Polymorphic response patterns under frequency-dependent selection

ARMANDO MACHADO
Duke University, Durham, North Carolina

In a discrete-trials procedure, a frequency-dependent schedule shaped left-right choice proportion toward various equilibrium values between 0 and 1. At issue was (1) whether pigeons match when the overall reinforcement probabilities for two responses depend inversely on their recent frequency, and (2) how pigeons meet the schedule constraint in terms of local responding. That is, do they respond quasi-randomly (Bernoulli mode), or do they learn the stable pattern of the schedule (stable-pattern mode)? Molar choice behavior always tracked the equilibrium solution of the schedule, but the molecular response patterns varied substantially. Markov chains applied to the data revealed that responding was generally intermediate between the memoryless Bernoulli mode, and the perfect memory stable-pattern mode. The polymorphism of molecular patterns, despite molar regularities in behavior, suggests that (1) in order to engender the Bernoulli or stable-pattern modes, the reinforcement rule must strongly discourage competing response patterns (e.g., perseveration), and (2) under frequency-dependent schedules, molar matching is apparently not the outcome of momentary maximizing.

Laboratory studies have shown that frequency-dependent selection, the differential reinforcement of infrequent responses, sometimes engenders high degrees of behavioral variability. For example, when Page and Neuringer (1985) rewarded pigeons for generating sequences of eight left or right pecks that differed from the sequences generated on the last 50 trials, they obtained highly variable, random-like behavior. That is, the birds not only pecked each key equally often, but also showed few or no sequential dependencies between consecutive responses (see also Hest, Haaren, & Van De Poll, 1989; Morgan & Neuringer, 1990; Machado, 1989, 1992, in press; Morris, 1987, 1989; Neuringer, 1991, 1992). However, frequency-dependent selection may also engender stereotypical behavior. In a previous study (Machado, 1992), a frequency-dependent schedule differentially reinforced the momentarily least likely of two responses—the more a pigeon preferred the right key, the less food it received from that key and the more food it received from the left key. Similar contingencies operated when preference shifted to the left; the probability of food for left (L) and right (R) pecks was equal only at indifference. Although the pigeons pecked each key on 50% of the trials, they did so by alternating rather than responding randomly.

The preceding results illustrate two points. First, pigeons can track the overall probabilities of reinforcement for each of two responses when these probabilities depend inversely on the subject’s recent choices. Good tracking is evidenced by an average choice proportion close to indifference, because in these studies, the payoff probabilities were equal for both responses only at indifference. In other words, at indifference, the response ratio matched the reinforcement ratio.

Second, in frequency-dependent schedules, pigeons match by responding locally in one of two ways: either they respond quasi-randomly, as if tossing an unbiased coin (e.g., Machado, 1989, 1992; Page & Neuringer, 1985) or they learn a simple pattern (e.g., alternation) that satisfies the schedule constraint (Machado, 1992, in press). These two response modes are predicted by different theories of choice. Thus, the random-like mode is predicted by stochastic models of choice (e.g., Bush & Mosteller, 1955) that assume that at the beginning of each trial, pigeons toss an imaginary coin whose bias, under the action of rewards and nonrewards, has converged to, and stabilized at, the matching equilibrium of the schedule. I refer to this response mode as the Bernoulli mode. Note that when behavior is stable, the Bernoulli mode predicts few or no sequential dependencies between consecutive responses.

The second response mode—perform a simple pattern such as alternation that meets the schedule constraint—is predicted by the decision rule called momentary maximizing (Shimp, 1966). By sampling from a pool of response patterns, pigeons learn the momentary changes in reinforcement probability brought about by each of these patterns and therefore select the alternative with the highest probability of payoff on the next trial. According to momentary maximizing, then, in frequency-dependent schedules, behavior should converge to a limit cycle. For example, when the matching equilibrium is at indifferent-
ence, behavior should converge to the sequence RLRL..., which automatically satisfies the schedule constraint and is therefore stable. Equivalent formulations hypothesize that the schedule shapes keypecking toward the alternation pattern RRLRL... or that switching is differentially reinforced (Machado, 1992). Another example: When the matching equilibrium requires a proportion of .25 for R pecks, the pattern RLLL... predicted by momentary maximizing is the simplest pattern that automatically satisfies the schedule constraint; when repeated continuously, no other pattern does better than RLLL... Hence, by analogy with the biological concept of evolutionarily stable strategy (e.g., Maynard Smith, 1982), I refer to this response mode as the stable-pattern mode. Note that in contrast to the Bernoulli mode, the stable-pattern mode predicts specific sequential dependencies between consecutive responses.

What determines the local response mode in frequency-dependent schedules? One possible factor is the number of previous responses that must be remembered in order to perform the stable pattern reliably. Thus, while alternation requires memory of only the last response, double alternation (i.e., RRLR...) requires memory of the last two responses, and the pattern RLLL... requires memory of the last three. When this number is large, behavior tends to approach the Bernoulli mode as in Neuringer's experiments. In fact, when the experimenter reinforces a sequence of eight pecks, provided this sequence did not occur during the last 50 trials (Page & Neuringer, 1985), the stable pattern requires cycling through 51 of the 256 possible eight-peck sequences, a feat far beyond the memory capacities of pigeons. Conversely, when the stable pattern requires memory of only a few responses, behavior approaches the stable-pattern mode (Machado, 1992, in press). I therefore hypothesize that as the stable pattern of the frequency-dependent schedule becomes longer, pigeons are more likely to switch from the stable-pattern to the Bernoulli mode.

The present investigation addresses the two foregoing issues, the molar tracking of overall reinforcement probabilities and the molecular patterning of responses, in a new frequency-dependent schedule. The payoff functions that relate the momentary probability of each response to its probability of reward do not always intersect at indifference. Therefore, matching sometimes requires a partial preference (e.g., a proportion of right pecks equal to .25). Do pigeons match under asymmetric, frequency-dependent selection? If so, how good is the matching relation? How do they respond locally (Bernoulli vs. stable-pattern mode)? Following the hypothesis, I predict that the matching equilibrium moves away from indifference, the Bernoulli mode will occur more frequently because the stable pattern of the schedule becomes longer, more asymmetric, and presumably harder to learn. For example, when the matching proportion for right pecks goes from .5 to .25, the stable pattern goes from alternation, RLRL..., to the pattern RLLLRLLLL... Hence, the Bernoulli mode will be more likely in the latter than in the former condition.

Suppose now that in a schedule with stable pattern RLLL..., a pigeon can remember only its last response but follows the rule "If the last peck was on the right, then peck left; if the last peck was on the left, then with probability 1/2 peck right, and with probability 1/2 peck left" (e.g., Heyman, 1979). On average, this bird will emit one right peck followed by three left pecks, the stable pattern of the schedule. How, then, can the experimenter tell the difference between this bird, who does not fit the Bernoulli stable-pattern framework, and another bird who remembers its last three responses but performs the stable pattern imperfectly? More generally, from a stream of pecks, how can the experimenter determine the local mode of responding, be it Bernoulli, stable pattern, or otherwise? As described below, I address this issue through the theory of Markov chains, which includes as special cases, but is not reduced to, the Bernoulli and stable-pattern modes.

METHOD

Subjects
Six pigeons (Columba livia) were maintained at 80% (±15 g) of their free-feeding body weights. Water and grit were continuously available in their home cages. All birds had participated in a short experiment (Machado, in press) similar to the first condition of the present study (see below); they had no other experimental history.

