The learning of response patterns in choice situations

ARMANDO MACHADO and RICHARD KEEN
Indiana University, Bloomington, Indiana

This study presents a theory by which to understand how pigeons learn response patterns in simple choice situations. The theory assumes that, in a choice situation, patterns of responses compete for the final common path; that the competition is governed by two variables, the overall reinforcement probability obtained by emitting the patterns, T, and the differences in reinforcement probabilities among the patterns, D; and that the ratio D/T determines the final strength of specific response patterns. To test these predictions, three experiments were run in which pigeons were more likely to receive food when they pecked the momentarily least-preferred of three response keys. On the basis of previous research, it was predicted that the birds would be indifferent among the keys (molar aspect) and would also acquire a response pattern that consisted of pecking each key once during three consecutive trials (molecular aspect). The present theory went further and predicted that the strength of that pattern would increase with the ratio D/T. In the first two experiments, D was manipulated while T remained constant, and in the third, T was manipulated while D remained constant. The results agreed with the theory, for the strength of the response pattern increased with D and decreased with T, whereas overall choice proportions were always close to the matching equilibrium.

Probably the most robust finding in the study of choice is Herrnstein's (1961, 1970) matching law: In a two-choice situation, the proportion of responses to one alternative equals the proportion of reinforcers obtained from that alternative. In symbols,

$$\frac{B_1}{B_1 + B_2} = \frac{R_1}{R_1 + R_2}, \tag{1}$$

where B_1 and B_2 are the number of choices of alternatives 1 and 2, respectively, and R_1 and R_2 are the number of reinforcers obtained from the same alternatives.

For some researchers, matching is a by-product of an animal's sensitivity to local variables—in particular, to the momentary probabilities of reinforcement (see, e.g., Hinson & Staddon, 1983; Shimp, 1982b; Silberberg, Hamilton, Ziriax, & Casey, 1978). For example, Shimp (1966, 1969) has argued that pigeons and rats match because they track the current probabilities of reinforcement associated with the choice alternatives and choose the alternative with the highest momentary reinforcement probability. By tracking local probabilities of reinforcement, the animal patterns its consecutive choices in

ways that generate the matching relation when single responses are aggregated. Because this viewpoint stresses the role of point variables, it has been termed a *molecular* approach.

Other authors disagree. For them, molar matching stems from the animal's direct sensitivity to the rates of reinforcement, the rates of responding, and the correlation between these two rates (see, e.g., Baum, 1973; Rachlin, Battalio, Kagel, & Green, 1981; Staddon, 1983; Staddon & Motheral, 1978). That is to say, the choice processes "operate on the same aggregations of behavior that show matching" (Williams, 1988, p. 189). According to this viewpoint, single responses occur in proportion to their overall strength, and strength is determined by reinforcement rate (e.g., Herrnstein, 1970; Heyman, 1979; Killeen, 1981; Nevin, 1969, 1979). Because this viewpoint stresses temporally extended variables as the key ingredients of the choice process, it has been termed a *molar* approach.

The preceding summary (see Williams, 1988, for a detailed analysis) highlights the fact that, to some extent, the difference between molecular and molar approaches revolves around the importance assigned to response patterns in choice situations. For the former approach, response patterns are the primary phenomenon, and matching is their derivative; for the latter approach, matching is primary, and response patterns are an occasional derivative. This difference in approach naturally raises the question, What do experiments show about the presence or absence of response patterns in choice situations?

The empirical evidence has been mixed. On the one hand, in some studies, animals have exhibited reliably the patterns of choices predicted by momentary maximizing

accounts, patterns that, when aggregated, yield the matching relation. For example, in a two-key, discrete-trials equivalent of a concurrent variable-interval variable-interval schedule (Conc. VI schedule, for short), with reinforcement probabilities in a 3 to 1 ratio, Shimp (1966) observed that pigeons acquired the predicted maximizing pattern, a sequence of two consecutive choices of the richer key followed by a single choice of the other key (see, also, Shimp, 1969; Silberberg et al., 1978; Silberberg & Williams, 1974). Moreover, studies that have shown that pigeons and rats can learn response patterns such as win–stay, lose–shift strategies (Hiraoka, 1984; Shimp, 1976; Williams, 1972, 1983, 1991) or response runs of specific length (Shimp, 1982a) may also be interpreted as (indirectly) supporting the claim that global regularities in behavior, such as matching, are a derivative of the local organization of behavior.

On the other hand, molar matching has also been obtained in the absence of the patterns predicted by the molecular accounts. For example, using a discrete-trials, Conc. VI schedule, Nevin (1969, 1979) found molar matching but no evidence that changeover probability increased with run length—as sensitivity to changes in local reinforcement probability would predict—or that his pigeons had learned the specific sequences of choices predicted by momentary maximizing. Similarly, in a free-operant study with Conc. VI schedules, Heyman (1979) also found no evidence of tracking momentary reward probabilities (see, also, Herrnstein, 1997, Chap. 3). Furthermore, studies that increased the interval between consecutive choices found that local response patterns were disrupted but matching still held (Silberberg & Williams, 1974; Williams, 1991). Finally, it is also the case that molecular theorists have advanced mainly computer simulations (e.g., Shimp, 1966; Shimp, Childers, & Hightower, 1990; Silberberg et al., 1978), but no direct evidence, to show that the processes they have identified underlie the molar regularities observed in Conc. VI schedules.

Molecular theorists have attempted to explain the preceding results in a variety of ways. For example, Shimp (1982b) has claimed that the tacit assumption of local randomness contained in most molar theories is contrary to fact and that Nevin's (1969, 1979) results are simply irrelevant to the molar–molecular issue because, in Conc. VI schedules, "there is a statistically forced time-allocation matching relation [the pingpong ball effect]" (pp. 114–115). On the other hand, Silberberg et al. (1978) argued that a slightly different decision rule—momentary maximizing with memory errors or momentary maximizing with response perseveration—may account for the molar and molecular data from a variety of procedures. These accounts have not remained unchallenged, however (see Shimp, 1990; Shimp et al., 1990; Williams, 1990; see, also, Hineline, Silberberg, Ziriax, Timberlake, & Vaughn, 1987; Machado, 1993), and 30 years after the controversy first surfaced, we are still far from its resolution.

A major difficulty in assessing the foregoing results and arguments is that we know very little about the conditions in which specific response patterns occur in choice situations, the rules of their acquisition, and the determinants of their asymptotic strength. Without a theory that could be conceived of as the equivalent of the matching law for response patterns—a quantitative formulation relating the strength of particular response patterns to the reinforcement contingencies—our accounts (no matter how reasonable they may be) will remain post hoc, and our intuitions (no matter how plausible they may sound) will remain disconnected and hard to refute. Hence, the major goal of the present study is to sketch and test an integrated, reasonably precise, and empirically sensitive theory of what determines the strength of response patterns in simple choice situations.

The remainder of this paper is organized as follows. First, we describe the assumptions of the theory and derive some quantitative consequences from these assumptions. Next, we report the results of three experiments designed to test the theory. And finally, in the General Discussion section, we relate the theory and experimental findings to the broader issues of matching and response patterning.

In developing a theory of response patterns, to avoid potential misunderstandings, we state first what the theory to be proposed is not about. It is not a general theory of choice, but a theory to be applied to situations in which two or more similar responses (e.g., left and right key-pecks in pigeons, or leverpresses in rats) are reinforced with qualitatively identical rewards (e.g., food); it is not a theory of the effects of single reinforcers, but of the cumulative effects of many reinforcers on response patterns; it is not a theory about the acquisition of response patterns, but about their asymptotic strengths; and, as will be obvious, it is not a fully articulated theory (e.g., it does not deal with temporal issues), but only the beginning of one.

Our guiding question is this: In a simple choice situation, what factors determine the strength of a response pattern? For classification purposes, we can divide the set of all factors into two subsets—the first including the factors related to the reinforcement contingencies, the second including the factors specific to the pattern under examination. To isolate the first subset, assume, for the moment, that all the patterns are equal in terms of the second subset of factors and that only six patterns, R_1, R_2, ... , R_6, compose the total population of patterns. This situation is illustrated schematically in the horizontal axis of Figure 1, top panel. At a particular moment, each pattern has an associated reinforcement probability, illustrated in the figure by the height of the vertical lines. The six reinforcement probabilities define a payoff function over the population of patterns, and our goal is to identify the properties of this function that may determine how a pattern's strength changes.

To that end, we conceive of the patterns as being in an arms race, competing among themselves for the final common path. Their competition is governed by two variables: the average reinforcement probability across the population of patterns (represented by the dashed horizontal line) and the differences in reinforcement probabilities across
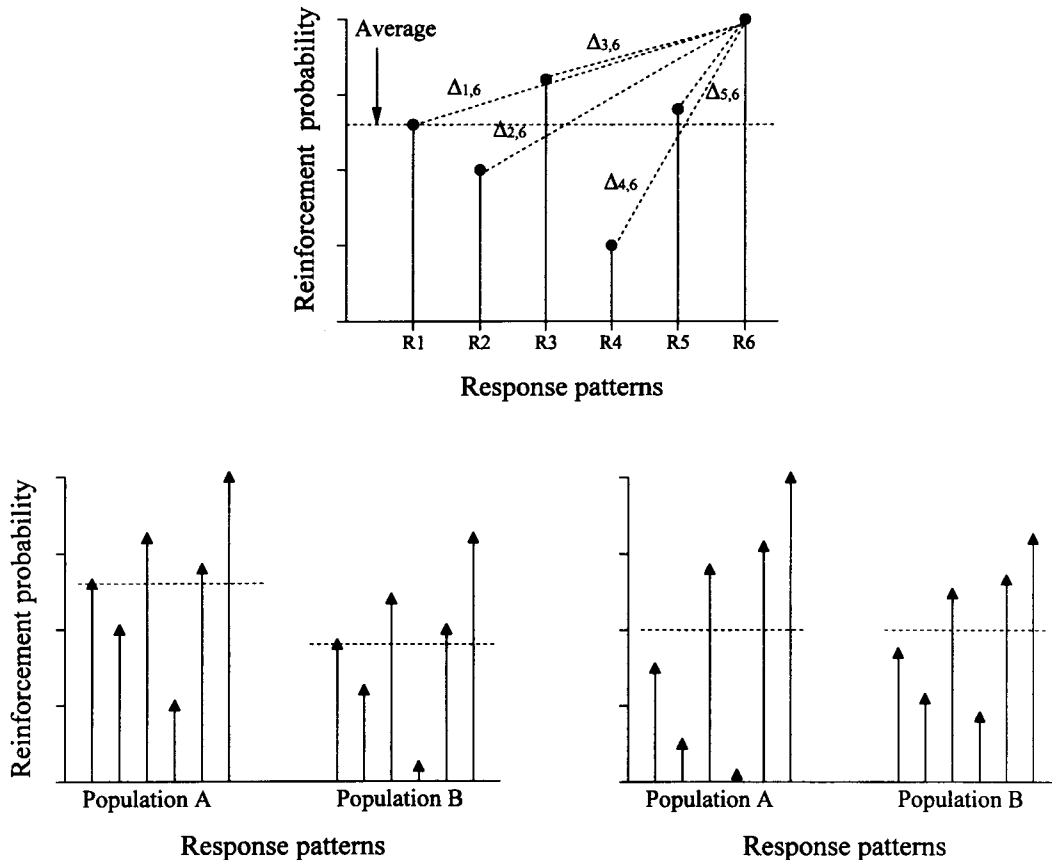
Figure 1. The top panel shows a population of response patterns, $R1$, $R2$, ..., $R6$, and the reinforcement probability associated with each pattern. The horizontal line is the average of the probability function, T. The symbols Δ show the differences among the reinforcement probabilities. The bottom panels show two populations of response patterns and their respective reinforcement function. In the left panel, the average reinforcement probability is higher for Population A than for Population B, but their variances are equal. In the right panel, the averages are equal, but the variance is greater for Population A. Throughout, it was assumed that the patterns had the same strength.

the patterns (represented by the Δ symbols). Henceforth, we will refer to this set of differences generically as the *variance* of the reinforcement function.

Consider now two populations, A and B, exposed to the reinforcement functions drawn in the bottom left panel of Figure 1. The functions have the same variance but different averages. We predict that the strength accrued to the pattern associated with the highest reinforcement probability (i.e., the rightmost pattern of each population) will be higher in population B than in population A. In other words, for constant variance, the lower the average, the stronger the winning pattern will be. There are two ways to understand intuitively this prediction. Functionally, *all else being equal*, when the overall reinforcement probability is higher, the *pressure* to strengthen the winning pattern is lower, for regardless of what the animal does, it obtains a greater amount of food. Mechanistically, lowering the overall reinforcement probability makes the discrimination among the differences in the payoff probabilities easier to learn.

The bottom right panel shows the effect of varying the set of differences in payoff while keeping the average payoff constant. In this case, we predict that the strength accrued to the pattern associated with the highest reinforcement probability will be higher in population A than in population B. The intuitive reason for this prediction is that, *all else being equal*, large fluctuations in the overall reinforcement probability increase the *pressure* to prefer the pattern reinforced with the highest probability, and larger differences are easier to discriminate than smaller ones.

Combining the two foregoing ideas suggests the hypothesis that the strength accrued to the winning pattern will increase—according to some function f—with the ratio of the variance to the average of the payoff function. This ratio is a sort of Weber fraction in the domain of reinforcement probabilities. Furthermore, what was just stated for the winning pattern applies also, but in the opposite direction, to the pattern associated with the lowest reinforcement probability: The larger the set of (negative)

differences, or the lower the average of the payoff function, the more that pattern should lose its strength.

