*CHOICE BEHAVIOR ON DISCRETE TRIALS: A
DEMONSTRATION OF THE OCCURRENCE OF A
RESPONSE STRATEGY¹*

ALAN SILBERBERG AND DAVID R. WILLIAMS

THE AMERICAN UNIVERSITY AND UNIVERSITY OF PENNSYLVANIA

Three pairs of pigeons were trained to peck at two keys presented simultaneously in discrete trials with intertrial intervals of 1, 22, or 120 sec. Left-key responses incremented the probability of reinforcement for the first right-key response and, conversely, right-key responses incremented the probability of reinforcement for the first left-key response. In terms of relative response rates, it was found that all birds' choices were described by a momentary maximizing strategy, but this fact was not reflected in the detailed sequential statistics for birds with the longer (22 or 120 sec) intertrial intervals. It was hypothesized that choice behavior, in general, may be accurately described by a momentary maximizing sequence, but that prior failures to demonstrate this were due to "errors" in executing the momentary maximizing sequence. These misappropriated responses, which are hypothesized to be randomly distributed among the responses defining the momentary maximizing sequence, caused successive choices to appear to be statistically independent when, in fact, they were not.

Herrnstein (1961) demonstrated that pigeons on concurrent variable-interval schedules (*conc VI VI*) equate (match) the relative response rate (responses to a key divided by the response total to both keys) to the relative reinforcement frequency (reinforcements to a key divided by the reinforcement total to both keys). Although this finding has been reproduced many times with *conc VI VI* schedules, two incompatible accounts have been offered in explanation of it.

Herrnstein argued that pigeons match "... not because they take into account what is happening on the two keys, but because they respond to the two keys independently (p. 270)." Consistent with his hypothesis is the view that pigeons are not controlled by such possible stimulus events as the loci of prior responses, but are sensitive only to the relative reinforcement frequency obtained for responding to the two keys. A prediction of this *frequency matching* model is that the probability

of a particular response is independent of prior choices.

Shimp (1966) offered a different account for the phenomenon of matching. He argued that an organism's choice behavior maximizes the time rate of reinforcement (*i.e.*, reinforcements per unit time). For example, a pigeon responding on concurrent fixed-ratio schedules would determine which key, with its associated schedule, offered the higher density of reinforcement in time and would respond exclusively to that key. In some choice situations, this concurrent response rule may be achieved by the specific response strategy of choosing that alternative that momentarily has the higher reinforcement probability. This particular strategy, which Shimp called momentary maximizing, approximates matching on *conc VI VI* schedules. Unlike the frequency matching model, momentary maximizing, defined at the level of individual response sequences, predicts sequential response dependencies—that is, the probability of a left or right response varies from choice to choice, depending on where the last response occurred.

Although each of these two models predicts matching on *conc VI VI* schedules, a momentary maximizing strategy, unlike frequency matching, also specifies that responses occur in that order which maximizes the reinforcement frequency at each instance a choice is made.

¹This research was conducted while the first author was an NIMH pre-doctoral fellow at the University of Pennsylvania. The costs of manuscript preparation and research were borne, respectively, by NIMH grant 1 RO1 MH22881-01 (Alan Silberberg, principal investigator), and NSF grants 7946X and GB 14055 (David R. Williams, principal investigator). Reprints may be obtained from Alan Silberberg, 321 Asbury Building, The American University, Washington, D.C. 20016.

By analyzing the individual response sequences composing relative response rates it may be possible to determine whether or not momentary maximizing adequately describes choice allocation on *conc VI VI* schedules. Two studies directed at this issue, one by Shimp (1966) and one by Nevin (1969), have produced contradictory results.

