invariance model and of the simple linear-operator model (Bush & Mosteller, 1955). They found that the ratio-invariance model predicts the opposite result—slower acquisition with larger probability ratios if the difference $\phi_1 - \phi_2$ is constant. The linear-operator model makes similar predictions, except that the predicted differences among conditions are much smaller. The original version of melioration theory (Herrnstein & Vaughan, 1980) predicts identical acquisition rates for all conditions with the same difference between reinforcement probabilities ($\phi_1 - \phi_2$). Thus, none of these theories predicts the pattern of results obtained by Bailey and Mazur. However, a more recent version of melioration theory (Vaughan, 1985) is much more general in its form, and this version makes no specific predictions about this matter.

Bailey and Mazur (1990) showed that their results were consistent with an analysis based on Weber's law. Weber's law states that for any stimulus dimension (probability of reinforcement in this case), two stimuli will be more discriminable if they differ by a larger percentage. Following this principle, the difference between the probabilities .12 and .02 should be much more discriminable than the difference between .40 and .30, because the first two probabilities differ by 600% and the second two probabilities by only 33.3% (treating the smaller of the two probabilities as 100% in each case). Bailey and Mazur therefore suggested that the acquisition of preference was faster in conditions in which ϕ_1/ϕ_2 was greater because the difference between the two probabilities was more discriminable in these conditions.

The present experiment used a procedure similar to that of Bailey and Mazur (1990), except that a free-operant procedure was used. This was done both to examine the applicability of the Weber's law hypothesis to the free-operant case and to test the predictions of Myerson's kinetic model. The kinetic model makes no predictions for the Bailey and Mazur experiment because that experiment used a discrete-trial procedure, with a 1-sec intertrial interval (ITI) after each trial. Myerson and Hale (1988) suggested that the factors that control choice in discrete-trial situations are substantially different from those in free-operant procedures, and the predictions of the kinetic model are limited to the latter. For free-operant choice, however, the kinetic model makes the same prediction as does the original version of melioration theory—acquisition rates should be identical for all conditions with the same difference between reinforcement probabilities ($\phi_1 - \phi_2$). Therefore, if the results of the present experiment were similar to those of Bailey and Mazur, they would be inconsistent with the predictions of the kinetic model. In addition, such results would tend to contradict the view that free-operant choice and discrete-trial choice are governed by different principles (Myerson & Hale, 1988).

The procedure of this experiment differed from that of Bailey and Mazur (1990) in several other ways. Because there were no ITIs, each session could include more responses, and sessions ended after 1,000 or 1,100 responses, as compared with 600 responses in the Bailey and Mazur experiment. Because of the larger number of responses, lower reinforcement probabilities were used to lessen the amount of satiation that might occur during a session. Four pairs of schedules with a .06 difference in reinforcement probabilities were used: .16, .10, .13, .07, .10, .04, and .07, .01. This experiment was therefore designed to determine whether higher ratios of ϕ_1/ϕ_2 would produce faster acquisition in a free-operant procedure, with a lower range of reinforcement probabilities than those used by Bailey and Mazur.

METHOD

Subjects

Eight White Carneau pigeons were maintained at 80% of their free-feeding weights. All had previous experience with a choice procedure similar to the one used in this experiment.

Apparatus

The experimental chamber was 30 cm long, 30 cm wide, and 30.5 cm high. Recessed in the right side wall of the chamber, 22.8 cm above the floor, were three response keys, separated by 5.6 cm, and each measuring 2.5 cm in diameter. The center key was not used in this experiment. Each of the two outside keys could be transilluminated by red or green light produced by two 2-W bulbs mounted behind the keys. A force of about 0.1 N was required to operate each key. A hopper that could present grain for a specified interval was situated 12 cm below the center key. When the grain was presented, the hopper was illuminated with white light from the two 2-W bulbs mounted above it. Two 2-W houselights were mounted above the Plexiglas ceiling of the chamber.

The chamber was enclosed in a sound-attenuating wooden box that had an exhaust fan for ventilation and for masking extraneous noises. An IBM-compatible personal computer was programmed in the Medstate language to control all stimuli and to record the responses.

Procedure

The pigeons all had considerable experience with a similar procedure, so no additional training was needed to get them to respond in the experiment. The experiment included 20 conditions, each consisting of two or three one-probability sessions, followed by one transition session and then three two-probability sessions. One-probability sessions and transition sessions ended after 1,100 responses or 30 min, whichever came first. Two-probability sessions ended after 1,000 trials or 30 min, whichever came first.

