Journal of Experimental Psychology
Vol. 59, No. 2, 1960

PARTIAL AND CORRELATED REWARD IN ESCAPE LEARNING [1]

GORDON H. BOWER [2]

Yale University

Previous studies (Campbell & Kraeling, 1953; Bower, Fowler, & Trapold, 1959) have demonstrated that magnitude of reinforcement operates as a variable in instrumental escape learning. In those experiments rats were run in a straight alley in which the startbox and runway provided a continuous and intense electric shock, and the goalbox provided a continuous but weaker shock. It was found that running speed was faster the greater the reduction in shock intensity at the goal. Moreover, these performance effects were reversible; when the amount of shock reduction was changed, performance shifted to the level appropriate to the new amount of shock reduction.

Such studies have involved conditions in which *S* receives the same reward on every trial; however, of equal interest are those conditions in which the reward varies from trial to trial. Thus, a large reward (i.e., large amount of shock reduction) may be given on some trials and a small reward on others; or reward may be given immediately on some trials and delayed on others. Experiment I is an investigation of partial reinforcement which may be considered as a limiting case of variable magnitude of reward. Of further interest are those conditions in which variations in the reward are contingent upon variations in some characteristic of the response (e.g., its speed). In Exp. II, one condition of correlated reinforcement is investigated.

EXPERIMENT I

In this experiment the reward was varied independently of *S*'s performance. Two goal events were used, 100% or 0% reduction in shock intensity. Subjects received random sequences of these goal events, with the sequences differing in the proportion of 100% shock reduction trials. These conditions are somewhat analogous to partial reinforcement using, say, food and no food as the goal events.[3]

The primary interest in this experiment was to ascertain the relationship between speed and percentage of reinforcement when the goal events were given in a random sequence. Four groups of *S*s were trained with sequences corresponding to 100, 75, 50, or 25% reinforcement. A second question of interest was whether partial reinforcement in this escape situation causes slow adjustment of performance when the conditions are changed to 0% rewards, which is analogous to extinction.[3]

[1] This experiment was conducted during the writer's tenure on a predoctoral fellowship, MF-8585, granted by the National Institutes of Mental Health, Bethesda, Maryland.

[2] Now at Stanford University.

[3] Since a trial in the alley required eventual removal of *S* from the goalbox, the 0% shock reduction event may be described as a long delay of reward (removal from charged goalbox) instead of "nonreward," and as such might maintain performance above operant level. This description in terms of variable delay in no way changes the logic of the experiment.

Method

Subjects.—The *S*s were 20 male albino rats, about 100 days old. They were assigned randomly to four groups of five *S*s and were maintained with free access to water and Purina Lab Checkers.

Apparatus.—The apparatus is described in detail elsewhere (Bower et al., 1959). A 4-ft. straight alley with start and goal compartments was used. Two L-shaped strips of galvanized sheet metal, separated by a 1-in. floor gap, formed the two walls and halves of the floor. Two independent shock systems were connected to the sides of the alley; one shock system provided shock to the startbox and middle section of the alley, the other to the goalbox. The *S* received shock as long as its feet touched the two strips of sheet metal forming the halves of the floor. The 3 in. width of the alley prevented *S* from escaping shock by running along only one side of the alley.

Raising the start door operated a microswitch which closed the shock circuits to the alley and started a Standard Electric clock. The clock stopped when *S* interrupted a photobeam 2 in. inside the goalbox. Running times were recorded to the nearest .01 sec. and were converted to speed scores (1/time).

Design and procedure.—The shock intensity in the startbox and runway was always 250 v. The shock in the goalbox was either 250 or 0 v. Twenty acquisition trials were given. The four groups differed in the proportion (.25, .50, .75, or 1.00) of trials on which the goalbox shock was 0 v.; within these restrictions, random sequences of goal shocks were used. After the initial 20 trials, two groups (100% and 50%) were given 20 additional trials with the 250-v. goal shock present on every trial.

The *S* was placed in the start box and when it oriented toward the goal the start door was raised, turning on shock. When *S* entered the goalbox and interrupted the photobeam, the goalbox door was lowered and *S* was left in the goalbox for 20 sec. The trials were massed; immediately following removal from the goalbox, *S* was reinserted into the startbox for the next trial.

Results

The results in Fig. 1 show group mean speeds plotted in running averages of four training trials. Speed on the later acquisition trials is clearly related to percentage of reinforcement. Analyses of variance of mean speeds on Trials 13–20 demonstrate significant differences between the groups ($F = 12.16$, $df = 3,16$, $P < .001$). Component analysis of the group differences reveal that speed was a linear function of percentage of reinforcement (linear $F = 36.05$, $df = 1,16$, $P < .001$; residual $F < 1.00$). This linear relationship is shown graphically in Fig. 2, where a 0% control group from a previous experiment is also presented.