Although Shimp's study consisted of several experiments, the basic procedure was as follows: a center-key response illuminated two side keys and occasionally assigned a reinforcement with unequal probability to one of the two keys. Once an assignment was made to a key, no further assignments could be made until the reinforcement was delivered. Each side-key response darkened the choice keys and reilluminated the center key. Shimp found that pigeons matched and that their choice behavior was dependent on changes produced in the relative probability of reinforcement by the loci of preceding responses. In other words, matching occurred as the product of a response strategy.

Nevin presented evidence that contradicts Shimp's momentary maximizing hypothesis as a description of choice behavior on *conc VI VI* schedules. Pigeons' response sequences on *conc VI 1-min VI 3-min* schedules were studied on a discrete-trials basis. A trial began with the illumination of both keys. A response to either key turned off both keylights for an intertrial interval (ITI) of 6 sec. Nevin's results conformed to a matching prediction without displaying systematic response dependencies. In other words, matching occurred, but it did not appear to be the product of a response strategy.

Unlike Shimp's experiment, in which the ITI was subject-controlled, the longer 6-sec ITI of the Nevin study may have contributed to his failure to find a response strategy underlying choice allocation. It is possible that the relatively longer time between successive choices in this study minimized the conspicuousness of the consequences of different response strategies. It is clear that the execution of a response strategy requires the recollection of the locus of at least the last response. For example, the momentary maximizing sequence in the Nevin study, which was three consecutive responses to the key associated with the VI 1-min schedule followed by a single response to the key associated with the VI 3-min schedule, required the pigeon to recall the loci

of its last three responses in order to respond without error. If errors were made in the execution of this response sequence, and these errors were randomly distributed among the choices defining the momentary maximizing sequence, matching would obtain, and it could still be the product of a momentary maximizing strategy; however, evidence for this strategy may not be clear in terms of sequential dependencies because of the sequential independence of randomly distributed errors. The hypothesis to be assessed by the present experiment is that choice behavior can be described by some reinforcement optimizing strategy, but that evidence for this based only on sequential statistics may be obfuscated by errors in its execution.

In the present study, the response rule governing choice allocation can be manifested not only by analysis of sequential statistics, which is the typical way of defining a response strategy (e.g., Shimp, 1966), but also by relative response rates. Different relative response rates are anticipated if choice allocation can be described by a frequency matching model or a momentary maximizing strategy. Thus, an advantage of the present paradigm is that if the demonstration of a momentary maximizing strategy is shrouded in statistical analyses by the occurrence of misappropriated responses, it may still be evidenced in the relative response rates. Moreover, in order to assess the effects of ITI duration on sequential dependencies, the present experiment used ITIs of length 1, 22, or 120 sec.

METHOD

Subjects

Six adult, male, Silver King pigeons, experimentally naive at the beginning of training, were kept at approximately 80% of their free-feeding weights for the duration of the experiment.

Apparatus

Subjects were placed in a column of cages to which a Lehigh Valley Electronics response panel consisting of three translucent response keys and a solenoid-operated grain hopper was sequentially presented, one cage per session, by an automatically programmed elevator device. A session began by the presentation of the response panel. At the end of a session, the

response panel moved back from the cage and advanced to the next cage. This process was repeated for all subjects.

When the response panel was flush against a cage, the left and right keys were transilluminated by green and red light respectively. The center key remained darkened and non-functional throughout the experiment. Transillumination was accomplished by I.E.E. multi-stimulus projectors mounted behind the response keys. Each key required a minimum force of 0.1 N to be operated. Masking noise was provided by a rotary fan mounted behind the response panel.

Procedure

Birds trained to respond to the left (green) and right (red) keys were pretrained on a VI 1-min VI 1-min schedule with a response-dependent ITI of 6 sec—that is, the illuminated keys were concurrently available until a response was made to either key; every response was followed by a blackout of 6 sec, either immediately or after reinforcement. Only one reinforcement (3-sec access to mixed grain) could be available on any trial, and once available, no other reinforcements could be assigned until it was obtained by a response to the appropriate key. Pretraining consisted of two daily sessions; each session terminated after 30 reinforcements.