At the beginning of a session, the houselights came on, and the two response keys were illuminated—the left key red and the right key green. The houselights remained on for the entire session, except during reinforcement periods. Every effective peck on either key produced a 0.1-sec darkening of both keys as feedback. If a response resulted in reinforcement, the houselights and the keylights were turned off, and the hopper was illuminated and raised to present grain for 1.6 sec. After reinforcement, the houselights and the keylights were again illuminated. At the end of a session, all chamber lights were turned off. This procedure was used throughout the experiment, and the only differences among conditions were the probabilities of reinforcement for the two keys, as described below.

In one-probability sessions, the computer operated a single probabilistic reinforcement schedule for both alternatives. Before each response, a reinforcer might be assigned to one of the two keys, with equal probability. Once a reinforcer was assigned to one key, the next response on that key delivered the reinforcer, but no additional reinforcers would be assigned to either key until that reinforcer was collected.

In two-probability sessions, two independent probabilistic schedules assigned reinforcers for the two keys. Each response on Key 1 was reinforced with a probability of ϕ_1 ; each response on Key 2 was reinforced with a probability of ϕ_2 . The left key had the higher probability of reinforcement (ϕ_1) in 10 of the conditions; the right key had the higher probability in the other 10 conditions. In transition sessions, the first 100 trials used the same procedure as the preceding one-probability session, and the remaining trials used the same procedure as the subsequent two-probability sessions.

In the 20 experimental conditions, five different pairs of probabilities appeared in the two-probability sessions, with each pair occurring in 4 different conditions. In two-probability sessions, the left key had the higher probability of reinforcement in odd-numbered conditions; it had the lower probability of reinforcement in even-numbered conditions. In Conditions 5, 10, 15, and 20, ϕ_1 was .19 and ϕ_2 was .01. These conditions were included to examine rates of acquisition when there was a large difference in reinforcement probabilities, and they served as a point of comparison with all of the other conditions, in which the difference between reinforcement probabilities was .06. The other four pairs of reinforcement probabilities (.16, .10, .13, .07, .10, .04, and .07, .01) each appeared once

in every block of four consecutive conditions that were separated by the .19, .01 conditions. The presentation order of these four pairs of probabilities was counterbalanced across subjects.

In all one-probability sessions, the probability that a reinforcer would be assigned for the next response (assuming that no reinforcer had already been assigned and not yet delivered) was the mean of ϕ_1 and ϕ_2 for the two-probability sessions that were to follow in that condition. For instance, in the .16, .10 conditions, the probability of assigning a reinforcer in the one-probability sessions was .13. The mean of ϕ_1 and ϕ_2 was used for the one-probability phase to minimize the change in the overall probability of reinforcement when the two-probability phase began during a transition session. If a subject was responding on the two keys about equally often when the two-probability phase began (after the first 100 responses of the transition session), the overall probability of reinforcement would be the mean of ϕ_1 and ϕ_2 .

RESULTS

For two-probability conditions, Key 1 will continue to refer to the key with the higher probability of reinforcement. In one-probability sessions, the reinforcement probabilities were the same for the two keys, but Key 1 will refer to the key that had the higher reinforcement probability later in the condition, during the two-probability phase that followed. The proportion of responses on Key 1 will be denoted P_1 . By the last one-probability session, P_1 remained close to .5 in all conditions. For the five different probability pairs, P_1 ranged from .46 to .51 in the last one-probability session, with a mean of .49.

Perhaps the most straightforward way to examine the rates of acquisition with different reinforcement probabilities is to plot the values of P_1 over time for the different conditions. Figure 1 shows the values of P_1 in consecutive blocks of 500 responses for the five different probability pairs used in the experiment, averaged across subjects and across the four replications for each probability pair. The first point of each curve (labeled 0 on the x -axis) is based on the last 500 responses of the one-probability phase of each condition for each subject (the last 400 responses of the final one-probability session and the first 100 responses of the transition session). Blocks 1 and 2 are based on the final 1,000 responses of each transition session, during which time the two-probability phase was in effect. Blocks 3–8 present the results of the three two-probability sessions, with two blocks of 500 responses from each of the three sessions. For a small number of sessions (9 of the 800 sessions represented in Figure 1), the percentages are based on fewer than 500 responses, because a subject failed to complete all of the possible responses.

