JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR VOLUME 9, NUMBER 2 MARCH, 1966

PROBABILITY OF RESPONSE AND PROBABILITY OF REINFORCEMENT IN A RESPONSE-DEFINED ANALOGUE OF AN INTERVAL SCHEDULE[1]

J. R. MILLENSON

CARNEGIE INSTITUTE OF TECHNOLOGY

Variable interval (VI) responding was hypothesized to be a function of differential reinforcement susceptibilities of various unspecified behavior chains that mediate interresponse times (IRTs). To test this hypothesis, probabilities of reinforcement were regulated for the lengths of chains of key pecking responses of pigeons, analogous to the way that VI regulates probabilities of reinforcement for IRTs. This procedure generated a number of VI-like effects, supporting the notion that VI behavior can be construed as a special case of an interaction between the organism's function relating reinforcement susceptibilities to chain length and the experimenter's function relating probabilities of reinforcement to chain length.

A variable interval (VI) schedule may be characterized as a procedure in which the probability of reinforcement for a response increases as a function of the time since the preceding response (Newman and Anger, 1954; Anger, 1956; Revusky, 1962; Millenson, 1963; Catania and Reynolds, *in press*). An interesting property of VI schedules is that, despite the increased reinforcement probability that they provide for longer interresponse times (IRTs), they tend to generate shorter IRTs, and thus higher rates of response, as exposure continues (Anger, 1956; Millenson, 1963). This process is self-limiting, however, for a steady response rate is eventually achieved which is appreciably lower than response rates characteristically associated with ratio schedules that yield corresponding frequencies of reinforcement.

This rate transition under VI is compatible with the hypothesis stated by various investigators (Anger, 1956; Malott and Cumming, 1964; Catania and Reynolds, *in press*), that different IRT regions may have different reinforcement susceptibilities. In particular if short IRTs are more susceptible to reinforcement than long IRTs, then even under VI where longer IRTs have a greater probability of reinforcement, it might be expected that short IRTs would tend eventually to predominate.

Malott and Cumming report data bearing on this possibility that different IRT regions are differentially susceptible to reinforcement. In one treatment using rat subjects and water reinforcement, they concurrently reinforced two IRT regions (an upper and a lower) with probabilities of 1.0 and 0.4 respectively. Under these conditions rats produced a greater proportion of IRTs in the lower region, even though they failed to optimize their overall rate of reinforcement. Further support is provided by Brandauer (1958), and Ray and McGill (1964) who reinforced all IRTs with a constant probability and found that very short IRTs came to predominate.

It appears plausible that the behavioral mechanisms responsible for such varying susceptibilities lie with the ongoing activities that occupy the time between two specified responses. During short IRTs there is little time for any behavior save the rather stereotyped chain of initiating and completing the specified response. During long IRTs, on the other hand, there is time for a much longer and more varied sequence of behavior. It follows that reinforcement of short IRTs is likely to catch a short, stereotyped behavior sequence whereas reinforcement of a long IRT is likely to catch a long, variable behavior sequence.

This analysis implies that differential reinforcement susceptibility of various IRT regions, and by implication the IRT effects of VI, may depend in part on differential susceptibility of the chains of behavior separating two responses. Such a topographical

---

[1] Supported in part by research grant MH-07722-01 from the NIH to H. A. Simon, principal investigator. Reprints may be obtained from the author, Dept. of Psychology, Carnegie Institute of Technology, Schenley Park, Pittsburgh 13, Pa.

analysis is speculative, however, because these interval schedules provide no direct experimental control over the behavior mediating an IRT. In the present study an attempt was made to make such mediating chains explicit by programming contingencies of reinforcement for specified mediating behavior. Since these contingencies were programmed analogously to the way that IRTs are programmed under one kind of VI, the procedure amounts to a response-defined analogue of VI.

EXPERIMENT I

Subjects

Two male homing pigeons (3,4), lacking previous experimental history, were maintained at 85 per cent of free-feeding weights.

Apparatus

The birds worked in individual 18 by 18 by 18 in. home cages. A removable response and reinforcement panel was inserted to form one wall during daily experimental sessions. The panel contained two 1-in. diameter keys, spaced 4 in. apart and symmetrical to a grain magazine located below them. The keys could be transilluminated by lights from inline display units. Reinforcement was 3 sec access to a tray of mixed grain. A switching circuit in conjunction with a geiger counter programmed the contingencies and set the various experimental probabilities of reinforcement. (For details see Millenson, 1963.) Responses were recorded on counters and a cumulative recorder. Experimental sessions were conducted daily providing both birds were within 15 g of their 85 per cent weights.