The birds were then introduced to the main experimental procedure. The response keys were presented with a response-dependent ITI of either 1 sec (Birds 1226, 1752), 22 sec (Birds 1168, 2561) or 120 sec (Birds 1216, 1710). When a response produced access to grain, the keys and houselight were darkened and the food hopper was illuminated.

Responding, which was recorded on print-out counters, was maintained by a schedule in which responses to a key incremented the reinforcement probability for the first response to the other key. Only changeover responses between keys could be reinforced; however, even if a changeover response was not reinforced, it reset the reinforcement probability incremented by responding to the other key to zero. Each response to the left key incremented the reinforcement probability for the first right response 0.12 per response. For example, a right response following five left responses would be reinforced 0.6 of the time. The first, second, and third responses to the right key in-

cremented the reinforcement probability for the first left response to 0.2, 0.8, and 1.0 respectively. Thus, two right responses followed by a left response would be reinforced 0.8 of the time.²

Except for a one-month delay between Sessions 50 and 51 for the 1-sec and 22-sec ITI birds, the study was conducted daily, one session per day for 100 days. Each session terminated after 100 changeover responses.

RESULTS

Figure 1 compares relative response rates and relative reinforcement frequencies separately for each bird in the experiment. The light bars show the relative frequency of reinforcement and the dark bars the relative frequency of choice for the left key. Frequency matching is indicated when the dark and light bars are of equal height. Data for the 1-sec ITI birds are presented in panels 1 and 2, for the 22-sec ITI birds in panels 3 and 4, and for the 120-sec ITI birds in panels 5 and 6. If it is assumed that successive choices are independent and that the relative response rate is stable throughout each session, frequency matching could be achieved only if the relative response rate and the corresponding relative frequency of reinforcement on the left key were both equal to 0.57. Since only changeover responses are reinforced, the prediction of the momentary maximizing rule is, of course, that relative response rates will equal 0.5.³

Panels 1 and 2 of Figure 1 show that both 1-sec ITI birds responded approximately equally on both keys, despite the unequal reinforcement frequencies that such responding produced. Further, across sessions, small and opposed changes in the relative response rates and relative reinforcement frequencies produced progressively larger differences in these measures. During the last 60 sessions, in fact, these birds showed a very small preference for the right key despite the high reinforcement frequency to the left key.

²An appendix presenting the mathematical derivations for the predictions of frequency matching and momentary maximizing is available from the senior author.

³The momentary maximizing rule does not, in fact, produce the highest time rate of reinforcement. The optimal response strategy in this experiment is to respond twice to the right key and once to the left key.

Fig. 1. Relative response rate (dark bar) and relative reinforcement frequency (light bar) to the left key as a function of sessions averaged over 20-session blocks. The top, middle, and bottom pairs of panels are, respectively, from each of the two subjects in the 1-sec, 22-sec, and 120-sec ITI groups.

The initial performance of the birds with 22-sec ITIs was markedly different from that of birds with 1-sec ITIs. During the first 20 sessions, Bird 1168 (Figure 1, panel 3) matched response and reinforcement frequencies almost exactly while Bird 2561 (panel 4) "over-matched". As the experiment progressed, however, both 22-sec ITI birds tended to distribute responses more equally on both keys, and during the last 20 sessions were similar to the 1-

Fig. 2. Conditional probability of alternation as a function of run length. The left- and right-hand columns refer, respectively, to the first and last 20-session blocks. The top, middle, and bottom rows are, respectively, from the 1-sec, 22-sec, and 120-sec ITI groups. Data points in the form of squares label left-key runs; right-key runs are labelled by circles.

sec ITI birds in terms of relative response rate. It should be noted that for Sessions 61 to 80 of panel 3, which followed a one-month hiatus in the experiment, Bird 1168 showed a tendency to match. For the last 20 sessions, however, its relative response rates and relative reinforcement frequencies diverged, conforming with the performance of the other birds.