Figure 1 shows that the fastest acquisition of preference for Key 1 occurred in the .19, .01 conditions. There were also substantial differences in the rates of acquisition among the four probability pairs with $\phi_1 - \phi_2$ equal to .06. Acquisition was fastest in the .07, .01 condition and slowest in the .16, .10 condition, with intermediate and similar acquisition rates in the other two conditions. A two-way repeated measures analysis of variance (ANOVA) was conducted with the data from the four

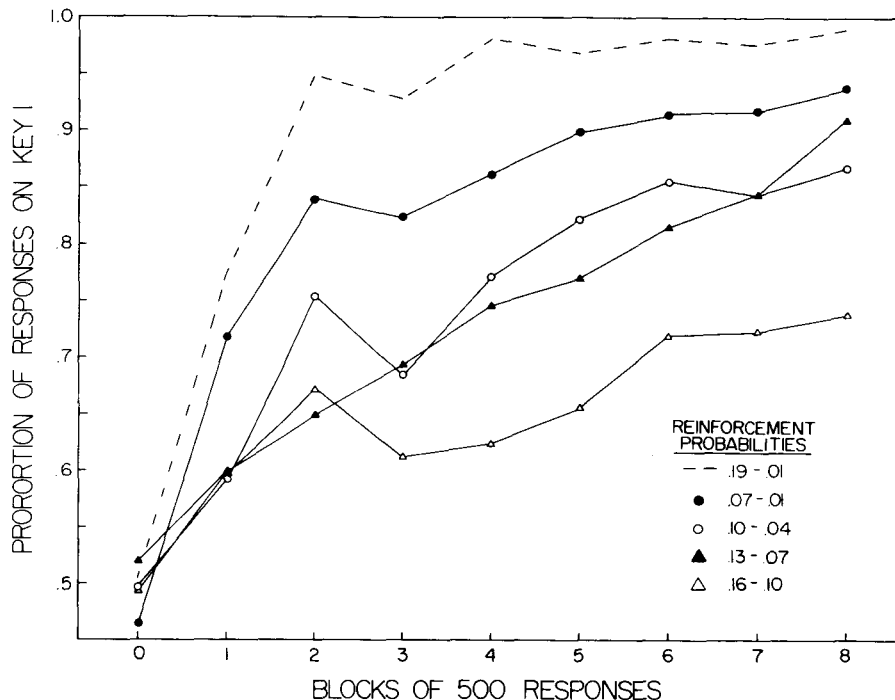


Figure 1. For each of the five probability pairs, the proportions of responses on Key 1 (the key with the higher probability of reinforcement) are shown in 500-response blocks. Block 0 includes the last 500 responses of the one-probability phase, and Blocks 1 through 8 are from the two-probability phase.

probability pairs other than .19, .01. The results from the .19, .01 conditions were excluded because we wished to test the hypothesis that acquisition rates would be identical for all conditions in which $\phi_1 - \phi_2$ was the same. The ANOVA revealed a significant effect of blocks of responses [$F(8,56) = 36.86, p < .01$], as well as a significant effect of reinforcement probabilities [$F(3,21) = 10.29, p < .01$]. The blocks by reinforcement probability interaction was also significant [$F(24,168) = 159.33, p < .01$].

To test this same hypothesis in a different way, we conducted an analysis similar to the one used by Bailey and Mazur (1990). The strategy involved fitting an equation to each set of acquisition data and then comparing estimates of a learning-rate parameter across conditions. A curve-fitting procedure was used to determine the value of the learning-rate parameter that yielded the best correspondence between the data and the following hyperbolic equation:

$$P_1(r) = [1 - P_1(0)] \left(\frac{r}{r+M} \right) + P_1(0). \quad (1)$$

$P_1(0)$ is the choice proportion for Key 1 at the start of the two-probability phase; $P_1(r)$ is the choice proportion after r responses in the two-probability phase. M is the learning-rate parameter: Larger values of M indicate slower rates of acquisition. This equation assumes that P_1 begins at a value of $P(0)$ and approaches an asymptote

of 1.0. Other equations could be used to describe the pattern of acquisition; in fact, Bailey and Mazur tried three other equations in addition to Equation 1 (exponential, logistic, and linear equations). We chose to use the hyperbolic equation because it has been successfully applied to other learning curves (Mazur & Hastie, 1978) and because it provided the most accurate fits for the results of Bailey and Mazur.

The curve-fitting procedure used a least-squares criterion to determine the best-fitting value of M for each of the 160 acquisition curves obtained in this experiment. For each curve, $P_1(0)$ was set equal to the proportion of responses on Key 1 during the last 500 trials of the one-probability procedure—the last 400 responses of the final one-probability session plus the first 100 trials of the transition session. Therefore, M was the only free parameter in Equation 1. M can be interpreted as an estimate of the number of responses needed for $P_1(r)$ to reach the midpoint between $P_1(0)$ and 1.0. In some cases, the estimated value of M either was very large or was not obtainable, either because there was little increase in preference for Key 1 or because a subject developed a preference for Key 2, with its lower probability of reinforcement. A large value for M is appropriate in these cases, because both of these behavior patterns indicated that the subject was slow to develop a preference for the key with the higher probability of reinforcement. However, a few very large values of M could have a large influence on the group means. To avoid this problem, we arbitrarily set

an upper limit of 10,000 on the value of M . In 14 of the 160 cases, either this limit was exceeded or no estimate of M could be obtained. For these 14 cases (10 of which occurred in the .16, .10 conditions), M was set at 10,000 for the data analyses described below.

