

Effects of varied and partial reward on discrete-trial patterning of rats*

J. M. BLOOM**, DONALD T. WILLIAMS†, and LEROY P. METZE††

Texas Research Institute, 1300 Moursund Avenue, Houston, Texas 77025

Rats were trained in the discrete-trial operant apparatus with single alternation of large reward and small reward or large reward and nonreward, followed by either transfer or extinction. The results showed that both groups acquired appropriate response patterning, that patterning is conditioned, and that the stimuli which control patterning are derived from reward conditions on immediately preceding trials. A modification of the sequential hypothesis of instrumental learning was proposed to account for the results.

It is well documented that rats can pattern their responses appropriately to a single alternation pattern of reward and nonreward, i.e., fast responding on rewarded trials and slow responding on nonrewarded trials. The effect has been demonstrated for the runway response (e.g., Bloom & Capaldi, 1961; Capaldi & Stanley, 1963; Bloom & Malone, 1968), for free-response barpressing (Bloom & Smith, 1965), and for discrete-trial barpressing (Wall & Goodrich, 1964; Gonzalez, Bainbridge, & Bitterman, 1966).

The purpose of this experiment was to determine if rats can also pattern discrete-trial bar responses with single alternation of large and small reward and, if so, to determine how such patterning is related to that with reward and nonreward.

METHOD

Subjects

The Ss were 32 naive female albino rats, approximately 100 days of age. They were purchased from the Holtzman Company, Madison, Wisconsin.

Apparatus

The apparatus consisted of two standard single-lever retractable-bar operant conditioning cages (Scientific Prototype No. A-100) with pellet dispensers, enclosed in blower-ventilated sound- and light-resistant chambers (Scientific Prototype No. SPC-300). Bar insertion and retraction time was approximately 2 sec, and the bars were modified so that they could not be depressed prior to full insertion or after initiation of bar retraction. Latency was automatically timed and printed out to the nearest tenth of a second.

Procedure

Pretraining consisted of 14 days of 23-h food deprivation and handling and 1 day of magazine training, consisting of 15

continuous reward one-pellet (Noyes 45 mg) barpresses with the bar always present.

In the acquisition phase, Group 5-0 received 21 trials each day for 21 days. Each trial consisted of bar insertion, a barpress or 15 sec time without a press, and bar retraction. Barpresses produced a five-pellet reward on odd-numbered trials and nonreward on even-numbered trials. Group 5-1 received identical training, except that on even-numbered trials a barpress produced a one-pellet reward.

Following acquisition, each group was divided into three matched subgroups on the basis of responding on each type of trial on the last acquisition day. Then Groups 5-0-0 and 5-1-1 (five Ss each) continued on the acquisition schedule; Group 5-0-1 (five Ss) was transferred from 5-0 to 5-1; Group 5-1-0 (five Ss) was transferred from 5-1 to 5-0; and Groups 5-0-E and 5-1-E (six Ss each) were given extinction, which consisted of 21 nonrewarded trials each day. The transfer and extinction phases continued for 9 days.

All training was carried out under 23-h food deprivation, with water always available in the home cage and the apparatus. Throughout training, there were two intertrial intervals of 15 and 45 sec, and each interval occurred equally often following each type of trial.

RESULTS

An inspection of the data revealed that, on the last day of acquisition, all Ss were alternating their responses in a remarkably stable manner. For analysis, the first daily trial was excluded and median response latency on each type of trial for each S was determined. The medians were then transformed to Log X. Group means are shown in Fig. 1. Appropriate response patterning developed in both groups, with response latency on large reward trials decreasing, then increasing, then decreasing again. Latency on small and nonreward trials decreased and then increased.

Analysis of variance on Days 1-10 showed that patterning developed faster in Group 5-0 than in Group 5-1 (Day by Group by Reward condition, $F = 2.24$, $df = 9/270$, $p < .05$). Analysis of the data of the last 3 acquisition days showed that the difference between responses on alternating trials was greater in Group 5-0 than in 5-1 ($F = 13.08$, $df = 1/30$, $p < .01$). There was no difference between groups on large reward trials ($F < 1$).

