PARTIAL REINFORCEMENT (ACQUISITION) EFFECTS WITHIN SUBJECTS

ABRAM AMSEL, JOHN R. MACKINNON, MICHAEL E. RASHOTTE,
AND C. THOMAS SURRIDGE¹

UNIVERSITY OF TORONTO

Acquisition performance of 22 rats in a straight alley runway was examined. The animals were subjected to partial reinforcement when the alley was black ($B\pm$) and continuous reinforcement when it was white ($W+$). The results indicated (a) higher terminal performance, for partial as against continuous reinforcement conditions, for starting-time and running-time measures, and (b) lower terminal performance under partial conditions for a goal-entry-time measure. These results confirm within subjects an effect previously demonstrated, in the runway, only in between-groups tests, where one group is run under partial reinforcement and a separate group is run under continuous reinforcement in the presence of the same external stimuli. Differences between the runway situation, employing a discrete-trial procedure and performance measures at three points in the response chain, and the Skinner box situation, used in its free-operant mode with a single performance measure, are discussed in relation to the present findings.

In the course of testing hypotheses concerning the effects of prediscrimination experiences on discrimination learning, a prediscrimination condition was run in which the same 22 subjects were partially reinforced for an approach response in the presence of a black stimulus ($B\pm$) and continuously reinforced for the same response in the presence of white ($W+$). This group of subjects, running under $B\pm W+$ conditions, was carried to 324 trials. An examination of the data revealed, for the first time, that the partial-reinforcement-acquisition effect previously shown by others (Goodrich, 1959; Haggard, 1959; Wagner, 1961; Weinstock, 1954) using separate partial- and continuous-reinforcement groups—i.e., as a between-group difference—can be demonstrated to occur within subjects. Animals running under the $B\pm W+$ condition, partially reinforced to black and continuously to white, eventually started and ran more rapidly in the presence of the stimulus signaling partial reinforcement ($B\pm$) than to the cue for continuous reinforcement ($W+$), but entered the goal more slowly to $B\pm$ than to $W+$.

METHOD

Subjects

Twenty-two albino rats, 11 male, 11 female, about 90 days old at the start of the experi-

ment were supplied by Woodlyn Farms, Guelph, Ontario. Subjects were maintained on a 22-hr food deprivation schedule, and were fed 10 g of laboratory chow per day, including the two 500 mg. pellets given as reward on each of the two reinforced trials each day, about 30 min after their daily trials.

Apparatus

A straight alley contained, in series, a 15 in. entry box, a 12 in. start box, a 50 in. runway, and a 12 in. goal box, all separated by guillotine-type doors. The inside width was 2½ in. throughout. Three performance measures were taken: (1) starting time was taken from raising of the start door to a point 6 in. into the runway; (2) running time was taken over the next 24 in. of the runway; and (3) goal-entry time covered the remainder of the runway and the first 4 in. of the goal box. These measurements were made by the interruption of infra-red beams of photocells placed at the required positions in the alley. Manipulation of the black and white stimulus conditions was accomplished by inserting sheets of the appropriate color between the walls of the apparatus and clear plexiglass

¹This research was supported by a grant (G-13895) from the National Science Foundation. Reprints may be obtained from Abram Amsel, Dept. of Psychology, University of Toronto, Toronto 5, Canada.

inner linings. The experiment was run under dim illumination to reduce any possible aversiveness of white, as compared with the black, stimulation.

Procedure

Three trials, one each of B+, B-, and W+, were run each day. Within a block of six consecutive days, the orders of these were randomized, with the restriction that one of each of the six possible orders occurred during every six-day block. The inter-trial interval for any given S on any day was about 20 min.

RESULTS AND DISCUSSION

The results of this study are shown in Fig. 1. Here are plotted mean median starting, running, and goal times against successive six-day blocks of trials for B± and for W+ trials separately. For each S a median score of each kind was computed for each six-day block. These were then averaged over all Ss for each block. For starting and running times, the curves separate after the first four or five blocks (72 to 90 trials), and the starting and running times to B± stay lower (faster) from then on. For the goal measure, the curves separate from the fourth block on, the W+ performance being faster after that. Statistical analysis of these data would seem superfluous since all within-graph and between-graph comparisons involve the same Ss, rather than different groups of Ss. The probabilities that successive differences between B± and W+ would remain in the same direction could be estimated by raising .5 to the 14th, 13th, and 15th powers, respectively, for starting, running, and goal times. These data are so reliable that they can be reproduced, approximately, for 20 of the 22 Ss individually. The phe-

nomenon does not depend on averaging over Ss.