Figure 2 shows the mean values of M for the five different probability pairs. As can be seen, the mean value of M increased as the ratio of ϕ_1/ϕ_2 grew larger. A one-way repeated measures ANOVA (which excluded the results from the .19, .01 conditions) found a significant effect of the probability pairs [$F(3,21) = 16.11, p < .01$]. In addition, the increasing trend in M that is predicted by the Weber's law hypothesis was supported by a significant linear contrast [$F(1,21) = 40.50, p < .01$]. The analysis based on the best-fitting hyperbolic curves therefore leads to the same conclusion as the analysis based on the raw data shown in Figure 1: With the difference $\phi_1 - \phi_2$ held constant, acquisition was faster in the conditions with the larger ratios of ϕ_1/ϕ_2 .

Besides measuring the overall rates of acquisition, Bailey and Mazur (1990) examined how individual trials ending with reinforcement or nonreinforcement affected the value of P_1 on the next several trials. To determine whether the patterns they observed in their discrete-trial procedure would also be found in this free-operant procedure, we conducted a similar analysis. Consider all of the responses made on Key 1. Each of these responses might or might not be reinforced. If response n was a reinforced response on Key 1, what was the probability that response $n+1$ would also occur on Key 1? Was this probability higher than if response n was not reinforced? We can ask similar questions about subsequent responses: If response n was a reinforced response on Key 1, was the probability of choosing Key 1 higher for response $n+2, n+3$, and so on, than if response n was not reinforced? The same types of questions can be asked about the effects of reinforcement and nonreinforcement for responses on Key 2. We will call any response that occurred after response n

a *same-key response* if it was made on the same key as response n (regardless of which key may have been pecked between response n and the later response in question).

To examine the effects of reinforcement on subsequent choices, the response-by-response data were analyzed from three different portions of each condition—the first 500 responses of the final one-probability session, the first 500 responses of the two-probability phase (i.e., responses 101–600 of the transition session), and the first 500 responses of the final two-probability session. These segments were selected because we reasoned that they would be representative of behavior late in the one-probability phase, early in the two-probability phase, and late in the two-probability phase, respectively. Figure 3 presents the results for responses on Key 1. The results are shown separately for each of the five probability pairs, averaged across subjects and replications. For all reinforced pecks on Key 1, the solid lines in Figure 3 show the proportion of same-key responses for each of the next 6 responses. To eliminate any potential effects of other reinforcers that sometimes occurred between response n and later responses, the calculations excluded all cases in which a reinforcer was delivered for any intervening response. For instance, in the upper left panel, the solid line for 6 responses ahead is plotted at .52. This means that after a reinforced response on Key 1, if there were no reinforcers for the next 5 response, the probability that Key 1 would be chosen as the 6th response was .52. The dashed lines in Figure 3 show the proportion of same-key responses for the 6 responses that followed all nonreinforced pecks on Key 1, again excluding all cases in which a reinforcer was delivered for any intervening response. The number of responses used to obtain the proportions varied widely for the different data points in Figure 3. For the reinforced responses (solid lines), more than 250 responses contributed to each data point in Figure 3. For the more numerous nonreinforced responses (dashed lines), more than 3,400 responses contributed to each data point.

The most consistent result shown in Figure 3 is the higher proportion of same-key responses on the first response after a reinforced response than after a nonreinforced response. The proportion of same-key responses after a reinforced response is higher in all 15 panels. In many cases, the difference between the two proportions was due to a decreased proportion of same-key responses immediately after a nonreinforced response. This low proportion of same-key responses one trial after a nonreinforced response, which was most pronounced in the one-probability phase and early in the two-probability phase, is indicative of a tendency to switch from one key to another after each nonreinforced response. Another sign of the tendency to switch keys after each nonreinforced response is the up-and-down pattern of the dashed lines in the first two columns of Figure 3. In all 10 panels of the first two columns, the same-key response proportions after nonreinforced responses (dashed lines) increased from one to two responses ahead, then decreased for three responses ahead, then increased again for four responses