*Supported by National Institute of Mental Health Grant 1 R03 MH-14174-01.

**Requests for reprints should be sent to J. M. Bloom, 9234 Wickford, Houston, Texas 77024.

†Now at Indiana State University.

††Now at Western Kentucky University.

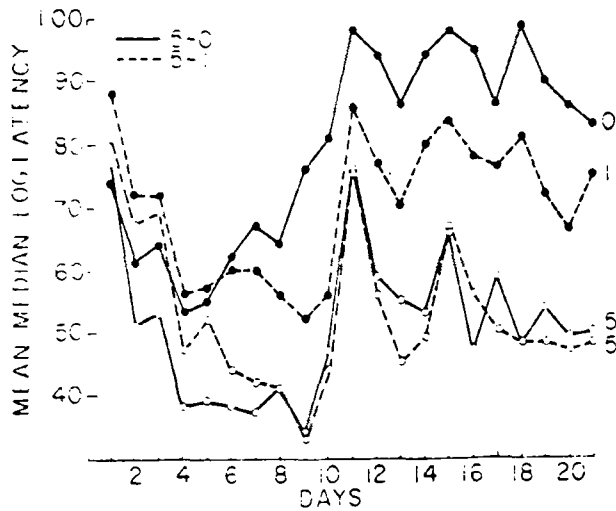


Fig. 1. Acquisition responses on large reward, small reward, and nonreward trials.

The transfer data showed no changes in any group from response levels at the end of acquisition on large reward trials ($F_s < 1$). However, as shown in Fig. 2, there were changes on small reward and nonreward trials in the transfer groups. A shift from nonreward to small reward in Group 5-0-1 produced a gradual decrease in latency to the level of the 5-1-1 control group. Likewise, a shift from small reward to nonreward in Group 5-1-0 produced a gradual increase in latency to the level of the 5-0-0 control group. Analysis of variance on data of the first transfer day for small and nonreward trials showed no difference between Groups 5-0-0 and 5-0-1, nor between Groups 5-1-1 and 5-1-0 ($F_s < 1$).

The extinction data are also shown in Fig. 2. Group 5-0-E responded slower than Group 5-1-E on the first 3 days of extinction but faster on the last 6 days. There was no difference between groups in level of extinction ($F = 1.44, df = 1/10$). However, Group 5-0-E extinguished at a slower rate than Group 5-1-E, as indicated by the significant Group by Trials interaction ($F = 3.36, df = 8/80, p < .01$).

A close inspection of Fig. 2 shows further that the extinction groups responded faster in the initial stage of extinction than did control and transfer groups on small and nonreward trials. Since the extinction and transfer phases of the experiment were carried out simultaneously, it was possible to test these differences by means of analyses of variance. The analyses showed that, over the first 4 days of transfer and extinction, Group 5-0-E responded faster on nonreward trials in extinction than combined Groups 5-1-0 and 5-0-0 on nonreward trials in transfer ($F = 15.18, df = 1/14, p < .01$). Likewise, Group 5-1-E responded faster on nonreward trials than combined Groups 5-0-1 and 5-1-1 on small reward trials ($F = 10.75, df = 1/14, p < .01$).

DISCUSSION

The acquisition results indicate that rats can pattern

their responses appropriately to a single alternation pattern of large and small reward and that the development of such patterning is quite similar to that with large reward and nonreward.

In both conditions, on large reward trials there was a decrease, an increase, and then a decrease in response latency. On small reward and nonreward trials, there was a decrease, then an increase in latency. The terminal acquisition differences in patterning appear to be related to the amount of reward on concurrent trials. Terminal response levels were fastest and quite similar on large reward trials in both conditions, slower on small reward trials, and slowest on nonreward trials.