Apparatus
The experimental chamber, enclosed in an outer box, was 32 cm along the sides and 45 cm high. The floor was wire mesh and the side walls and the ceiling were Plexiglas. A ventilating fan attached to the outer box provided air circulation and masking noise.

The front, aluminum wall of the chamber was equipped with two response keys, 23 cm above the floor, 2.2 cm in diameter, and 5 cm apart, center to center. Each key could be illuminated from behind with a red light. Directly below the keys, a hopper opening (4.5 x 7 cm and 7 cm from the floor) permitted access to mixed grain. A white light illuminated the hopper from above when grain was delivered. The houselight on the left wall of the chamber was always on except when food was delivered. All events were controlled and data were recorded by an IBM PC.

Procedure
After one session of autoshaping to reinstate keypecking, the experiment began. Each session consisted of 500 trials. At the beginning of each trial, both keylights were turned on. A peck on either key turned both keys off for either a 1-sec intertrial interval (ITI) or a 3-sec food delivery. Pecks during the ITI reset the timer for the interval. After the ITI or the food delivery, a new trial began immediately.

The reinforcement schedule was always frequency dependent: As the tendency to peck one key increased, the probability of reinforcement for pecks on that key decreased, while the probability of reinforcement for pecks on the other key increased. The point at which the payoff probabilities for left and right pecks were equal defined the matching equilibrium of the schedule. This equilibrium value changed across the six conditions of the experiment.

During each condition, one of the response keys was defined as the target key, and the proportion of pecks on the target key during
the last 40 trials estimated the momentary tendency to peck that key. This estimate determined the reinforcement probabilities on the next trial according to the following equations:

\[
\begin{align*}
\Pr(\text{Reinf}|T) &= \frac{k(a-0.5)s + k(1-a)}{a + 0.5} \\
\Pr(\text{Reinf}|T) &= k(s-a),
\end{align*}
\]

where \(\Pr(\text{Reinf}|T)\) and \(\Pr(\text{Reinf}|N)\) are the payoff probabilities on the target and nontarget keys, respectively, \(s\) is the estimate of the probability of pecking the target key, and \(a\) and \(k\) are two schedule parameters with \(-0.5 \leq a \leq 0\) and \(0 \leq k \leq 1\). When choice proportion \(s\) equals \(a + 0.5\), the two functions intersect, matching between response and reinforcement ratios occurs, and choice proportion is at equilibrium. The constant \(k\) determines the expected probability of reinforcement per trial, which is maximal at equilibrium and equal to \(k/2\).

**Figure 1.** Probabilities of reinforcement on the next trial as a function of the proportion \(s\) of pecks on the target key during the last 40 trials. "Overall" is the expected probability of food on the trial. The point at which the curves intersect \((s = 0.5 + a)\) defines the equilibrium value of the schedule.

**Figure 2.** Expected probability of reinforcement on the L and R keys as a function of the last three responses. Overall proportion of right pecks is assumed at equilibrium, .25 in this case. \(\Pr(\text{Reinf}|R)\) is always less than \(\Pr(\text{Reinf}|L)\), except after LLL. If the subject always picks the alternative with the highest payoff probability, then its behavior will converge to the stable pattern RLLLRLLL. Appendix A shows the derivations.

Figure 1 shows the three instances of Equation 1 used in the experiment. In the top panel, equilibrium occurs at \(s = 0.5\); in the middle at \(s = 0.33\), and in the bottom at \(s = 0.25\). In all cases, \(k = 0.2\). Three aspects of Figure 1 are critical. (1) The curve for the overall probability of reinforcement is almost flat near the equilibrium value. Consequently, the proportion of reinforced trials in a session should always be close to \(s = k/2\). (2) By manipulating \(a\) and/or the target key, the equilibrium proportion for right pecks can take any value between zero and one. For example, when left is the target key and \(a = -0.25\) (bottom panel), equilibrium occurs when the proportion of right pecks, \(\Pr(R)\), equals .75. If the target is the right key, equilibrium would be at \(\Pr(R) = 0.25\). From here on, the equilibrium value of the schedule is always identified in terms of proportion of right pecks. (3) Each experimental condition corresponded to one of the following equilibrium values: \(.25, .33, .5, .67,\) or \(.75\). These values were selected because each is associated with a simple stable pattern: RLLL when \(\Pr(R) = 0.25\), RLL when \(\Pr(R) = 0.33\), RL when \(\Pr(R) = 0.5\), RRL when \(\Pr(R) = 0.67\), and RRRL when \(\Pr(R) = 0.75\).

The stable patterns are the limit cycles predicted by momentary maximizing. An analytical proof of this statement for the case \(\Pr(R) = 0.25\) is derived in Appendix A, but Figure 2 shows the main result. The figure plots the expected payoff probabilities for a left and a right peck after each sequence of three responses. If the subject always chooses the alternative with the highest momentary probability of reinforcement, then its behavior will be attracted to the limit cycle RLLLRLLL..., the stable pattern of the schedule. The result generalizes to other equilibrium values and corresponding stable patterns.

For up to Trial 40 of each session, the estimate of the probability of pecking the target key also depended on the last responses of the previous session. For example, at the beginning of Trial 11, the \(s\) value in Equation 1 was based on the preceding 10 responses and on the last 30 responses of the previous session.

Half of the birds started with the right key, and the other half with the left as the target key. The birds were then exposed to the three conditions depicted in Figure 1 in the top-to-bottom order. Next, the target key changed for each bird, and the three conditions were repeated in the same order. Each condition lasted a minimum of 10
sessions and until the proportion of pecks on the right key appeared stable for 5 consecutive sessions. Table 1 shows the order of experimental conditions and the number of sessions in each condition.

Data analysis. In the analysis of the molar results, standard techniques were used; the molecular analysis was partly based on the theory of Markov chains, which requires a brief description (for more details, see, e.g., Gottman & Roy, 1990).

Suppose the probability of a right peck depends exclusively on the location (L or R) of the last response. This response rule can be modeled by the first-order Markov chain illustrated in Figure 3. When performance is stable, the model is characterized by two parameters—the probability of pecking the right key given that the last response was on the right key, $P\{R|R\}$ (right light arrow), and the probability of pecking the left key given that the last response was on the left key, $P\{L|L\}$ (left light arrow). The complements of $P\{R|R\}$ and $P\{L|L\}$ are the switching probabilities, $P\{L|R\}$ and $P\{R|L\}$, respectively (dark arrows).

In order to test the hypothesis that the data were generated by a first-order chain, the experimenter compares via a chi-square test the observed proportion of response triples against the values predicted by the hypothesis. The predicted values are calculated as follows: Let Obs and Exp stand for observed and expected, respectively. Then,

$$P\{RLL\}^{\text{Exp}} = P\{RL\}^{\text{Obs}} \times P\{L|RL\}^{\text{Obs}}$$

$$= P\{RL\}^{\text{Obs}} \times P\{L|L\}^{\text{Obs}}$$

where the second equality follows because, according to the hypothesis, only the last response affects the next one. Note that the probabilities of the response pairs RL, LL, and LR, each estimated from the data, are used to predict the probabilities of a response triple (the example generalizes easily to the remaining seven triples).

If the first-order model is rejected, a similar test for a second-order model is then conducted, but if the first-order model is not rejected, a zero-order model is tested next. In summary, the experimenter attempts to determine the smallest chain that fits the data well according to the chi-square criterion; the order of the chain is an index of the extent of sequential dependencies present in the data, or, colloquially, of how many responses the bird remembers at the beginning of each trial.

