

## Sequences of reward magnitude and runway performance\*

EDWARD L. WIKE and DENNIS D. KING

*University of Kansas, Lawrence, Kansas 66044*

Three groups of rats underwent 24 days of training and 12 days of extinction (three trials per day) in a runway under conditions of increasing (I), decreasing (D), and random (R) sequences of reward magnitudes (0, 45, and 500 mg). The I Ss ran faster over the daily trials, the D Ss slowed down, and the R Ss ran at approximately equal speeds on each trial. The patterned running observed in training persisted in extinction, with the R Ss running fastest and the I Ss next. The results were discussed in terms of Capaldi's sequential theory and Amsel's frustration theory.

It has become clear in recent years that the *sequence* of rewarding events is an important parameter influencing discrete-trial instrumental conditioning (Capaldi, 1967). One principle that has emerged from sequential investigations is the following: The reward sequence of a less preferred reward followed by a more preferred reward produces greater resistance to extinction than the reverse reward sequence. Both the generality of this principle and the consistency of the supporting evidence are impressive. It is valid for partial reinforcement (Spivey & Hess, 1968; Mackintosh & Little, 1970), for delayed reward (Campbell, 1969; Capaldi, Godbout, & Ksir, 1968), for duration of delayed reward (Wike & Atwood, 1970), for quality of reward (Russin, Boyer, & Cross, 1970; Boyer, Russin, & Cross, 1971), and for magnitude of reward (Bowen, 1968; Leonard, 1969a, b; Leonard, Albin, & Lebowitz, 1969; Wike & Atwood, 1970).

While the scope and consistency of the results supporting the foregoing sequential principle are evident, the processes involved are not. In a previous runway study (Wike & Atwood, 1970), increasing (I) vs decreasing (D) daily schedules of reward magnitude (one, two, and four units of food) and delay duration (0, 20, and 40 sec) were manipulated. The I schedule of reward magnitude and the D schedule of delay duration resulted in superior performance in both acquisition and extinction. More importantly, it was observed that the reward schedules produced patterned running. For example, the Ss with I reward magnitude speeded up over the daily trials, while the D Ss slowed down, and these patterns of responding continued during extinction.

The general aim of the present study was to investigate more precisely and in a more detailed fashion the effects of I and D sequences of reward magnitude upon runway acquisition and extinction. The experiment differed in four important respects from the previous investigation of Wike and Atwood. (1) The range of reward magnitudes (0, 45, and 500 mg) was wider than that employed in the Wike-Atwood study (one, two, and four sugar cereal pellets) in an effort to

reveal magnitude effects more clearly. (2) In contrast to the original study in which two-thirds of the acquisition trials involved delayed reward and two-thirds of the extinction trials involved delay-box confinements, all rewards were given without delay. Immediate rewards were imposed in order that the effects of reward magnitude might be assessed without the possible confounding by delay in training and confinement in extinction. (3) A control group with random (R) sequences of reward magnitudes was included in the design to verify the impression from previous studies that invariant sequences like I and D generate less resistance to extinction than an R sequence. (4) Most importantly, the resulting performance was examined on a trial by trial basis in order to determine whether or not patterned running occurred in the I and D groups in acquisition and extinction. Capaldi (1970, p. 383f) has asserted that a primary assumption of sequential theory is that what is learned in acquisition is determined by the specific sequences in which trials are experienced by the organism. If different things are learned as a result of different sequences, it follows that under optimal reward conditions these different learnings should be revealed when the data are subjected to detailed scrutiny. Furthermore, if extinction is a transfer of training paradigm, then it follows that the different learnings in acquisition should be evident also in extinction. The assessment of these implications from sequential theory by means of a trial by trial examination of performance in acquisition and extinction was the main purpose of the present study. Finally, it should be noted that, since the smallest reward magnitude was no reward, the study could be conceptualized as a partial reinforcement experiment in which nonreward was presented on the first or last daily trial or occurred randomly.

### METHOD

#### Subjects

The Ss were 36 naive male Holtzman albino rats, about 75 days old at the beginning of acquisition. They were housed in standard individual Wahmann cages and were randomly and equally assigned to the three reward schedules.

