

REINFORCEMENT OF LEAST-FREQUENT SEQUENCES OF CHOICES¹

CHARLES P. SHIMP

STANFORD UNIVERSITY

When a pigeon's choices between two keys are probabilistically reinforced, as in discrete trial probability learning procedures and in concurrent variable-interval schedules, the bird tends to maximize, or to choose the alternative with the higher probability of reinforcement. In concurrent variable-interval schedules, steady-state matching, which is an approximate equality between the relative frequency of a response and the relative frequency of reinforcement of that response, has previously been obtained only as a consequence of maximizing. In the present experiment, maximizing was impossible. A choice of one of two keys was reinforced only if it formed, together with the three preceding choices, the sequence of four successive choices that had occurred least often. This sequence was determined by a Bernoulli-trials process with parameter p . Each of three pigeons matched when p was $\frac{1}{2}$ or $\frac{1}{4}$. Therefore, steady-state matching by individual birds is not always a consequence of maximizing. Choice probability varied between successive reinforcements, and sequential statistics revealed dependencies which were adequately described by a Bernoulli-trials process with p depending on the time since the preceding reinforcement.

Invariant relationships between reinforcement probability and response probability are especially important to both the theorist and the experimenter concerned with behavioral control. Recently, some interest has centered on one possible relationship, equality, which in discrete trial probability learning experiments is sometimes called probability matching. There is evidence that a pigeon may match probabilities (Graf, Bullock, and Bitterman, 1964), but recent data suggest that if training is adequately extended, a pigeon tends to maximize, *i.e.*, to choose consistently an alternative having the higher probability of reinforcement (Shimp, 1966).

Behavior produced by concurrent variable-interval reinforcement schedules provides further evidence that a pigeon may match the relative frequency of a choice to the relative frequency of reinforcement of that choice (Catania, 1963; Herrnstein, 1961, 1964; Reyn-

olds, 1963). However, it has been shown that in these schedules, matching may result as a by-product of maximizing. In other words, if a subject maximizes, then as a mathematical consequence the relative frequency of a choice sometimes approximates the relative frequency of reinforcement of that choice. In a series of experiments with pigeons (Shimp, 1966), sequential statistics and other data in addition to matching were successfully predicted from an assumption of maximizing. It might be concluded from these earlier studies that if an experimenter wishes to use a probabilistic reinforcement schedule and to know in advance what a pigeon's steady-state choice probability will be, then he should compute the maximizing behavior for the schedule and take as his prediction the resulting average choice probability.

But what if the mathematically optimal strategy were made so complicated that a pigeon could not possibly follow it? Would a pigeon ever match the relative frequency of a choice to its relative frequency of reinforcement? This paper describes the behavior produced by a reinforcement schedule that gave a pigeon an opportunity to match but none to maximize.

This schedule can be described briefly as follows. A reinforcer was presented after a choice of one of two keys only if the choice,

¹This work was done at Brown University and was supported in part by USPHS Grant MH-02456 (D. S. Blough, Principal Investigator) and in part by the author's NSF Graduate and Postdoctoral Fellowships. The author wishes to acknowledge his indebtedness to Dr. D. S. Blough, and to thank Mr. Lloyd Marlowe, who helped with the computer programming. Reprints may be obtained from the author, Institute for Mathematical Studies in the Social Sciences, Ventura Hall, Stanford University, Stanford, California 94305.

together with the preceding three choices, formed the sequence of four successive choices that had occurred least often. This sequence was determined by the immediately preceding choices and by a rule derived from a Bernoulli-trials process. Thus, the schedule, to some extent, continually shaped choice behavior towards the Bernoulli-trials process. If a bird's choices were adequately described by this random process, then the following would be true: (1) successive choices would be independent; (2) choice probability would be constant; and (3) a bird would match, but matching would not be a consequence of maximizing.

METHOD

Subjects

Three two-year old male White Carneaux pigeons were kept at 80% of their free-feeding weights. They were selected from seven which had served in an earlier probabilistic reinforcement study (Shimp, 1966). The three birds which appeared to have had the shortest and most stable latencies in that earlier study were chosen.

Apparatus

Three two-key Lehigh Valley Electronics pigeon chambers were connected to a LINC on-line computer (Clark and Molnar, 1964).