The Bernoulli and stable-pattern response modes also fit the Markovian framework. The former is a zero-order Markov chain, whereas the latter can be modeled by Markov chains of the first, second, or third order, according to the number of responses that need to be remembered in order to perform the stable pattern. The former describes the stable pattern when equilibrium is .33 or .67 (stable patterns LLR and RRL, respectively), and the latter describes the stable pattern when equilibrium is .25 or .75 (stable patterns LLLR and RRRR, respectively). For example, when the stable pattern RRL is performed reliably, the dark arrows of the second-order chain are close to 1.

RESULTS

Molar Analysis

Figure 4 shows the proportion of pecks on the right key and the proportion of reinforced trials on each session of the experiment. The horizontal lines represent the equilibrium value for right pecks during each condition.

When performance stabilized, the proportion of right pecks was always close to the equilibrium value of the
schedule. The few deviations from equilibrium can be classified in three categories: (1) a preference for one key when the schedule equilibrium was at .5 (asterisks in Figure 4); (2) a choice proportion closer to indifference than the equilibrium value ("x" signs in figure); and (3) a choice proportion more extreme than the equilibrium value ("+" sign in figure). Although the deviations shown by Birds 371 and 374 in Conditions 4 and 5 were probably due to the initial differential reinforcement of left pecks, for the remaining birds, the deviations were not clearly related to previous conditions.

Figure 4 also shows that, as predicted, reinforcement frequency was roughly constant across all sessions. A between-within analysis of variance (ANOVA) showed that the average number of reinforced trials did not differ significantly ($\alpha = .01$) across conditions (within-subjects variable, $F(5,20) = .53$), or between the two groups of birds (left or right initial target key as the between-subjects variable; $F(1,4) = .20$). It follows that the gradient of reinforcement across responses rather than the overall frequency of reinforcement determined the steady-state choice proportions.

In order to see how observed choice proportions related to obtained reinforcement proportions, Figure 5 plots the log transform of the generalized matching equation (Baum, 1974; Herrnstein, 1970; Staddon, 1969). That is,

$$\log \left[ \frac{N_R}{N_L} \right] = a \log \left[ \frac{R_R}{R_L} \right] + \log \left[ b \right].$$

where $N_R/N_L$ is the ratio of right to left response frequencies, $R_R/R_L$ is the ratio of right to left reinforcement frequencies, and $a$ and $b$ are the bias and sensitivity param-
Molecular Analysis

The preceding analyses show that pigeons track well overall reinforcement probability in asymmetric, frequency-dependent schedules. This finding raises the second set of issues of the study: How did the birds respond locally? Did they flip an imaginary coin, or did they learn the stable pattern? Were there other response modes? Did the Bernoulli mode become more frequent as the stable pattern became longer? To answer these questions, I determined the smallest Markov chain that fitted the data well and then interpreted its parameters.

The responses emitted by each bird during the last four sessions of each condition were combined in a single string of 2,000 pecks. All subsequent analyses were based on the frequency of each single, pair, triple, and, in general, n-tuple of responses in that string. The frequency of each n-tuple was computed by using a moving window, such that in the substring RLRRLR, for example, the first two pairs are RL and LR, and the first two triples are RLR and LRR.

The first analysis determined the order of the Markov chain that fitted the data of each condition. If the birds responded randomly, the Markov chains would be of order zero; if the birds learned the stable pattern, the order of
the chains would depend on the pattern. Also, if the Bernoulli mode became more likely than the stable-pattern mode as preference became more biased, then zero-order chains should increase in frequency from Conditions 1 to 3 and 4 to 6.

Table 2 displays the results of the chi-square tests for Markov chains of Orders 1-4. To illustrate how the table is read, look at the results of Bird 163 during Condition 1. The first column indicates that the proportion of right and left pecks did not predict well the frequencies of the response pairs; this finding rejects a zero-order chain and suggests a first-order chain. However, a first-order chain was also rejected because, as the second column shows, response triples were not predicted well from pair data; hence, we tentatively accept a second-order chain. The third column indicates that the frequency of 4-tuples was well predicted by the frequency of response triples, and this result confirms the second-order model ("2" in Summary column). If the Markov model is adequate (which, so far, I have assumed rather than proved), the performance of this bird can be described succinctly by the four parameters of the second-order chain: P{R|LL}, P{R|LR}, P{R|RL}, and P{R|RR} (see Figure 3).

The Summary column of Table 2 shows that the data clearly reject the Bernoulli mode (zero-order chain) on all but two occasions. Thus, the Bernoulli mode did not become more likely as preference became more biased. On the other hand, in 21 cases, the order of the chain was either equal to or greater than the minimum required to learn the stable pattern (i.e., 1 for RLL... , 2 for RRL... and LRR... , and 3 for RRRL and LLLL). This means that on 21 occasions, the stable pattern could have been performed reliably because sufficient sequential dependencies were developed; whether the birds did in fact display the stable pattern can be determined from the parameters of the chain fitted to the data.
Table 2
Chi-Square Results for Markov-Chain Models of Orders 1, 2, 3, and 4