Procedure

(I) To establish a stable mediating sequence of behavior between two specified responses, the birds were trained initially on a procedure first described by Mechner (1958). In the present version of that procedure, reinforcement was delivered for a left key response (R_2) if and only if a sequence of *n* or more successive right key responses (R_1) preceded it. No discriminative stimulus (S^D) was given when *n* was reached, and R_2's occurring before a sequence of *n* or more R_1's resulted in a brief (½ sec) time-out (during which the key lights were extinguished) and reset the R_1 requirement to *n*. In preliminary training *n* was gradually raised from 0 to 19, where it was fixed for 30 consecutive sessions. By that time systematic changes in response rates and length of R_1 run distributions (Mechner, 1958) had ceased.

(II) After session 30 a new procedure was instituted for 25 sessions. The first reinforcement of each session was programmed as above, but for the remainder of the session each R_1 had a fixed probability (P) of changing the probability of reinforcement for R_2 from 0 to 1. Once the probability of reinforcement for R_2 became 1 it remained so until R_2 was emitted. As in the first phase above, there was no S^D given when the probability of reinforcement for R_2 changed. Occurrences of R_2 resulted in either reinforcement or ½ sec time-outs. In either case the probability of reinforcement for R_2 immediately after an R_2 was zero, and remained zero until an average of $1/P$ R_1's had occurred to return it to 1. Thus, at least one R_1 had to occur between any two consecutively reinforced R_2's, and the probability of reinforcement for a given sequence of consecutive R_2's was an exponential function of the length of the sequence. This technique simulates the contingencies of a random interval (RI) schedule where each quantal 'pause' of some duration has a fixed probability of making reinforcement possible for a response, and where the probability of reinforcement for any given IRT is an exponential function of its length (Millenson, 1963). In the present study, the role played in RI by pausing is taken by right key pecks (R_1's). The longer the pause in RI the greater the probability of reinforcement; the longer the consecutive run of R_1's in the present experiment the greater the probability that R_2 will be reinforced. In this phase P was fixed at 0.04, a value calculated (upon assumption of interaction with the observed behavior of the previous phase) to give about the same overall reinforcement rate as the previous phase.

(III) In phase II the first reinforcement was delivered according to Phase I contingencies in order to delay schedule-behavior interaction until a characteristic behavioral warmup phenomenon associated with the procedure had dissipated. In the warmup, sequences of successive R_1's occurring before the first reinforcement were systematically shorter than sequences after the first reinforcement. In Phase III the consequences of permitting this warmup portion to interact with the probabilistic contingencies were evaluated by programming the entire session under the rules described in Phase II above. Thus phases II and III differed only in the way they scheduled the first reinforcement of the session. This portion of the experiment lasted 20 sessions.

(IV) Finally, 20 sessions were run with the value of P at 0.02, all other features being identical to the previous phase.

Sessions were 20 min long except during the last phase where they lasted 40 min.

RESULTS

A succession of consecutive right key responses (R_1's) before a shift to the left key (R_2) defines a run, and the number of R_1's in a run is termed the run length. Run length in the present experiment is analogous to IRT in interval procedures. The present data are therefore appropriately analyzed in terms of fluctuations in run length correlated with the experimental treatments.

Figure 1 shows a main effect of the experiment. Mean run length is plotted over the course of the four phases. Only the last few sessions of the first phase in which runs of 19 or more consecutive R_1's were required for reinforcement (labeled 19-1) are shown. At the end of Phase I (session 30) the mean run length hovers near 19 for both birds. The probabilistic contingencies of Phase II take hold very quickly for bird 4, and its mean run length drops to about one quarter of its Phase I value in just six sessions. Thereafter through session 75 run length shows little systematic change. Bird 3 reacted more slowly but by session 45 its average run length had also declined to a low asymptote. The introduction of Phase III after session 55, permitting warmup effects to interact with the probabilistic schedule, does not appear to have affected either bird's mean run length. In view of the lack of difference between the results of Phases II and III they are merged in further commentary as simply the $P = .04$ treatment. At session 76, when P was halved, a shift to an even lower mean run length was generated by bird 4. Bird 3 showed a milder reaction to this treatment though a slight and significant (no overlap when means from the last five sessions of the two treatments respectively are compared) decline occurred.