Procedure

Both keys were illuminated with white light only after a blackout ended. The blackout lasted 1 sec if a bird did not peck during that time; it lasted $(n + 1)$ sec if a bird pecked at least once during the n -th sec. For example, if a bird pecked at least once during the first second of the blackout, then the blackout lasted 2 sec. This postponement contingency presumably punished responses near the beginning of the blackout; it was used to prevent short bursts of several pecks. A choice response was defined as any peck at one of the two lighted keys, left or right, and was correspondingly labeled A_0 or A_1 . Various measures of the relative frequency of A_1 will be denoted by $\text{Pr}(A_1)$. Which measure is used will be clear from context. A choice response turned off both key lights. Then the blackout followed, either immediately or after reinforcement. A reinforcement was the presentation of mixed grain for 1.3 sec.

The programming of reinforcements was based on the 16 possible sequences of four successive choices, such as the following: A_0, A_0, A_1, A_0 ; A_1, A_0, A_0, A_1 ; and so on. The computer kept a running count of the number of times each of the 16 sequences had appeared in the preceding N choices. After each choice, each of these 16 frequencies was multiplied by a corresponding weight. The sequence corresponding to the smallest of the 16 weighted frequencies was denoted the least-frequent sequence. The next choice was reinforced if it was the fourth choice of the least-frequent sequence. For example, if the computer found that the sequence A_0, A_0, A_0, A_1 was momentarily the least-frequent sequence, then reinforcement followed the next choice only if the choice was an A_1 and if the preceding three choices were A_0 's.

The weights were derived from a Bernoulli-trials process, with parameter p equal to the theoretical probability of an A_1 choice. When p was 0.5, each weight equaled unity and N equaled 96. When p was 0.25, the weight assigned to a sequence containing a number i of A_0 's and $4-i$ A_1 's was $p^i(1-p)^{4-i}N$, where N was 256. Thus, when p was 0.25, sequences with three or four A_0 's had the smaller weights and were more often followed by reinforcement. For each bird, the last N consecutive choices from the end of one session were stored for use at the beginning of the next session.

If there was no unique least-frequent sequence, then a reinforcement was programmed for the fourth choice in the tied sequence that had been reinforced least often in the session. The least frequently reinforced sequence was obtained in the same way as the least-frequent sequence.

On the last two days for which p was 0.5, reinforcement was never presented after an A_1 response. Reinforcements were programmed as before but were simply omitted after A_1 's. In all other ways the procedure remained the same. On those two days a bird had to continue to make A_1 responses to get food; otherwise, no sequence ending in A_0 would have been the least frequent.

Reinforcement was not always available because of the least-frequent sequence criterion: the three choices preceding any given choice could differ from the corresponding three choices in the momentarily least-frequent sequence. In addition, reinforcement was never

available after any of the first three postreinforcement choices except during 10 sessions of pretraining when p was 0.75. In other words, reinforcement was not available until a bird generated a sequence of four choices that did not overlap the preceding reinforced sequence.

The present reinforcement schedule allowed a bird to match but did not force it to do so. For example, if a bird chose according to a Bernoulli-trials process with $\text{Pr}(A_1)$ equal to 0.8, while the schedule parameter p was 0.5, then $\text{Pr}(A_1)$ would be greater than the relative frequency of reinforcement of A_1 . Furthermore, it is possible to show that even if $\text{Pr}(A_1)$ did approximate p , $\text{Pr}(A_1)$ did not have to approximate the actual relative frequency of reinforcement of A_1 .

Each session ended after the 100th reinforcement. The three birds were run simultaneously, but independently, at the same hour every day, seven days a week. Table 1 lists the programmed parameter p , the reinforced responses, and the number of days for each stage of the experiment.

Table 1

Sequence of Experimental Conditions

Parameter p of the theoretical Bernoulli-trials process	Number of sessions	Reinforced choices
0.75	10	A_0, A_1
0.50	14	A_0, A_1
0.50	2	A_0
0.25	23	A_0, A_1

The parameter p equals the theoretical probability of an A_1 choice.

RESULTS

Figure 1 shows that when p was 0.5, $\text{Pr}(A_1)$ was not quite constant over successive postreinforcement choices. As revealed in Fig. 2, neither was the relative frequency of reinforcement constant; it was necessarily zero for the first three choices, approximately 0.09 on the fourth choice, and about 0.05 on later choices.

The sequential statistics in Table 2 indicate that $\text{Pr}(A_1)$ was different after different preceding response sequences; that is, successive responses did not seem to be independent. The sequential data from the different birds were much alike.