However, reinforcement contingencies are not the sole determiners of the strength of a response pattern. Another subset of factors needs to be considered, for this subset includes factors that modulate the amount by which the reinforcement contingencies change the strength of a pattern. Consider, for example, memory constraints, a topic also referred to as stimulus control. Variables such as the type, the number, and the age of the cues that must be remembered to perform a pattern may restrict the effects of reinforcement on the strength of that pattern. To illustrate, as the number of cues that must be remembered increases, variability in the performance of the pattern typically increases (see, e.g., Gallistel, 1990; Mechner, 1958; Shimp, 1982a). As another example, intuition suggests and experiments confirm that, as the interval between responses increases, the strength of a response pattern decreases (Shimp, 1976; Silberberg & Williams, 1974; Williams, 1983, 1991). In the extreme, memory limitations may also explain why pigeons simply cannot learn complex response patterns, such as the sequence RRRLRLLL (Machado, 1992).

In addition to memory constraints, the cost of performing a pattern is also likely to modulate the effects of reinforcement. Compare, for example, the pattern of alternating between two keys with the pattern of repeating the previous response. For the same frequency of occurrence, alternation is probably more costly than repetition, which means, more generally, that cost is likely to differ across patterns. As evidence, consider that, when pigeons receive food for completing sequences of eight choices distributed over two keys, regardless of the specific sequence of the choices, they generally show a strong preference for the sequences *all pecks on the left*, or *all pecks on the right* (Hunziker, Saldana, & Neuringer, 1996; Machado, 1997). Furthermore, the costs entailed by a pattern may also increase with the pattern's frequency of occurrence. This much is suggested by Machado's (1997) findings: Pigeons were rewarded when they completed sequences of eight two-key choices. Sequences that contained three or four changeovers were always rewarded, whereas sequences with fewer or more changeovers were rewarded with lower probabilities. The results showed that, although most sequences contained the optimal numbers of three or four changeovers, these sequences did not occur exclusively. It seems reasonable to conclude that the cost of changeovers may have increased as the bird switched more and more between the keys.

The preceding arguments lead us to represent, in a single function, the combined effects of all the sources that are likely to modulate the effects of the reinforcement contingencies. We refer to this function generically as the cost function g and assign to it two properties. First, the cost function may differ across response patterns. And second, the cost of a pattern increases monotonically with its frequency of occurrence—that is, with its strength. Although a full theory of response patterns would need to specify the exact shape of function g and

to include in g the effects of other variables and processes (e.g., response induction), such a detailed theory is beyond our current knowledge.

Finally, we arrive at the main hypothesis: The difference between functions f and g drives the change in a pattern's strength. Function f describes the selective effects of reinforcement in the absence of any constraints; function g adds the effects of the constraints. Hence, each function addresses one of the two subsets of factors that are likely to determine a pattern's asymptotic strength.

The preceding hypothesis and its implications can be expressed more rigorously as follows. Let the strength of a pattern at the beginning of time epoch ε equal S. Furthermore, let the average and the variance of the payoff function during ε equal T and D, respectively. We assume that, at the end of the time epoch, S will have changed by an amount determined by the ratio D/T minus an amount determined by the cost of the pattern. That is,

$$\Delta S = \alpha [f(D/T) - g(S)], \qquad (2)$$

where both functions, f and g, are assumed to be monotonically increasing and $\alpha > 0$ is a proportionality constant. Asymptotically, $\Delta S = 0$, and this occurs when $f(D/T) = g(S)$. Solving for S yields

$$S(\infty) = g^{-1}[f(D/T)], \qquad (3)$$

where $S(\infty)$ is the asymptotic value of S and g^{-1} is the inverse function of g. Because f and g were assumed to be monotonically increasing, g^{-1} and $g^{-1}\circ f$ are also monotonically increasing. Hence, $S(\infty)$ increases with the ratio D/T. Note that the asymptotic strength of a pattern is not affected by the duration of the time epoch ε, although its rate of acquisition may be.

Our major hypothesis may be stated as follows: In simple choice situations, the asymptotic strength of a response pattern depends on the Weber-like ratio of the variance to the average of the reinforcement function.

GENERAL METHOD

Before we can test the preceding theory, two problems need to be solved: how to control for the likely possibility that the cost function differs across patterns, and how to manipulate *independently* the variance and the average of the reinforcement function. The first problem occurs when we attempt to apply the theory to a single experimental condition, because, on the basis of the ratio D/T, we could predict that pattern x would be stronger than pattern y, but the prediction would not hold necessarily if pattern x had a higher cost than pattern y. More generally, without knowing the cost function for each pattern, nothing definitive can be stated about asymptotic strengths in a single experimental condition. This difficulty is not as troublesome as it may seem, however, for instead of attempting to predict the strengths of the various patterns during a single experimental condition, we can try to predict how one pattern's strength changes as we vary D or T across conditions. By focusing on a single pattern across conditions and assuming that its cost does not change as D and T are varied, the aforementioned difficulty is eliminated.

The second problem, how to manipulate D and T independently, can be solved by using a reinforcement schedule that was developed on the basis of previous work by Blough (1966), Shimp (1967), and Machado (1992). (We will show later on that the pres-

ent schedule is remarkably similar to a Conc. VI schedule.) In a dis-crete-trials situation, a pigeon chooses one of three response keys: left, center, or right. The probability of reinforcement for each peck depends on the number of similar pecks emitted during the last, say, 10 trials. We refer to this number, critical in what follows, as the sample size M. The exact reinforcement function is illustrated in the top panel of Figure 2 and has the equation

$$P(S^+) = \frac{1}{27\frac{i}{M}+1}, \qquad (4)$$

where $P(S^+)$ is the probability of reinforcement following a left, a center, or a right keypeck, and i is the number of left, right, or cen-ter keypecks produced during the last M trials. For example, if a bird has pecked the left key five times, the center key three times, and the right key two times on the last 10 trials, on the next trial, food will follow a choice of the left key with probability $1/(27 \times .5 + 1) = .069$, of the center key with probability .110, and of the right key with probability .156. The number 27, a sort of penalty factor, determines how fast the function decreases with the proportion of responses on one key.

If we restrict our attention to the population of patterns composed by left, right, and center keypecks (i.e., single responses) and as-sume that their cost functions are equal, Equation 3 predicts that, at equilibrium, the bird will be indifferent among the three keys and obtain food on 10% of the trials (see the equilibrium point in Fig-ure 2). This prediction stems from the fact that Equation 3 subsumes

under variable D the dynamics of melioration (see, e.g., Herrnstein, 1997).

But what sorts of patterns will occur in this choice situation? And how strong will they be at the steady state? To answer these ques-tions, we need to examine the payoff function for response patterns induced by Equation 4. The filled circles on the bottom left panel of Figure 2 show this function when molar choice proportions equal 1/3 (see Appendix A for the derivation). Only a subset of all pairs and triplets is shown, because the curve is the same for the remain-ing ones. The function for pairs LL, LR, and LC shows that repeat-ing is less likely to be reinforced than switching. Hence, if the costs of all the pairs were identical, Equation 3 would predict that switch-ing should be stronger than repeating. In addition, if we also con-sider the function for triplets, LLL, LLR, ... , LRC, we see that, after the bird switches from one key to another (e.g., from L to R), the reinforcement probability is highest on the third key (i.e., C; compare LRC with LRL and LRR). Hence, if the costs of all the patterns were identical, Equation 3 would predict that the patterns that involve one peck on each key during three consecutive trials should win the competition. On the other hand, the functions for pairs and triplets also reveal that, under cost equality, the patterns that consist of repeating the same response (e.g., LLL) should have the lowest strength, for they are the least likely to be reinforced.

So far, we have simply applied Shimp's (1966) momentary max-imizing analysis to the present situation. However, the theory being proposed goes beyond Shimp's analysis, because it states that the asymptotic strength of the various patterns depends on the ratio of
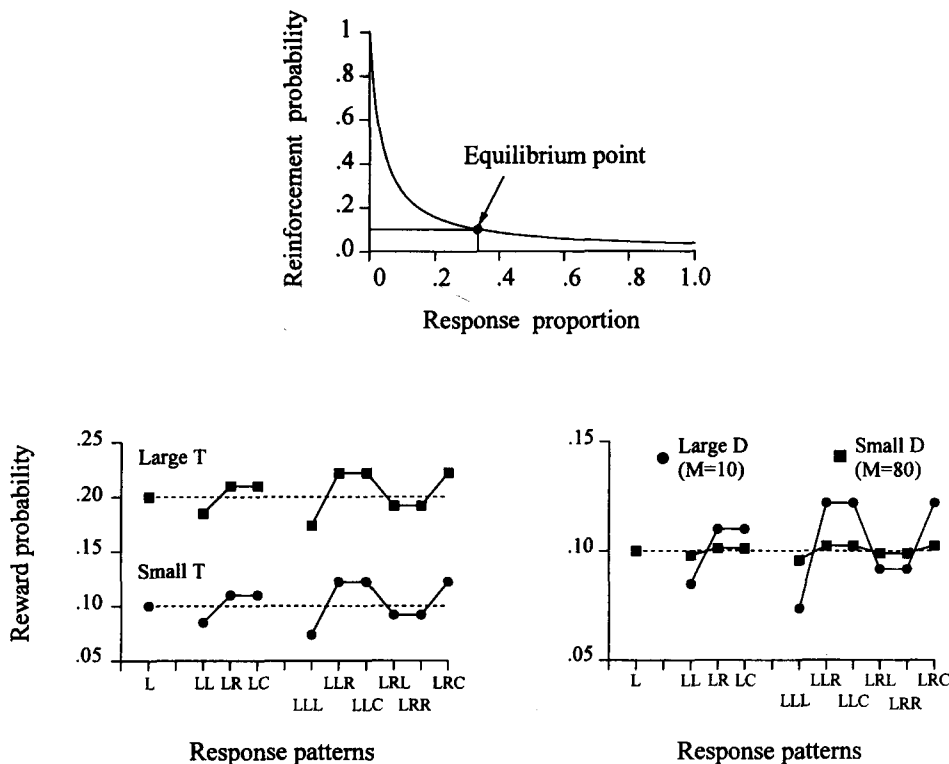

Figure 2. The top panel shows the reinforcement function used in the three experiments. The prob-ability of reinforcement for one response decreases with the proportion of similar responses emitted during the last M trials. At the matching equilibrium, all responses are equally likely, and the overall probability of reinforcement is .1. The bottom left panel shows the effect of adding a constant to the re-inforcement function. The average (dotted line) changes, but the variance remains the same. The bot-tom right panel shows the effect of the sample size M on the probabilities of reinforcement for various response patterns. A smaller sample induces higher variance but the same average (dotted line). In both cases, it was assumed that the molar proportions equaled 1/3.

the variance to the average of the reinforcement function for patterns. Hence, to test the theory, we need to manipulate these two variables independently.

To change the average, but not the variance, we fix the sample size M and simply add a constant to Equation 4. The effect on the probabilities of reinforcement following pairs and triplets is illustrated by the squares in the bottom left panel of Figure 2. In the example, the shift equaled .1. The average reinforcement probability changed (see dashed lines), but the set of differences in the reinforcement probabilities remained constant. Equation 3 predicts that the pattern *one peck on each key during three consecutive trials*—which from now on we call *stable patterns*[1]—will be stronger, the smaller the value of T. Conversely, the *repeat patterns*, LLL, RRR, and CCC, will be weaker, the smaller the value of T (note that, for these patterns, $D < 0$).

To change the variance but not the average, we change the number of trials from which the response proportion is obtained—that is, the sample size M. The effects on the reinforcement contingencies for pairs and triplets are illustrated in the bottom right panel of Figure 2. Note that the variance of the reinforcement function is inversely related to the sample size. The overall reinforcement probability, however, remains constant (dashed line). The effects of the sample size on D and T are intuitively reasonable: When the sample is small, a few consecutive pecks on the same key cause large changes in the reinforcement probabilities, but when the sample is large, the same number of consecutive pecks on one key causes only small changes in the reinforcement probabilities. On the other hand, the average reward probability remains constant, because, regardless of the sample size, sample proportions are unbiased estimates of the *true* pecking probabilities (Appendix A may be consulted for a more rigorous version of this argument). Equation 3 predicts that the stable patterns will be stronger but the repeat patterns weaker, the smaller the value of M or, equivalently, the higher the value of D.

Using the new schedule, three experiments were conducted to test how the average and the variance of the payoff function affect the asymptotic strength of response patterns. In the first two, the variance was manipulated across conditions, while the average remained constant; in the third, the average was manipulated, while the variance remained constant. After analyzing the results of each experiment separately, we combined them to see whether the ratio D/T was correlated with the asymptotic strength of the response patterns.

EXPERIMENT 1

The first experiment manipulated the variance of the payoff function by changing the sample size across experimental conditions. We predict that, regardless of the value of D, molar choice proportions will agree with the matching law—the birds will choose each key equally often. However, as D increases, the stable patterns will be stronger, and the repeat patterns will be weaker.

It is important to realize that, to a large extent, the stable patterns may change in strength without that determining the strength of the repeat patterns, and vice versa. In other words, the predicted changes are not forced. A hypothetical example clarifies this point. A bird could respond in runs of three pecks during one condition and, thus, generate a high proportion of repeat patterns but no stable patterns; in another condition, the bird could respond in runs of only two responses and, thus, generate fewer repeat patterns but still no stable patterns. In this example, a substantial change in the proportion of the repeat patterns occurred without any significant change in the proportion of the stable patterns.