A more detailed analysis of run lengths is available in Fig. 2 which indicates how their actual relative frequencies were distributed on the final session or sessions at each of the three treatments. The distributions labeled 19-1 are similar to previously published data on this procedure with rats (Mechner, 1958) and mice (Millenson, 1962). They are symmetrical with their mode just above the criterion of 19. The $P = .04$ distributions are shifted markedly towards short runs and their variances are narrowed compared with the 19-1 distributions. There is only slight overlap between the 19-1 and the P distributions. The distributions for $P = .02$ move still further into the short run length region and continue the variance compression begun by the $P = .04$ treatment.

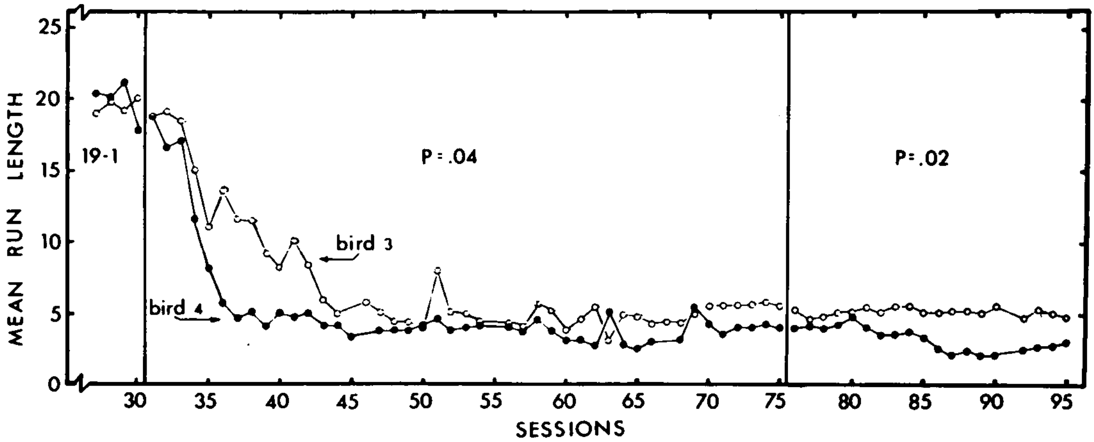

Fig. 1. Changes in mean run length over consecutive sessions for birds 3 and 4. Only the last four sessions of treatment 19-1 are shown. Occasional points are missing due to recording failures.

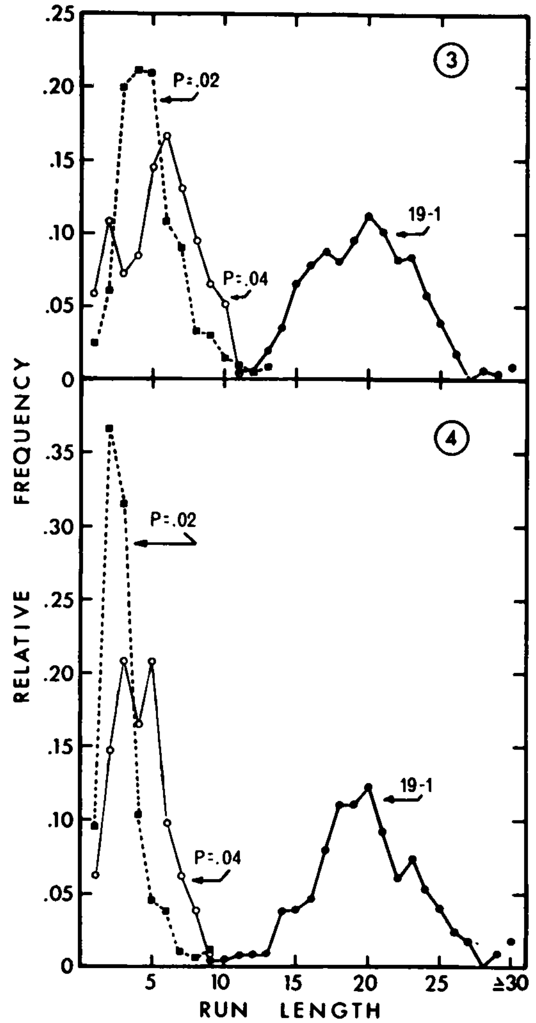

Fig. 2. Relative frequency distributions of run lengths associated with each of the three experimental treatments. The distributions for the 19-1 treatment are composites of the last four sessions (27-30) of that procedure. The distributions for P = .04 and .02 treatments represent data from the final session of that treatment.

