

Intertrial reward (ITR) and delay of reward: Effects on resistance to extinction

LEE E. OVERSTREET, JR. and PATRICK E. CAMPBELL, *Southern Methodist University, Dallas, Tex. 75222*

Six groups of rats ($N = 10$) were trained to traverse a straight runway. Two of the groups were trained under regular consistent- and partial-reward procedures, i.e., RRRRR and RNRNR. Two additional groups were trained under the partial-reward schedule above but, in addition, received ITR(r), i.e., RNrRNrR. One such group received ITR after a wait of 15 sec in the intertrial interval box, another after a wait of 30 sec. The remaining two groups were trained under either 15-sec or 30-sec partial delay of reward, i.e., RDRDR. The results of an extinction phase showed that ITR and delayed reward are not identical processes. ITR was found to decrease resistance to extinction only in the 30-sec wait condition.

Intertrial reward (ITR) refers to a procedure developed by Capaldi, Hart, & Stanley (1963) to test several implications of Capaldi's (1967) modified-aftereffects hypothesis. According to that theory, the commonly observed partial-reinforcement-extinction effect (PRE) is caused by the conditioning of nonrewarded aftereffects (SN) to the instrumental response (R_I) on rewarded trials during acquisition. Such conditionings occur when nonrewarded trials (N) are followed by rewarded trials (R). As a result of several N - R transitions, the SN - R_I connection receives considerable habit and, according to the principle of stimulus generalization, supplies generalized habit to the aftereffects of successive N trials ($SN_1, SN_2, SN_3 \dots SN_n$) that occur during extinction. The ITR procedure was developed to test the necessity of the SN - R_I connection as the determinant of the PRE. ITR consists of placing an S directly into the baited goalbox during the intertrial interval (ITI), separating N from R trials. The rationale for this procedure is that the SN - R_I connection will not be strengthened since R_I does not occur, and the result of ITR will be the replacement of SN with the aftereffects of reward (SR). The following R trial will then result in an increment in the SR - R_I habit and not in the SN - R_I habit. Partial-reward (PR) schedules that include ITR between N and R trials should show no more resistance to extinction than a continuously rewarded (CRF) control

group. Under either schedule, SR is the only aftereffect conditioned to R_I . Capaldi et al (1963) tested this hypothesis and reported that Ss trained with ITR between N and R trials showed no more resistance to extinction than a CRF group, while Ss trained with ITR between R and N trials showed the typical PRE.

The hypothesis that ITR replaces SN with SR seems to fit the above data quite adequately. Nevertheless, Capaldi and his associates (Capaldi, Hart, & Stanley, 1963; Capaldi & Poynor, 1966) suggested that ITR and partially delayed reward (PD) procedures may be identical or at least similar. On purely operational grounds, the two procedures do have something in common. In both procedures, reward is forthcoming after some specified waiting period. When PD schedules are employed, reward is given immediately on some trials but only after a delay on other trials. This, of course, also occurs in PR schedules using ITR. The major difference between the two procedures is that the entire waiting period is spent in the goalbox in PD schedules, while under ITR, a portion of the wait occurs in the goalbox, but the remainder of the wait occurs in a waiting box. The similarity in these procedures has resulted in the alternative hypothesis that ITR may replace SN with the aftereffects of delay (SD) and not necessarily with SR . Although adequate parametric data are surprisingly lacking on this point, the assumption is that relatively low temporal values of SD will not result in increased resistance to extinction (Renner, 1964). This alternative delay hypothesis of ITR effects was quite attractive to the present authors since several unpublished experiments from our laboratory, and from others as well, have suggested that ITR may only partially reduce the magnitude of the PRE. The present experiment was designed to test the delay hypothesis of ITR effects.

SUBJECTS

The Ss were 60 experimentally naive female albino rats of the Sprague-Dawley strain, supplied by the Holtzman Co. of Madison, Wis. Ss were approximately 73 days old at the beginning of the experiment.