Method

Birds. Four experimentally naive pigeons (*Columba livia*) participated in the experiment. The birds were housed in individual home cages, with water and grit continuously available, but with no dark–light cycle in effect. Throughout the experiment, the birds were maintained at 80% of their free-feeding body weight.

Apparatus. A standard experimental chamber for pigeons from Med Associates was used. The front aluminum panel contained three keys placed at the vertices of an inverted isosceles triangle. The top two keys were 10 cm apart and 27 cm from the floor; the bottom key was 8.5 cm from the top two keys. The keys could be illuminated from behind with red light. Directly below the center key and at 4 cm from the floor was a 6×7 cm hopper opening. The bird had access to mixed grain when the hopper was raised and illuminated with a 7.5-W white light. On the back wall of the chamber, another 7.5-W houselight provided general illumination. An outer box equipped with a ventilating fan enclosed the experimental chamber. All the events were controlled by a microcomputer.

Procedure. After the birds had learned to peck the keys through autoshaping, they were trained to respond under low probabilities of reinforcement per peck. Sessions were divided into trials, and at the beginning of each trial, one randomly selected key was illuminated. A peck on that key turned the keylight off, and with probability s, it activated the hopper. If the peck did not activate the hopper, a 0.4-sec intertrial interval (ITI) followed. Pecks during the ITI reset the timer for the interval. After the ITI or the food delivery, a new trial began. The houselight was always illuminated, except during reinforcement. Across five sessions of 50 trials each, the value of s decreased from .3 to .1.

During the experiment proper, all the procedural details remained the same, except that all the keylights were illuminated at trial onset and the probability of reinforcement after each peck was given by Equation 4. Thus, the following sequence of events occurred on each trial: (1) The computer determined the number of times the bird pecked each key during the last M trials, using, when necessary, the last trials of the previous session (for the very first session of the experiment, the computer generated M random choices); (2) from the number of left, center, and right keypecks, the computer determined the reward probabilities for the next trial, using Equation 4; (3) the three keylights were then illuminated, and when the bird pecked one of the keys, either a reinforcer or the ITI followed. The cycle was then repeated until the bird had obtained 50 reinforcers.

The experiment was divided into three experimental conditions, each one defined by the sample size M and two transition phases that tried to minimize any carryover effects from the previous condition. Table 1 shows the order of the conditions and the number of sessions for each one. Because the data on response patterns were not analyzed after each session, the duration of the experimental conditions was determined by the stability of the molar proportions. Each condition lasted a minimum of 25 sessions and until molar proportions showed no systematic trend for 5 consecutive sessions. During the transition phases, only one key delivered reinforcers, with a constant probability of .1 per peck; responses on the other two keys were extinguished. The key that delivered reinforcers was determined by looking at the last 5 sessions of the previous condi-

Table 1
Order of Conditions in Experiment 1

Bird	Condition 1			Condition 2			Condition 3	
	M	N	Trans.	M	N	Trans.	M	N
6112	20	30	4	10	35	6	80	30
8261	20	25	3	10	35	4	80	30
2581	20	30	3	80	35	3	10	35
5263	20	35	3	80	35	4	10	30

Note—M, sample size; N, number of sessions; Trans., number of sessions in transition phase.

tion and selecting the least-preferred key. For example, at the end of Condition 1, Bird 6112 preferred the center, the left, and the right keys, in that order; hence, the right key was the target key during the first transition phase. Transition sessions continued until at least 80% of all the choices were on the target key.

Results and Discussion

Figure 3 shows the proportions of left, center, and right keypecks observed at the end of each session of each experimental condition. In general, molar proportions converged rapidly to the 1/3 matching equilibrium (represented by the dotted line). This rapid approach to matching, also reported by Mark and Gallistel (1994), is typical of frequency-dependent schedules (Machado, 1992, 1993, 1994) and may be due to the fact that these schedules contain a strong negative feedback loop from choice proportion to reinforcement probability. Birds 6112 and 2581 showed some oscillations during the first condition ($M = 20$), but these oscillations disappeared during the last sessions. There were also some clear in-

stances of bias (e.g., Bird 6112, $M = 10$), but overall, the quality of matching did not seem to differ appreciably or systematically across conditions.

It was critical to the theory under examination that the proportion of reinforced trials remain constant across conditions, for otherwise the effect of D would be confounded with the effect of T. Figure 4 (lines without symbols) shows that this was indeed the case. Although T tended to be slightly lower when $M = 10$ than in the other two conditions, the differences were quite small and probably did not affect the results described below. The average across birds and sessions equaled .09, .11, and .10 for $M = 10$, 20, and 80, respectively.

Figure 4 also shows the proportion of the stable patterns for each session of the experimental conditions. (Although the theory being tested deals only with asymptotic performance, these results are included, because they inform about within- and between-subjects variability in performance.) The proportion of the stable patterns was obtained by counting the number of instances in which
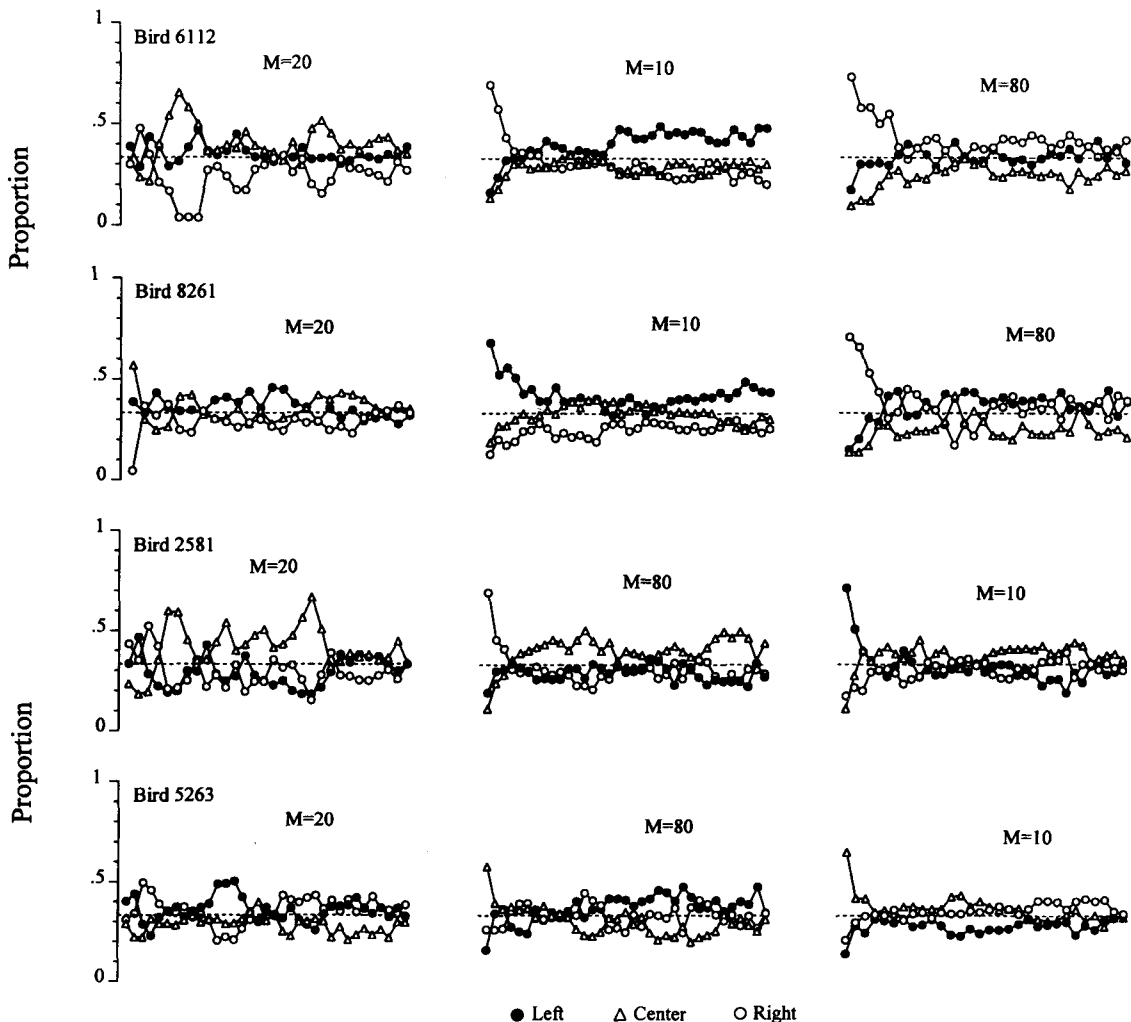

Figure 3. Molar results from the three conditions of Experiment 1. M refers to the sample size for each condition. The dashed horizontal line shows the 1/3 matching equilibrium.
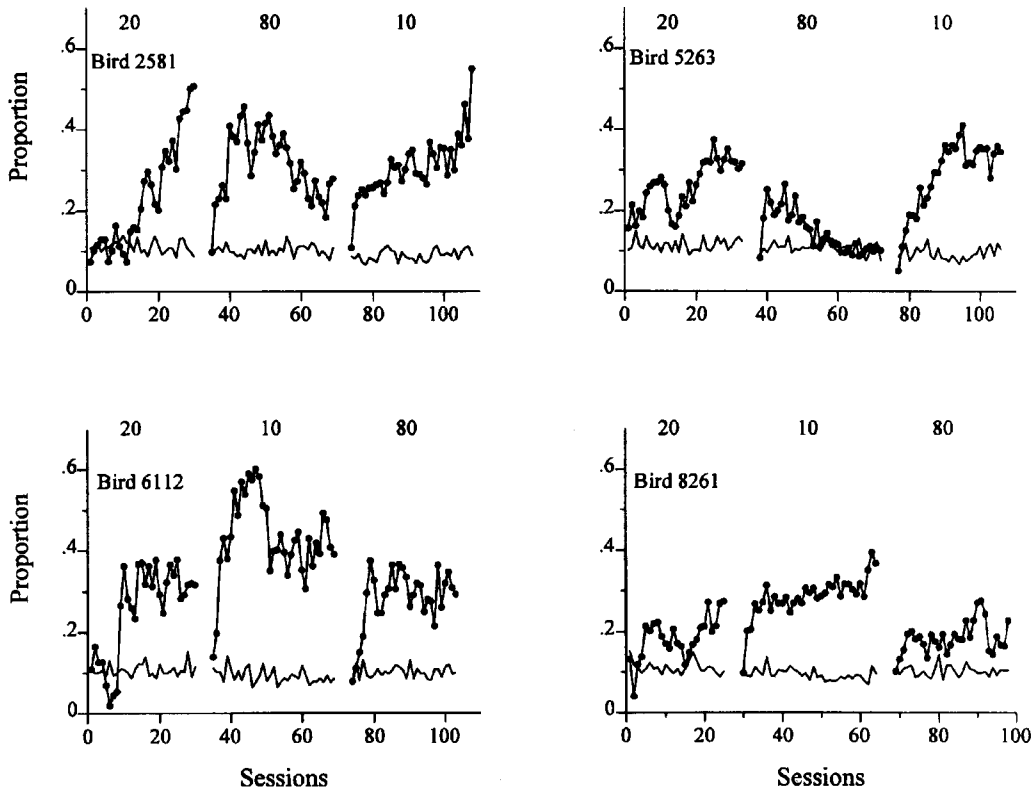
Figure 4. Proportions of the stable patterns (line with dots) and reinforcers (simple line) during each session of Experiment 1. The numbers at the top identify the sample size for each condition.

the three keys were pecked consecutively and then dividing that number by the total number of pecks. The counting was done using overlapping sequences of choices, so that, in the string CCRLCRR, for example, we count three stable patterns, CRL, RLC, and LCR.

The results show that, for Birds 2581 and 5263 (top panels), the stable patterns were stronger when $M = 10$ or 20 than when $M = 80$. For the other 2 birds, the stable patterns were stronger when $M = 10$ than in the remaining conditions. It is also clear from the figure that there was no correlation between the strength of the stable patterns and the overall probability of reinforcement. There was some indication, in particular for Bird 2581, that the final values, when $M = 20$ and 10, had not yet stabilized.

Figure 5 summarizes the data from all the triplets during the last five sessions of each condition. The triplets are separated into three sets, one containing the stable patterns and represented by 1X, 1Y, and 1Z (i.e., one response of each type occurred during three consecutive trials), another containing the repeat patterns and represented by 3X (i.e., three identical responses during three consecutive trials), and another containing the remaining triplets and represented by 2X and 1Y. To compare the sets directly, the proportion of each set was divided by the number of its elements (i.e., 6 for the stable, 3 for the repeat, and 18 for the remaining patterns). Finally, the re-

sults for the 4 birds were averaged, and the standard error of the mean proportion was computed. For reference purposes, the dotted line shows the 1/27 value that would be expected if the three choices occurred randomly.

The results indicate that, as the sample size increased, the strength of the stable patterns decreased, whereas the strength of the repeat patterns increased. One-way, repeated measures analyses of variance (ANOVAs) showed that these effects were significant $[F(2,6) = 8.79, p = .02,$ for the repeat patterns, and $F(2,6) = 8.11, p = .02,$ for the stable patterns]. These results agree with the predictions based on Equation 3. The strength of the remaining patterns did not change appreciably with the sample size $[F(2,6) = 0.63].$

In summary, the data showed that (1) molar choice proportions always approached the matching equilibrium; (2) deviations from matching were small and did not seem to be related to the experimental condition; and (3) the strength of the response patterns changed in the direction predicted by the theory (when D increased, the strength of the stable patterns increased, whereas the strength of the repeat patterns decreased).