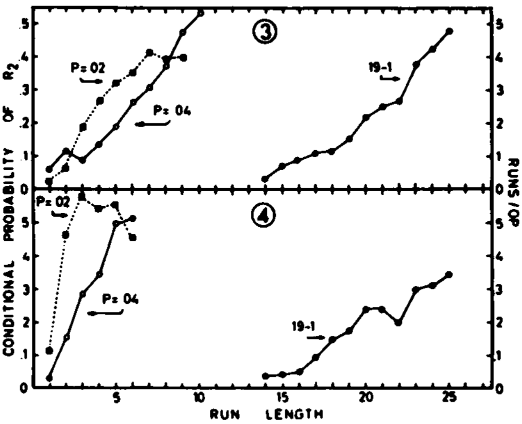

Fig. 3. Probability of R_2 given that a run of R_1's of length x have just occurred. (The right ordinate is just another name for this same dependent variable.) The data are those of Fig. 2. Points plotted are based on a minimum of 10 opportunities.

Figure 3 is a transformation of the distributions of Fig. 2 in which the conditional probability of terminating a run (*i.e.*, making R_2) is plotted against the length of the R_1 sequence already achieved. The dependent variable of Fig. 3 is a measure of run probability but takes into account the unequal opportunities for runs of different lengths. These run/op curves are analogous to IRTs/op (Anger, 1956). Under the 19-1 treatment these run probability functions show a continuous increase throughout the reliable range (15-25). The curves associated with the 19-1 phase may be described as follows: as the yet unterminated run gets longer, the chances of a shift to R_2 and hence its termination become more and more likely. The shape of these 19-1 curves resembles IRT/op functions obtained by Kelleher, Fry, and Cook (1959) from rats on DRL procedures. The $P = .04$ treatment produces a similar continuous increase in the probability that a run will be terminated as it gets longer. But the slopes are considerably steeper than those of the 19-1 curves, indicating that in addition to moving the range of the run distribution to a short run length region, the probabilistic contingency affects the form of the distribution as well. The $P = .02$ treatment maintains a still steeper slope in these runs/op functions at shortest run lengths. At $P = .02$ however, a maximum is reached after which the probability of terminating the run decreases.

Representative cumulative records for R_2 at each treatment are shown in Fig. 4. Unlike conventional cumulative response records, in Fig. 4 the metric of time has been replaced by the cumulation of successive instances of R_1 along the abcissa. This kind of presentation is appropriate because each R_1 occurrence corresponds to unit time in interval schedules. The 19-1 curves of Fig. 4, showing the low slope and the characteristic reinforcement spacing of DRL, substantiate the similarity of some of the present behavioral effects to DRL. The cumulative curves for the two P values

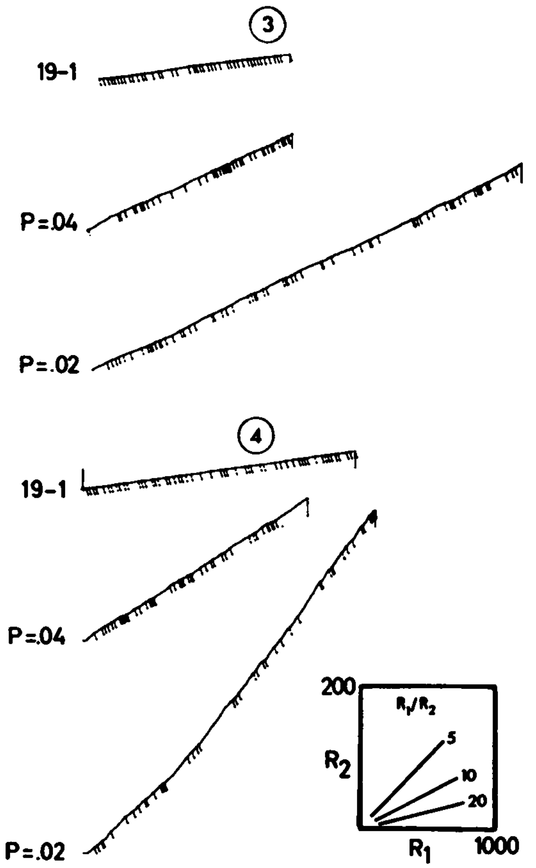

Fig. 4. Frequencies of R_2 cumulated vertically against frequencies of R_1 cumulated horizontally. Data are for the final sessions at 19-1 and $P = .04$, and for session 90 at $P = .02$. Pips indicate reinforcements.