<table>
<thead>
<tr>
<th>Bird</th>
<th>Condition</th>
<th>Equil. Value</th>
<th>Stable Pattern</th>
<th>Order of the Markov Chain (df)</th>
<th>Summary (Chain Order)</th>
</tr>
</thead>
<tbody>
<tr>
<td>96</td>
<td>.50</td>
<td>RL</td>
<td>18.72* 4.22</td>
<td>2.23 11.47</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>.67</td>
<td>RRL</td>
<td>91.35* 7.60</td>
<td>6.72 10.53</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>.75</td>
<td>RRRR</td>
<td>120.58* 2.43</td>
<td>2.17 5.61</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>.50</td>
<td>RL</td>
<td>187.39* 0.55</td>
<td>4.54 2.32</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>.33</td>
<td>RLL</td>
<td>50.52* 2.24</td>
<td>1.32 6.63</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>.25</td>
<td>RLLL</td>
<td>17.75* 5.43</td>
<td>6.57 10.11</td>
<td>1</td>
</tr>
<tr>
<td>346</td>
<td>.50</td>
<td>RL</td>
<td>34.92* 3.83</td>
<td>1.96 10.15</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>.33</td>
<td>RLL</td>
<td>12.71* 11.74</td>
<td>3.50 10.55</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>.75</td>
<td>RRRL</td>
<td>120.58* 2.43</td>
<td>2.17 5.61</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>.50</td>
<td>RL</td>
<td>187.39* 0.55</td>
<td>4.54 2.32</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>.33</td>
<td>RLL</td>
<td>50.52* 2.24</td>
<td>1.32 6.63</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>.25</td>
<td>RLLL</td>
<td>17.75* 5.43</td>
<td>6.57 10.11</td>
<td>1</td>
</tr>
<tr>
<td>172</td>
<td>.50</td>
<td>RL</td>
<td>0.34 33.56*</td>
<td>6.73 6.85</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>.67</td>
<td>RRL</td>
<td>25.04* 7.71</td>
<td>1.50 5.18</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>.75</td>
<td>RRRL</td>
<td>42.56* 1.56</td>
<td>9.69 6.46</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>.50</td>
<td>RL</td>
<td>6.68* 12.23*</td>
<td>5.11 11.1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>.33</td>
<td>RLL</td>
<td>14.45* 0.11</td>
<td>5.40 4.53</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>.25</td>
<td>RLLL</td>
<td>64.37* 9.31*</td>
<td>1.39 8.27</td>
<td>2</td>
</tr>
<tr>
<td>163</td>
<td>.50</td>
<td>RL</td>
<td>64.24* 15.27*</td>
<td>1.33 4.72</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>.67</td>
<td>RRL</td>
<td>94.03* 13.79*</td>
<td>1.23 10.22</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>.75</td>
<td>RRRL</td>
<td>81.78* 7.57</td>
<td>14.85* 11.25</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>.50</td>
<td>RL</td>
<td>65.68* 20.87*</td>
<td>5.58 3.98</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>.33</td>
<td>RLL</td>
<td>31.20* 15.83*</td>
<td>4.26 7.12</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>.25</td>
<td>RLLL</td>
<td>8.76* 1.32</td>
<td>9.23 7.33</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>.33</td>
<td>RLL</td>
<td>45.52* 4.50</td>
<td>8.90 20.48*</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>.25</td>
<td>RLLL</td>
<td>7.00* 25.72*</td>
<td>14.32* 10.72</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>.50</td>
<td>RL</td>
<td>1.75 5.89</td>
<td>2.10 7.37</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>.67</td>
<td>RRL</td>
<td>4.44* 4.90</td>
<td>13.55* 14.50</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>.75</td>
<td>RRRL</td>
<td>5.18 33.66*</td>
<td>2.58 18.48</td>
<td>2</td>
</tr>
<tr>
<td>374</td>
<td>.50</td>
<td>RL</td>
<td>38.18* 36.45*</td>
<td>9.58 2.56</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>.33</td>
<td>RLL</td>
<td>55.32* 20.04*</td>
<td>3.05 13.99</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>.25</td>
<td>RLLL</td>
<td>67.21* 36.92*</td>
<td>43.23* 15.72</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>.50</td>
<td>RL</td>
<td>109.14* 59.18*</td>
<td>14.32* 10.72</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>.67</td>
<td>RRR</td>
<td>2.95 44.28*</td>
<td>15.78* 9.79</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>.75</td>
<td>RRRL</td>
<td>26.66* 38.77*</td>
<td>17.06* 12.87</td>
<td>3</td>
</tr>
</tbody>
</table>

Note—For a significance level of .01, the critical values are 6.63, 9.21, 13.28, and 20.09, respectively. Equil. Value = equilibrium value. *p < .01

Table 2 gives the complete set of transition probabilities, from zero to third order. Column "Unc" corresponds to the unconditional probability of a right peck; the remaining columns are the probabilities of a right peck, following the pattern at the top of the column. The most important values, printed in bold, are the parameters of the chain fitted to the data. Given that a fourth-order chain was suggested only twice—Bird 371, Condition 2, and Bird 374, Condition 4—chains of Orders 1 and 3, respectively, were used in these two instances (these values were suggested by the chi-square results shown in Table 2).

When equilibrium was at .5, four response modes were observed: (1) random-like behavior—the zero-order chain for Bird 346, Condition 4, with P[R] = .5; (2) alternation, the stable pattern—first-order chain for Bird 96, Conditions 1 and 4, in which P[R|R] < P[R] and P[R|L] > P[R]; (3) double alternation—second-order chain for Birds 172 and 374, Condition 1, in which P[R|R] is much less, and P[R|LL] is much greater than P[R]; and (4) response perseveration—second-order chain of Bird 163, in which P[R|LL] < P[R|LR] < P[R|RR]; that is, each R increases the tendency to peck R again.

When equilibrium was at .33 or .67, Birds 96 and 172 developed first-order dependencies, switching and perseveration, respectively, which are insufficient to perform the stable pattern. The remaining birds displayed second-or third-order dependencies, but the molecular patterns varied. For example, when the stable pattern was LLR, Bird 163 showed response perseveration—the probability of a right peck increased with each additional right peck; Bird 346 showed a decreased tendency to peck R after LR—P[R|LR] = .29 < P[R] = .38—as the stable pattern predicts, but a second R in a row increased this tendency slightly—P[R|RR] = .41. Bird 346 showed a third pattern, namely, the first R increased the tendency to peck R from .32 = P[R] to .52 = P[R|LR], but another R reset this probability to its baseline value—P[R|RR] = P[R]. This transient perseveration effect also occurred when equilibrium was .25.

The pattern of transition probabilities of the remaining third-order chains shows that, in general, performance
of these sequential dependencies was not consistent with the stable-pattern mode. The probability of playing the stable pattern, however, this conclusion presupposes that the Markov chains are appropriate models of the data. To check whether this is the case, we have to determine how well the chains describe properties of the data not directly used to estimate their parameters. For this purpose, I analyzed the run structure of the data, a statistic frequently used in choice studies (e.g., Nevin, 1969). Another reason to study the run structure is that it sheds further light on the molecular structures of responding, and on how they relate to the Bernoulli and stable-pattern modes.

### Distribution of Response Runs

Figure 7 shows a representative sample of the distribution of response runs on the majority and minority keys (filled and open circles, respectively). The first data point of each graph corresponds to the relative frequency of runs of Length 1, the second data point to runs of Length 2, and the last point to runs of Length 10 and above. To illustrate, if left is the majority key, then runs of Lengths 1, 2, and 3 are observed.

### Table 3

<table>
<thead>
<tr>
<th>Order of the Markov Chain</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bird Condition</td>
<td>Equil. Value</td>
<td>Unc</td>
<td>L</td>
<td>R</td>
</tr>
<tr>
<td>96</td>
<td>.50</td>
<td>.50</td>
<td>.41</td>
<td>.50</td>
</tr>
<tr>
<td>50</td>
<td>.50</td>
<td>.50</td>
<td>.41</td>
<td>.50</td>
</tr>
<tr>
<td>172</td>
<td>.50</td>
<td>.50</td>
<td>.41</td>
<td>.50</td>
</tr>
<tr>
<td>5</td>
<td>.50</td>
<td>.50</td>
<td>.41</td>
<td>.50</td>
</tr>
<tr>
<td>6</td>
<td>.50</td>
<td>.50</td>
<td>.41</td>
<td>.50</td>
</tr>
<tr>
<td>37</td>
<td>.50</td>
<td>.50</td>
<td>.41</td>
<td>.50</td>
</tr>
<tr>
<td>374</td>
<td>.50</td>
<td>.50</td>
<td>.41</td>
<td>.50</td>
</tr>
</tbody>
</table>

Note—Unc = unconditional probability of a right peck. Equil. Value = equilibrium value. Values printed in bold are the parameters of the Markov chain fitted to the data. Data are based on the last four sessions of each condition.
1, 2, and 3 correspond to the sequences RLR, RLLR, and RLLLR, respectively. In order to obtain relative frequencies, the absolute frequency of each run was divided by the total number of runs for that key. When equilibrium equals .5, the majority/minority classification is arbitrary. The distributions of run length show that as choice proportion became more extreme, runs of Length 1 on the minority key increased in frequency and the distribution of runs on the majority key shifted to the right, although the shape of that distribution differed across birds. Some birds (e.g., Bird 96) showed a geometric-like distribution with the mode always at Length 1, whereas other birds (e.g., Bird 374) showed in some conditions either a bimodal distribution or a single-mode distribution, but with the mode at runs of Lengths 2 or 3.