APPARATUS

The flat-gray runway was 58 x 4 x 6 in. high. The initial 10-in. section served as a startbox and the final 18 in. as a goalbox. The sections of the runway were separated by guillotine doors and covered with

hinged-hardware cloth tops. The final 3.5-in. portion of the goal section consisted of two identically appearing food-cup compartments, mounted on runners so that they could be individually aligned with the goalbox. One of these sections was always used on immediate- or delayed-reward trials and the other only on nonrewarded trials. An additional guillotine door was located immediately in front of the reward compartments and was used to allow immediate or delayed access to the food compartments. Response times were recorded by a system of two photoelectric relays, wired to control the pulsing of a Grason-Stadler printout counter. One measure (run time) was measured from the opening of the start door until S interrupted a photobeam located 24 in. from the start door. A second measure (goal time) was recorded for an 18-in. section between the first photobeam and a second located 2 in. prior to the terminal guillotine door. On all trials, the terminal door remained closed until S had broken the final photobeam, so that S could not use the position of the door as a cue for R , D , or N trials.

DESIGN AND PROCEDURE

Six Ss , chosen randomly from the 60 Ss , were weighed daily for 5 days while on ad lib food and water. Their mean median weight for the last 3 days was taken as an ad lib weight estimate for all Ss . Ss were then reduced to and maintained at 85% of this estimate by makeup feedings given once each day and at least 15 min after experimental training. Following 14 days of handling and weight stabilization, the Ss were allowed to explore the runway in groups of four, in pairs and individually over a 3-day period. Wet mash (Purina rat meal and tap water) was continuously available in the runway during exploration. The final 2 days of pretraining consisted of five immediately rewarded runway trials each day. During these and all subsequent runway trials, reward consisted of 15 sec access to wet mash, and the intertrial interval (ITI) was 45 sec.

Following the above pretraining, Ss were divided randomly into six groups of 10 Ss each and given 5 days of acquisition training at five trials a day. Two of the groups were controls and were given "traditional" CRF and PR reward training, i.e., RRRRR and RNRNR, respectively. Nonreward confinement time was 15 sec on all N trials. Two additional groups were given response training as in PR above but, in addition, were given ITR(r) following each N trial, i.e., RNrRNrR. One of these groups (PR-15) was given ITR after waiting 15 sec in the ITI box. The other group (PR-30) spent 30 sec in the ITI box prior to being placed in the goalbox for ITR.

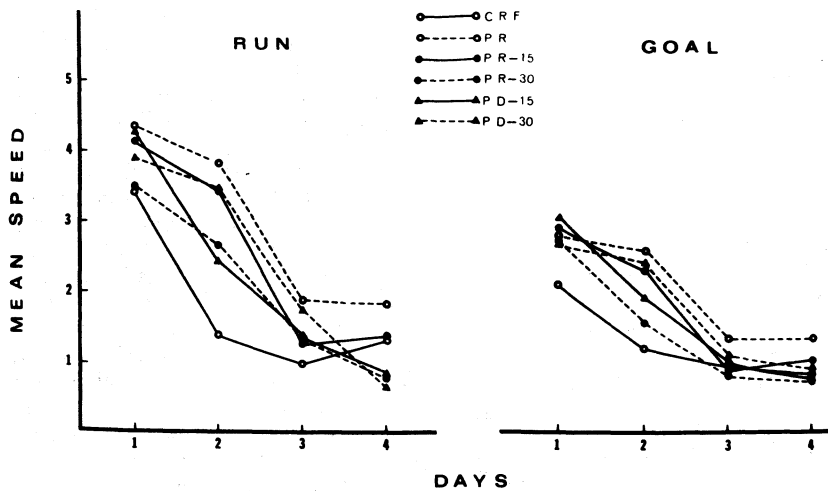


Fig. 1. Mean median run and goal speeds in extinction for each reward schedule.

Since the ITI was held constant at 45 sec, the PR-30 Ss were started on R trials immediately after their 15-sec ITR reward confinement on Trials 3 and 5. The PR-15 Ss were returned to the ITI box for 15 sec prior to beginning R trials on Trials 3 and 5. It should be mentioned, at this point, that the two temporal variables (time between N confinement and ITR and the time between ITR and R trials) were confounded in the PR-15 and PR-30 groups. Any attempt to remove this source of confusion, however, would have confounded the ITI with treatments. The remaining two groups were given 15 sec (PD-15) or 30 sec (PD-30) delayed reward on Trials 3 and 5, i.e., RDRDR.

The Ss were run in squads of six, with one S from each treatment represented. The orders of running squads and Ss within squads were randomized each day. Following the 5 days of acquisition training, all Ss were given 4 days of extinction at five trials a day. During extinction, the procedure for running Ss remained the same as in acquisition, except that no R, D, or ITR trials were given.