appear qualitatively similar to records of RI behavior, with reinforcements irregularly spaced, and a near linear slope generated. The reinforcement frequencies in the curves indicate that the attempted equation of reinforcement rates for 19-1 and $P = .04$ was reasonably successful, and show that the reinforcements per R_1 at $P = .02$ were approximately halved from their value at $P = .04$.

To evaluate further the similarity of the behavioral effects of the P contingencies to those of RI, several determinations were made of the average length of runs as a function of their ordinal post-reinforcement run position. Between any two reinforcements a number of runs may intervene, only the last of which is reinforced. Averaging runs in corresponding ordinal positions from a large number of inter-reinforcement samples will determine whether the length of a run depends in any way on its ordinal position between two consecutive reinforcements. This analysis, performed on data from late sessions at $P = .04$ and .02, consistently showed run length to be independent of ordinal post-reinforcement position.

EXPERIMENT II

Experiment I indicated that RI-type contingencies applied to run lengths led to a predominance of short runs. Because long runs had a greater associated probability of reinforcement, this finding suggests that short runs, and in general short chains, are more susceptible to reinforcement. In Exp II a more direct test was made. Runs greater than a certain value were reinforced with probability of 1.0, while shorter runs were associated with a low probability of reinforcement.

Procedure

The subjects from Exp I, augmented by two additional experimentally naive pigeons, male bird 1, and female bird 5, were used. The apparatus and general details were those of Exp I.

The four birds were exposed to 65 sessions of a procedure identical to phase 1 of Exp I, with the run length requirement set at 25. (Bird 1 failed to work consistently at this value and its requirement was reduced to 20 midway through this training.) Following this treatment, runs between 5 and 25 (5 and 20 for bird 1) continued to be reinforced with probability = .05, and runs $\geqq 25$ ($\geqq 20$ for bird 1) continued to be reinforced with probability = 1. This final treatment was in effect for 12 sessions. Throughout the experiment, sessions lasted until 45 reinforcements had been obtained.

Results

The stabilized distributions obtained after 65 sessions of 25-1 (20-1 for bird 1) are shown as solid histograms in Fig. 5. These are similar to the 19-1 histograms of Fig. 2. The intervening history with the RI analogue given birds 3 and 4 apparently did not prevent the recovery of typical behavior for this procedure. The change in behavior that occurred for three of the birds when a small probability of reinforcement was applied to runs between 5 and 25 (5 and 20 for bird 1) is shown by the shaded histograms of Fig. 5. The distributions for birds 1, 4 and 5 show a marked shift leftward to the region of very short runs. By the last session of this treatment none of these three birds showed runs in the region where they would have been reinforced each time, *i.e.*, $\geqq$ 25 ($\geqq$ 20 for bird 1). In fact, the distribution of bird 4 shows little overlap with its former 25-1 distribution. This shift in run lengths is correlated with a loss in efficiency in terms of total response output per reinforcement. Birds 1, 4, and 5 emitted approximately two to three times as many total responses ($R_1 + R_2$) per reinforcement as they did before their shift.

Bird 3 failed to react to the .05 probability condition showing little change in run length distribution when $P = .05$ was instituted between 5 and 25.

DISCUSSION

When sequences (runs) of responses were reinforced by contingencies analogous to those that RI sets for IRTs (Exp I), a characteristic pattern and distribution of runs emerged. The stabilized form of this behavior consisted principally of very short runs. Furthermore, over an appreciable range, the conditional probability of occurrence of a run (runs/op) increased as a function of run length. Though short runs came to be favored by the birds, these runs had a small probability of reinforcement; hence, relatively few of them were reinforced compared with longer runs.

The emergence of short runs as the dominant behavioral pattern supports the conclusion that the reinforcement susceptibilities of the two sorts of runs are different. This conclusion is supported by the results of three birds in Exp II.