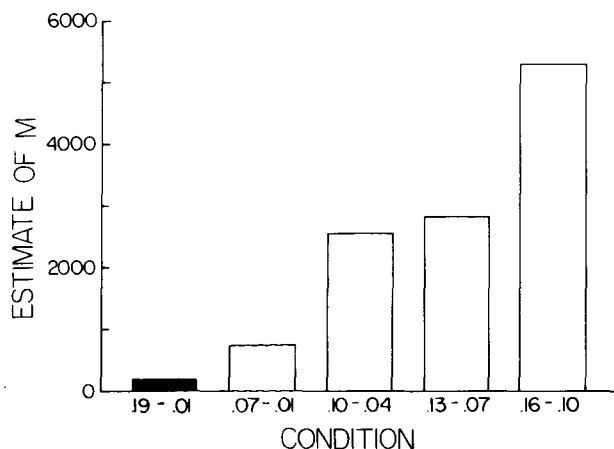


Figure 2. For each of the five probability pairs, the mean estimate of M , the learning-rate parameter in Equation 1, is shown. Larger values of M indicate slower acquisition.

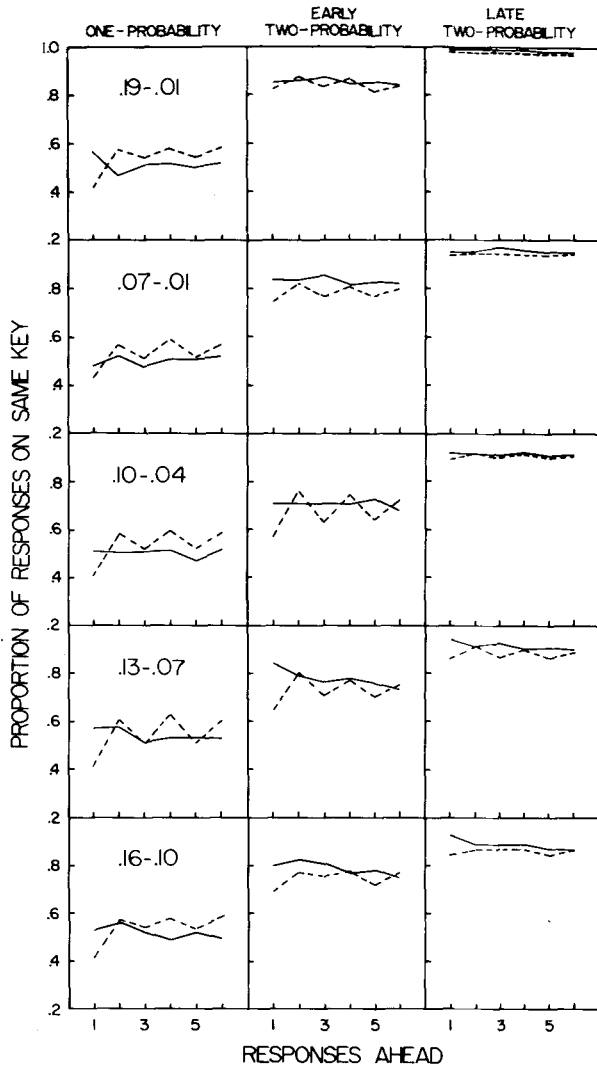


Figure 3. For the five probability pairs used in this experiment, the proportion of responses on Key 1 (the key with the higher probability of reinforcement) are shown for the 6 responses that followed reinforced pecks (solid lines) and nonreinforced pecks (dashed lines) on Key 1. The three columns show the results for the first 500 responses of the final one-probability session, the first 500 responses of the two-probability phase, and the first 500 responses of the final two-probability session, respectively.

ahead, and so on. By the end of the two-probability phase, this up-and-down pattern had largely disappeared (right column). No such pattern was found for the same-key response proportions after reinforced responses during any phase of the experiment.

Figure 3 also shows that late in the two-probability phase (right column), when there was much less of a tendency to switch keys after each nonreinforced response, the probability of a same-key response was slightly, but consistently, higher for several trials after a reinforced response than after a nonreinforced response. Averaged across all panels of the right column, the same-key response proportion was .946 for the first response after

a reinforced response and .903 after a nonreinforced response. By six trials ahead, the two response proportions were much closer, averaging .920 and .914 for reinforced and nonreinforced responses, respectively. Although the differences in response proportions were small, they were consistently higher after reinforced responses: In 28 of the 30 comparisons shown in the right column of Figure 3, the same-key response proportion was at least slightly higher after reinforced responses [$\chi^2(1) = 22.53, p < .01$].