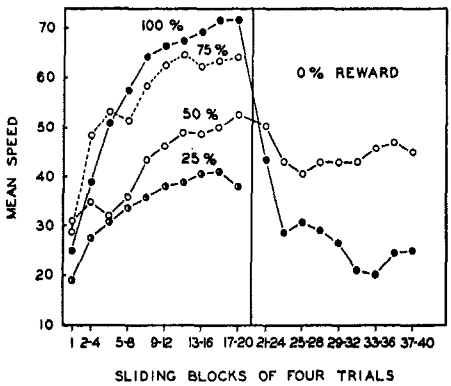

FIG. 1. Group mean speeds in sliding blocks of four trials. After Trial 20, the 100% and 50% groups were switched to a 0% reward condition.

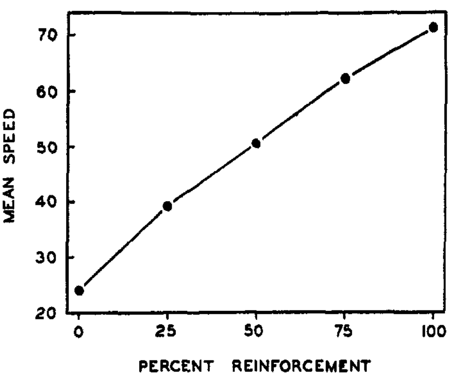

FIG. 2. Mean speed as a function of percentage or reinforcement. Each point is a group average speed over Trials 13–20.

The results of shifting to 250-v. goal shocks on every trial may be seen in the right half of Fig. 1. The *S*s trained under 50% reinforcement show considerable resistance to change in performance, whereas *S*s trained under 100% reinforcement show an abrupt decrement in speed when switched to 0% reward. The speed curves for the two conditions cross after only four trials at the new shock condition. A trend analysis over Trials 17–40 shows a reliable groups-by-trials interaction ($F = 7.05$, $df = 5{,}40$, $P < .001$), which demonstrates the difference in rate of shift.

Experiment II

In Exp. II, variations in reward were correlated with running speed, reward being given when *S* responded slowly. In the present study, the cutoff selected was a speed of 40, or running time of 2.5 sec. and 100% shock reduction was given for speeds below 40 and 0% shock reduction for speeds above 40. This condition is analogous to the DRL schedule studied extensively in the Skinner box (Ferster & Skinner, 1957).

Logan (1959) has proposed an equilibrium model for conditions of correlated reinforcement. The equilibrium performance is defined as that speed which, on the average, gets that percentage of reinforcement which will maintain just that average speed. The most direct test of this model is to compare the performance of the correlated reward *S*s with that of matched control mates. The matched control *S* receives the same sequence of rewards as that received by its experimental partner; however, these rewards are given irrespective of the control *S*'s speed. According to the equilibrium model, performance depends only upon the sequence of rewards and not upon the correlation of reward with speed. Because the sequence of rewards is equated, the equilibrium prediction is that correlated *S*s and their matched controls will perform similarly.

Method

Four pairs of rats were used, with the correlated and matched control *S*s of the pairs designated randomly. The apparatus and general procedure was the same as in Exp. I. The hair was shaved from *S*s' backs to provide better electrical contact. Each *S* received 50 massed trials. For an *S* in the correlated group, the 250-v. goal shock was turned off as soon as it broke the photobeam provided its running time exceeded 2.50 sec.; otherwise, the goal shock was left on. In either case, *S* was removed from the goalbox after 20 sec. The matched control procedure insured that a correlated *S* and its control partner were equated on the entire sequence of goal events.

Results

Group mean speeds in blocks of five trials are shown in Fig. 3. The curves start and rise together for 10 trials, but thereafter show progressive separation. The separation on Trials 41–50 was statistically significant ($F = 55.25$, $df = 1{,}6$, $P < .001$). The mean speed for correlated *S*s rises initially above the cutoff speed

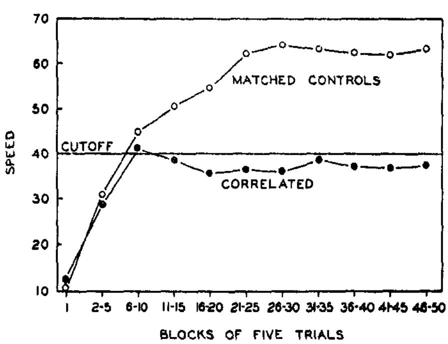

Fig. 3. Group mean speeds in blocks of five trials. The horizontal line represents the cutoff speed of 40.

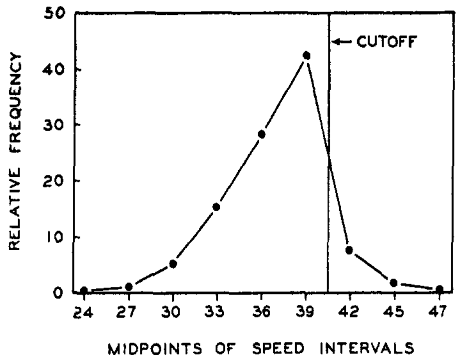

FIG. 4. Pooled speed distribution for the correlated group; plotted by intervals of three speed points.

of 40, but then declines to a relatively stable average speed just below the cutoff. The mean speeds over Trials 41–50 of the four *S*s in the correlated group were 37, 36, 36, and 38; their respective matched controls had mean speeds of 71, 65, 56 and 58. Apparently, the correlated condition reduced between-*S* variability to a large extent.