It might be supposed that the birds exposed to probabilistic reinforcement of their runs followed the law of least effort and optimized work output per reinforcement. If some arbitrary cost or effort is assigned to emitting R_1, and another arbitrary cost or effort is assigned to switching to R_2, there will exist an optimum run length that will minimize the average effort per reinforcement. The optimum run length will be a function of the respective response costs (presumably determined by the relative proximity of the two operandi, their intensive requirements, *etc.*) and the value of P. In the present experiments, optimization as a description of the behavioral effects seems to be ruled out because the birds' failure to increase average run lengths when P was reduced from .04 to .02 is contrary to an optimization prediction. Similarly in Exp II, optimization again fails to predict the results for three of the four birds that declined in efficiency.

In Exp I similarities between the responding at the 19-1 treatment and responding on DRL were noted. These similarities suggest that in certain schedules the metrics of response and time may be interchangeable. Experimentally, the ability to substitute responses for time would be desirable since behavior, unlike time, provides a set of topographical and intensive variables which lend themselves to manipulation and measurement.

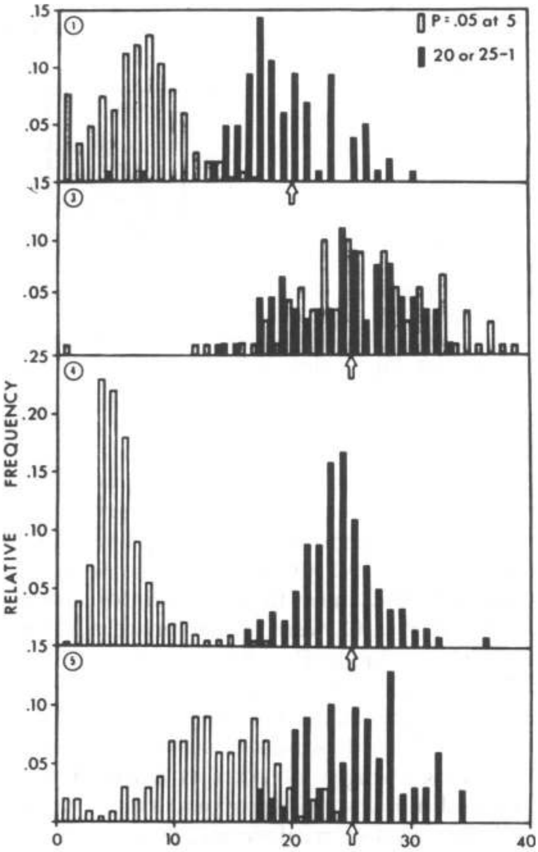

Fig. 5. Relative frequency distribution of run lengths with probability of reinforcement = 1 for runs $\geqq$20 or $\geqq$25 (solid bars), and with a treatment (shaded bars) that reinforced, in addition, runs between 5 and 25 (or 5 and 20) with $P = .05$.

Whenever the substitution is successful the temporal effects of schedules become special cases of the behavioral effects of reinforcing chains of behavior in various ways. In particular, the interpretation of the present experiment is that long and short homogeneous chains have different reinforcement susceptibilities. How far can this formulation be extended to account for the IRT characteristics of VI discussed in the introduction? The application depends critically on the extent to which the probabilistic contingencies of Exp I match in their behavioral effects the IRT effects of VI, or more accurately, RI. Five correspondences may be extracted from the results.

(1) Mean run length, like mean IRT under VI, declines upon first exposure to the probabilistic contingencies.

(2) On both VI schedules and the procedure of Exp I this decline stops well short of its limiting value. On VI, IRTs remain considerably longer than IRTs generated by random ratio contingencies and other schedules which differentially reinforce high rates. Similarly, at both P values, in Exp I, stabilized mean run length stopped well short of its limiting value of 1.

(3) IRT/op functions from RI and runs/op functions (Fig. 3) exhibit similar forms. Distributions of IRTs were obtained by Farmer (1962, 1963) at various mean RI values. A number of Farmer's distributions were perturbed by a periodic artifact (Farmer, 1963, p 615) which makes them unsuitable for comparison, but two in particular (bird 12, RI 960 sec and bird 5, RI 120 sec) were unimodal and appeared free from such artifacts. IRT/op functions calculated from these data rise systematically from short IRTs to the modal IRT, then level off for the remainder of the reliable IRT range, closely resembling the form of the runs/op functions for $P = .02$ in the present experiment.