These distributions differ from the Bernoulli mode predictions because, in general, the proportion of the two runs of Length 1, RLR plus LRL, do not add up to one.
For example, Bird 96 alternated too much (RLR + LRL \( \neq 1 \)), while Birds 163 and 374 showed various degrees of response perseveration (RLR + LRL \( \neq 1 \)). The distributions also differ from the stable-pattern mode because, in general, there are too many short runs on the majority key and/or too few runs of Length 1 on the minority key. Moreover, with few exceptions, when equilibrium differed from .5, the distribution of runs on the majority key did not peak at the value predicted by the stable-pattern mode—that is, 2 when \( P(R) \) equaled .33 or .67 and 3 when \( P(R) \) equaled .25 or .75. The exceptions occurred during the last condition, performance on both keys was close to the stable-pattern mode, that is, a high relative frequency of runs of Length 1 on the minority key and a mode at Length 3 on the majority key, although runs of Length 2 on that key were also frequent.

The solid and dotted lines in Figure 7 are the predictions of the Markov chains. For each condition, the order of the chain was suggested by Table 2 and is indicated by the number at the top. The Markov chains fitted the run data very well. Table 4 shows chi-square statistics that measure the goodness-of-fit of the model for the data of Figure 7. (Appendix B derives the distribution of runs from a Markov chain and determines the degrees of freedom of the chi-square test.) For a significance level of .01, only one test was significant. The results for the remaining subjects were identical—of the 67 tests performed, 61 (91%) were not significant. I conclude that the Markov-chains model provided a good description of the data.

Despite similar performance at the molar level, the analyses done so far reveal individual differences in the molecular structure of responding. These differences, together with the fact that the Markov-chains model describes the data reasonably well, indicate that the response mechanism was generally intermediate between the memoryless Bernoulli mode and the perfect memory stable-pattern mode. However, one could still argue that the birds did learn the stable pattern but failed to display it on every trial because of memory errors (Silberberg et al., 1978). One way to test this stable-pattern-with-errors hypothesis is through a lag analysis (Gottman & Roy, 1990; Machado, 1992).

**Lag Analysis.** Let \( X_n \) be the response on trial \( n \); then, \( P(X_{n+k} = R|X_n = R) \) is the probability of pecking the right key \( k \) trials later given a right peck on the current trial, and, similarly, \( P(X_{n+k} = L|X_n = L) \) is the probability of pecking the left key \( k \) trials later, given a left peck on the current trial. A plot of \( P(X_{n+k} = R|X_n = R) \) and \( P(X_{n+k} = L|X_n = L) \) against \( k \), the lag value, yields two profiles of conditional probabilities (these profiles are linear transforms of the autocorrelation function).

Figure 8 shows the profiles predicted by various response modes. The first datapoint of each profile is the overall proportion of right pecks, filled circles, or the proportion of left pecks, open circles; the next datapoints represent the successive values of \( P(X_{n+k} = R|X_n = R) \) (filled circles) and \( P(X_{n+k} = L|X_n = L) \) (open circles). As the top left panel illustrates, the memoryless Bernoulli mode predicts flat lag profiles, that is, no sequential dependencies. The top right panels illustrate the predictions of a first-order Markov chain (Appendix C shows how to derive a lag profile from a Markov chain). Only two profiles are possible—either a zigzag profile indicating a tendency to alternate, or a geometrically decaying profile indicating response perseveration.

The bottom panels show perfect performance of the stable pattern in three different conditions. Take, for example, the RRL pattern (equilibrium at .67). When repeated continuously (i.e., RRLRL...), the following holds: (1) The overall proportion of right pecks, \( P(X_n = R) \), equals \( \frac{2}{3} \); (2) \( P(X_{n+1} = R|X_n = R) = \frac{1}{2} \) because after a right peck, another right peck follows only on half of the trials; (3) \( P(X_{n+2} = R|X_n = R) = \frac{1}{4} \) because only half of the time does a right peck follow another right peck two trials later; (4) \( P(X_{n+3} = R|X_n = R) = 1 \) because after a right peck, another right peck always follows three

<table>
<thead>
<tr>
<th>Condition</th>
<th>Key</th>
<th>Chain</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>Chain</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>Chain</th>
<th>( \chi^2 )</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L</td>
<td>R</td>
<td>7.59</td>
<td>7</td>
<td>8.22</td>
<td>6</td>
<td>2</td>
<td>4.38</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>L</td>
<td>R</td>
<td>3.79</td>
<td>2</td>
<td>5.63</td>
<td>5</td>
<td>2</td>
<td>5.93</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>L</td>
<td>R</td>
<td>9.05</td>
<td>7</td>
<td>4.96</td>
<td>6</td>
<td>2</td>
<td>3.20</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>L</td>
<td>R</td>
<td>13.65</td>
<td>7</td>
<td>2.00</td>
<td>2</td>
<td>3</td>
<td>14.34</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>L</td>
<td>R</td>
<td>1.15</td>
<td>4</td>
<td>5.76</td>
<td>6</td>
<td>3</td>
<td>6.99</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>L</td>
<td>R</td>
<td>5.06</td>
<td>3</td>
<td>7.83</td>
<td>6</td>
<td>3</td>
<td>2.51</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.64</td>
<td>7</td>
<td>6.67</td>
<td>5</td>
<td>3</td>
<td>0.54</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.63</td>
<td>2</td>
<td>6.52</td>
<td>5</td>
<td>3</td>
<td>7.52</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Note—Chain = order of the Markov chain. A blank space indicates that not enough degrees of freedom (df) existed to perform the test. \( *p < .01 \).
The stable-pattern mode makes strong predictions concerning the shape of the lag profile, a shape that remains similar in form even when the stable pattern is performed with errors. To see this, assume, as Silberberg et al. (1978) have argued, that the bird either remembers where it is in the stable pattern and responds correctly, or forgets where it is and therefore "guesses." For example, when the stable pattern is RRL after pecking RR, the bird always pecks L if it remembers its position in the sequence; if the bird forgets, then it pecks L with probability \( \frac{1}{2} \). The lag profile predicted by the stable-pattern-with-errors mode is therefore a weighted average of the "pure" stable-pattern profile illustrated in the bottom panels of Figure 8 (0% errors) and the flat profile of the Bernoulli process (100% errors), with the weight corresponding to the probability of forgetting. What is critical, however, is that the shape of the profile remains the same as the horizontal line is approached.

Figure 9 presents the lag profiles generated by the birds in the course of the study (circles) and the predictions of the corresponding Markov chains (lines). The numbers at the top are the order of the chain suggested by Table 2 for each condition.

As expected from the molar analysis, the first datapoint of each profile is close to the equilibrium value of the schedule. However, a comparison of Figures 8 and 9 readily shows that, with few exceptions, this molar adapta-
tion was not achieved by "flipping a (sometimes biased) coin" or "displaying the stable pattern." Also, in most conditions, the lag profile was clearly not intermediate between the stable-pattern and the flat Bernoulli profiles. In several instances (Bird 172, Conditions 2-5; Bird 163, Conditions 1-5; Bird 374, Conditions 2-4), the profiles show the second datapoint clearly above the first, a perseveration tendency not anticipated by any hypothesis discussed in this study.

The few exceptions mentioned above are the profiles of Birds 371 and 346 (Condition 1) and 96 (Conditions 1-4), in which alternation is very strong, and the last condition of Bird 374, in which the profile for the minority key, open circles, shows the second, third, and fourth points below, and the fifth point above the first one, as the stable-pattern mode predicts (cf. bottom right profile in Figure 8).