EXPERIMENT 2

Experiment 2 was a systematic replication of the preceding study. It was motivated by the fact that Experi-
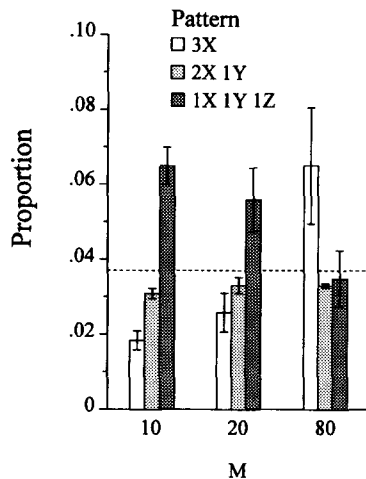
Figure 5. Mean proportion of the stable patterns (one peck on each key during three consecutive trials, i.e., 1X, 1Y, 1Z), the repeat patterns (three consecutive pecks on the same key, i.e., 3X), and the remaining triplets (2X, 1Y) as a function of the sample size M. The data are from the last five sessions of each condition of Experiment 1 (excluding the replication). The dotted line shows the expected proportion if the bird responded randomly. The error bars show the standard error of the means.

ment 1 included two features that may have obscured the effects of the variance of the reinforcement function on the strength of the stable patterns. First, during the transition phase that separated the experimental conditions, only one key delivered rewards, and the birds quickly learned to peck that key on most trials. As a consequence, the stable patterns may have been insufficiently extinguished during the transition sessions, and some carryover effect from one condition to the next may have occurred. Second, the order of the experimental conditions was not fully counterbalanced, since all the birds were initially exposed to a sample size of 20. To minimize the first problem, during the transition phase of Experiment 2, the birds received food only when they emitted one of the repeat patterns, LLL, RRR, or CCC. By reinforcing repetitions, we attempted to strengthen response patterns that, in a sense, are the opposite of the stable patterns. We reasoned that, at the beginning of the new condition, the stable patterns would be weaker than their competitors and, therefore, their final strength would depend more clearly

on current conditions. To eliminate the second problem, the order of the conditions was fully counterbalanced.

Method

Birds and Apparatus. Four pigeons (*Columba livia*), 1 naive (Bird 6447) and 3 with previous experience in frequency discrimination (Machado & Cevik, 1997) and response variability (Machado, 1997) tasks, participated in the experiment. All the housing details and the experimental apparatus remained the same as those in Experiment 1.

Procedure. The preliminary training and the experiment proper were similar to those in Experiment 1, except for the following changes. The sample size varied from 10 to 50 to 160, and, by subtracting .02 from Equation 4, the average reward probability at equilibrium was set at $T = .08$. Table 2 shows the order of experimental conditions and transition phases and the number of sessions for each one. Note that, unlike that in Experiment 1, the first experimental condition was also preceded by a transition phase, and a fourth condition replicated one of the first two. Each experimental condition lasted for a minimum of 40 sessions and until the proportion of the stable patterns showed no systematic trend for 5 consecutive sessions. Sessions ended after 48 reinforcers, and the first two choices after each reinforcer were never reinforced. This last modification attempted to reduce any control that food itself might exert over the keypecks.

During the transition sessions, the computer chose one of the three keys at random, but with the constraint that each key was chosen 16 times per session. When a variable interval ($\bar{x} = 60$ sec) elapsed, three consecutive choices of that key were followed by food. Afterwards, the computer selected a new key, and the cycle was repeated. The transition phase continued until the proportion of the stable patterns was below .10 for five consecutive sessions.

Results and Discussion

Figure 6 shows the proportions of single responses for each session and the experimental condition. Again, the proportions were always close to the 1/3 matching equilibrium, and, with the exception of a few instances of bias, no large or systematic differences were visible across conditions.

Figure 7 shows the proportions of reinforced trials and stable patterns across sessions and conditions. The proportion of reinforced trials remained approximately constant throughout the experiment but slightly below the scheduled value of .08. The averages across birds and conditions equalled .066, .071, and .071, for $M = 10, 50$, and 160, respectively. Concerning the stable patterns, the figure shows that, in general, their proportion was higher when $M = 10$ than in the remaining conditions. The differences between the $M = 50$ and the $M = 160$ condi-

Table 2
Order of Conditions in Experiment 2

| Bird | Trans. | Condition 1 | | Trans. | Condition 2 | | Trans. | Condition 3 | | Trans. | Condition 4 | |
		M	N		M	N		M	N		M	N
6447	7	50	49	7	160	50	6	10	55	8	160	55
2738	5	160	42	5	50	43	5	10	77	7	50	47
5291	5	50	46	6	10	48	6	160	46	5	10	83
9882	5	10	42	7	160	60	6	50	51	5	10	77

Note—Trans., number of sessions in transition phase; M, sample size; N, number of sessions.
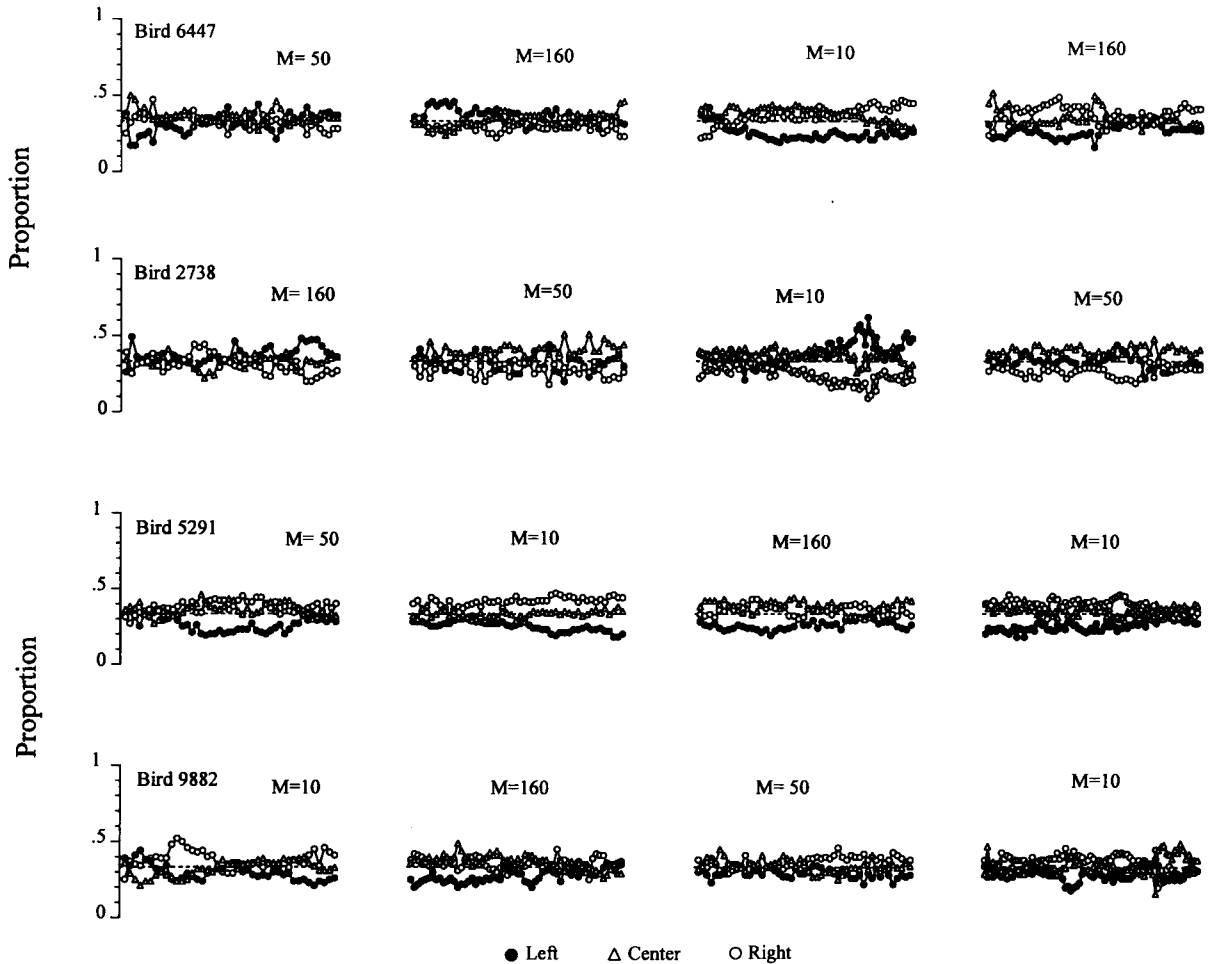
Figure 6. Molar results from the four conditions of Experiment 2. *M* refers to the sample size for each condition.

tions, however, were not consistent across birds. The last condition, which replicated one of the first two, produced the expected results for 3 birds. However, for Bird 6447, there was an unexpectedly high proportion of stable patterns, despite the high value of $M = 160$. This last finding suggests that, even though the transition sessions brought the strength of the stable patterns below .1, some carry-over effects between conditions may have been present.

The triplet data for the last five sessions are presented in Figure 8. As in Experiment 1, the proportion of the stable patterns was highest and the proportion of the repeat patterns was lowest when the sample size was 10, but there were no substantial differences between the $M = 50$ and the $M = 160$ conditions. The strength of the remaining patterns, 2X 1Y, did not change appreciably across conditions. One-way, repeated measures ANOVAs yielded a significant difference for the stable $[F(2,6) = 10.67, p = .01]$ and the repeat $[F(2,6) = 6.68, p = .03]$ patterns.

In summary, (1) molar choice proportions were always close to the matching equilibrium, and the occasional deviations from matching did not seem to be related to the experimental condition; (2) as was predicted, the largest variance (i.e., the smallest sample size) yielded the strongest stable patterns and the weakest repeat patterns; but (3) for small variance (i.e., relatively large samples), there were no major differences in the strength of the response patterns.

EXPERIMENT 3

Experiment 3 varied the average T of the reinforcement function, while keeping its variance D constant. Given the theory presented above, we predict that (1) regardless of the value of T, molar choice proportions will follow the matching law; (2) the stable patterns will be stronger as the average reinforcement probability is lower; and (3) conversely, the repeat patterns will be stronger as T increases.

Method
Birds and Apparatus. Four pigeons (*Columba livia*), 1 naive (Bird 10371) and 3 with previous experience in frequency discrimination (Machado & Cevik, 1997) and response variability (Machado, 1997) tasks, participated in the experiment. All the housing details and the experimental apparatus remained the same as those in Experiment 2.
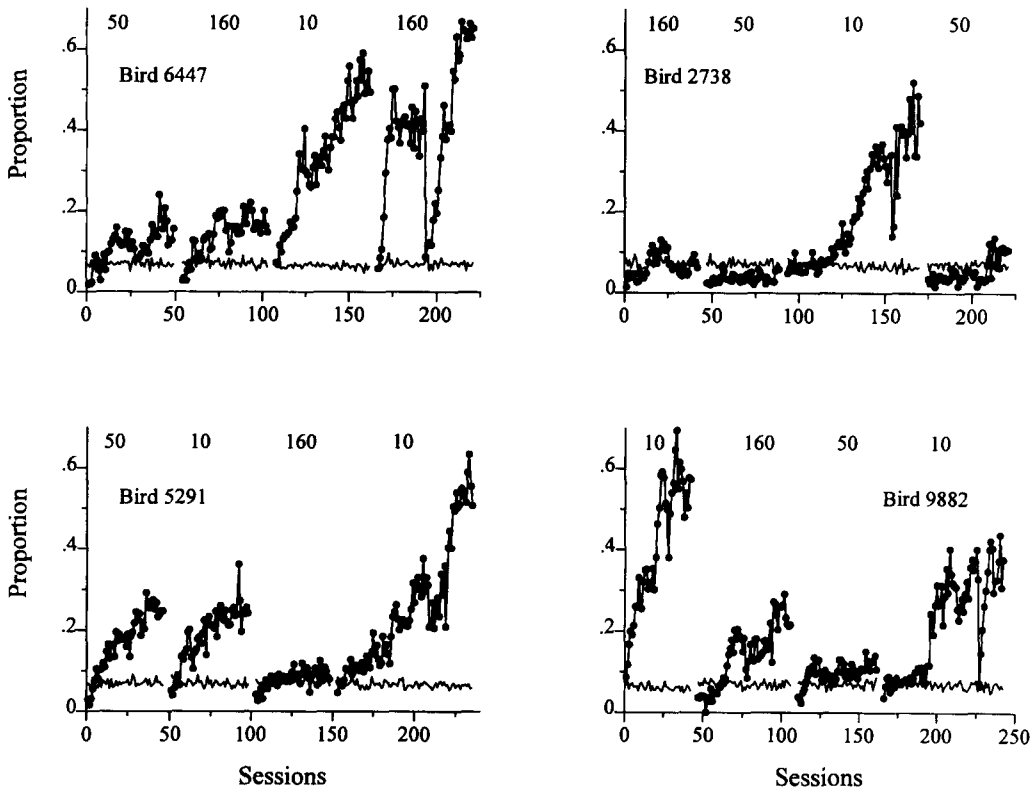
Figure 7. Proportions of the stable patterns (line with dots) and reinforcers (simple line) during each session of Experiment 2. The numbers at the top identify the sample size for each condition.

Procedure. The general procedure, including the transition phases, remained the same as that in Experiment 2. The only difference was that, in Experiment 3, the sample size remained constant at $M = 40$, whereas the overall probability of reinforcement T varied across experimental conditions. By adding the constants −.05, 0, and +.05 to Equation 4, the scheduled value for T changed from .05 to .10 to .15. Table 3 shows the details.