(4) The independence of mean run length from its post-reinforcement run position parallels the IRT effect in RI. Catania and Reynolds (*in press*) have shown that on a schedule closely approximating RI, the durations of IRTs produced are approximately independent of post-reinforcement time.

(5) The irregular spacing of reinforcements between mediating sequences of R_1's seems to resemble closely the kind of irregular spacing in time between reinforcements produced by VI schedules.

The parallels cited between IRTs and run lengths suggest again that in certain respects the metrics of time in schedules and responses in probabilistic chains may be interchangeable. The properties of mediating chains can account for a number of the characteristic features of variable interval performance.

A major disparity between run length effects of the probabilistic contingencies and actual VI performance is associated with the effect of decreasing P from .04 to .02 in Exp I. This operation is analogous to doubling the mean VI on interval contingencies and comparable to halving P in RI (*cf.* Farmer, 1963). In both VI and RI the behavioral effect of such a change is typically to lower the rate, or put another way, to raise the mean IRT (Catania and Reynolds, *in press*; Farmer, 1963). In Exp I however, the mean run length slightly declined from $P = .04$ to .02. To assess in detail differences between the effects of P in RI and P in the present procedures will require exploration of a wider range of P values than was studied here. If the decline in mean run length between $P = .04$ and .02 is genuine, a non-monotonic function must relate P to mean run length. Such a function would be necessary in view of the result from simple chaining ($P = 1$) where it is well established that the probability of the run going beyond a single emission of R_1 is virtually zero (Millenson and Hurwitz, 1961). One possibility is that the relations between P and IRT, or P and run length, are not independent of the path by which they are obtained. In the present study the probabilistic contingencies were reached by way of a previous history with a required 19-1 sequence.[2] Interval contingencies are generally reached by way of a previous history of CRF. Since the 19-1 procedure appears very like DRL, the question is raised as to whether VI after CRF and VI after DRL produce identical effects.

REFERENCES

Anger, D. The dependence of interresponse times upon the relative reinforcement of different inter-

---

[2] In preliminary work, attempts to institute the probabilistic contingencies after shaping the simple chain of $R_1 + R_2 \rightarrow S^R$ were without effect on the fixed sequence of R_1 and R_2 which had been welded fast by the history of CRF.

response times. *J. exp. Psychol.*, 1956, **52**, 145-161.

Brandauer, C. M. The effects of uniform probabilities of reinforcement on the response rate of the pigeon. Unpublished doctoral dissertation, Columbia Univ., 1958.

Catania, A. C. and Reynolds, G. S. A quantitative analysis of the behavior maintained by interval schedules of reinforcement. *J. exp. Anal. Behav.*, in press.

Farmer, J. Properties of behavior under random interval reinforcement schedules. Unpublished doctoral dissertation, Columbia Univ., 1962. An abbreviated report with the same title appears in *J. exp. Anal. Behav.*, 1963, **6**, 607-616.

Kelleher, R. T., Fry, W., and Cook, L. Interresponse time distribution as a function of differential reinforcement of temporally spaced responses. *J. exp. Anal. Behav.*, 1959, **2**, 91-106.

Malott, R. W. and Cumming, W. W. Schedules of interresponse time reinforcement. *Psychol. Rec.*, 1964, **14**, 211-252.

Mechner, F. Probability relations within response sequences under ratio reinforcement. *J. exp. Anal. Behav.*, 1958, **1**, 109-121.

Millenson, J. R. Acquired counting behavior in mice maintained under two reinforcement procedures. *Animal Behav.*, 1962, **10**, 171-173.

Millenson, J. R. Random interval schedules of reinforcement. *J. exp. Anal. Behav.*, 1963, **6**, 437-443.

Millenson, J. R. and Hurwitz, H. M. B. Some temporal and sequential properties of behavior during conditioning and extinction. *J. exp. Anal. Behav.*, 1961, **4**, 97-106.

Newman, E. B. and Anger, D. The effect upon simple animal behavior of different frequencies of reinforcement. Report PLR-33, Office of the Surgeon General, 1954.

Ray, R. C. and McGill, W. Effects of class-interval size upon certain frequency distributions of interresponse times. *J. exp. Anal. Behav.*, 1964, **7**, 125-127.

Revusky, S. H. Mathematical analysis of the durations of reinforced interresponse times during variable interval reinforcement. *Psychometrika*, 1962, **27**, 307-314.

Received December 21, 1964