The generality of these findings might be questioned by raising the possibility that they depend upon the particular conditions of the present experiment. For example, the partially-reinforced stimulus is always black, the continuously-reinforced, white. Are the results, then, an artifact of black versus white? In addition, the nature of the experiment, as performed here, demands that the partially-reinforced stimulus be exposed twice as frequently as the continuously-reinforced. Do the results in any way depend on this factor? Obviously, the best answers to these questions await further research; however, our data suggest that the phenomenon has some generality. First of all, there is the reversal of findings for the starting and running measures, on the one hand, and the goal measure, on the other. If either a black-white preference or number of exposures to black as against white produced the effect, it should show up in all three measures. The largest differences found are in the goal measure, where performance to white is superior to performance to black. Secondly, a trial-by-trial—rather than block-by-block—examination of the first 18 days (54 trials) shows no initial tendency to run faster to black, but some such tendency to white, the stimulus associated with continuous reward (Fig. 2). This shows up particularly in the starting measure where Ss ran faster to W+ than to B± on nine out of the last 10 of these first 18 days. Examination of later segments shows that more vigorous performance in the presence of B± emerges in the latency and running measures at about the time the opposite tendency emerges in the goal measure—after about 72 trials.

Fig. 1. Starting, running, and goal measures for all 108 days (324 trials) showing the reversal of the difference between B± and W+ between the first two measures and the third.

Fig. 2. Starting, running, and goal measures for the first 18 days (54 trials) showing initial relationships between B± and W+ performance.

An initial difference in performance favoring the continuous condition has been a finding characteristic of experiments where separate partial and continuous groups are employed. In these latter kinds of studies, equal numbers of trials are usually run under partial and continuous conditions, meaning fewer reinforcements for the former than for the latter. In the present experiment numbers of reinforcements, and not numbers of trials, have been equalized between the B± and W+ conditions, and the initial superiority of the continuous condition is still demonstrable.

One other question should be asked about these data. Since there were two B trials on each day, one a reward trial and the other nonreward, was it possible for Ss to form a discrimination on the basis of the first B trial as to what the second would be? For example, if the first of these trials were B+, might S anticipate that the second would be B- and run slower to it? This would have the effect of exaggerating the difference in goal times in Fig. 1 while reducing the differences in starting and running times, as shown. An internal analysis of the data indicated there was some discrimination of B+ and B- in a few of the Ss. But it is not at all clear that it was based on the earlier B trial within a day's session, since it showed up in the same Ss on the first B trial of a day. This suggests the operation of a cue these few Ss picked up from the experimental situation, probably from the experimenter's differential activity before a rewarded or nonrewarded run. It also suggests that runway experiments claiming to show learning on the basis of carried-over after-effects or traces of reinforcement and nonreinforcement from one trial to the next (*e.g.*, Capaldi and Stanley, 1963) must exercise great control over such possible influences. In any event, the effect, as it showed up in the present study, was only in some Ss and almost restricted to the goal measure. If the data are plotted for the goal measure eliminating such Ss, the difference between B± and W+ is still there. Again, if for all Ss response level on B+, B-, and W+ trials are plotted separately, the B+ curve remains completely separated from the W+ curve as in Fig. 1, but with the difference slightly reduced. The starting and running data are not appreciably affected by this factor, and

the phenomenon described in Fig. 1 remains a genuine within-subjects partial reinforcement acquisition effect.

These results confirm and extend findings previously only demonstrated between groups. They indicate that the mechanisms of partial versus continuous reinforcement as applied to runway acquisition behavior in rats, at least, can exist and operate simultaneously, in the same organism, in situations differing only in a single dimension. That is to say, organisms in simple learning situations can behave in a manner appropriate to partial or to continuous (intermittent) reinforcement, at different times, depending on the environmental stimuli. The results are in no way different from those obtained when one group is run under partial and another under continuous reinforcement in what is otherwise the same experimental situation. In such studies, also, the asymptotic performance shows, very characteristically, greater vigor of approach under partial than under continuous reinforcement in the starting and running measures, and lesser vigor of approach under partial than under continuous reinforcement in the goal measure.