Figure 4 has the same format as Figure 3, except that it shows the proportion of same-key responses for the 6

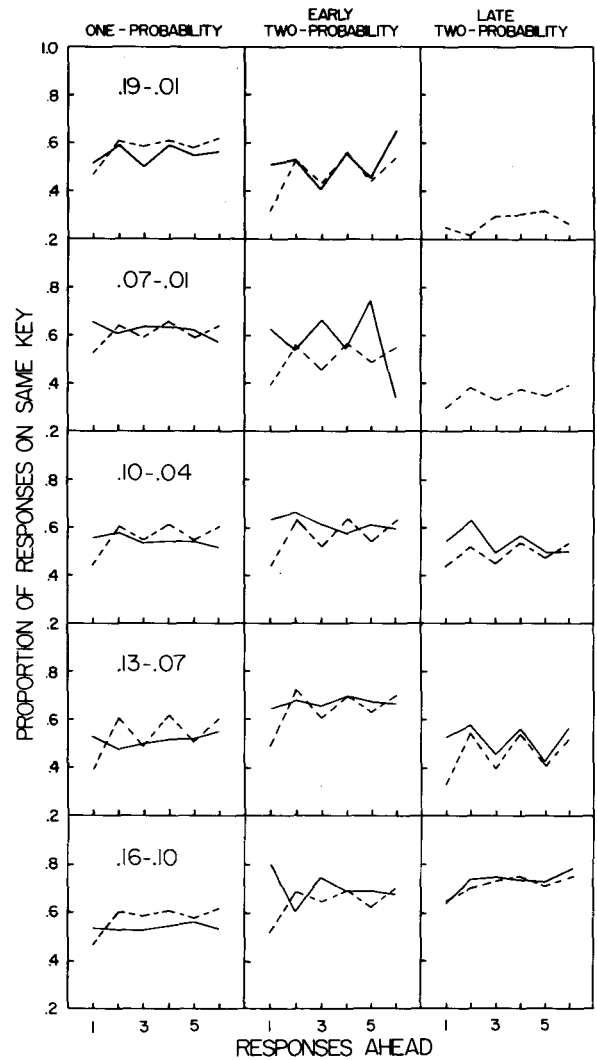


Figure 4. For the five probability pairs used in this experiment, the proportion of responses on Key 2 (the key with the lower probability of reinforcement) is shown for the 6 responses that followed reinforced pecks (solid lines) and nonreinforced pecks (dashed lines) on Key 2. The three columns show the results for the first 500 responses of the final one-probability session, the first 500 responses of the two-probability phase, and the first 500 responses of the final two-probability session, respectively. Two panels have no solid lines because there were fewer than 10 reinforced responses, so no reliable proportions could be obtained.

responses that followed all reinforced and nonreinforced pecks on Key 2. Two panels in the right column have no solid lines, because, in these 2 cases, there were fewer than 10 reinforced responses on Key 2, so no reliable proportions could be obtained. Overall, the results in Figure 4 are similar to those of Figure 3. In 12 of 13 cases, the proportion of same-key responses was higher on the first response after a reinforced peck than after a nonreinforced peck. The dashed lines in the two left columns show the same up-and-down pattern after nonreinforced responses—this pattern is found without exception in all 10 panels. For the 3 probability pairs that allowed a comparison of reinforced and nonreinforced responses late in the two-probability phase, same-key response proportions tended to be slightly higher after reinforced responses: In 15 of the 18 comparisons shown in the right column of Figure 4, the proportions were higher after reinforced responses [$\chi^2(1) = 8.0, p < .01$].

DISCUSSION

The main finding of this experiment was that when the difference between two reinforcement probabilities ($\phi_1 - \phi_2$) was held constant, the development of preference was faster when the ratio of the probabilities (ϕ_1/ϕ_2) was larger. The development of preference for the higher probability of reinforcement was fastest in the .07, .01 conditions, intermediate in the .10, .04 and .13, .07 conditions, and slowest in the .16, .10 conditions. Two different methods were used to quantify the rate of acquisition. The first method simply compared the mean response proportions across successive portions of the two-probability phase from each condition (Figure 1). In the second method, a hyperbolic equation (Equation 1) was fitted to the data from each of the 160 acquisition curves obtained in this experiment, and the parameter M was used as a measure of the rate of acquisition (Figure 2). Both methods of data analysis led to the same conclusion: Acquisition was faster when the ratio ϕ_1/ϕ_2 was larger.

This result, obtained in a free-operant procedure, is consistent with that of Bailey and Mazur (1990), who used a discrete-trial procedure that included a 1-sec ITI after each choice response. As we discussed in the introduction, these findings conflict with the predictions of several theories of acquisition. The simple linear-operator model (Bush & Mosteller, 1955) and Staddon's (1988) ratio-invariance model both predict slower acquisition with larger probability ratios. The original version of melioration theory (Herrnstein & Vaughan, 1980) predicts equal acquisition rates for all conditions in which the quantity $\phi_1 - \phi_2$ is the same. The kinetic model (Myerson & Miezin, 1980) makes no predictions for discrete-trial procedures, but, for the present experiment, its predictions are the same as those of the original version of melioration theory—equal acquisition rates with equal values of the quantity $\phi_1 - \phi_2$. Thus, none of these four models correctly predicts the results of this experiment.