Across sessions, the 120-sec ITI birds (panels 5 and 6) also showed a tendency to diverge from matching; however, this divergence was neither as great nor as rapid in developing as in the 22-sec ITI birds. Bird 1710 in panel 6 showed a strong bias for the left key, which

disappeared after the first 20 sessions. During the last 20 sessions, this bird also tended to respond equally to both keys.

Each panel of Figure 2 shows the conditional probability of an alternation between keys as a function of run length (responses to a key before a changeover response). These conditional probability curves are based on the summed data from Table 1 for each pair of birds in the three ITI manipulations for the first and last 20 sessions only. All plots on the left side of the figure are the mean conditional probability curves during the first 20 sessions for each experimental manipulation. Plots on

Table 1

Number of response runs (successive responses to a key before a changeover) of lengths 1, 2, 3, or 4 or more to the left (L) and right (R) keys for each of the six subjects by 20-session blocks.

Bird #	Length 1		Length 2		Length 3		Length 4 or More	
	L	R	L	R	L	R	L	R
Sessions 1 to 20								
1226	811	805	134	173	34	19	21	3
1752	819	866	159	123	20	9	2	2
1168	469	657	276	227	131	68	124	48
2561	437	551	256	263	114	115	193	71
1216	421	533	234	227	133	129	212	111
1710	357	595	201	220	118	86	324	99
Sessions 21 to 40								
1226	865	791	79	179	38	20	18	10
1752	890	867	103	132	6	0	1	1
1168	543	757	237	209	126	24	94	10
2561	524	596	259	264	118	82	99	58
1216	433	535	261	252	132	108	174	105
1710	412	570	206	218	121	94	261	118
Sessions 41 to 60								
1226	956	847	32	145	9	4	3	4
1752	775	858	169	116	49	19	7	7
1168	590	799	251	162	84	30	75	9
2561	569	619	254	258	102	90	75	33
1216	495	500	234	269	135	122	136	109
1710	446	544	220	230	123	97	211	129
Sessions 61 to 80								
1226	910	815	76	178	12	5	2	2
1752	954	842	43	157	2	1	1	0
1168	586	790	249	155	97	39	68	16
2561	660	697	201	202	81	69	58	32
1216	476	516	237	261	127	127	160	96
1710	456	532	230	259	131	133	183	78
Sessions 81 to 100								
1226	900	871	86	126	13	2	1	1
1752	942	864	53	128	4	5	1	3
1168	662	803	249	160	61	29	28	8
2561	692	585	189	274	59	97	60	44
1216	491	529	228	254	116	114	165	103
1710	521	572	230	235	100	97	149	96

the right side of the figure are of the last 20 sessions. The top, middle, and bottom rows present data from the 1-sec, 22-sec, and 120-sec ITI birds respectively. Left-key runs are labelled by squares and right-key runs by closed circles.

Each conditional probability curve specifies a relative frequency defined by the number of runs of length n divided by the number of runs equal to or greater than n . For example, if there are four response runs each of lengths 1, 2, and 3, the relative frequency of each run is $4/12$. The conditional probability, however, is $4/12$, $4/8$, and $4/4$ respectively for each run. If such a conditional plot of the relative frequency of various run-lengths has a zero slope, it is assumed to represent a random distribution of alternations across response runs (see Anger, 1956).

The Y-intercepts of these conditional probability curves specify the probability of alternating responses between keys. If these curves are below 0.5 on the ordinate, they represent a tendency to perseverate responses to a key before alternating. Curves above 0.5 represent a tendency to alternate responses between keys. Although the data were analyzed in terms of runs of lengths 1, 2, 3, and 4 or more, only run lengths up to three are presented because the conditional probability of alternation for runs of four or more is always 1.0.