Another noteworthy finding is the consistency of the profiles of some of the subjects across consecutive conditions (e.g., Bird 96, all conditions; Bird 163, first five conditions; Bird 374, four middle conditions). For these birds, despite substantial changes in overall choice proportion, the pattern of local responding persisted.

Figure 9 also shows that the Markov chains fitted the lag data reasonably well. A quantitative assessment of the goodness-of-fit was not attempted, however, for the following reason: A chain of order $n$ fits perfectly the first $n+1$ datapoints of each profile, and, once these points are fixed, the remaining points are automatically constrained to specific ranges of values. For example, if the

![Figure 9](image-url)
overall proportion of right pecks is fixed at \( \frac{1}{2} \), then \( P\{L|L\} \), the second point in the lag profile for left responses, can take any value in the interval \((0,1)\) [1 is excluded]), but \( P\{R|R\} \), the second datapoint in the profile for right pecks, cannot be less than \( \frac{1}{2} \). Similarly, once \( P\{R|R\} \) and \( P\{L|L\} \) are fixed, the third point of each profile is immediately constrained. For this reason, the fit of the model in Figure 9 must be interpreted cautiously.

What variables might account for the most salient finding of this study, namely, the polymorphism in the molecular response patterns despite the molar matching equilibrium? Overall reinforcement frequency remained constant (Figure 4), a result that rules out this variable as an explanation. Also, because all the birds approximated the equilibrium value of the schedule, differences in response patterning were not mere reflections of differences in the absolute frequencies of left and right pecks. Another possibility is the previous history of the subjects. In fact, 4 birds (Birds 371, 172, 163, and 374) maintained the response pattern developed in the previous study (see Machado, in press, for details) during Condition 1. However, the 2 remaining birds developed a switching pattern not present before and, more importantly, the patterns of Condition 1 were neither clearly related to performance in subsequent conditions nor particularly stable. For example, during Condition 4, Birds 346, 172, 371, and 374 did not display the same pattern of Condition 1. Although experimental history cannot be ruled out as the origin of the individual differences, it does not explain why these differences were maintained in the present study. The following analysis provides a possible answer.

Produced and Reinforced Patterns

Figure 10 displays the data from two “extreme” subjects, Birds 96 and 374, during some experimental conditions. The behavior of Bird 96 was always well described by a first-order chain; the behavior of Bird 374 was more complex and required chains of Orders 2, 3, and even 4.

**Patterns**

Figure 10. Comparison between the proportion of each response pattern (filled circles) and the proportion of the reinforcers that followed that pattern (open circles). When an open and a filled circle overlap, matching occurs. Each panel corresponds to one experimental condition identified by the equilibrium (Eq) for R pecks. 1-tuples are the L and R single responses, with L corresponding with the leftmost datapoint; 2-tuples are the 4 pairs, from left to right, LL, LR, RL, and RR. The proportions for 1- and 2-tuples are read in the left-hand y-axis. The 3- and 4-tuples are the 8 triples (LLL, LLR, ..., RRL, RRR) and the 16 quadruples (LLLL, ..., RRRR), respectively. The values for 3- and 4-tuples are read on the right-hand y-axis. Data are from Birds 96 and 374.
Each panel corresponds to one experimental condition, which is identified by the equilibrium value for right pecks. In each panel, the two leftmost filled circles are the proportions of the two single responses (L and R; 1-tuples); the next four filled circles are the proportions of the four pairs LL, LR, RL, and RR (2-tuples); the next eight filled circles are the proportions of the eight triples, LLL, LLR, LRL, ..., and RRR (3-tuples), and the 16 rightmost filled circles are the 4-tuples LLLL, ..., and RRRR. The open circles represent the proportion of the reinforcers that followed the respective pattern. The values for single responses and pairs of responses are read in the left-hand y-axis; the values for triples and 4-tuples are read in the right-hand y-axis.

When a filled and an open circle overlap, the corresponding response pattern is in matching equilibrium; that is, the proportion of the pattern matches the proportion of reinforcers received from the pattern.

In most cases, reinforcers matched response patterns irrespective of the frequency distribution of these patterns. For example, compare the two top panels, each with equilibrium at .5. Although the distribution of 2-, 3-, and 4-tuples differed, matching was reasonably good in both instances. The same holds for Bird 374, whose data are displayed in the middle panels. Again, the distribution of response patterns differs widely, but visual inspection suggests that matching was of similar magnitude. The two bottom panels confirm this finding in a condition in which equilibrium was not at .5. The local response mode differed markedly (cf. Conditions 3 vs. 6 for Bird 374 in both Figures 7 and 9), but matching was still obtained.

By ensuring matching at all subsequent levels of responding whenever a response mode yielded matching at the level of single responses, the schedule “tolerated” various idiosyncratic response organizations. I return to this issue in the Discussion section.

**DISCUSSION**

The molar analyses showed that pigeons can track the overall probabilities of reinforcement associated with two responses when these probabilities are frequency dependent and equal at choice proportions other than indifference. Averages both over sessions (Figure 4) and blocks of trials (Figure 6) showed that the proportion of right responses always converged to the equilibrium solution of the schedule; once at equilibrium, response ratios matched reinforcement ratios (Figure 5). These findings extend previous results (e.g., Machado, 1989, 1992; in press; Page & Neuringer, 1985) that had shown that pigeons track frequency-dependent schedules with equilibrium at indifference.

The second issue, how these molar results were generated, is more difficult to answer. Previous research suggested two local response modes, the Bernoulli and the stable-pattern modes, and I hypothesized that as the stable pattern of the schedule becomes longer, the Bernoulli mode is more likely to replace the stable-pattern mode.

The molecular analyses revealed many sequential dependencies between responses, a result that clearly rejects the Bernoulli mode. Moreover, in the two instances in which performance was random-like, equilibrium was at .5, which means that alternation, the stable pattern, was easy to learn. Thus, the hypothesis that random-like behavior is more likely to occur as the stable pattern becomes longer was not supported.

On the other hand, there was some evidence that pigeons are more likely to learn the stable pattern when it is short. During at least one of the two conditions in which equilibrium was at .5, Birds 96, 346, and 371 alternated reliably; when equilibrium was at other values, only once was the stable pattern displayed reliably (Bird 374, last condition).

In most instances, the Bernoulli stable-pattern framework did not accommodate the molecular patterns generated by the birds. However, these patterns were well described by a more powerful framework, the theory of Markov chains. The chains described well the frequency of various response tuples (singles, pairs, triples, etc.), the run structure of the data, and the serial correlation between responses (lag profiles). The good fit of the model legitimizes interpretations of the data based on the parameters of the chains.

The transition probabilities of the chains confirmed that the regularity at the molar level hid a polymorphic array of molecular structures of responding (e.g., response perseveration, strict alternation, double alternation, and other patterns not easily labeled). These results have some implications for the study of operant variability and choice behavior.