RESULTS

Acquisition

Each S's daily median run and goal times were converted to comparable speed units (24/time for run and 18/time for goal). Analyses of variance on these data revealed a significant effect for days in both the run ($F = 24.24$, $df = 4/216$, $p < .001$) and goal measures ($F = 3.11$, $df = 4/216$, $p < .05$). Groups, however, were significant only in the goal measure ($F = 8.19$, $df = 5/54$, $p < .01$). Subsequent Newman-Keuls contrasts showed that CRF ran faster than all other groups that did not differ among themselves. These and all additional Newman-Keuls contrasts were calculated using the .05 significance level.

Extinction

The run and goal times were converted to speed measures as in acquisition and are

presented in Fig. 1. Inspection of the figure suggests that extinction was asymptotic for all groups by Day 3. An analysis of variance of Day 4 showed no effect for schedules on asymptotic extinction ($F = 1.37$, $df = 5/54$, $p > .05$ for run, and $F < 1$ for goal). Analyses of variance were then performed over the first 3 days for each response measure and revealed that days were highly significant ($p < .001$) for both measures. Groups were also significant for both the run ($F = 2.61$, $df = 5/54$, $p < .05$) and the goal sections ($F = 43.17$, $df = 5/54$, $p < .001$). Newman-Keuls contrasts for the run measure revealed that, using CRF as a reference, all groups except PD-15 and PR-30 showed increased resistance to extinction. However, when the PR control is used as a reference, only CRF and PR-30 show slower running speeds, i.e., PD-15 did not run significantly faster than CRF or slower than PR. The rank orders of the groups were identical in the two measures. Newman-Keuls contrasts of the goal speeds, however, showed one difference not found in the run data. The PD-15 schedule resulted in faster speeds than did CRF but, as in the run section, failed to show slower speeds than did PR. The remaining contrasts were the same as those found in the run speeds.

DISCUSSION

It is apparent that ITR and partially delayed reward are not identical procedures. The delay hypothesis of ITR effect would have been supported if PR-30 and PD-30 had shown greater resistance to extinction than did CRF and if PR-15 and PD-15 had not. The actual results, however, suggest that ITR was effective in decreasing resistance to extinction only in the PR-30 schedule. The delayed-reward hypothesis of ITR effect, however, would predict negligible ITR effects for PR-30 since the stimulus, SD30 sec, would be conditioned to R_I , and SD30 sec is, of

course, capable of increasing resistance to extinction. It should be mentioned, at this point, that the value of the "delay" stimulus following PR-15 and PR-30 could be calculated in several ways. If the duration of delay is calculated from the termination of the runway trial, PR-15 would result in SD15 sec and PR-30 in SD30 sec. If delay begins with the completion of the running response (which is the usual meaning), the stimuli would be SD30 sec and SD45 sec for PR-15 and PR-30, respectively.

Capaldi, Hart, and Stanley, as early as 1963, suggested that ITR and partial delay might be similar procedures. A careful reading of their paper suggests that they defined a delay interval by the second method mentioned above. They also expressed doubts that ITR could be accounted for in terms of delayed reward. By 1966, however, Capaldi and Poynor seemed more confident, but their definition of the delay interval was not made completely clear. At one point in their paper, they write, "The major difference between the two procedures appears to be that under delay the entire wait period is spent in the goal box while under ITR only a portion of the total wait period is spent in the goal box, the remainder of the wait occurring in the intertrial-interval box." At a later point, however, they conclude that ITR may change S^N to S^D15 sec. These values, of course, could not be calculated using either of the methods reviewed here. Capaldi and Poynor had apparently used the nonreward goalbox-confinement time as the duration of delay in their calculations. Unfortunately, no amount of our juggling of these intervals could account for the effectiveness of ITR at PR-30 in the present experiment. The method apparently intended by Capaldi and Poynor would result in S^D15 sec (the N-confinement time) being conditioned in both PR-15 and PR-30. Two considerations fail to support this hypothesis: (1) PR-15 and PR-30 did not show similar extinction rates; and (2) PD-15, i.e., S^D15 sec, resulted in increased resistance to extinction in the goal area.