The results are, however, consistent with our hypothesis based on Weber's law. This hypothesis states that acqui-

sition will be faster when the two reinforcement probabilities are more easily discriminated and that discriminability depends on the ratio ϕ_1/ϕ_2 . It is worth noting that this result was obtained in the present experiment despite the fact that as the ratio ϕ_1/ϕ_2 increased, the total number of reinforcers per session decreased (due to the lower reinforcement probabilities in these conditions). Because the presentation of reinforcers is what allows a subject to learn which alternative has the higher probability of reinforcement, it might be expected that acquisition would be slower in conditions with the lower probabilities of reinforcement. However, in both this study and that of Bailey and Mazur (1990), the ratio of reinforcement probabilities was evidently a more important factor than the number of reinforcers delivered, and acquisition was faster with higher reinforcement ratios.

Throughout this discussion, our terminology has implied that the rate of acquisition depended on a subject's ability to discriminate between two different reinforcement probabilities. However, it is not certain that reinforcement probability was the dimension that controlled the rate of acquisition, because reinforcement probability was correlated with other variables, such as the number of responses per reinforcer and the time between reinforcers. An example of one alternative interpretation is that the subjects learned to discriminate the difference in the mean number of responses required per reinforcer. For example, in the .16, .10 condition, one key required about 6 responses per reinforcer and the other about 10 responses per reinforcer, a difference of only 4 responses. In the .07, .01 condition, however, one key required about 14 responses per reinforcer and the other about 100 responses per reinforcer, a difference of 86 responses. Perhaps the faster acquisition in the latter condition was due to the greater difference in the mean number of responses required per reinforcement between the two keys. Additional research with appropriately chosen reinforcement probabilities might help to determine whether the controlling dimension in this type of choice situation is reinforcement probability, responses per reinforcer, or perhaps some other factor. The present study does tell us, however, that the controlling dimension is *not* the difference between the two probabilities ($\phi_1 - \phi_2$).

In several ways, the results of the trial-by-trial analyses of responses after reinforcement and nonreinforcement were similar to those of Bailey and Mazur (1990). As in the Bailey and Mazur study, there was a tendency for subjects to switch response keys after each nonreinforced response during the one-probability phase and early in the two-probability phase. Because both keys were controlled by a single reinforcement schedule during the one-probability phase, switching keys after each response was, in fact, the optimal strategy in this phase. In the two-probability phase, however, switching between keys was not optimal, and exclusive preference for Key 1 would have yielded the greatest number of reinforcers per session. A comparison of the results from early and late portions of the two-probability phase showed that the switching tendency decreased as this phase progressed.

Another similarity to the results of Bailey and Mazur (1990) was the finding that the proportion of same-key responses was higher after reinforced responses than after nonreinforced responses. In both studies, evidence for this effect was clearest late in the two-probability phase, after the switching tendency just mentioned had subsided. With a 1-sec ITI after each response, Bailey and Mazur found a higher response proportion of same-key responses for roughly two responses after a reinforcer, and they called this a *positive recency effect*. This term may not be particularly appropriate for the results obtained in the present study, because in many cases shown in Figures 3 and 4, the difference between reinforced and nonreinforced trials was due to a decreased probability of a same-key response immediately after a nonreinforced response, not to an increased probability after a reinforced response. Yet regardless of what the effect is called, late in the two-probability phase of this experiment, a slightly higher response proportion persisted for at least six responses after a reinforced response. Other studies on choice have found evidence for similar effects that lasted for several responses after a reinforced response (e.g., Menlove, 1975). The effects of reinforcement are not always in this direction, however, and some studies have found *negative recency effects*—lower proportions of same-key responses after a reinforcer (e.g., Fantino & Royalty, 1987; Killeen, 1970). Fantino and Royalty (1987) presented evidence that whether positive, negative, or no recency effects are observed may depend on specific features of the choice procedures that are employed.

Bailey and Mazur (1990) raised the question of whether the short-term effect of reinforcement (as just described) and the long-term effect (a gradual increase in preference for the key with the higher probability of reinforcement) were the products of a single process or of two distinct processes. The same question can be raised about our results, although we cannot answer it. Perhaps both the short-term and long-term effects can be explained by a single, familiar principle—that reinforcement strengthens responses. Alternatively, it is possible that a process often called “short-term memory” or “working memory” was responsible for the positive recency effect and “long-term memory” or “reference memory” was involved in the gradual development of preference for one key over many hundreds of trials.