The results presented in Figure 2 show three consequences of the experimental procedures employed:

(1) Throughout the experiment, there was an increasing tendency toward alternation of responses between keys—the strategy suggested by the momentary maximizing principle. This change can be seen in the increase in the value of the Y-intercepts of the conditional probability curves from Sessions 1 to 20 to Sessions 81 to 100. The only obvious reversal of this generalization occurs in the curve for right-key runs in the 1-sec ITI group.

(2) Across groups during both the initial and terminal 20-session blocks, the value of the Y-intercept is inversely related to the length of the ITI—that is, the longer the ITI, the lower the tendency to alternate between keys. Further, only the 1-sec ITI birds showed any strong sequential dependencies in their conditional probability plots of response runs to a key. With regard to the 22-sec and 120-sec ITI birds, these curves suggest that choice behavior

is independent of preceding choices across runs of lengths 1, 2, and 3 to either key.

(3) Curves for right-key runs have a higher Y-intercept value than left-key runs, especially in the 22-sec and 120-sec ITI groups. This difference was established during the first 20 sessions for all groups and continued throughout the experiment.

DISCUSSION

A discrete-trials choice paradigm was used in which responses incremented the reinforcement probability of a changeover. Successive choices were separated, for three pairs of pigeons, by an ITI of 1, 22, or 120 sec. The general findings were: (1) all pigeons came to respond equally to both keys despite unequal reinforcement frequencies on the two keys. This tendency toward equal relative rates of responding developed at different rates for each of the three groups, however. The 1-sec ITI group responded with a 0.5 relative response rate during the first 20-session block. The 22-sec and 120-sec ITI groups, on the other hand, tended toward matching during the first few 20-session blocks, and only after many trials tended to respond equally to both keys. (2) With regard to the conditional probability of a changeover, the 1-sec ITI birds rapidly developed a tendency to alternate responses between keys. The longer ITI birds, however, acquired this tendency less rapidly. In fact, the 120-sec ITI group, which during the initial 20-session block tended to perseverate responses to a key (Y intercepts below 0.5), never developed a tendency to alternate responses between keys.

When considering each of the ITI groups individually, the 1-sec ITI birds reflected in an unambiguous fashion that momentary maximizing was the concurrent response rule defining their choice allocations. Since only changeovers could be reinforced, the momentary maximizing strategy was strict alternation of responses between keys. The fact that the relative response rates equalled 0.5, and the conditional probability curves had high Y-intercept values, establishes the primacy of this response rule for this group.

With regard to the longer ITI groups, however, the evidence for momentary maximizing based on sequential statistics alone, is ambiguous. During the first 20 sessions for the 22-sec

ITI group and the first and last 20 sessions for the 120-sec ITI group (Figure 2), the Y-intercepts of the conditional probability curves approximated 0.5, even though the values of the Y-intercepts must exceed 0.5 to define the underlying response strategy as momentary maximizing by this statistic alone. Although successive choices appear statistically independent by this measure, particularly for the 120-sec ITI group, it is noteworthy that both the 22-sec and 120-sec ITI groups' steady state relative response rates were approximately 0.5—the value conforming to the prediction of the momentary maximizing sequence. This contradiction between an assay for strategies based on relative response rates *versus* the conditional probability curves suggests that momentary maximizing can occur without apparent sequential dependencies.

The fact that momentary maximizing, defined either in terms of equal relative response rates or of sequential dependencies in the response data, can account for choice allocation in the present study does not imply that frequency matching on *conc VI VI* schedules may be described in the same way. Even though matching occurred pre-asymptotically during several 20-session blocks, the fact that in the present study each choice produced dramatic shifts in local reinforcement probability made the finding of sequential dependencies and equal relative response rates likely events. It remains to be established whether or not on conventional *conc VI VI* schedules, where changes in local reinforcement probability are less abrupt, matching is the product of a response strategy. A conclusion supported by the present study, however, is that Nevin's demonstration of matching without sequential dependencies does not mean that momentary maximizing was not the rule for choice allocation in his study: the finding of equal relative response rates for all birds, despite the absence of sequential dependencies for the 120-sec ITI group, suggests that momentary maximizing can occur *without* obvious sequential dependencies in choice data.