Results and Discussion

Figure 9 shows that molar choice proportions were always close to the 1/3 matching equilibrium. However, unlike the previous experiments, the deviations from matching were systematically related to the value of T. Thus, when $T = .15$, the deviations from 1/3 were large, but when $T = .05$, the deviations were quite small; when $T = .10$, the deviations from matching were intermediate. Observed in all the birds, these findings indicate that molar choice proportions came closer to the matching equilibrium as the overall probability of reinforcement decreased.

The top lines in Figure 10 show the proportion of reinforced trials for all the sessions of each experimental condition. To improve readability, all the values were shifted upwards by .4. The obtained values were slightly less than the scheduled ones, particularly during the condition $T = .15$. The averages across birds and sessions equalled .047, .086, and .117 for $T = .05$, .10, and .15, respectively. This discrepancy between scheduled and

obtained values, owing mainly to the decision not to reinforce the first two choices after each reinforcer, does not jeopardize the interpretation of the other findings, because the curves remained relatively constant within conditions and clearly distinct across conditions.

Figure 10 also shows the proportion of the stable patterns. In general, the strength of these patterns increased as the value of T decreased. In particular, the stable patterns attained their highest strength for all the birds when $T = .05$. Also, the last condition, a replication, yielded results roughly consistent with those from the first two conditions. Finally, by comparing Figures 9 and 10, we can see that, as the strength of stable patterns increased, molar choice proportions came closer to the 1/3 matching equilibrium.

Figure 11 shows the triplet data for the last five sessions of each condition. As the overall reinforcement probability increased, the proportion of the stable patterns decreased, whereas the proportion of the repeat patterns increased. The proportion of the remaining patterns did not change appreciably across conditions. One-way, repeated measures ANOVAs yielded a significant effect for the stable patterns [$F(2,6) = 9.71, p = .01$], but the result for the repeat patterns only approached significance [$F(2,6) = 3.61, p = .09$].

In summary, (1) choice proportions were always reasonably close to the 1/3 matching equilibrium, but the
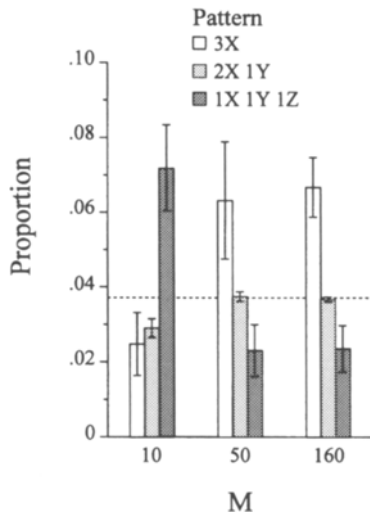
Figure 8. Mean proportion of the stable patterns (1X, 1Y, 1Z), repeat patterns (3X), and remaining patterns (2X, 1Y) as a function of the sample size *M*. The data are from the last five sessions of each condition of Experiment 2 (excluding the replication). The dotted line shows the expected proportion if the bird responded randomly. The error bars show the standard error of the means.

quality of matching improved when the average reinforcement probability decreased; (2) as the theory predicted, the strength of the stable patterns decreased with T, whereas the strength of the repeat patterns increased with T; this latter effect, however, did not reach statistical significance; and (3) there was a clear correspondence between the strength of the stable patterns and the quality of matching.

The correspondence between the strength of the stable patterns and the quality of molar matching, observed in Experiment 3, was not observed in Experiments 1 and 2. The significance of this asymmetry is unclear, for one might have expected that choice proportions would *always* come closer to perfect matching when the strength of the stable patterns increased. Although the present theory does not attempt to describe how each reinforcer affects responding—only how a reinforcement function affects the terminal strength of response patterns—the following observation may help explain this unexpected finding. Figures 5, 8, and 11 show that, when D changed

in Experiments 1 and 2, the strengths of the stable and repeat patterns changed by comparable amounts, albeit in opposite directions, but when T changed in Experiment 3, the strength of the repeat patterns changed by a much larger amount. Thus, it may be that T affects primarily the strength of the repeat patterns and secondarily the strength of the stable patterns, whereas D may affect both types of patterns equally. The effects of D and T may, thus, correspond to the two distinct ways to strengthen a pattern—reinforce it more (the main effect of increasing D?) or extinguish its competitors (the main effect of decreasing T?).

GENERAL DISCUSSION

The present study attempted to answer the following question: In simple choice situations, what factors determine the strength of a response pattern? To that end, we have conceived of a choice situation in the following way: An animal chooses from a population of response patterns; in contact with the contingencies of reinforcement, the strength of the various members of the population changes (i.e., the patterns replicate differentially, and the population evolves, as it were); at the steady state, one or more members are dominant. The process of change may depend on two general sets of factors, one related to the reinforcement function defined over the population (e.g., its average and variance), the other related to the population itself (e.g., memory constraints, the effort associated with the various patterns, induction across the patterns). We then sketched a quantitative theory for predicting how these two sets of factors may determine the outcome of the patterns' arms race. The theory predicts that, all else being equal, the asymptotic strength of a response pattern varies directly with the variance and inversely with the average of the reinforcement probability function. The data from the three experiments tend to agree with these predictions. In Experiments 1 and 2, the variance was manipulated by changing the sample size while the average remained constant. It was found that, when the variance of the payoff function increased, the strength of the patterns for which $D > 0$ increased, whereas the strength of the patterns for which $D < 0$ decreased. In Experiment 3, the average of the reinforcement function was manipulated while its variance remained constant. The data showed that the strength of the response patterns changed in the predicted direction.

Table 3
Order of Conditions in Experiment 3

| Bird | Trans. | Condition 1 | | Trans. | Condition 2 | | Trans. | Condition 3 | | Trans. | Condition 4 | |
		T	N		T	N		T	N		T	N
10405	6	.05	67	6	.15	67	6	.10	65	6	.15	65
10490	6	.10	44	7	.05	61	9	.15	45	5	.05	67
10371	5	.10	45	6	.15	45	6	.05	50	7	.15	45
5269	6	15	44	7	.10	62	6	.05	63	7	.10	56

Note—Trans., number of sessions in transition phase; T, overall reinforcement possibility; N, number of sessions.
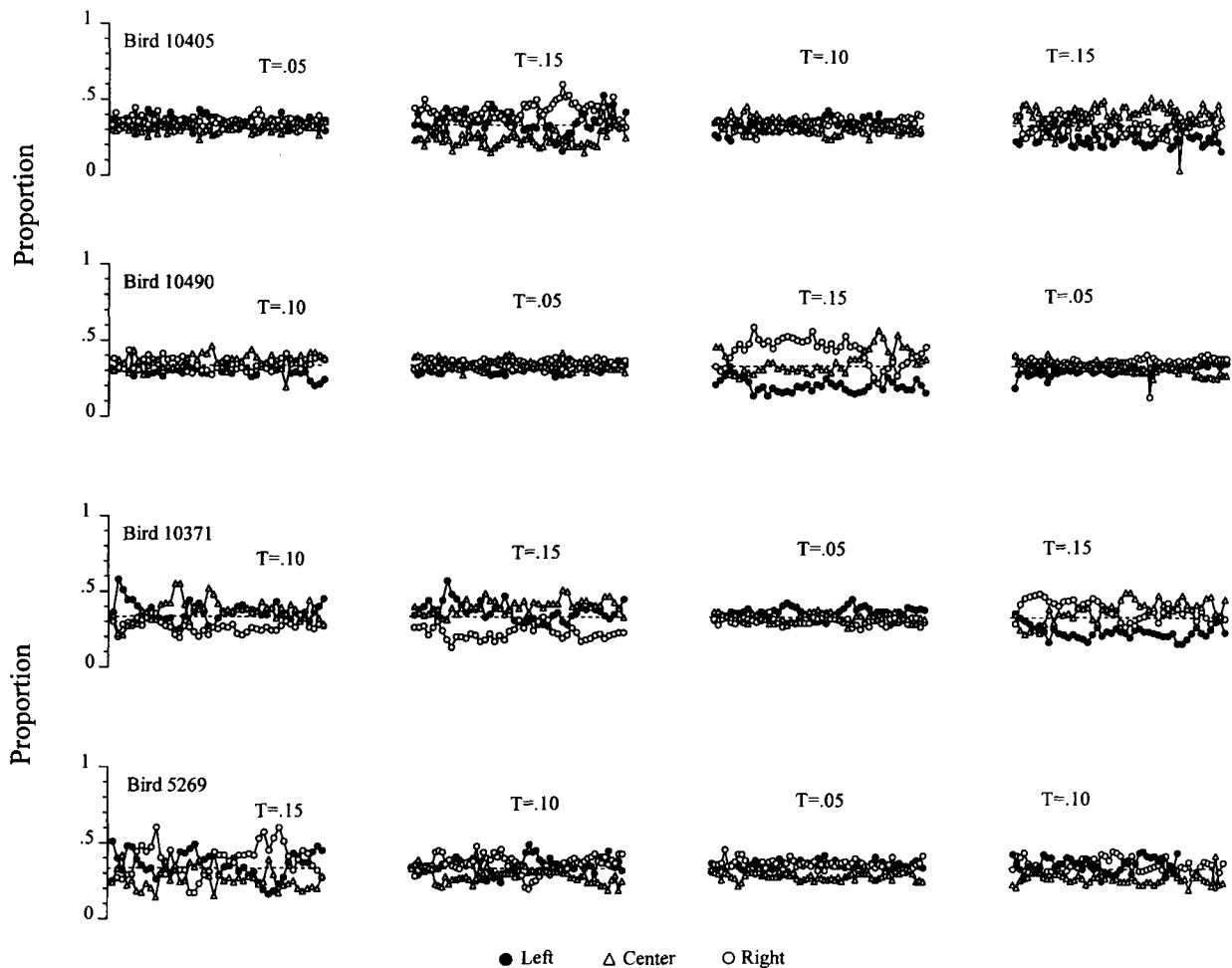
Figure 9. Molar results from the four conditions of Experiment 3. *T* refers to the (scheduled) average probability of reinforcement for each condition.

The theory not only identified two variables, *D* and *T*, but also proposed a way to combine them, the ratio *D/T*. We can now assess how well this ratio was correlated with the asymptotic strength of the response patterns. Figure 12 shows the average strength of the stable and repeat patterns, plotted against the ratio *D/T*. The value of *D* was set to $1/\sqrt{M}$, where *M* was the sample size.[2] The top panels show the data from Experiments 1 and 2, the middle panels show the data from Experiment 3, and the bottom panels show the data from all three experiments. The lines are best-fitting regression lines. Although the relation between *D/T* and the strength of a pattern cannot be linear—for strength must remain between 0 and 1, whereas *D/T* has no bounds—in most cases, a linear relation described reasonably well the major trend in the data. The proportions of the repeat patterns in Experiment 3, however, showed some curvature, which is also apparent in the collapsed data displayed in the bottom right panel. Be that as it may, we conclude that the ratio *D/T* is a good first-order predictor of the strength of a response pattern.

The Learning of Response Patterns

Our experimental findings concerning the role of *D* are consistent with previous studies. Williams (1972), for example, found that when win–stay and win–shift response patterns were reinforced with probabilities .65 and .35, respectively, pigeons did not learn the win–stay pattern, but they learned it when the probabilities changed to .8 and .2 (see, also, Williams, 1991). In terms of the preceding theory, the pattern was stronger when *D* was larger. Similarly, Silberberg and Williams (1974) found strong alternation (which, in their experiment, violated the matching law) when "each choice produced dramatic changes in local reinforcement probabilities" (p. 321)— that is, when *D* was large. Our theory and findings add to these studies one proviso: The effect of the differences in reinforcement probability are modulated by the overall context of reinforcement. In other words, the correct predictor of the strength of a pattern may be the ratio *D/T*, not *D* alone. We know of no other studies on response patterning that manipulated the variance or the average of the reinforcement function.
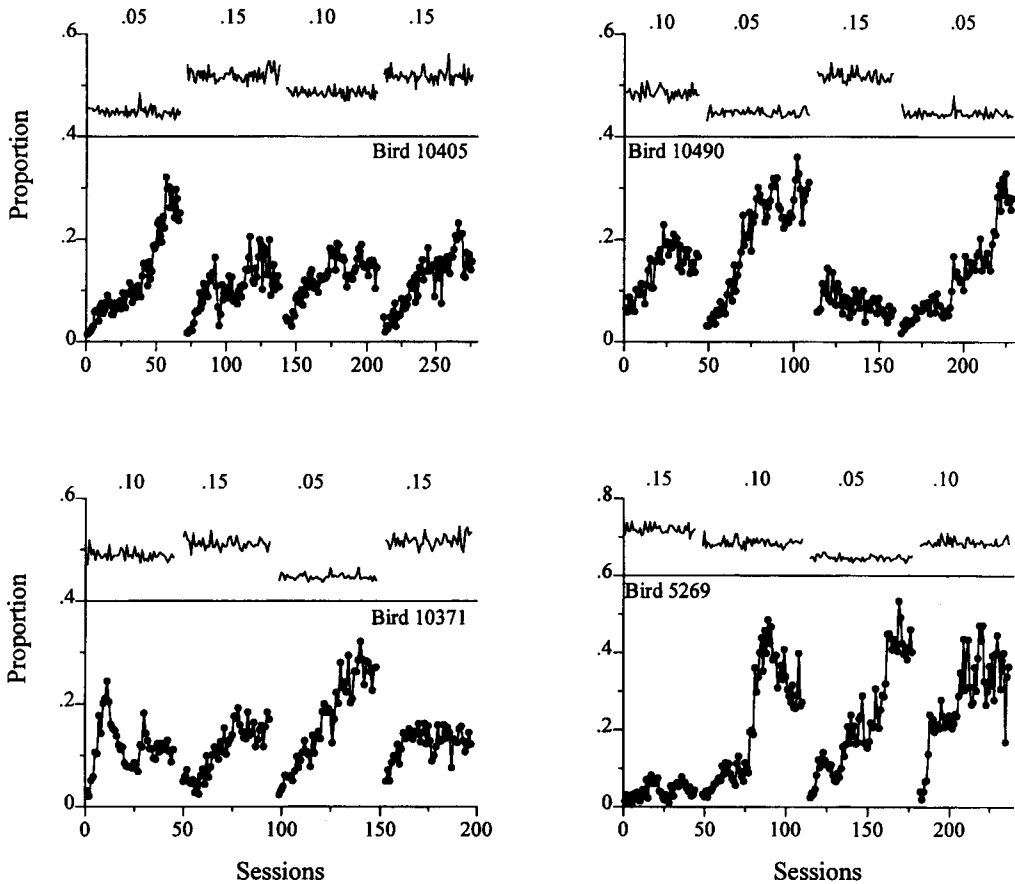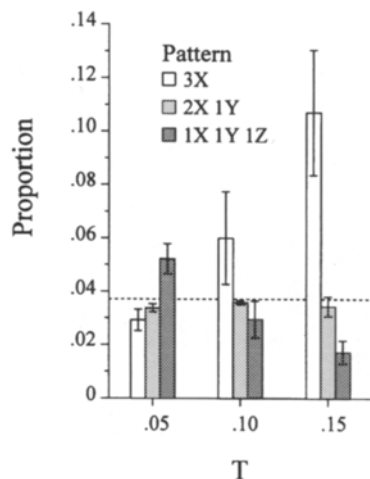
Figure 10. Proportion of the stable patterns (line with dots) and reinforcers (simple line) during each session of Experiment 3. The reinforcement proportions were shifted upwards by .4. The numbers at the top identify the average probability of reinforcement for each condition.