#### Apparatus

The apparatus was a 92.5-in.-long black runway that was 4 in.

\*This research was supported by a grant from the General Research Fund of the University of Kansas.

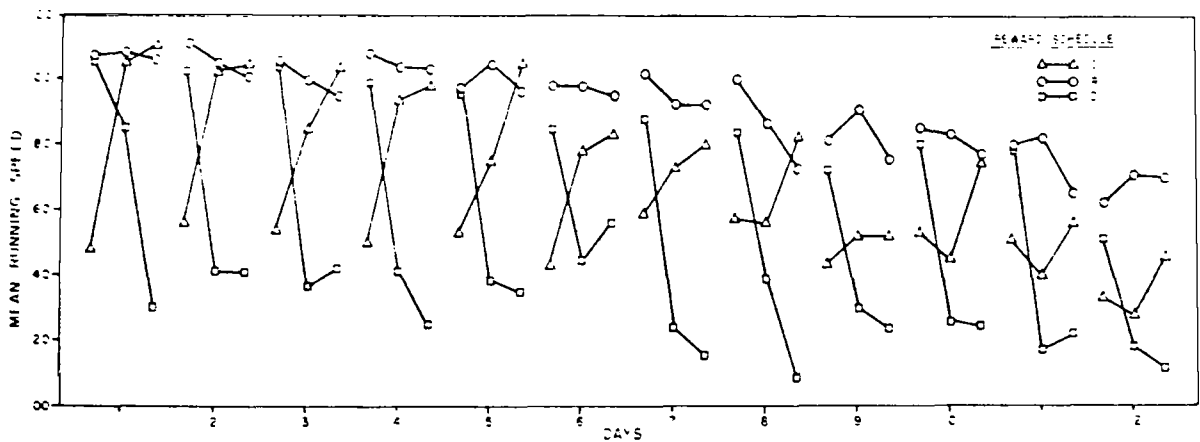


Fig. 1. Mean running speeds on the three daily trials for the three reward groups during extinction.

wide and 7.5 in. high. The first 14 in. served as a startbox and the runway was attached to a 9 x 10 in. black goalbox. Starting and running times were measured by two Hunter Klockcounters. The starting time was the period from the elevation of the startbox door to S's breaking a light beam 1 ft from the door; running times were recorded over the next 51 in. of the runway.

#### Procedure

In the first 7 days of pretraining, the Ss were on an ad lib feeding and watering schedule and were handled daily. Over the next 2 weeks, the Ss' body weights were reduced to 80% of their normal levels. On the following day, the Ss explored the runway for 10 min in groups of six. During exploration, the small metal reward pan was filled with 45-mg Noyes pellets. On the next 2 days, each S had two timed runway trials. On the trials the reward pan contained one 500-mg and one 45-mg pellet.

Following pretraining, the Ss had 24 days of training, with three trials per day. The trials were rotated among six Ss, with two Ss from the I, D, and R groups constituting a squad. The intertrial interval in training was approximately 5-6 min. The Ss in the I group received no reward on Trial 1 (but were confined to the goalbox for 30 sec), a 45-mg pellet on Trial 2, and a 500-mg pellet on Trial 3. On Trials 2 and 3 the Ss were removed from the goalbox immediately after eating the food pellet. The Ss in the D group received the same rewards in the reverse sequence. The Ss in the R group had the six possible orders of the three rewards in four different randomly selected sequences.

Extinction consisted of three trials per day for 12 days. The Ss were confined to the goalbox for 30 sec on every extinction trial and were run in squads of six Ss as in acquisition. The reward pan was removed from the goalbox during extinction. The Ss had a 60-sec criterion for entry into the goalbox and were maintained at 80% of their normal body weights throughout training and extinction.