It is, of course, well known that large differences between performances under several systems of reinforcement, each related to a distinctive stimulus, can be obtained within individual Ss by operant conditioning procedures (*e.g.*, Ferster and Skinner, 1957). While the present result shows this same kind of thing in the runway apparatus, these may not be coextensive phenomena. The runway situation employing discrete trials, often widely separated in time, and the Skinner-box situation used in its free-responding (operant) mode, reflect different kinds of behavior and often produce apparently quite different kinds of results. For example, in the former situation, resistance to extinction is a nonmonotonic function of number of rewards (Ison, 1962; North and Stimmel, 1960). In the latter, the monotonic function suggested by the data of Perin (1942) and Williams (1938) apparently holds up even after a great many reinforcements (Dyal and Holland, 1963; Harris and Nygaard, 1961). The question here is whether the lever box, operant conditioning situation (granted its precision as a measuring instrument) provides a counterpart to the different response measures (start-

ing, running, and goal) which are now taken fairly routinely in runway studies. Has the present result, duplicating the interaction between direction of the partial-continuous difference and response measure employed, been established with the typical operant procedure? If it has not, what would be its counterpart in the operant procedure?

There are studies of fixed-interval and fixed-ratio behavior which provide data on time to the first response (starting time?) along with other measures of response rate early (running time?) and late (goal time?) in the fixed-interval or fixed-ratio sequence. We are not aware of an experiment, in which these three measures are taken, where one discriminative stimulus ($S+$) signals reinforcement at the end of every fixed-ratio run, while a different stimulus ($S\pm$) signals reinforcement only at the end of every other fixed-ratio, on the average. This would seem to be one area where a significant integration could be made between runway and lever-box studies. There is a study (Zimmerman, 1960) which attempts to establish a link between intermittent reinforcement in the free-operant situation and partial reinforcement in the runway. The part of this experiment in which FR runs are sometimes reinforced and sometimes not to an S^D seems equivalent to the $S\pm$ portion of such an experiment. Zimmerman finds that time to first response (starting time) scores increase and become more variable under intermittent than they were under continuous reinforcement. If Zimmerman's procedure were extended so that there was one stimulus ($S+$) signalling reinforcement at the end of each FR run, and another stimulus ($S\pm$) signalling reinforcement at the end of every other run, on the average, would the S 's time to first response be shorter to $S\pm$ than to $S+$? And would its inter-response times (rate) be shorter to $S\pm$ early in the run and longer late in the run?

This study lends further support to an interpretation of discrete-trial, partial reinforcement acquisition, as well as extinction, phenomena in terms of an anticipatory frustration (r_F) approach (Amsel, 1958, 1962; Spence, 1960). To explain the differences between the starting and running measures and the goal measure, it needs only to be assumed that in the first two segments, where r_F and therefore avoidance tendencies are weak, the

r_F contributes drive strength to the still-dominant approach tendency; whereas in the goal segment a point is reached where approach and avoidance tendencies are more nearly equal, and the facilitating aspects of r_F are counteracted by increased conflict.

REFERENCES

- Amsel, A. The role of frustrative nonreward in non-continuous reward situations. *Psychol. Bull.*, 1958, **55**, 102-119.
- Amsel, A. Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychol. Rev.*, 1962, **69**, 306-328.
- Capaldi, E. J. and Stanley, L. R. Temporal properties of reinforcement after-effects. *J. exp. Psychol.*, 1963, **65**, 169-175.
- Dyal, J. A. and Holland, T. A. Resistance to extinction as a function of number of reinforcements. *Amer. J. Psychol.*, 1963, **76**, 332-333.
- Ferster, C. B. and 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**, 57-63.
- Haggard, D. F. Acquisition of a simple running response as a function of partial and continuous schedules of reinforcement. *Psychol. Rec.*, 1959, **9**, 11-18.
- Harris, P. and Nygaard, J. E. Resistance to extinction and number of reinforcements. *Psychol. Rep.*, 1961, **8**, 233-234.
- Ison, J. R. Experimental extinction as a function of number of reinforcements. *J. exp. Psychol.*, 1962, **64**, 314-317.
- North, A. J. and Stimmel, D. T. Extinction of an instrumental response following a large number of reinforcements. *Psychol. Rep.*, 1960, **6**, 227-234.
- Perin, C. T. Behavior potentiality as a joint function of amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93-113.
- Spence, K. W. *Behavior Theory and Learning*. Englewood Cliffs: Prentice-Hall, 1960.
- Wagner, A. R. Effects of amount and percentage of reinforcement and number of acquisition trials on conditioning and extinction. *J. exp. Psychol.*, 1961, **62**, 234-242.
- Weinstock, S. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, 1954, **47**, 318-323.
- Williams, S. B. Resistance to extinction as a function of number of reinforcements. *J. exp. Psychol.*, 1938, **23**, 506-521.
- Zimmerman, D. W. Intermittent reinforcement of discriminatively controlled responses and runs of responses. *J. exp. Anal. Behav.*, 1960, **3**, 83-91.

Received July 2, 1963