In some conditions of Experiments 1 and 2, the stable patterns were very strong, but no bird ever showed an exclusive preference for them. The absence of exclusive preference calls into question some attempts to apply the matching law directly to response patterns. Thus, after conceding that local contingencies of reinforcement sometimes produce response patterns, Herrnstein (1997) went on to say that "when a subject discriminates shifting local probabilities, ... the matching law applied locally predicts *local exclusive preferences*, presumably for the alternative that is locally reinforcing with the higher probability" (p. 70, italics added). In contrast, our theory suggests that the nonexclusive preference for the stable patterns is due to a cost function that increases with the frequency of these patterns. According to the theory, the strength of the stable patterns increased until an equilibrium was reached wherein the tendency to a further increase, $f(D/T)$, was offset by the additional cost of performing the patterns, $g(S)$. Without the concept of a cost function—with cost being broadly conceived—it is not clear how to explain the fact that the stable patterns in the present experiments or the momentary maximizing pat-

terns in previous studies (e.g., Shimp, 1966; Silberberg et al., 1978) did not occur with a probability of 1.

Not all the data from individual birds agreed with the theory, however. In some cases, the absolute strength of the stable patterns did not conform to the value of D or T; in other cases, the results from an early condition could not be replicated. These two results suggest the presence of carryover effects across conditions, the transition phases notwithstanding. Other studies have also reported order effects when the differences in reinforcement probabilities are manipulated. Williams (1991), for example, reported that, in a choice situation, one group of rats learned to shift levers reliably when the difference in reward probabilities for shifting and staying were large and that they maintained that behavioral pattern when, subsequently, the difference was reduced. However, a group of rats that did not benefit from an initial exposure to the large difference did not learn to shift levers with the smaller difference. The author concluded that "stimulus control is more likely to be established with larger differences in local reinforcement probability [i.e., large D], and, once established, may persist even after the differences in re-

Figure 11. Mean proportion of the stable patterns (1X, 1Y, 1Z), repeat patterns (3X), and remaining patterns (2X, 1Y) as a function of the overall reinforcement probability T. The data are from the last five sessions of each condition of Experiment 3 (excluding the replication). The dotted line shows the expected proportion if the bird responded randomly. The error bars show the standard error of the means.

inforcement probability are reduced to a level that would not enable the establishment of stimulus control if presented from the beginning of training [i.e., small D]" (Williams, 1991, p. 471). Similarly, Machado (1993) found that, once a pigeon learned to double alternate when the variance in the reinforcement function was large, it maintained that response pattern when the variance later decreased (in that study, the variance was also manipulated by changing the sample size).

There is a striking similarity between the theory proposed here and Nevin's (e.g., 1988) behavioral momentum theory. Using an analogy from mechanics, Nevin proposed that the magnitude of a behavioral change depends on two variables—a disruptor, which acts like an external force, and resistance to change, which acts like mass (the analogy is with Newton's second law, $F/m = a$, where F, m, and a represent force, mass, and acceleration, respectively). According to Nevin, the differences in reward probability (D) are one type of disruptive force, and the overall reinforcement rate in the situation (T) is one determinant of behavioral mass. Hence, Nevin's behavioral momentum theory, applied to response patterns, also suggests the critical role of the ratio D/T.

The Local Structure of Choice Behavior

As we mentioned in the introduction, some studies have found that the patterns and the sequential dependencies between consecutive responses are predicted by molecular accounts (e.g., Shimp, 1966; Silberberg et al., 1978), whereas others have not found such dependencies (Heyman, 1979) or, worse, have found the opposite ones

(Nevin, 1969, 1979). The sequential dependencies observed in the former studies correspond to the stable patterns in our experiments, because, when molar choice proportions are at equilibrium, the stable patterns are the momentary maximizing sequences. The opposite dependencies, observed in the latter studies, correspond to the repeat patterns in our experiments, because, when molar choice proportions are at equilibrium, the repeat patterns are the momentary minimizing sequences. What is interesting to note here is that our experiments reproduced the two opposite types of sequential dependencies by varying the ratio D/T: As the ratio increased, the repeat patterns decreased and the stable patterns increased.

To show more clearly the relationship between the ratio D/T and the local structure of behavior, we reanalyzed our data in the following way. For each reinforced response, we determined whether the same response occurred during the trials that immediately preceded and the trials that immediately followed the reinforced trial. The analysis proceeded in both directions simultaneously and stopped when another reinforced trial was found or a total of 5 trials in each direction had been included. This procedure ensured that the analysis was based on the same number of (reinforcer-free) trials before and after the reinforced trial. The totals were converted into probabilities, and the probabilities were first averaged across the three responses and then across the 4 birds. Figure 13 shows the results for the last five sessions of each condition.

On each panel, Trial 0 corresponds to a reinforced trial. The right curve is the probability of emitting the just reinforced response during each of the next five trials, and the left curve is the same probability for each of the preceding five trials. The nine panels (3 experiments × 3 conditions per experiment) are arranged from left to right and from top to bottom according to the ratio D/T of the corresponding condition. For example, in one condition of Experiment 2, $M = 50$ and $T = .08$; hence, $D = 1/\sqrt{50}$, and $D/T = 1.77$. This condition is represented by the middle panel.

Consider the left curves first. As we move from lower to higher D/T ratios, the left curves first increase and then decrease. The change from a positive to a negative slope is quite orderly, for no decreasing curve precedes an increasing one. Moreover, the three curves for each individual experiment show the same orderly relation. A positive slope indicates response perseveration (i.e., strong repeat patterns); a negative slope indicates a tendency to switch keys (i.e., strong stable patterns). In particular, when a bird responds LCR or RCL frequently, the reinforced response generally differs from the two preceding responses—hence the dip in the left curve on trials -2 and -1 in the two rightmost bottom panels. In summary, as the ratio D/T increased, the tendency to persevere gave way to the tendency to switch keys—the momentary maximizing pattern replaced its opposite. This
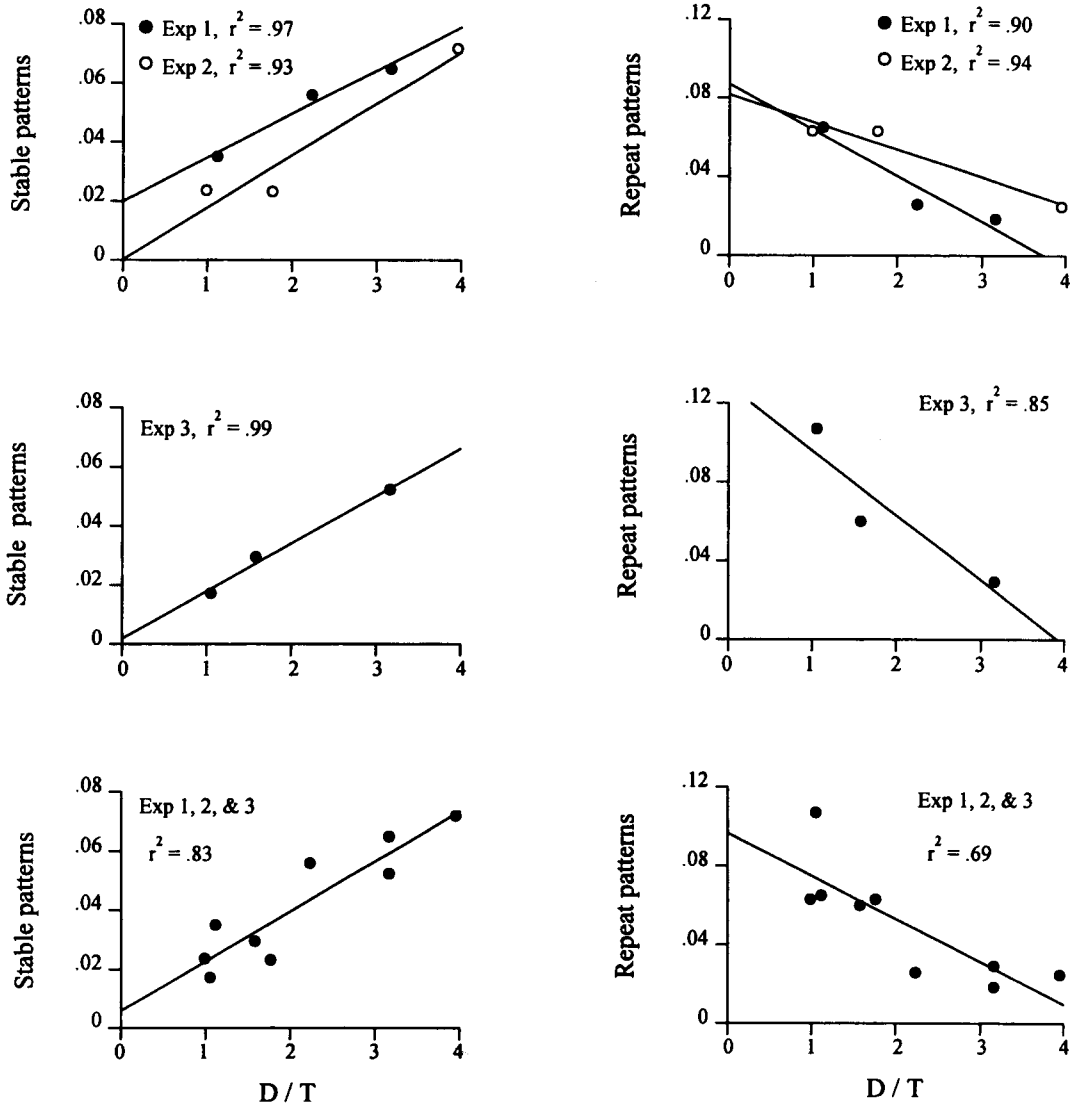
Figure 12. The left and right panels plot the average proportions of the stable and repeat patterns, respectively. The top panels show the data from Experiments 1 and 2, the middle panels show the data from Experiment 3, and the bottom panels show the data from all three experiments. The data are from the last five sessions of each condition (excluding replications). Also shown are the best-fitting regression lines. The value of D equaled $1/\sqrt{M}$, where M was the sample size.

result reproduces the two extreme forms of sequential dependency found in previous studies, but it accounts for them in terms of a single variable, the ratio D/T.

Consider now the right curves. They are included because they allow us to answer the question of whether, at the steady state, reinforcers strengthened individual responses (see, e.g., Bailey & Mazur, 1990; Mazur, 1992). The question is important, because if that were the case, our findings about the effects of T could be explained in a simpler way. When the number of rewards (i.e., T) increases, the cumulative strengthening effect of the rewards also increases. It follows that the frequency of the repeat patterns should increase with T, as our experiments showed.