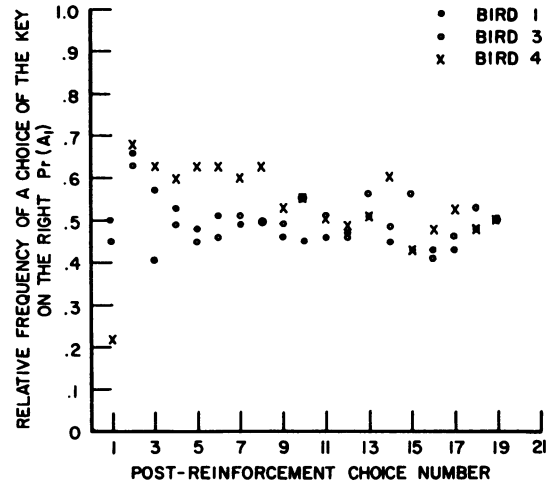


Fig. 1. Inconstancy of the probability of a choice of A_1 , $\text{Pr}(A_1)$, as a function of the number of choices since the preceding reinforcement, when p was 0.5. The frequencies on which the points are based decrease towards the right because the reinforcement of earlier choices reduced the number of opportunities for later postreinforcement choices. Data averaged over sessions 10, 11, and 12.

Figure 3 shows $\text{Pr}(A_1)$ for each bird on each session of the experiment. In all cases when p was 0.5, $\text{Pr}(A_1)$ closely approximated p . The averages of $\text{Pr}(A_1)$ over all 12 sessions for birds 1, 3, and 4 were 0.50, 0.50, and 0.50; and the corresponding relative frequencies of reinforcement were respectively 0.46, 0.48, and 0.50. Thus, the relative frequency of a choice approximated the relative frequency of reinforcement of that choice and equaled the theoretical p -value.

The mean numbers of responses per session over sessions 1 to 5 and 8 to 12 were 2369 and 2072 respectively for bird 1, 2746 and 2437 for

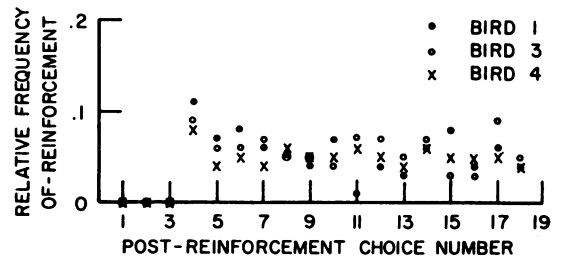


Fig. 2. The probability of reinforcement, or the number of reinforced choices divided by the number of opportunities for that choice, for postreinforcement choices 1 to 18, when p was 0.50. Reinforcement was never presented until a bird made at least four choices after the preceding reinforcement. Data averaged over sessions 10, 11, and 12.

Table 2
Sequential Statistics ($p = \frac{1}{2}$)

	Sessions 8, 9, 10		
	Bird 1	Bird 3	Bird 4
Pr(A ₁)	0.50	0.50	0.50
Pr(A ₁ A ₀)	.51	.54	.52
Pr(A ₁ A ₁)	.49	.47	.49
Pr(A ₁ A ₀ A ₀)	.48	.52	.51
Pr(A ₁ A ₀ A ₁)	.45	.41	.44
Pr(A ₁ A ₁ A ₀)	.54	.56	.53
Pr(A ₁ A ₁ A ₁)	.53	.53	.54
Pr(A ₁ A ₀ A ₀ A ₀)	.43	.48	.49
Pr(A ₁ A ₀ A ₀ A ₁)	.46	.42	.43
Pr(A ₁ A ₀ A ₁ A ₀)	.52	.58	.53
Pr(A ₁ A ₀ A ₁ A ₁)	.51	.55	.48
Pr(A ₁ A ₁ A ₀ A ₀)	.54	.55	.53
Pr(A ₁ A ₁ A ₀ A ₁)	.45	.40	.44
Pr(A ₁ A ₁ A ₁ A ₀)	.56	.53	.53
Pr(A ₁ A ₁ A ₁ A ₁)	.56	.52	.59
Pr(E ₁)	.49	.48	.55

The probabilities of an A₁ choice after different preceding sequences of choices. For example, Pr(A₁|A₀A₀A₁) is the probability of an A₁ after a preceding sequence of two A₀'s followed by an A₁. E₁ stands for a presentation of the feeder after an A₁.

bird 3, and 2493 and 2796 for bird 4. Thus, birds 1 and 3 improved in efficiency while bird 4 did less well. Each day the longest of the three sessions was noted and their average, over all days when p was 0.50, was 107 min.