The pooled distribution of speeds over Trials 26–50 for correlated *S*s is shown in Fig. 4. A distinct mode appears in that speed interval just below the cutoff. Over the last 25 trials about 90% of their runs were rewarded.

DISCUSSION

The results of Exp. I show that speed is a linear function of the percentage of reinforcement; that is, the speed produced by random variations between the two rewards (100% and 0% shock reduction) is approximately the average of the speeds produced by the two rewards separately. Although this result was obtained with escape learning, it is consistent with the results of Logan[4] for variable magnitude and delay of food reinforcement. The result is contrary to the results obtained by Weinstock (1958) and Goodrich (1959) by varying percentage of food reward and nonreward. However, a number of procedural differences (e.g., number and massing of trials, nonrewards vs. long delays, the part of the response chain being measured, etc.) make serious comparisons difficult. That performance changes are slower following training with variable reward in this situation is consistent with the general findings for instrumental appetitive conditioning

The results of Exp. II clearly disconfirm the equilibrium prediction. Performance depends *not only* upon percentage of reinforcement but also upon the correlation of reward with speed. It should be noted that the performance of the uncorrelated, matched control *S*s is close to the speed expected on the basis of previous findings. Over-all, the percentage of reinforcement for the matched controls was about 80–85%. From Fig. 2, this percentage of reinforcement, when given uncorrelated with speed, would be expected to produce an average speed of about 64. This figure compares favorably with the average of 63 observed for the matched control group.

The fact that correlated *S*s learned to run slowly is consistent with the micromolar theory proposed by Logan (1956; 1959). In this approach, different speeds are treated as *different* responses, and the probability of a particular speed depends upon its reward relative to the rewards for competing speeds. The reinforcement dimension most relevant to the present results is the interval of reinforcement, defined as the time between the initiation of the response and the receipt of reward. The incentive for a particular response speed is assumed to be a decreasing function of its interval of reinforcement. For the present contingencies, this principle clearly favors speeds below 40 since they receive the reward immediately. Furthermore, although speeds of 38 and 20, for example, receive the same goal events, speeds of 38 have higher incentive because they have a shorter interval of reinforcement. In other

[4] Results reported in a forthcoming book, to be titled *Incentive*.

words, speeds of 20 are associated with 5 sec. of shock in the alley, whereas speeds of 38 are associated with only 2.65 sec. of alley shock. Thus, it is expected that *S* will learn to make predominately those response chains which minimize shock time, and the predicted mode of the speed distribution in Fig. 4 is at that speed interval just below the cutoff. Because the controlling contingencies "drive" each *S* into this optimal response region, the between-*S* differences are negligible.

Summary

Rats were trained to run down a straight alley to escape a 250-v. shock, the shock in the goalbox being either 0 or 250 v. In Exp. I the percentage of reinforcement was either 25, 50, 75, or 100, and it was found that speed was a linearly increasing function of this variable. When the 50% and 100% groups were switched to 0% reward, the 100% group extinguished faster than the 50% group.

In Exp. II reward was correlated with slow speeds, the goal shock being 0 v. for speeds below 40 and 250 v. otherwise. Matched control *S*s received the same sequence of goal events as that received by correlated *S*s, but reward was uncorrelated with the control *S*'s speed. It was found that *S*s in the correlated group learned to run at speeds just below the cutoff, receiving about 90% reward, whereas the matched controls ran significantly faster. This latter result was related to a micromolar theory which considers interval of reinforcement as one component of incentive motivation.

References

Bower, G. H., Fowler, H., & Trapold, M. A. Escape learning as a function of amount of shock reduction. *J. exp. Psychol.*, 1959, **58**, 482–484.

Campbell, B. A., & Kraeling, D. Response strength as a function of drive level and amount of drive reduction. *J. exp. Psychol.*, 1953, **45**, 97–101.

Ferster, C. B., & Skinner, B. F. *Schedules of reinforcement.* New York: Appleton-Century-Crofts, 1957.

Goodrich, K. P. Performance in different segments of an instrumental response chain as a function of reinforcement schedule. *J. exp. Psychol.*, 1959, **57**, 57–64.

Logan, F. A. A micromolar approach to behavior theory. *Psychol. Rev.*, 1956, **63**, 63–73.

Logan, F. A. The Hull-Spence approach. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. II. New York: McGraw-Hill, 1959.

Weinstock, S. Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. *J. exp. Psychol.*, 1958, **56**, 151–158.

(Received March 20, 1959)