A model that can account for the data of this study is based on the hypothesis that the choice behavior of all birds is consistent with a momentary maximizing strategy, but the strategy frequently cannot be observed because the birds can "recall" the loci of their last few choices only a small portion of the time. The

likelihood of using their past choices as discriminable events is affected by the value of the ITI—the longer the ITI, the less likely it is that prior choices can serve as discriminative stimuli. Although the longer the ITI the higher the likelihood that the loci of prior responses will be "forgotten", it is assumed in this model that instances of forgetting will be equiprobable after all response sequences. Therefore, on those occasions when prior choices are correctly recalled, the momentary maximizing sequence will, on *conc VI VI* schedules, have a higher likelihood of reinforcement than all other response sequences. The fact that the acquisition of the steady state relative response rate of 0.5 was slower for the longer ITI groups than for the 1-sec ITI group (Figure 1) is consistent with this model: since the probability of correctly recalling prior choices' loci is inversely related to ITI value, the momentary maximizing sequence, as reflected by equal responding to both keys, should be learned more rapidly by subjects with short ITI durations.

Once the momentary maximizing sequence has been acquired, this model predicts that on those occasions when recollection of prior choices' loci is correct, conditional probability curves will show strong sequential dependencies. It is hypothesized, however, that when prior choices' loci are forgotten, pigeons will "guess" what response would have been next in the momentary maximizing sequence. Choices governed by guessing will be randomly distributed among the response alternatives that define the momentary maximizing sequence. These statistically independent responses will tend to minimize the underlying response dependencies that characterize choice behavior when errors do not occur. The relative response rate will be unchanged, however, because each choice in the momentary maximizing sequence will, over time, receive an approximately equal number of misappropriated responses. Therefore, based on this model, the equal response rates obtained across all three ITI groups reflects control by the momentary maximizing strategy. Conditional probability plots provided further evidence for this response strategy only for the 1-sec ITI group because the longer ITIs of the other two groups increased the number of misappropriated responses in the execution of the momentary maximizing sequence.

Aside from explaining the present data, this "forgetting" model offers a means of resolving the apparent absence in the Nevin study of a reinforcement optimizing strategy such as had been found on Shimp's paradigm. In the Shimp study, each choice trial began with a center-key response. Therefore, the ITI was functionally defined by the pigeon's response latency to the center key. Although data concerning the absolute response rates to that key were not provided, it is reasonable to assume that the mean response-dependent ITI was comparable to that of the 1-sec ITI group in the present experiment. As with the 1-sec ITI birds, Shimp found evidence for momentary maximizing not only at the level of relative response rate, but also in terms of individual response sequences. In the Nevin study, on the other hand, the longer ITI used no doubt played a large role in increasing the difficulty of recalling the loci of preceding responses. This difficulty may have been a factor obscuring the demonstration of a momentary maximizing strategy based on response se-

quences. As with the 22-sec and 120-sec ITI birds in the present study, momentary maximizing may have been the principle governing choice behavior in the Nevin study, but it may not have been reflected in the sequential statistics due to errors in following the momentary maximizing sequence.

REFERENCES

- Anger, D. The dependence of interresponse times upon the relative reinforcement of different interresponse times. *Journal of Experimental Psychology*, 1956, 52, 145-161.
- Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1961, 4, 267-272.
- Nevin, J. A. Interval reinforcement of choice behavior in discrete trials. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 875-885.
- Shimp, C. P. Probabilistically reinforced choice behavior in pigeons. *Journal of the Experimental Analysis of Behavior*, 1966, 9, 443-455.

Received 25 April 1973.

(Final Acceptance 31 October 1973.)