To understand what the right curves should look like if reinforcers strengthened single responses, consider the following argument. At equilibrium a bird pecks the left, right, and center keys with roughly constant probabilities p, q, and r, respectively. The average of these probabilities is, obviously, $1/3$ because $p + q + r = 1$. If when a left response is reinforced, its strength increases, p should increase to p_1, at least during the next few trials; the same argument for right and center pecks implies that, when these responses are reinforced, q and r should increase to q_1 and r_1. Therefore, the average of the new values p_1, q_1, and r_1 should be greater than $1/3$. By comparing the right curves in the figure with the dotted lines (equal to $1/3$), we see that any strengthening effects of rein-
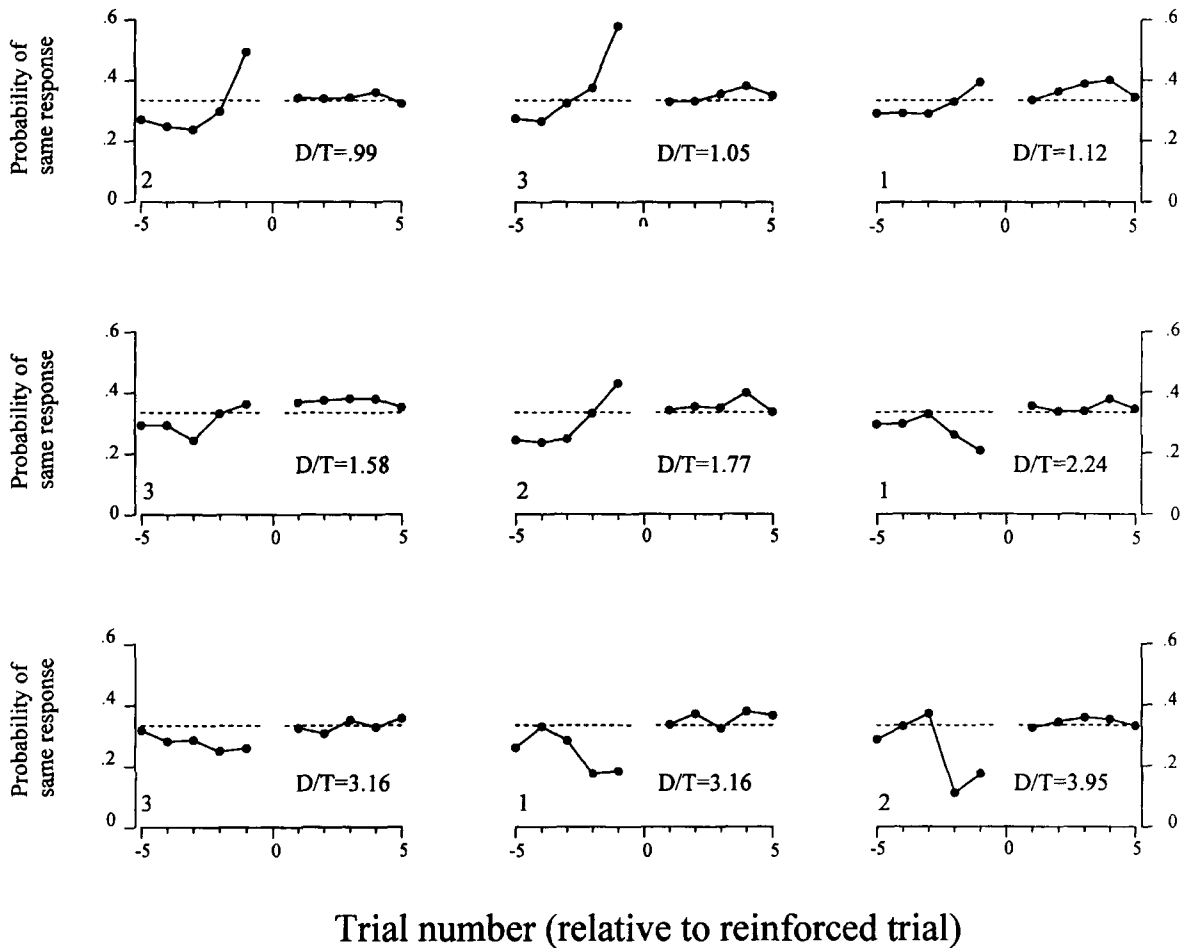
Trial number (relative to reinforced trial)

Figure 13. Assume that a response was reinforced on Trial 0. The right curve on each panel shows the probability of emitting the same response during the next five trials, and the left curve shows the probability of emitting the same response during the preceding five trials. Each panel corresponds to a different experiment (identified by the number at bottom left) or experimental condition (identified by the corresponding ratio D/T, where $D = 1/\sqrt{M}$).

forcement were minimal at best, nonexistent at worst. Only when $D/T = 1.12$ and 1.58 was there some evidence of a (weak) strengthening effect of reinforcement.

Furthermore, a close analysis of the behavior of each bird revealed that, instead of tending to repeat the just-reinforced response, most birds tended to emit a particular sequence when they resumed pecking after consuming the reinforcer. For example, during the $M = 10$ condition, Bird 6112 emitted the sequence CLC after most reinforcers. Because that sequence occurred regardless of which response had just been reinforced, the postreinforcement probabilities averaged to the expected value of 1/3. Interestingly, the postreinforcement sequences were not always the dominant sequence; they occurred regardless of whether the repeat or the stable patterns were dominant, and they occurred even when, as in Experiments 2 and 3, the two choices that followed a reinforced trial were never reinforced. In summary, in contrast with some previous studies (but see Mazur, 1992, for a summary of conflicting findings on this issue), we found no evidence of an immediate strengthening effect of rein-

forcers. Instead, reinforcers seemed to cue idiosyncratic response patterns. It follows that our results concerning T cannot be explained by a simple strengthening effect of rewards on individual responses.

Another Look at Concurrent Variable-Interval Schedules

The application of the proposed theory to Conc. VI schedules presents some difficulties, but it also offers some new insights. To see the difficulties, consider the following situation. A pigeon chooses among three keys, left, center, and right, that set up rewards according to three independent VI 20-sec schedules. Assume that the bird responds at an average rate of λ pecks per second and allocates a proportion p of its choices to the left, a proportion q to the right, and a proportion r to the center keys, with $p + q + r = 1$ (see, e.g., Gibbon, 1995). The probability that, say, a left peck is rewarded is then equal to

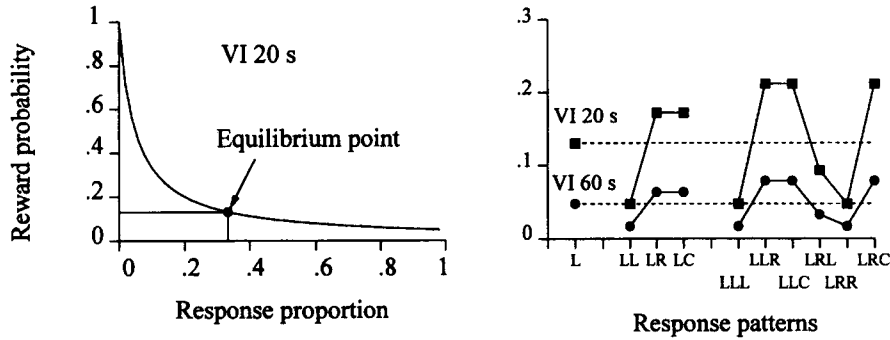$$P(S^+ | L) = \frac{\gamma_L}{\gamma_L + \lambda p}, \tag{5}$$

Figure 14. Assuming that the bird pecks at an overall rate of $\lambda = 1$ pecks per second and allocates proportions p, q, and r to the left, right, and center keys, respectively, the left panel shows the probability of reinforcement for a response as a function of its proportion. The right panel shows the reward probability for single responses, pairs, and triplets when $p = q = r = 1/3$. The squares are for the variable-interval (VI) 20-sec schedules, and the circles are for the VI 60-sec schedules. The dotted lines represent the average reinforcement probabilities.

where γ_L is the reciprocal of the VI parameter. This hyperbolic function is plotted in the left panel of Figure 14, and it shows that the more the bird prefers the left key, the lower the probability of reinforcement for left keypecks. If we restrict our attention to the population of patterns composed by left, right, and center keypecks and assume that their cost functions are equal (i.e., no bias), Equation 2 predicts that, at equilibrium, the bird will be indifferent among the keys and matching will occur.

The similarity between Equation 5 for VI schedules and the equation used in the present experiments (Equation 4) predicts that the reinforcement functions for response patterns should also be similar in the two schedules. The right panel of Figure 14 shows that this is indeed the case.[3] The squares represent the reward probabilities induced by the VI 20-sec schedules when molar choice proportions are at equilibrium (i.e., $p = q = r = 1/3$; see Appendix B for the derivation). As in Figure 2, the function for pairs shows that repeating is less likely to be reinforced than switching, and the function for triplets shows that, after the bird switches from one key to another, the reinforcement probability is highest on the third key. Hence, if the costs of all the patterns were identical, Equation 3 would predict that the stable patterns should win the competition. For similar reasons, the repeat patterns should lose the competition.

The problem, however, is that cost differences may invalidate these predictions. Equation 3 states only that the strength of some specific patterns should increase, whereas that of others should decrease, with the ratio D/T. Hence, to test this prediction with Conc. VI schedules, the ratio D/T should be manipulated across conditions. Here one is tempted to change the VI parameters, from 20 to 60 sec, say. But the second function in the right panel of Figure 14 (see circles) shows that changing the VIs to 60 sec decreases both the average and the variance of the payoff function by approximately the same factor. The ratio D/T remains basically unchanged.

Two consequences follow from the preceding analysis, both with negative overtones. First, if the theory is correct, the asymptotic strengths of the various response patterns should not change with the VI parameters. We know of no data set that could be used to test this prediction. Second, and perhaps more important, VI schedules are not suited to test the roles of D and T, because they do not allow the independent manipulation of these variables. In fact, the only way to change the ratio D/T in Conc. VI schedules is to vary the animal's overall response rate λ. But when λ changes, the interval between consecutive responses and, therefore, the cuing function of previous responses also change. As a consequence, the experimenter would be unable to determine which variable, the ratio D/T or the interresponse interval, caused the effect.

Be that as it may, our theory suggests new ways to think about response patterns in Conc. VI schedules. For example, in typical VI schedules, overall response rate is much higher than overall reinforcement rate, and, therefore, a run of responses on one key changes only slightly the reward probabilities on the other keys. It follows that typical Conc. VI schedules may be functionally equivalent to the high M schedule used in some conditions of Experiments 1 and 2. If this interpretation is correct, the theory proposed in this study would predict weak stable patterns in the typical Conc. VI schedules. Conversely, to obtain strong stable patterns, the theory states that Conc. VI schedules need to be modified so that each response causes a larger change in the local reward probabilities, as in the low M conditions of our experiments, while simultaneously, the overall reward probability remains low.

The theory also predicts that, without a changeover delay, undermatching is likely to occur. Consider, for example, a two-key, Conc. VI 20 sec VI 60 sec. If we plot the reinforcement function for response pairs and triplets, using the equations provided in Appendix B, we see that the reward probabilities for switching (LR and RL) are higher than the reward probabilities for staying (LL

and RR). Hence, without a changeover delay, the animal is likely to switch frequently and to undermatch. Note that the high switching rate is explained by invoking, not adventitious or accidental reinforcement (see, e.g., Davison & McCarthy, 1988, p. 10), but the relatively higher rate of *direct* reinforcement for switching than for staying responses. More generally, the theoretical analysis of the payoff function for response patterns and the empirical study of how its properties (e.g., D and T) affect responding may help us to achieve a deeper quantitative understanding of the animal's behavior in Conc. VI schedules.

The theory may also be extended to types of patterns other than sequences of keypecks. For instance, if a pigeon chooses between two interresponse times, one with duration t_1 and rewarded according to schedule r_1, another with duration t_2 and rewarded according to schedule r_2, the theory would predict that, asymptotically, the strength of these response units should depend on the ratio D/T. To derive specific predictions, the experimenter would need to obtain the functions relating D and T to (1) the schedule parameters r_1 and r_2, (2) the average and variance of the durations of the two interresponse times t_1 and t_2, and (3) the proportion of choosing t_1—say, p. Once these functions are obtained, either analytically or by computer simulations, the theory can be tested.

Matching and Response Patterns

The picture that emerges from the studies summarized above, as well as from our experimental results, is that molar matching does not seem to be reducible to the learning of any specific response patterns. For matching occurs when the animal responds in short or long runs, by learning the momentary maximizing or the momentary minimizing patterns, with strong sequential dependencies between consecutive responses or in a quasi-random way (see, e.g., Heyman, 1979). Furthermore, matching seems to be generated by a fast-acting, reward-following process—of the sort of melioration, for example—whereas response patterns seem to be generated by a slower and noisier process. And yet, response patterns do occur in choice situations, and under some conditions, matching may even be violated because the animal learned a response pattern (see, e.g., Silberberg & Williams, 1974). And so, one is left with the thorny issue of how the molar and molecular structures of behavior are interrelated.

We distinguish at least three viewpoints on the interrelation of the molar and molecular aspects of behavior in choice situations. Some authors seem to imply that distinct and independent processes may be involved: "Orderly effects at the molecular level, though perhaps of interest in their own right, do not shed light on molar processes" (Nevin, 1979, p. 300). Other authors seem to believe that local processes shape behavior into new response units and, then, global processes strengthen these units (e.g., Nevin, 1982; Shimp, 1982b; Williams, 1990). Finally, still other authors argue that global and local processes jointly control behavior: "The results . . . suggest

that control by local reinforcement contingencies and by molar reinforcement contingencies are in competition" (Williams, 1991, p. 457). The first viewpoint does not seem tenable, because, as Silberberg and Williams (1974) have shown, sometimes the matching law is violated because the animal has learned a response pattern incompatible with matching. The second and third viewpoints, although more promising, have not been sufficiently elaborated. Thus, it is unclear how we can separate the strengthening of a response unit from the shaping of that unit when the only evidence available for the unit is its strength. And without a model for competition between contingencies, the sense of William's metaphor remains elusive. The last two viewpoints, however, can be integrated with the theory proposed here and can generate testable hypotheses.