The two sessions in which p was 0.50, but in which no A₁'s were reinforced, gave an esti-

mate of how far choice behavior was determined by the relative frequency of reinforcement. On the two days when only A₀'s were reinforced, Pr(A₁) decreased from 0.50 to about 0.40 on the first day (session 13) and to about 0.35 on the second day (session 14). All three birds ceased to respond on both days. The reinforcement rate, or number of reinforcements per response, was reduced by considerably more than half. Non-reinforcement of A₁ would be expected to reduce the reinforcement rate by at least half. The average probability of reinforcement became even lower the more a bird responded exclusively on the only key on which reinforcement was ever programmed.

On the first day for which p was 0.25, Pr(A₁) was more nearly constant than later, when, as shown in Fig. 4, it was a negatively accelerated, monotonically increasing function of the number of choices since the preceding reinforcement.

Figure 5 illustrates that the relative frequency of reinforcement was comparatively high on the fourth choice, decreased abruptly on the fifth, and thereafter decreased slowly. Just after reinforcement a bird tended to produce sequences with many A₀'s, and these sequences were often reinforced; if they were not reinforced, a bird was then more likely to make A₁'s.

Sequential dependencies shown in Table 3

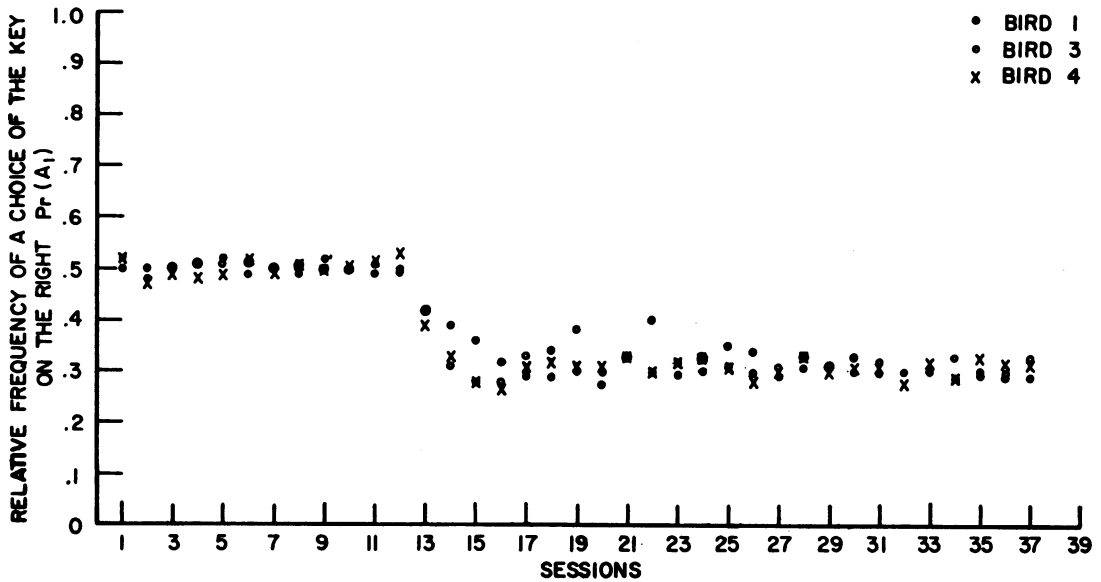


Fig. 3. The relative frequency of A₁, Pr(A₁), for each session. The parameter p was 0.5 on sessions 1 to 14 and 0.25 on sessions 15 to 39. On sessions 13 and 14, reinforcements were omitted after A₁'s although reinforcements were programmed as before.

appeared to be stable by the end of training and usually disclosed response perseveration.

The mean values of $Pr(A_1)$ over the last 11 sessions for birds 1, 3, and 4 were 0.30, 0.31,

and 0.31 respectively. The corresponding relative frequencies of reinforcement were 0.28, 0.22, and 0.33. Thus, $Pr(A_1)$ approximated the relative frequency of reinforcement or $Pr(A_1)$ approximated p .