In summary, there was more rapid acquisition with higher probability ratios, and this pattern is inconsistent with the predictions of four different quantitative theories of acquisition. Exactly how damaging the results are for these theories is difficult to assess. Vaughan (1985) has already proposed a more general version of melioration theory that makes no specific predictions for our experiment. The linear-operator model can be modified and elaborated in many ways (see Kacelnik et al., 1987), and it is quite possible that some variation of the model could predict the pattern of results obtained in this experiment. Furthermore, all four of the models make other predictions about the characteristics of choice behavior (for either acquisition or steady-state performance) that have

received empirical support. Nevertheless, the question addressed by this experiment was a very basic one: When will the acquisition of preference be rapid, and when will it be slow? It seems to us that this is one of the most fundamental questions a theory of acquisition must answer. Acquisition rates varied greatly across the conditions of this experiment. We suggest that a viable theory of choice acquisition must account for these different acquisition rates, which were found even though the difference in reinforcement probabilities ($\phi_1 - \phi_2$) was constant.

REFERENCES

- BAILEY, J. T., & MAZUR, J. E. (1990). Choice behavior in transition: Development of preference for the higher probability of reinforcement. *Journal of the Experimental Analysis of Behavior*, *53*, 409-422.
- BUSH, R. R., & MOSTELLER, F. (1955). *Stochastic models for learning*. New York: Wiley.
- CATANIA, A. C. (1969). Concurrent performances: Inhibition of one response by reinforcement of another. *Journal of the Experimental Analysis of Behavior*, *12*, 731-744.
- FANTINO, E., & ROYALTY, P. (1987). A molecular analysis of choice on concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior*, *48*, 145-159.
- HARLEY, C. B. (1981). Learning the evolutionarily stable strategy. *Journal of Theoretical Biology*, *89*, 611-633.
- HERRNSTEIN, R. J., & VAUGHAN, W. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143-176). New York: Academic Press.
- HORNER, J. M., & STADDON, J. E. R. (1987). Probabilistic choice: A simple invariance. *Behavioural Processes*, *15*, 59-92.
- KACELNIK, A., KREBS, J. R., & ENS, B. (1987). Foraging in a changing environment: An experiment with starlings (*Sturnus vulgaris*). In M. L. Commons, A. Kacelnik, & S. J. Shettleworth (Eds.), *Quantitative analyses of behavior: Vol. 6. Foraging* (pp. 63-87). Hillsdale, NJ: Erlbaum.
- KILLEEN, P. (1972). A yoked-chamber comparison of concurrent and multiple schedules. *Journal of the Experimental Analysis of Behavior*, *18*, 13-22.
- LESTER, N. (1984). The feed:feed decision: How goldfish solve the patch depletion problem. *Behaviour*, *89*, 175-199.
- MAZUR, J. E., & HASTIE, R. (1978). Learning as accumulation: A reexamination of the learning curve. *Psychological Bulletin*, *85*, 1256-1274.
- MENLOVE, R. L. (1975). Local patterns of responding maintained by concurrent and multiple schedules. *Journal of the Experimental Analysis of Behavior*, *23*, 309-337.
- MYERSON, J., & HALE, S. (1988). Choice in transition: A comparison of melioration and the kinetic model. *Journal of the Experimental Analysis of Behavior*, *49*, 291-302.
- MYERSON, J., & MIEZIN, F. M. (1980). The kinetics of choice: An operant systems analysis. *Psychological Review*, *87*, 160-174.
- STADDON, J. E. R. (1988). Quasi-dynamic choice models: Melioration and ratio invariance. *Journal of the Experimental Analysis of Behavior*, *49*, 303-320.
- STADDON, J. E. R., & HORNER, J. M. (1989). Stochastic choice models: A comparison between Bush-Mosteller and a source-independent reward-following model. *Journal of the Experimental Analysis of Behavior*, *52*, 57-64.
- STUBBS, D. A., & PLISKOFF, S. S. (1969). Concurrent responding with fixed relative rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, *12*, 887-895.
- VAUGHAN, W. (1985). Choice: A local analysis. *Journal of the Experimental Analysis of Behavior*, *43*, 383-405.
- VAUGHAN, W., & HERRNSTEIN, R. J. (1987). Stability, melioration, and natural selection. In L. Green & J. H. Kagel (Eds.), *Advances in behavioral economics: Vol. 1* (pp. 185-215). Norwood, NJ: Ablex.