## RESULTS

Each S's starting and running times in training and extinction were transformed to reciprocals. These transformed scores were subjected to analyses of variance, with the reward schedules as a between factor and days and trials as within factors. To simplify the results, only the running speeds will be presented. The results for starting speeds were in the same directions as those for running speeds but the starting speeds were generally more variable. The running speeds for the three

groups in training are displayed on a trial to trial basis over days in Fig. 1. An analysis of variance of these data revealed five significant sources of variation: reward schedules, days, trials, Reward Schedules by Trials, and Reward Schedules by Trials by Days ( $F_s = 13.56, 29.41, 10.85, 43.86,$  and  $2.52$ , respectively). Sources of variation involving the within factors of days, trials, or both were evaluated by both regular and Box's conservative F tests (Wike, 1971). The first four effects were significant at less than the .005 level; the three-way interaction of Reward Schedules by Trials by Days was significant at less than the .001 level by a regular F test and at less than .10 by a conservative F test. Since Box's conservative F test was designed to compensate for maximum heterogeneity of variances and covariances, we may safely reject the null hypothesis in the case of the three-way interaction. The nature of the Reward Schedules by Trials interaction is apparent from Fig. 1: The I Ss speeded up over the daily trials, the D Ss slowed down, and the R Ss ran at approximately the same speed on all trials. These patterns of running developed early in training and remained stable thereafter. The interaction of Reward Schedules by Trials by Days, an effect of smaller magnitude, was the consequence of the lack of patterned running on the first day and the emergence of different patterns of running by the three groups on subsequent days. While these two interactions are the most important findings in acquisition, it should be noted that Tukey b tests of the overall group means revealed that the R and I groups ran significantly ( $p < .01$ ) faster than Group D but did not differ reliably from each other.

The running speeds in extinction are displayed in Fig. 2. An analysis of variance of these data revealed six significant effects: reward schedules, days, trials, Reward Schedules by Trials, Days by Trials, and Reward Schedules by Trials by Days ( $F_s = 29.75, 39.68, 14.15, 71.07, 3.36,$  and  $3.14$ , respectively). The first four effects were significant at less than the .001 level; the latter two effects were significant at less than the .001 level by regular F tests and at less than the .10 level by

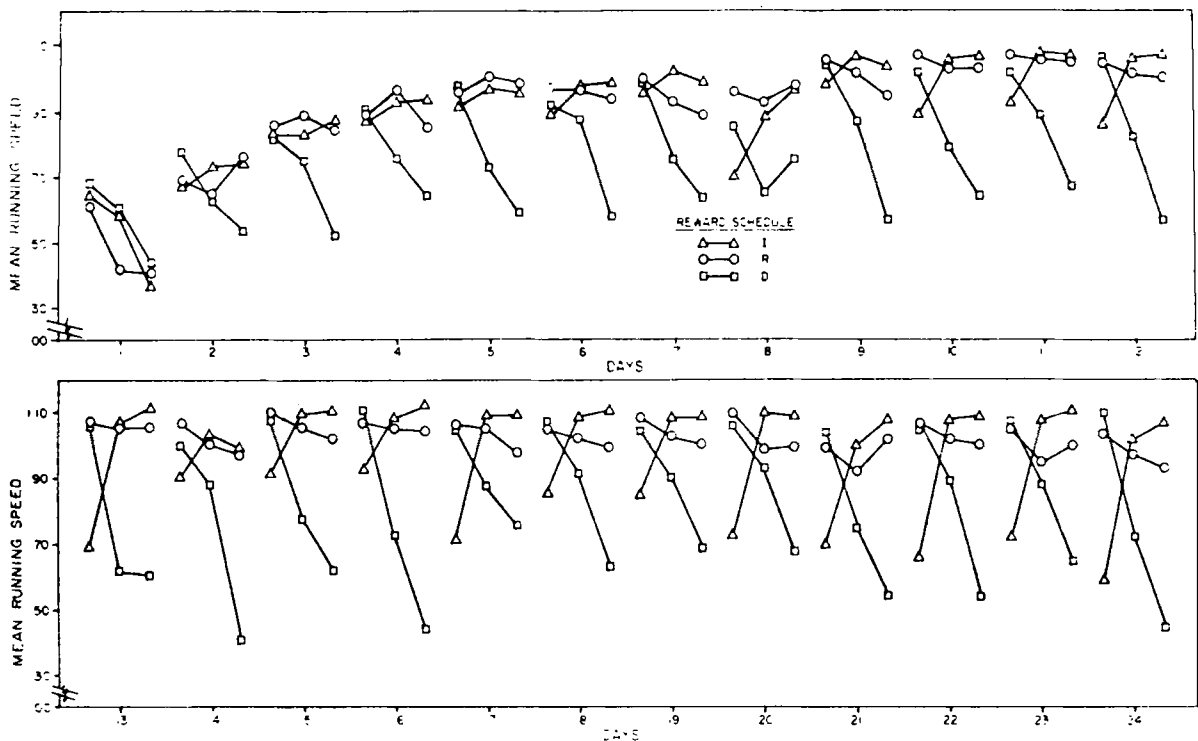


Fig. 2. Mean running speeds on the three daily trials for the three reward groups during training.