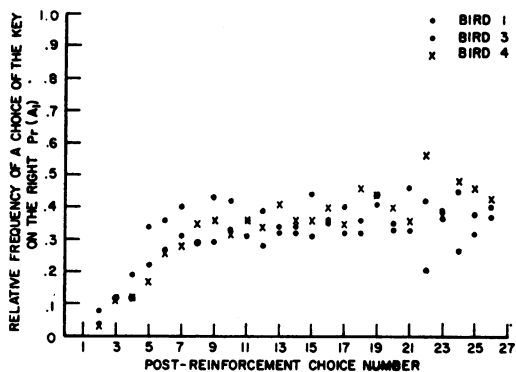


Fig. 4. Inconstancy of the probability of a choice of A_1 , $Pr(A_1)$, as a function of the number of choices since the preceding reinforcement, when p was 0.25. The frequencies on which the points are based decrease towards the right because the reinforcement of earlier choices reduced the number of opportunities for later postreinforcement choices. Data averaged over sessions 34, 35, and 36. See Table 3 for the data for the first postreinforcement choice.

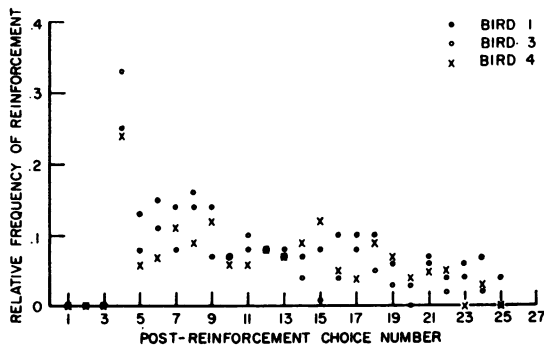


Fig. 5. The probability of reinforcement, or the number of reinforced choices divided by the number of opportunities for that choice, for postreinforcement choices 1 to 25, when p was 0.25. Reinforcement was never presented until a bird made at least four choices after the preceding reinforcement. Data averaged over sessions 34, 35, and 36.

Table 3
Sequential Statistics ($p = 0.25$)

	Real Data				Stat-Data		
	Sessions 34-36				Sessions 31-33	Based on $Pr(A_1)$ curve	Based on pseudo-maximizing lists
	Bird 1	Bird 3	Bird 4	Birds 1,3,4	Birds 1,3,4		
$Pr(A_1)$	0.29	0.31	0.31	0.30	0.31	0.30	0.31
$Pr(A_1 A_0)$.28	.31	.31	.30	.29	.28	.30
$Pr(A_1 A_1)$.31	.32	.32	.32	.34	.34	.33
$Pr(A_1 A_0A_0)$.25	.30	.27	.27	.29	.27	.28
$Pr(A_1 A_0A_1)$.28	.27	.31	.29	.31	.31	.28
$Pr(A_1 A_1A_0)$.36	.31	.40	.36	.30	.31	.35
$Pr(A_1 A_1A_1)$.36	.43	.36	.38	.41	.39	.43
$Pr(A_1 A_0A_0A_0)$.24	.27	.25	.25	.27	.26	.31
$Pr(A_1 A_0A_0A_1)$.28	.24	.25	.26	.29	.29	.32
$Pr(A_1 A_0A_1A_0)$.35	.33	.38	.35	.29	.27	.37
$Pr(A_1 A_0A_1A_1)$.37	.41	.30	.36	.40	.39	.40
$Pr(A_1 A_1A_0A_0)$.27	.37	.32	.32	.32	.32	.21
$Pr(A_1 A_1A_0A_1)$.28	.34	.39	.34	.35	.36	.20
$Pr(A_1 A_1A_1A_0)$.39	.27	.41	.36	.35	.39	.30
$Pr(A_1 A_1A_1A_1)$.36	.46	.48	.43	.43	.40	.47
Average no. of responses per session	1389	1427	1724	1513	1484	1490	975
$Pr(E_1)$.29	.25	.32	.29	.28	.34	.24
$Pr(A_1 E_0)$.02	.15	.00				
$Pr(A_1 E_1)$.22	.86	.10				

The probabilities of an A_1 choice after different preceding sequences of events. For example, $Pr(A_1|A_1A_1A_0)$ is the probability of an A_1 after a preceding sequence of two A_1 's followed by an A_0 . E_1 stands for a presentation of the feeder after an A_1 . How the stat-data were obtained, and their significance, are described in the text.