Assume that (1) one-response and multiresponse patterns are members of the same population of competing patterns (i.e., the arms race is among patterns of different lengths, including single responses, runs of responses, and patterns involving changeovers), (2) the reinforcement contingencies change the strength of all types of patterns according to Equation 3, and (3) one-response patterns have faster acquisition rates than multiresponse patterns (or, in other words, in Equation 2, α is greater for single responses). A few consequences follow from these general assumptions. First, when the multiresponse patterns favored by the reinforcement contingencies yield molar matching in the aggregate—as the stable patterns in our experiments yield the 1/3 matching proportions—and the ratio D/T is high, response patterning and overall matching will both occur. Matching is predicted because, for single responses, Equation 2 makes the same predictions as melioration. However, because of the differences in the rate of learning, molar matching will typically precede the response patterning. Second, when the patterns favored by the reinforcement contingencies are incompatible with matching, molar matching may be observed initially, but after prolonged training, it will be violated. Third, regardless of which patterns are favored by the reinforcement contingencies, if these patterns are associated with small ratios of D/T or are difficult to learn (because of long ITIs, e.g.), matching is predicted even after prolonged training.

The foregoing consequences seem to be consistent not only with our experimental findings, but also with those of Silberberg and Williams' (1974) study. In the latter, three groups of pigeons obtained food only when they switched keys, but because switching in one direction was reinforced more often than switching in the other, the overall reinforcement probabilities for the two keys differed. When the ITI was 20 sec, the birds initially matched choice proportions with reinforcement proportions and did not switch very often; with additional training, however, switching became stronger, and matching was violated (by switching frequently between the keys, choice proportions were close to .5, but one key delivered much more food than the other); when the ITI was 5 sec, switching developed rapidly, and matching was vi-

olated (unfortunately, because the authors did not report the data from the first sessions, we do not know whether matching was also initially observed when the ITI was 5 sec); when the ITI was 120 sec, switching was not acquired, and molar choice remained close to matching.

In conclusion, this study advanced the hypothesis that the asymptotic strength of response patterns in simple choice situations may be a function of the ratio D/T. The results of three experiments agreed with the hypothesis, for the strength of the response pattern increased with D and decreased with T. The theory that generated the hypothesis is clearly incomplete—it has not detailed the specific form of functions f and g, for example, or how temporal discrimination processes affect the learning of response patterns. Whether it will be able to grow and to survive further empirical and theoretical challenges, remains to be seen.

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.

BAUM, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, **20**, 137-153.

BLOUGH, D. S. (1966). The reinforcement of least-frequent interresponse times. *Journal of the Experimental Analysis of Behavior*, **9**, 581-591.

DAVISON, M., & McCARTHY, D. (1988). *The matching law: A research review*. Hillsdale, NJ: Erlbaum.

GALLISTEL, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

GIBBON, J. (1995). Dynamics of time matching: Arousal makes better seem worse. *Psychonomic Bulletin & Review*, **2**, 208-215.

HERRNSTEIN, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, **4**, 267-272.

HERRNSTEIN, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, **13**, 243-266.

HERRNSTEIN, R. J. (1997). *The matching law: Papers in psychology and economics*. Cambridge, MA: Harvard University Press.

HEYMAN, G. M. (1979). A Markov model description of changeover probabilities on concurrent variable interval schedules. *Journal of the Experimental Analysis of Behavior*, **31**, 41-51.

HINELINE, P., SILBERBERG, A., ZIRIAX, J., TIMBERLAKE, W., & VAUGHAN, W., JR. (1987). Commentary prompted by Vaughan's reply to Silberberg and Ziriax (1987). *Journal of the Experimental Analysis of Behavior*, **48**, 341-346.

HINSON, J. M., & STADDON, J. E. R. (1983). Hill-climbing in pigeons. *Journal of the Experimental Analysis of Behavior*, **39**, 25-47.

HIRAOKA, K. (1984). Discrete-trial probability learning in rats: Effects of local contingencies of reinforcement. *Animal Learning & Behavior*, **12**, 343-349.

HUNZIKER, M. H. L., SALDANA, R. L., & NEURINGER, A. (1996). Behavioral variability in SHR and WKY rats as a function of rearing environment and reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, **65**, 129-143.

KILLEEN, P. (1981). Averaging theory. In C. M. Bradshaw, E. Szabadi, & C. F. Lowe (Eds.), *Recent developments in the quantification of steady-state operant behavior* (pp. 21-34). New York: Elsevier.

MACHADO, A. (1992). Behavioral variability and frequency-dependent selection. *Journal of the Experimental Analysis of Behavior*, **58**, 241-263.

MACHADO, A. (1993). Learning variable and stereotypical sequences of response: Some data and a new model. *Behavioral Processes*, **30**, 103-130.

MACHADO, A. (1994). Polymorphic response patterns under frequency-dependent selection. *Animal Learning & Behavior*, **22**, 53-71.

MACHADO, A. (1997). Increasing the variability of response sequences in pigeons by adjusting the frequency of switching between two keys. *Journal of the Experimental Analysis of Behavior*, **68**, 1-25.

MACHADO, A., & CEVIK, O. (1997). The discrimination of relative frequency by pigeons. *Journal of the Experimental Analysis of Behavior*, **67**, 11-41.

MARK, T. A., & GALLISTEL, C. R. (1994). Kinetics of matching. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 79-95.

MAZUR, J. E. (1992). Choice behavior in transition: Development of preference with ratio and interval schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 364-378.

MECHNER, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, **1**, 109-121.

NEVIN, J. A. (1969). Interval reinforcement of choice behavior in discrete trials. *Journal of the Experimental Analysis of Behavior*, **12**, 875-885.

NEVIN, J. A. (1979). Overall matching versus momentary maximizing: Nevin (1969) revisited. *Journal of Experimental Psychology: Animal Behavior Processes*, **5**, 300-306.

NEVIN, J. A. (1982). Some persistent issues in the study of matching and maximizing. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. II: Matching and maximizing accounts* (pp. 153-165). Cambridge, MA: Ballinger.

NEVIN, J. A. (1988). Behavioral momentum and the partial reinforcement effect. *Psychological Bulletin*, **103**, 44-56.

RACHLIN, H., BATTALIO, R., KAGEL, J., & GREEN, L. (1981). Maximization theory in behavioral psychology. *Behavioral & Brain Sciences*, **4**, 371-417.

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.

SHIMP, C. P. (1969). Optimal behavior in free-operant experiments. *Psychological Review*, **76**, 97-112.

SHIMP, C. P. (1976). Short-term memory in the pigeon: The previously reinforced response. *Journal of the Experimental Analysis of Behavior*, **26**, 487-493.

SHIMP, C. P. (1982a). Choice and behavioral patterning. *Journal of the Experimental Analysis of Behavior*, **37**, 157-169.

SHIMP, C. P. (1982b). Reinforcement and the local organization of behavior. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. II. Matching and maximizing accounts* (pp. 111-130). Cambridge, MA: Ballinger.

SHIMP, C. P. (1990). Theory evaluation can be unintentional self-portraiture: A reply to Williams. *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 217-221.

SHIMP, C. P., CHILDERS, L. J., & HIGHTOWER, F. A. (1990). Local patterns in human operant behavior and a behaving model to interrelate human and animal performances. *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 200-212.

SILBERBERG, A., HAMILTON, B., ZIRIAX, J. M., & CASEY, J. (1978). The structure of choice. *Journal of Experimental Psychology: Animal Behavior Processes*, **13**, 292-301.

SILBERBERG, A., & WILLIAMS, D. (1974). Choice behavior in discrete trials: A demonstration of the occurrence of a response strategy. *Journal of the Experimental Analysis of Behavior*, **21**, 315-322.

STADDON, J. E. R. (1983). *Adaptive behavior and learning*. Cambridge: Cambridge University Press.

STADDON, J. E. R., & MOTHERAL, S. (1978). On matching and maximizing in operant choice experiments. *Psychological Review*, **85**, 436-444.

WILLIAMS, B. [A.] (1972). Probability learning as a function of momentarily reinforcement probability. *Journal of the Experimental Analysis of Behavior*, **17**, 363-368.

WILLIAMS, B. A. (1983). Effects of intertrial interval on momentary maximizing. *Behaviour Analysis Letters*, **3**, 35-42.

WILLIAMS, B. [A.] (1988). Reinforcement, choice, and response strength. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (2nd ed., Vol. 2, pp. 167-244). New York: Wiley.

WILLIAMS, B. A. (1990). Enduring problems for molecular accounts of

operant behavior. *Journal of Experimental Psychology: Animal Behavior Processes, 16*, 213-217.

WILLIAMS, B. [A.] (1991). Choice as a function of local versus molar reinforcement contingencies. *Journal of the Experimental Analysis of Behavior, 56*, 455-473.

NOTES

1. The reason for not calling them *momentary maxiziming* patterns is that, when molar choice proportions are far from the matching equilibrium, the contingencies of reinforcement do not favor the stable patterns. For example, if a pigeon chooses the left key during a large number of trials, with *large* being defined with respect to the sample size M, most rewards will follow pecks at the right and center keys, and several *consecutive* pecks on one of these two keys are likely to be reinforced. In this case, the stable patterns would not be differentially reinforced. Eventually, however, the bias for the left key will be reduced, and the strengths of the three choices will approach equality. Then, and only then, will the stable patterns be the momentary maximizing patterns.

2. So far we have used the term *variance* to refer to the set of differences in the payoff probabilities, D, without implying a specific way to measure those differences. This vagueness was acceptable as long as we made only ordinal predictions. However, when a specific metric is needed, the standard deviation seems preferable to the variance because, if nothing else, it is commensurable with T. Appendix A shows that the variance of the payoff function is proportional to $1/M$. The standard deviation is, therefore, proportional to $1/\sqrt{M}$.

3. Another similarity is that, in both types of schedules, the overall reward probability tends to be fairly constant (see Figures 4, 7, and 10).

APPENDIX A

Assume that overall choice proportions equal the 1/3 matching equilibrium. One way to derive expressions that approximate the average and the variance of the reinforcement function is to use a Taylor series expansion. Starting with Equation 4,

$$h(i) = \frac{1}{27\frac{i}{M}+1} = \frac{M}{27i+M},$$

where i is the number of (left, right, or center) choices on the last M trials; we expand function h around the average value of i, called μ,

$$h(i) \approx h(\mu) + h'(\mu)(i - \mu),$$

where $h'(\mu)$ is the derivative of h evaluated at $i = \mu$.
Then,

$$E[h(i)] \approx h(\mu) = \frac{1}{10},$$

because, at the matching equilibrium, $\mu = M/3$. Also,

$$E[h(i)^2] \approx f(h)^2 + h'(\mu)^2\sigma_i^2,$$

where σ_i^2, the variance of i, equals $M \times 1/3 \times 2/3$ (variance of a binomial distribution). It follows that the variance of h will be approximately

$$\sigma_h^2 \approx h'(\mu)^2\sigma_i^2$$

$$\approx \left[\frac{-27M}{(27\mu+M)^2}\right]^2 \frac{2M}{9}$$

$$\approx \frac{k}{M},$$

where k is a constant. In conclusion, the variance of h changes inversely with M, and therefore, the standard deviation of h changes inversely with $\sqrt{M}$.

APPENDIX B

To determine the probability of reinforcement following a particular response pattern in Conc. VI schedules, we assume the following conditions. The bird responds at an overall rate of λ pecks per second (more specifically, according to a Poisson process with parameter λ). A proportion p of these pecks is directed at the left key, a proportion q at the right key, and a proportion r at the center key, with $p + q + r = 1$. The three schedules are independent and set up a reward with rates γ_L, γ_R, and γ_C. By conditioning on the appropriate exponentially-distributed interresponse times, we obtain the following probabilities of reinforcement:

$$P(S^+ \mid L) = 1 - \frac{\lambda p}{\gamma_L + \lambda},$$

$$P(S^+ \mid LL) = 1 - \frac{\lambda}{\gamma_L + \lambda},$$

$$P(S^+ \mid LR) = 1 - \left(\frac{\lambda q}{\gamma_R + \lambda q}\right)\left(\frac{\lambda}{\gamma_R + \lambda}\right),$$

$$P(S^+ \mid LLR) = 1 - \left(\frac{\lambda q}{\gamma_R + \lambda q}\right)\left(\frac{\lambda}{\gamma_R + \lambda}\right)^2,$$

and so on. Consider the third case, the probability of reinforcement following the left–right response pattern, $P(S^+ \mid LR)$. This probability depends on the interval from the preceding R response to the current R. We divide this interval into two parts, that between the last L and the last R responses, call it τ_1, and that between the last L and the *preceding* R response, call it τ_2:

$$\underbrace{R \to \text{non-}R \text{ responses} \to L}_{\tau_2} \underbrace{\to R.}_{\tau_1}$$
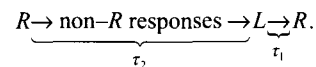
The interval τ_1 comes from an exponential distribution with parameter λ; the interval τ_2 comes from an exponential distribution with parameter λq. Hence,

$P(S^+ \mid LR)$

$$= \int_0^\infty \int_0^\infty \left(1 - \exp^{-\gamma_R(\tau_1 + \tau_2)}\right)\left(\lambda \exp^{-\lambda\tau_1}\right)\left(\lambda q \exp^{-\lambda q\tau_2}\right) d\tau_1 d\tau_2$$

$$= 1 - \left(\frac{\lambda q}{\gamma_R + \lambda q}\right)\left(\frac{\lambda}{\gamma_R + \lambda}\right).$$

Another way to derive the preceding result is to consider that the LR pattern will *not* be reinforced if a reward was not set up during τ_2 and it was not set up during τ_1. The first event occurs with probability

$$\frac{\lambda q}{\gamma_R + \lambda q},$$

because this is the probability that, of two Poisson processes, one with parameter λq (i.e., R pecks), the other with parameter γ_R (i.e., reinforcers on R key), the former occurs before than the latter. For similar reasons, the probability that no reward is set up during interval τ_1 equals

$$\frac{\lambda}{\gamma_R + \lambda}.$$