conservative *F* tests. For the reason offered above, the null hypothesis regarding the last two interactions can be rejected with some assurance. Again, the basis for the interaction of Reward Schedules by Trials is clearly revealed in Fig. 2; the unique patterns of running that developed in training in the three groups continued to be exhibited throughout extinction. The smaller magnitude interaction of Reward Schedules by Trials by Days resulted from some changes in the running patterns by the three groups over days; for example, the differences between R and I and D increased and the patterning in the I group became less distinctive with continued extinction. It should be pointed out also that on the average the R Ss ran faster ( $p < .01$ ) in extinction than both the I and D Ss and the I Ss were faster than the D Ss ( $p = .01$ ). Finally, a result that was not evaluated statistically but that is obvious by inspection of both Figs. 1 and 2 is that the performance of the R group in both acquisition and extinction was far less variable than that of the I and D groups.

## DISCUSSION

The finding that the I Ss ran significantly faster in extinction than D Ss offers additional support for the sequential principle stated above, i.e., that the reward sequence of a less preferred reward followed by a more preferred reward produces greater resistance to extinction than the reverse reward sequence. Capaldi's theory (1967) provides an explanation for this extinction result. The theoretical situation for I vs D

reward magnitude for two daily trials of training has been described by Leonard (1969a). In the case of I reward, the instrumental response to the aftereffect of Trial 1, i.e.,  $S^{R_s}$ , is conditioned on Trial 2 by a large reward. On the other hand, for D reward the response to the aftereffect of Trial 1, i.e.,  $S^{R_L}$ , is followed by a small reward. Since habit (*H*) is assumed by Capaldi to be a function of reward magnitude,  $S^{R_s}$ , an aftereffect on the right-hand side of the Capaldi stimulus continuum, should acquire greater *H* than  $S^{R_L}$ , an aftereffect on the left of the stimulus continuum. Assuming linear gradients of generalization with the same slope, the I group with greater generalized habit strength (*H*) should run faster in extinction. Extension of this account to the present experiment with three trials per day entails an assumption that in the I group the aftereffects of  $S^N$  and  $S^{R_s}$ , conditioned by S and L rewards, respectively, acquire greater summated *H* than an  $S^{R_L}$  conditioned by S rewards in the D group. Such an assumption does not appear unreasonable since (1) for the D group an aftereffect on the left of the stimulus continuum,  $S^{R_L}$ , and for the I group one to the right,  $S^N$ , possess equal *H* and (2) the aftereffect,  $S^{R_s}$ , is conditioned by L rewards in the I group but not conditioned in the D group.

Since the R group encountered a mixture of sequences, its superior performance in extinction is a more difficult problem to explain. Bowen (1968) and Leonard et al (1969) have proposed that an R group should run faster in extinction than I and D groups, because when sufficient training is given to produce

asymptotic performance the summation of Hs in the R group would be greater than for the H for two groups with invariant sequences. Inspection of Fig. 1 reveals that the R group did, in fact, run rapidly and at approximately equal speeds on all trials. The determination of which of the six sequences or combinations of sequences employed in the R group are critical in producing superior performance can only be learned by fractionation experiments. This matter obviously requires further study, since Leonard et al found that an R group responded faster in extinction than I and D groups after only 20 training trials. These interpretations of our results with Capaldi's theory also make it evident that, in the fuller development of the theory and in its extension to sequences of rewarding events involving more than two daily trials, the question of the validity of Hull's rule for summation of Hs will become crucial.