Table 4

Relative frequencies of sequences of four successive choices ($p = 0.25$).

Sequences of four choices	All Sequences		Reinforced Sequences only
	Sessions 34-36	Stat-data based on	Sessions 34-36
	Birds 1,3,4	$Pr(A_1)$ curve	Birds 1,3,4
$A_0A_0A_0A_0$	0.27	0.25	0.37
$A_0A_0A_0A_1$.10	.09	.10
$A_0A_0A_1A_0$.09	.10	.09
$A_0A_0A_1A_1$.04	.04	.05
$A_0A_1A_0A_0$.10	.10	.08
$A_0A_1A_0A_1$.04	.05	.04
$A_0A_1A_1A_0$.04	.04	.05
$A_0A_1A_1A_1$.03	.02	.01
$A_1A_0A_0A_0$.10	.09	.06
$A_1A_0A_0A_1$.04	.05	.04
$A_1A_0A_1A_0$.04	.05	.04
$A_1A_0A_1A_1$.02	.02	.02
$A_1A_1A_0A_0$.04	.04	.04
$A_1A_1A_0A_1$.02	.02	.02
$A_1A_1A_1A_0$.03	.02	.00
$A_1A_1A_1A_1$.02	.02	.02

In Table 4 it can be seen that the relative frequency of each of the 16 sequences, except for the most common sequence of four A_0 's, roughly approximated the corresponding relative frequency of reinforcement. The relative frequencies of the 16 sequences, as well as the relative frequency of A_1 , were stable over sessions.

The mean numbers of responses per session over the first five and the last five sessions when p was 0.25 were 1537 and 1560 for bird 1, 1948 and 1344 for bird 3, and 1938 and 1586 for bird 4. Although for each of the birds the number of choices per session was substantially reduced as soon as p changed from 0.50 to 0.25, only birds 3 and 4 showed a sizable improvement while p was 0.25. Each day the longest of the three sessions was noted and their average, over the last 10 sessions when p was 0.25, was 55 min.

DISCUSSION

Three characteristics of the data are of prime interest: independence, constancy, and matching. For the first of these, independence, sequential statistics gave ample evidence that the relative frequency of a choice depended on preceding sequences of choices. It is worthwhile to consider the significance of these ap-

parent dependencies. Did the probability of a choice depend on stimulus traces of preceding choices? Or did apparent dependencies result as by-products of responding that was independent of specific choices, but dependent on the length of time or on the number of choices since the preceding reinforcement?

The following argument suggests that much of the data can be described as the result of responding that was independent of preceding choices. Consider the case when p was 0.25. Suppose that successive postreinforcement responses were emitted independently and that $Pr(A_1)$ on the n -th choice after the preceding reinforcement was equal to the corresponding obtained probability, as for example in Fig. 4. Suppose, in other words, that choices were distributed as in a Bernoulli-trials process with p given by the data (see p. 205 in Feller, 1957).

A way to test this hypothesis of independence is to study the results of computer simulation of independent responding. If the resulting stat-data (Bush and Mosteller, 1955, p. 129) closely resemble the real data, then the hypothesis of independence is supported. The computer program for the experiment, together with a table of random numbers, (RAND, 1955) was used to simulate independent responding with changing p -values. The p -value on a postreinforcement choice equaled $Pr(A_1)$, averaged over sessions 28 to 33 and all three birds, on that postreinforcement choice. Table 3 shows that the stat-data disclose dependencies and in general approximate the real sequential statistics. Also, there is agreement between the stat-data and the real data in respect to $Pr(A_1)$, the relative frequency of reinforcement of A_1 , the number of responses per session, and the relative frequencies of the 16 sequences.

There is no adequate goodness-of-fit test to apply to all differences between the real data and the stat-data. The corresponding computer simulation could not be performed when p was 0.50 because data were not collected for enough choices that were adequately remote from the preceding reinforcement. Perhaps the dependencies in the data for which p was 0.50 also could be described by a Bernoulli-trials process with changing p -values if it were known what p was a function of. Although the magnitudes of the dependencies when p was 0.50 were not large, they were large enough

to ensure that a Bernoulli-trials process will not describe the data unless p does vary.