**Operant Variability**

Variability of response sequences is partly a matter of eliminating through reinforcement the sequential dependencies between responses, that is, of preventing subjects from using past responses as cues to subsequent responses. I have suggested that this Bernoulli-like response mode may be achieved through frequency-dependent selection (see also Blough, 1966; Neuringer, 1986; Rapoport & Budescu, 1992). Previous results (Machado, 1992, in press) indicated that, although possibly necessary, frequency-dependent selection is not sufficient because pigeons sometimes learn the stable pattern of the schedule—hence, the Bernoulli stable-pattern framework. What remained unclear after these studies was whether the same framework would fit the data when the reinforcement schedule enforces a preference away from indifference. Specifically, I asked whether performance could be biased toward one key without simultaneously engendering sequential dependencies. Metaphorically speaking, can pigeons behave like a biased coin and thus generate random-like but not highly variable behavior? Alternatively, if birds develop sequential dependencies, do these dependencies also match the stable pattern of the schedule? If performance fitted the Bernoulli stable-pattern framework, then both parsimony and integration would be achieved—parsimony.
because the idea that a single mechanism generates both variable and nonvariable behavior would be strongly supported (Machado, 1992), and integration because variability studies could then be cast in the same framework as the more extensive studies of choice behavior.

The present results do not vindicate the Bernoulli stable-pattern framework and, therefore, the following question remains: What are the critical conditions for generating (1) highly variable behavior, as in Neuringer’s studies, (2) the stable pattern, as in Machado’s earlier studies, or, as in the present study, (3) none of the above in spite of molar regularities? A possible answer is suggested by experiments that show that the magnitude of the difference in reinforcement probability across competing response patterns determines whether subjects learn the momentary maximizing sequence of the schedule (e.g., Williams, 1972, 1991). Then consider how, in the three items listed above, a response pattern changes the momentary payoff probabilities. First, in Neuringer’s schedules, a pattern is reinforced, provided it did not occur during the last N trials. Thus, the probability of reinforcement for a given pattern drops to zero once the pattern occurs, remains at zero until N trials without that pattern elapse, and then equals one until the pattern recurs. A step function describes the reinforcement contingencies for each pattern. Second, in Machado’s earlier studies, the probability of reinforcement for each pattern changed hyperbolically, decreasing when the pattern occurred and increasing when the pattern did not occur. Third, in the present study, the payoff probabilities changed linearly; together with a large window of 40 responses, linearity means that each response produces only a small change in the probability of reinforcement (e.g., in the top panel of Figure 1 each peck changes the payoff probability by .005 only; see also Appendix A).

In summary, the changes in the local probabilities of reinforcement are large in Neuringer’s variability schedules, intermediate in Machado’s hyperbolic schedules, and small in the present linear schedules. Therefore, response modes other than the Bernoulli and the stable pattern are not “tolerated” in Neuringer’s schedule, are somewhat “tolerated” in hyperbolic schedules, and are most “tolerated” in linear schedules.

The high “tolerance” of linear schedules to patterns other than the stable pattern is well illustrated in Figure 10. Although the distribution of response patterns differs markedly across both conditions and subjects, response proportions match reinforcement proportions reasonably well. Consequently, when the molar equilibrium of the schedule is reached, differential reinforcement of the stable pattern is weak or nonexistent. As some authors have stressed (e.g., Crow, 1977; Staddon & Simmelhag, 1971), a release of environmental pressure or, equivalently, an increase in environmental tolerance tends to promote or maintain behavioral diversity. The situation may be analogous to that described by Darwin: “Variations neither useful nor injurious would not be affected by natural selection [reinforcement], and would be left a fluctuating element, as perhaps we see in the species [molecular patterns] called polymorphic” (1859/1987, p. 131).

If the foregoing analysis is correct, it follows that in order to generate the Bernoulli mode, a schedule must define a difficult-to-learn stable pattern and implement contingencies that strongly counteract sequential dependencies. These conditions are met by large variability requirements (e.g., Page & Neuringer, 1985; Pryor, Haag, & O’Reilly, 1969; Machado, 1989, 1992), and/or by large ITIs (e.g., Bryant & Church, 1974; Neuringer, 1991). In order to generate the stable pattern, the schedule must define a stable pattern commensurate with the subject’s memory capacities (e.g., in pigeons, double alternation, but not, e.g., LLLRLRRR), differentially reinforce it, and strongly counteract any other response biases by, for example, increasing the difference in reward probability between the stable pattern and competing patterns (e.g., Fetterman & Stubbs, 1982; Silberberg & Williams, 1974). The present experiment failed to achieve the last condition, which may explain the polymorphism in the molecular patterns of responding.

Choice Behavior

Many investigators believe that molar matching is the outcome of a molecular process, momentary maximizing (Shimp, 1966; Silberberg et al., 1978; Staddon, 1983). Why, then, did most of the birds in the present study fail to learn the equivalent of the momentary maximizing sequence, the stable pattern of the schedule? In what follows, I discuss several possibilities, some of which future studies will have to address.

1. In order to account for the discrepancies between observed and predicted sequential dependencies, momentary maximizing (a deterministic rule) has always been supplemented with ad hoc hypotheses. Hence, could the present findings be accounted for by, say, momentary maximizing with response perseveration, or momentary maximizing with memory errors (Silberberg et al., 1978)? There are two problems with these hypotheses, the first empirical, the second, theoretical. First, none of the hypotheses account for the full range of results because the lag profiles rejected maximizing with errors, and response perseveration occurred in some, but not all of the birds (e.g., in Bird 163, but not in Bird 96). Second, hypotheses that emphasize steady-state behavior should not assume what may turn out to be incompatible with the (skipped) learning process. Consider, for example, momentary maximizing with errors. In order to fit the data, Silberberg et al. assumed that their subjects momentary maximized on 55% of the trials, but responded randomly on the remaining 45% because, on those trials, the birds forgot where they were in the sequence. Notwithstanding the reasonable fit of the model, the question that naturally comes to mind is whether such a high frequency
of forgetting would not undermine learning of the momentary maximizing sequence in the first place (see Machado, in press).

2. In the present study, the order of experimental conditions was from simple to complex stable patterns. Consequently, I cannot rule out the hypothesis that patterns developed during preceding conditions somehow prevented the birds from learning the complex stable patterns of subsequent conditions. The more general issue here is that, as in most operant studies, I assumed reversibility of behavior, that is, negligible effects of history, whereas the present findings indicate that when the schedule is highly "tolerant," molecular patterns may be irreversible. Future studies should elucidate how experimental history interacts with "tolerant" schedules.

3. Another possibility is that the pigeons did not learn the stable patterns because frequency-dependent schedules are intrinsically challenging. To see this, consider the discrete-trials schedules commonly designed to test momentary maximizing (e.g., Hiraoka, 1984; Shimp, 1966; Silberberg & Williams, 1974; Williams, 1972, 1991). In most of these schedules, a salient event such as food or a response occurrence resets the reinforcement contingencies (Staddon, Hinson, & Kram, 1981), whereas in frequency-dependent schedules there is no such event. In fact, the only informative event in frequency-dependent schedules is the number of left and right responses emitted on recent trials. Furthermore, because in frequency-dependent schedules the local probabilities of food change very gradually, pigeons must integrate these changes over extended periods of time. These two differences together with the present results suggest that frequency-dependent schedules may pose a greater challenge to the learning of the momentary maximizing sequence than the schedules used in previous studies.

The good fit of the Markov-chains model (see also Heyman, 1979) also raises a methodological issue related to how momentary maximizing has been tested. In the past, part of the evidence for the rule came from a comparison between the observed frequencies of the stable patterns against the frequencies predicted from the Bernoulli, null hypothesis. Take, for example, the data from Bird 96 when equilibrium was at .67. The observed proportion of the stable pattern (sequences RRL, RLR, and LRR) equaled .55, while the predicted value, based on the unconditional probability of a right peck, equaled .44; the two proportions differ (binomial, p < .01). Was the bird momentary maximizing, then? The Markov model indicates that the bird was "remembering" only the last response, a result that shows that frequencies above those predicted by the null hypothesis do not automatically indicate sequential dependencies isomorphic with the stable pattern.