The transfer results indicate that these response patterns are learned. Changes in amount of reward produced slow, gradual changes in response latency. A transfer from nonreward to small reward resulted in a gradual adjustment of responding approaching that of the small reward control level, and transfer from small reward to nonreward produced a similar adjustment. This type of result has traditionally been taken to indicate changes in learning rather than changes in performance.

Capaldi's (1967) sequential hypothesis provides the

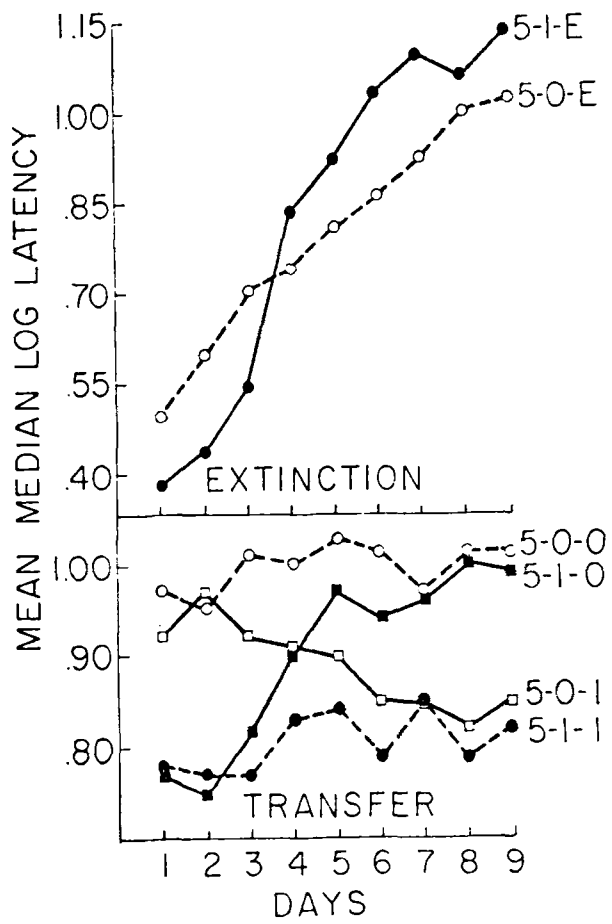


Fig. 2. Extinction responses (upper panel) and transfer responses on small reward and nonreward trials (lower panel).

basis for pattern learning. Large reward, small reward, and nonreward are assumed to give rise to distinctive stimuli, S^R , S^r , and S^N , respectively. These stimuli remain functional until the next trial and are available for conditioning by reinforcement to the instrumental response. The amount of habit strength (H) is a positive function of amount of reward, and H generalizes to stimuli that are similar to the conditioned stimulus. Nonreward following S^R or S^r produces growth of a negative inhibitory process (I), which, by implication, is also conditioned and which also generalizes to other stimuli. Primary and generalized H and I summate algebraically to produce effective strength of response in the presence of a particular stimulus. Finally, S^N is modified by successive nonrewards to produce S^{N2} , S^{N3} , S^{N4} , etc., according to the number of successive nonrewards in the chain.

Conceptually, the present results support a sequential interpretation of pattern learning in terms of differential reinforcement. In Group 5-0, S^N is consistently followed by large reward, which eventually produces fast responding on trials following S^N ; S^R is consistently followed by nonreward, which produces slow responding on trials following S^R . In Group 5-1 the process is similar, except that small reward and S^r are involved rather than nonreward and S^N . However, terminal acquisition response levels are different on trials following S^R in the two groups, because S^R is followed by small reward in Group 5-1 and by nonreward in Group 5-0. Terminal levels are quite similar in the two groups on trials following S^r and S^N because each is followed by the same large reward.

In transfer conditions, the changes in response levels on small and nonreward trials represent changes in response strength on trials following S^R , since S^R is now followed by a different amount of reward than occurred in acquisition. Note that a major assumption is that the effective stimuli for responding on a given trial in patterning derive not from the trial itself, but from the reward conditions of the immediately preceding trial.