If sequences were emitted with a minimum of repetition of sequences already emitted since the preceding reinforcement, then the reinforcement rate would be relatively high. This process, which may be called "pseudo-maximizing", describes the sequential statistics nearly as well as does independent responding. For example, imagine lists of choices having the following properties: (1) each of the 16 sequences occurs at least once in each list; (2) each list contains about 20 choices; (3) if p equals 0.25, sequences with higher probabilities of reinforcement tend to occur earlier in the lists. The second and third of these criteria are not meant to be defined precisely.

Computer simulation using 18 such lists, instead of random numbers restricted by a curve like Fig. 4, provided reasonable fits to first and second order sequential statistics, as shown in Table 3. However, the predicted number of responses per session was too small; that is, a real bird's behavior was not so efficient as the pseudo-maximizing lists. Moreover, the third order sequential statistics discriminated in favor of the Bernoulli-trials process with varying p . In short, the pseudo-maximizing hypothesis fails to describe the choice behavior as well as does a Bernoulli-trials process with changing p -values. An important tentative conclusion is that the probability of a given choice did not depend on sequences of preceding choices.

The second important characteristic of the data was the inconstancy of $\Pr(A_1)$ as a function of time elapsed or of the number of choices made since the preceding reinforcement. When p was 0.25, sequences with higher relative frequencies of reinforcement, especially the sequence of A_0, A_0, A_0, A_0 , usually occurred first. Thus, the relative frequency of reinforcement was high soon after reinforcement, and notably so on the fourth postreinforcement choice. It seems probable that the inconstancy of $\Pr(A_1)$ and the absolute reinforcement rate interacted, but the precise nature of this relationship cannot be stated now. The agreement among the curves of the three birds suggests the presence of some powerful controlling variables. When p was 0.50, the high relative frequency of reinforcement on the fourth postreinforcement choice may have resulted from the bird's tendency to repeat

a reinforced sequence which, if its frequency were low enough, could be reinforced several times in succession.

The third characteristic to be discussed is matching. The birds approximately matched and the closeness of the approximation depended on p : the relative frequency of A_1 typically approximated the relative frequency of reinforcement of A_1 , both when p was 0.50 and when p was 0.25. But the approximation of $\Pr(A_1)$ to p was closer when p was 0.50 than when p was 0.25. In addition, the relative frequencies of the 16 sequences roughly matched the corresponding relative frequencies of reinforcement at least when p was 0.25. These data were not recorded when p was 0.50.

The preceding paragraph raises the following question: in general, what matches what and why? It is becoming apparent that several different measures of response probability may approximate several different measures of reinforcement probability. In support of this statement there are at least two lines of evidence: (1) rate of response measures in single key studies, and (2) choice measures in two-key studies. In the first case there are the data of Anger (1956) and of Blough (1966). Anger suggested that with a variable-interval reinforcement schedule, the conditional probability of an interresponse time is roughly proportional to its relative time rate of reinforcement. Blough performed an experiment analogous to the present one. In Blough's schedule, a response was reinforced if it terminated an interresponse time that had occurred least often relative to a theoretical exponential interresponse time distribution. This distribution, as well as the theoretical binomial distribution of choices used in the present schedule, results from independent random events that occur with a constant probability. Blough discovered that response rate was approximately constant and, further, that it approximated the theoretical rate; however, successive interresponse times did not seem to be independent.

The results of two-key studies have shown that a pigeon may match, at least near the beginning of training, in probability learning experiments (Graf *et al.*, 1964; Shimp, 1966); or may match asymptotically, as in the present experiment and in concurrent variable-interval schedule experiments (Catania, 1963; Herrnstein, 1961, 1964; Reynolds, 1963).

Obviously, it is valuable to examine the reasons for matching in different procedures. The results of a previous two-choice probability learning study (Shimp, 1966) suggested that matching in probability learning experiments is only a transient phenomenon, and that matching in concurrent variable-interval schedules is a consequence of maximizing. The tentative conclusion there was that the time rate of reinforcement determined most of the data. In that study, the maximizing strategy consisted of a single repetitive pattern of choices, such as right, right, left, right, right, left, and so on. An analysis of maximizing in terms of such sequences of choices seems to assume considerable ability on the part of the subject to discriminate preceding sequences and the corresponding reinforcement probabilities. But in the light of evidence from the present experiment that choices may not have been controlled by preceding response sequences, it seems desirable to formulate for the earlier study a different rule to determine the maximizing choice. Such a rule, which depends on temporal controlling variables instead of on response sequences, is given below. Assume that the probability of reinforcement on each key increases linearly with time, and call the time since the last A_1 or A_2 response t_{A_1} or t_{A_2} respectively. Assume further that the reinforcement probability increases K times as fast for response A_1 as for response A_2 . The probabilities of reinforcement on the two keys are equal if $t_{A_1} = Kt_{A_2}$. Then, to maximize, a subject needs only to decide whether the ratio t_{A_1}/t_{A_2} is greater or smaller than the fixed constant K . If the ratio is greater than K , the subject should make an A_1 , if otherwise, an A_2 . It should be noted that the two assumptions are only approximately satisfied in the earlier study by Shimp (1966) and in the practice usually followed in concurrent variable-interval schedules.