In conclusion, the challenge raised by the present findings is to discover, in situations in which the informative event—choice proportion in recent trials—is neither as salient as food nor as simple as the time since the last response, what determines the extent and nature of the sequential dependencies. When the data are well described by a Markov chain, the issue boils down to what sets the order of the chain. The answer is likely to come not from static analyses, but from considerations of the learning process and its dynamics (e.g., Machado, in press; Shimp, Childers, & Hightower, 1990). For example, when previous responses cue the next response, we need to understand how this cue function develops; when a reinforcer is delivered in variability and in frequency-dependent schedules, we need to understand what property of behavior receives the credit for the event (Staddon & Zhang, 1991), or, equivalently, what is strengthened by the reinforcer. I suspect that the answer to these dynamic questions will explain why and when polymorphic molecular patterns coexist with molar behavioral regularities. In the meantime, I conclude, as Darwin once did, that "our ignorance of the laws of [behavioral] variation is profound" (Darwin, 1859/1987, p. 202).

REFERENCES

BEHAVIORAL POLYMORPHISMS 69


Some data and a new model. Behavioral Processes.


APPENDIX A

Local Reinforcement Probabilities

This appendix determines the changes in payoff probabilities produced by an arbitrary response pattern. We assume that choice proportion ($s$) equals the equilibrium value of the schedule. To illustrate the computations, assume that R is the target key, the equilibrium is at .25 (see bottom panel of Figure 1), and the subject has just produced the pattern RLL. The payoff probabilities on the next trial will depend on the number of L and R pecks on the preceding 37 trials (remember that choice proportion on the last 40 trials set the payoff probabilities according to Equation 1). But the average number of R pecks on the last 37 trials is $s^*37$ and, therefore, the average number of Rs and Ls on the last 40 trials equals $s^*37 + 1$ and $(1-s)^*37 + 2$, respectively. Replacing these values in Equation 1 yields the average probability of reward for an R and an L peck. The results for the remaining three-response patterns are calculated in similar ways and are shown in Figure 2.

APPENDIX B

Distribution of Response Runs and Markov Chains

This appendix shows how to derive the distribution of response runs from a Markov chain. The example takes a second-order chain and derives the distribution of runs of R pecks. The extension to chains of different order and to runs of L pecks is straightforward.

A run of R pecks of length $k$ is a string of $k$ consecutive right keypecks preceded and followed by a left keypeck (e.g., a run of Length 2 is LRRL).

Let S be the set of states associated with the chain. In a second-order chain, the last two responses define the current state. Hence,

$$S = \{LL, LR, RL, RR\}. \quad (B1)$$

Let $T$ be the transition matrix of the chain. For a second-order chain, we have

$$T = \begin{bmatrix}
q & 1-q & 0 & 0 \\
0 & 0 & 1-r & r \\
s & 1-s & 0 & 0 \\
0 & 0 & 1-p & p
\end{bmatrix}. \quad (B2)$$

where the four rows correspond to the states LL, LR, RL, RR, in this order, and similarly for the four columns. Thus, if the last two responses were LL, then with probability, $r$, the next response will be a right keypeck and the new state of the chain will be RR, that is, $P(\text{state RR given state LL}) = r$. More generally, the number in row $i$ and column $j$ of the matrix $T$, that is, $T_{ij}$, is the probability of going from state $i$ to state $j$ in one trial. Then,

$$P(\text{Run of Length } 1) = P[LR](1-r)$$

$$P(\text{Run of Length 2}) = P[LR]^2(1-p)$$

$$P(\text{Run of Length } n) = P[LR]^{n-1}(1-p) + P[LR]^{n-2} \quad (B3)$$

where the parameters $P[LR]$, $r$, and $p$ are obtained from the data. Parameter estimation uses 3 degrees of freedom; if each of the 10 classes (i.e., runs of Length 1 up to $\geq 10$) has an expected frequency of at least one, the total degrees of freedom of the chi-square test would be $6 = 10 - 3 - 1$. When the expected frequency is less than one, adjacent classes are combined until that frequency is at least one.

APPENDIX C

Lag Profiles and Markov Chains

The lag profile for R pecks corresponds to the set of probabilities $P(X_{n+k} = R | X_n = R)$, $k \geq 1$, where $X_n$ is the response on trial $n$. This appendix shows how to derive from a Markov...
Consider a chain with statespace $S$ and transition matrix $T$. It can be shown (e.g., Karlin & Taylor, 1975) that when the number of trials $n$ is large, the probability of finding the chain in state $j$ by trial $n$, $P\{s_n(j)\}$, is given by the solution of the following linear equations:

$$P\{s_n(j)\} = \sum_i T_{ij} P\{s_n(i)\}$$

subject to:

$$\sum_j P\{s_n(j)\} = 1. \quad (C1)$$

Now, let $V_k(j)$ be the probability of an R peck $k$ trials later, given that the chain is in state $j$ on the current trial, that is,

$$V_k(j) = P\{X_{n+k} = R | s_n(j)\}. \quad (C2)$$

By conditioning on the response on the next trial, $X_{n+1}$, we get

$$V_k(j) = \sum_i T_{ij} V_{k-1}(i)$$

or, in vector notation,

$$V_k = TV_{k-1}. \quad (C3)$$

where $V_0(j) = 1$ if state $j$ ends with an R response, and $V_0(j) = 0$ if state $j$ ends with an L response. By iterating Equation C4, we can determine $V_k(j)$ for any $k$ value.

On the other hand,

$$P\{X_{n+k} = R | X_n = R\} = \sum_{s_k \in S^*} P\{X_{n+k} = R | s_n(j)\} P\{s_n(j)\}$$

where $S^*$ is the subset of states that end with an R peck (e.g., LR and RR in the second-order case). Therefore,

$$P\{X_{n+k} = R | X_n = R\} = \sum_{s_k \in S^*} V_k(j) P\{s_n(j)\}$$

or, in vector notation,

$$P\{X_{n+k} = R | X_n = R\} = \sum_{s_k \in S^*} V_k(j) P\{s_n(j)\}$$

where the $P\{s_n(j)\}$ values are given by the solution of Equation C1.

For a first-order Markov chain, the state space is $S = \{L, R\}$ and the matrix $T$ is

$$T = \begin{pmatrix} q & 1-q \\ 1-p & p \end{pmatrix}, \quad (C7)$$

that is, $q = P(L | L)$ and $p = P(R | R)$. The lag profiles for R and L pecks are then

$$P\{X_{n+k} = R | X_n = R\} = \frac{1-q}{2-p-q} + \frac{1-p}{2-p-q} (p+q-1)^k$$

and

$$P\{X_{n+k} = L | X_n = L\} = \frac{1-p}{2-p-q} + \frac{1-q}{2-p-q} (p+q-1)^k - 1. \quad (C8)$$

As illustrated in the top middle panel of Figure 8, when $p+q-1 < 0$, the lag profiles have a dampened zigzag shape; when $p+q-1 > 0$, the profiles are similar to the top right panel in the figure; when $p+q-1 = 0$, the first-order chain reduces to the Bernoulli process, whose lag profile is shown in the top left panel of Figure 8.