The extinction results support this assumption. Initial extinction response level in Group 5-0 was quite similar to responding on large reward trials at the end of acquisition and much faster than responding on nonreward trials during the transfer phase. This seemingly anomalous result—faster responding without reward in extinction than with reward in transfer—is understandable on the basis that the effective stimulus for responding in extinction is occasioned by the previous trial amount of reward. In this case, fast responding was conditioned to occur on trials following S^N , and this tendency was maintained during extinction. For Group 5-1, fast responding was conditioned to S^r , producing only generalized response strength for responding in the presence of S^N . Thus, the greater resistance to extinction in Group 5-0 than in 5-1 is also understandable, since the latter group was extinguished with a stimulus not previously conditioned to the instrumental response.

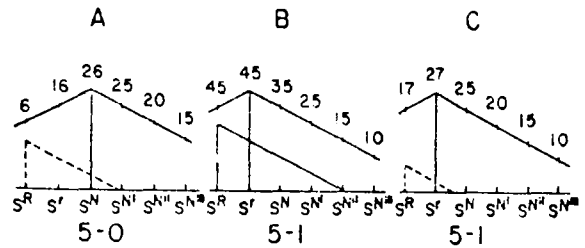


Fig. 3. Hypothetical gradients of habit strength (solid lines) and inhibition (dotted lines). The numbers above the gradients represent units of effective habit strength. Initial, intermediate, and terminal stages of extinction are represented as $S^{N'}$, $S^{N''}$, and $S^{N'''}$, respectively.

The results are also relevant to predictions derived from models based on assumptions of the sequential hypothesis, as shown in Fig. 3. In addition to the assumptions stated previously, some principles were followed in constructing the models which are explicit in Capaldi's (1967) depiction of such models. They are (1) that sufficient habit strength (H) generalize to S^R in the 5-0 condition to provide residual H for responding on trials following S^R , (2) that S^r lie somewhere on the stimulus continuum between S^R and S^N , and (3) that all processes have consistent effects across treatment conditions. Thus, in the 5-0 condition the gradient of generalized H from S^N must overlap and exceed inhibition (I) at S^R , and in the 5-1 condition generalized H from S^r must overlap and summate with effects associated with S^R .

Figure 3A shows a model for the 5-0 condition that is a straightforward translation of models shown in Capaldi (1967) for single alternation of reward and nonreward. Although a model for single alternation of large and small reward has not been presented by Capaldi, Fig. 3 also shows two possible ways in which the sequential view may be represented for the 5-1 condition.

Figure 3B is a logical extension of Capaldi's (1967) assumptions in relation to varied magnitude of reward. Large reward following S^r produces more H than small reward following S^R , and these effects generalize. Thus, Fig. 3B predicts exactly the same effective H on trials following S^R and S^r , and the S^s in Group 5-1 should not have patterned their responses. Further, the shift from 5-1 to 5-0 should have produced immediate inappropriate patterning, since effective H on trials following S^N is less than effective H on trials following S^R .

Figure 3C is based upon assumptions suggested by the data, namely, that S^R small reward sequences produce I but that I grows at a slower rate and to a lesser extent than I produced by S^R nonreward sequences. Patterning does occur in Group 5-1, but it appears at a later stage of training than in Group 5-0. Also, on small reward trials, the decrease and subsequent increase in response latency suggests the growth of an active decremental process (but less than that associated with nonreward), and the nonmonotonicity of the curve for response latency on large reward trials suggests a generalization effect.

Consequently, Figs. 3A and 3C were used in deriving predictions for terminal acquisition, transfer, and extinction responding for Groups 5-0 and 5-1 as follows:

(1) Both groups should exhibit appropriate response patterning, with fastest and about the same level of responding on trials following S^r in Group 5-1 and following S^N in Group 5-0, slower responding on trials following S^R in Group 5-1, and slowest responding on trials following S^R in Group 5-0.