Unlike those from the earlier two-key experiments, the present data indicate that the time rate of reinforcement was not the sole variable controlling choice behavior. Here, two reasons exist for suspecting that the relative frequency of reinforcement also controlled choice behavior. First, it is unlikely that a bird approximately matched merely by chance, while the way in which $\text{Pr}(A_1)$ changed between reinforcements varied over sessions. Second, if the relative frequency of reinforce-

ment had not controlled $\text{Pr}(A_1)$, then $\text{Pr}(A_1)$ might have been expected to remain at about 0.50 on sessions 13 and 14 when p was 0.50 but only A_0 's were reinforced. Otherwise, $\text{Pr}(A_1)$ presumably would have decreased from 0.50. The latter actually happened. It is not surprising that all three birds ceased to respond on each of these two days, since the reinforcement rate was of necessity halved abruptly from the rate prevailing in the foregoing sessions. And, as it turned out, when $\text{Pr}(A_1)$ departed from 0.50, the reinforcement rate was decreased further. Herrnstein (1958) studied a loosely analogous condition and found that a pigeon would continue to respond on a modified concurrent fixed-ratio schedule where responses on only one key were reinforced. The behavior was determined almost completely by the time rate of reinforcement. The relative frequency of reinforcement had only a slight effect.

Seemingly, in the present experiment a bird's "solution" to a problem for which the optimal strategy was impossible was a compromise ruled both by the absolute rate of reinforcement and by the relative frequency of reinforcement. And in fact, the compromises were restricted to those solutions which produced matching of $\text{Pr}(A_1)$ to p , of $\text{Pr}(A_1)$ to the relative frequency of reinforcement of A_1 , or of the relative frequencies of sequences of choices to the corresponding relative frequencies of reinforcement.

REFERENCES

- Anger, D. The dependence of interresponse times upon the relative reinforcement of different interresponse times. *J. exp. Psychol.*, 1956, **52**, 145-161.
- Blough, D. S. The reinforcement of least-frequent interresponse times. *J. exp. Anal. Behav.*, 1966, **9**, 581-591.
- Bush, R. R. and Mosteller, F. *Stochastic models for learning*. New York: Wiley, 1955.
- Catania, A. C. Concurrent performances: reinforcement interaction and response independence. *J. exp. Anal. Behav.*, 1963, **6**, 253-263.
- Clark, W. A. and Molnar, C. E. The LINC: a description of the Laboratory Instrument Computer. *Ann. N. Y. Acad. Sci.*, 1964, **115**, 653-668.
- Feller, W. *An introduction to probability theory and its applications*, second edition. New York: Wiley, 1957.
- Graf, V., Bullock, D. H., and Bitterman, M. E. Further experiments on probability matching in the pigeon. *J. exp. Anal. Behav.*, 1964, **7**, 151-157.
- Herrnstein, R. J. Some factors influencing behavior in a two-response situation. *Trans. N. Y. Acad. Sci.*, 1958, **21**, 35-45.

- Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *J. exp. Anal. Behav.*, 1961, 4, 179-184.
- Herrnstein, R. J. Secondary reinforcement and rate of primary reinforcement. *J. exp. Anal. Behav.*, 1964, 7, 27-36.
- RAND. *A million random digits*, The RAND Corp., Glencoe, Illinois: Free Press, 1955.
- Reynolds, G. S. On some determinants of choice in pigeons. *J. exp. Anal. Behav.*, 1963, 6, 53-59.
- Shimp, C. P. Probabilistically reinforced two-choice behavior in pigeons. *J. exp. Anal. Behav.*, 1966, 9, 443-455.

Received February 21, 1966