The most impressive result in the present study was the occurrence of patterned running in training and its persistence during extinction. This patterning was more clear-cut and stable in the present study than in the Wike-Atwood study (1970). In all likelihood, this difference was the consequence of using a much wider range of reward magnitudes than that employed in the first investigation. Most of the previous studies of reward sequences have not reported patterning (see Mackintosh & Little, 1970 for an exception). In part, such absences of patterning may be the outcome of limited training, the use of limited ranges of rewards, or changes in trial distribution from training to extinction. In most cases, however, whether or not patterns occurred is simply indeterminate in the published reports, because the investigators have aggregated trials over days or blocks of days rather than presenting their data on a trial by trial basis. Our data suggest that different reward schedules do not produce faster or slower speeds of running in general but rather distinctive *patterns* of running that are stable and that carry over into extinction. While an explanation of these behaviors was offered above in terms of Capaldi's theory, an alternative explanation is possible with Amsel's frustration theory (1972). Patterned running in training is a reflection of different magnitudes of frustration and incentive motivation on different training trials, e.g., decreases in reward magnitude or nonreward generate frustration and increases in reward magnitude generate increases in incentive motivation. Once patterned running is strongly conditioned, the patterns continue in extinction, being mediated by the frustration stimuli attendant upon the complete omission of reward in extinction. There are two difficulties with this interpretation. (1) If the slowing down of the D Ss over the daily trials is attributed to increasing  $r_F$ , then the D group with

greater frustration should have greater resistance to extinction than the I group. (2) If the patterned running is explained in terms of  $r_G$  and  $r_F$ , then the conditioning of these mediating reactions must be more rapid than has been supposed in order to account for the appearance of patterning so early in training.

## REFERENCES

- Amsel, A. Behavioral habituation, counterconditioning and a general theory of persistence. In A. H. Black and W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research*. New York: Appleton-Century-Crofts, 1972.
- Bowen, J. Effect of ascending, descending, and irregular order of varied reward magnitude. *Psychonomic Science*, 1968, 12, 209-210.
- Boyer, W. N., Russin, R., & Cross, H. A. Extinction in a straight alley as a function of varied quality and quantity of reward sequence. *Psychonomic Science*, 1971, 23, 92-93.
- Campbell, P. E. Sequences of delayed reward and extinction confinement: Effects of pattern running and extinction performance. *Psychonomic Science*, 1969, 17, 34-35.
- Capaldi, E. J. A sequential hypothesis of instrumental learning. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation*. Vol. 1. New York: Academic Press, 1967.
- Capaldi, E. J. An analysis of the role of reward and reward magnitude in instrumental learning. In J. H. Reynierse (Ed.), *Current issues in animal learning*. Lincoln: University of Nebraska Press, 1970.
- Capaldi, E. J., Godbout, R. C., & Ksir, C. A comparison of two delay of reward procedures, pre-reinforcement delay vs post-reinforcement delay. *Psychonomic Science*, 1968, 13, 279-280.
- Leonard, D. W. Amount and sequence of reward in partial and continuous reinforcement. *Journal of Comparative & Physiological Psychology*, 1969a, 67, 204-211.
- Leonard, D. W. Varied reward sequence and resistance to extinction in the straight alley. *Psychonomic Science*, 1969b, 16, 129-130.
- Leonard, D. W., Albin, R., & Lebowitz, M. Performance under massed or spaced extinction following different sequences of varied reward training. *Psychonomic Science*, 1969, 16, 130-132.
- Mackintosh, N. J., & Little, L. Effects of different patterns of reinforcement on performance under massed and spaced extinction. *Psychonomic Science*, 1970, 20, 1-2.
- Russin, R., Boyer, W. N., & Cross, H. A. Extinction as a function of varied quality and quantity of reward sequence in a straight alley. *Psychonomic Science*, 1970, 21, 274-275.
- Spivey, J. E., & Hess, D. T. Effect of partial reinforcement trial sequences on extinction performance. *Psychonomic Science*, 1968, 10, 375-376.
- Wike, E. L. *Data analysis: A statistical primer for psychology students*. Chicago: Aldine-Atherton, 1971.
- Wike, E. L., & Atwood, M. E. The effects of sequences of reward magnitude, delay, and delay-box confinement upon runway performance. *Psychological Record*, 1970, 20, 51-56.

(Received for publication November 6, 1972;  
revision received March 16, 1973.)