(2) The transfer from 5-1 to 5-0 should produce no initial change but gradually slower responding on trials following S^R , representing increased I produced by S^R nonreward sequences. The transfer from 5-0 to 5-1 should produce no initial change on trials following S^R , an immediate decrement on trials following S^r as compared to trials following S^N , and gradually faster responding on trials following S^R and trials following S^r , representing decreased I and generalized I produced by S^R small reward sequences.

(3) The initial stage of extinction should produce faster responding in both groups than on trials following S^R and S^r in control and transfer conditions. Initial extinction responding should be about the same in the two groups, followed by greater resistance to extinction in intermediate and terminal stages of extinction in Group 5-0-E than in Group 5-1-E.

The results support these predictions, with two exceptions. The 5-0 to 5-1 transfer produced no change rather than an immediate increase in response latency on trials following S^r , and initial extinction responding was faster in Group 5-1-E than in Group 5-0-E. The transfer results were affected by a procedural error. An inspection of the data revealed that, on a large number of small reward trials on the first transfer day, no response was made during the 15-sec response interval, which is probably functionally equivalent to nonreward. Thus, the transfer condition was not properly initiated.

There are at least two processes which might have influenced initial extinction responding. Note that the difference resulted from a decrease in response latency from terminal acquisition level on large reward trials in the 5-1 group, while the 5-0 group began extinction at about the same level. This type of result may be interpreted as a frustration effect (cf. Amsel, 1967). That is, frustrative arousal conditioned during acquisition with varied reward may be increased by the experience of nonreward during extinction, causing the S s initially to respond more rapidly. In this respect, the data are remarkably similar to those of Beier (1958), where rats were trained in the runway on irregular

partial (14-0) and varied (10-4) reward schedules.

The other possibility is that the patterning process itself may have influenced the extinction results. Campbell et al (1970) have shown that patterning reduces resistance to extinction. In the present study, the extent of patterning, even though limited by other processes, may have produced transient differences in extinction such that the greater the extent of patterning, the less the resistance to extinction.

If these explanations are valid, then the proposed modification of the sequential hypothesis appears to fit the data quite well. The predictive strength of the model, of course, would have to be tested by means of experimental designs generated by the model and is, thus, yet to be determined.

REFERENCES

- Amsel, A. Partial reinforcement effects on vigor and persistence. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1967.
- Beier, E. M. Effects of trial-to-trial variation in magnitude of reward upon an instrumental running response. PhD dissertation, Yale University, 1958. In F. A. Logan, *Incentive*. New Haven: Yale University Press, 1960. P. 193.
- Bloom, J. M., & Capaldi, E. J. The behavior of rats in relation to complex patterns of partial reinforcement. *Journal of Comparative & Physiological Psychology*, 1969, 54, 261-265.
- Bloom, J. M., & Malone, P. Single alternation patterning without a trace for blame. *Psychonomic Science*, 1968, 11, 335-336.
- Bloom, J. M., & Smith, N. F. Stimulus aftereffects of bar pressing. *Psychonomic Science*, 1965, 3, 23-24.
- Campbell, P. E., Knouse, S. B., & Wroten, J. D. Resistance to extinction in the rat following regular and irregular schedules of partial reward. *Journal of Comparative & Physiological Psychology*, 1970, 72, 210-215.
- Capaldi, E. J. A sequential hypothesis of instrumental learning. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1967.
- Capaldi, E. J., & Stanley, L. R. Temporal properties of reinforcement aftereffects. *Journal of Experimental Psychology*, 1963, 65, 169-175.
- Gonzalez, R. C., Bainbridge, P., & Bitterman, M. E. Discrete-trials lever pressing in the rat as a function of pattern of reinforcement, effortfulness of response, and amount of reward. *Journal of Comparative & Physiological Psychology*, 1966, 61, 110-122.
- Wall, A. M., & Goodrich, K. P. Differential responding on reinforcement and non-reinforcement trials occurring in fixed repeating patterns. *Psychonomic Science*, 1964, 1, 193-194.

(Received for publication September 19, 1972;
revision received December 19, 1972;
second revision received March 15, 1973.)