The present data are also relevant to two additional hypotheses that have been advanced to account for the effects of ITR (Surrige & Amsel, 1966; Lobb & Runcie, 1967). These hypotheses are related in that both assume that ITR reduces resistance to extinction by interfering with the development of anticipatory frustration. According to this view, ITR given very

shortly after an N trial converts the N trial to an R trial (mechanism unspecified, Surridge & Amsel, 1966) or converts N to R by interfering with the consolidation of the frustration reaction (Lobb & Runcie, 1967). The important point for our purposes is that both hypotheses assume that the effectiveness of ITR will decrease as the time between nonreward and ITR increases. Our results, of course, show just the opposite effect. The effectiveness of ITR increased with increasing time between nonreward and ITR. Surridge & Amsel (1966) also suggest that ITR effects may be due, in part, to a design artifact. They point to the fact that some ITR studies include a PR control group that receives ITR following R trials. They assume that this procedure increased the reward magnitude of R trials, and that ITR effects may be due to an increase in resistance to extinction for the PR control rather than to (or in addition to) whatever decremental effects are caused by ITR following N trials. This hypothesis cannot account for the present results since our PR control did not include ITR.

Only one hypothesis that might account for the effectiveness of ITR at PR-30 and the failure of ITR at PR-15 has occurred to the present authors. It will be recalled that these two groups differed not only in the waiting period following the N trial and preceding ITR, but they also varied in the time following ITR and the start of R trials (15 sec wait for PR-15 and immediate in PR-30). This latter interval may have facilitated a discrimination between R and ITR trials in the PR-15 group. These Ss received ITRs that were "set off" from N or R trials by a 15-sec interval both before and after the ITR placement. The PR-30 Ss were placed in the startbox for R trials immediately after ITR placement (this has been the common procedure in successful ITR studies). Perhaps an increased discriminability between R and ITR trials in PR-15 Ss caused the decrease in the effectiveness of ITR. A similar discrimination hypothesis has been suggested by Capaldi & Oliver (1967) in accounting for the decreased effectiveness of ITR following its repeated use.

REFERENCES

- 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. Pp. 67-156.
- CAPALDI, E. J., HART, D., & STANLEY, L. R. Effect of intertrial reinforcement on the aftereffect of nonreinforcement and resistance to extinction. *Journal of Experimental Psychology*, 1963, 65, 70-74.
- CAPALDI, E. J., & OLIVER, W. P. Effect of intertrial reinforcement following a substantial

- number of consistently rewarded trials. *Journal of Experimental Psychology*, 1967, 75, 135-138.
- CAPALDI, E. J., & POYNOR, H. Aftereffects and delay of reward. *Journal of Experimental Psychology*, 1966, 71, 80-88.
- LOBB, H., & RUNCIE, D. Intertrial reinforcement as interference with consolidation. *Psychonomic Science*, 1967, 9, 25-26.
- RENNER, K. E. Delay of reinforcement: A historical review. *Psychological Bulletin*, 1964, 61, 341-361.
- SURRIDGE, C. T., & AMSEL, A. Acquisition

and extinction under single alternation and random partial-reinforcement conditions with a 24-hour intertrial interval. *Journal of Experimental Psychology*, 1966, 72, 361-368.

NOTE

1. This research was partially supported by Grant No. 8469 from the National Science Foundation to Southern Methodist University. The study was conducted according to the APA statement of "Principles for the care and use of animals," June 26, 1968. Reprints may be obtained from Patrick E. Campbell, Department of Psychology, Southern Methodist University, Dallas, Tex. 75222.

Suppressing an avoidance response by a pre-aversive stimulus¹

H. M. B. HURWITZ and A. E. ROBERTS,
University of Tennessee, Knoxville, Tenn.
37916

Subjects trained under a free operant-avoidance schedule were exposed to a 1-min clicker (CS) followed by an unavoidable shock (US). During the preaversive stimulus, overall response rates declined and frequency of shock increased. During this period, the response-frequency distribution was U-shaped. When the US was removed, response rates increased and shock rates decreased.

Hurwitz & Black (1968) reported that when a signaled shock was superimposed on responding maintained by a free-operant-avoidance schedule (FOA) (Sidman, 1953), response rates during the signal (CS) were often only marginally affected (when compared to nonsignaled response rates). On the other hand, shock rates during CS increased dramatically, compared to shock rates during nonsignaled periods. It should be noted that under the FOA schedule increases in shock rates during CS can only result from reduced responding, so that the distribution of responses during CS must be different from response distributions during nonsignaled periods.

The purpose of this study was to investigate the distribution of avoidance responses during periods prior to the signal (pre-CS), during the pre-aversive signal (CS), and following termination of the signal (post-CS).

SUBJECTS

Three female hooded rats, who were experimentally naive, were employed as Ss.

PROCEDURE

An FOA schedule was programmed so that 0.1-sec shocks of 0.8 mA were given at 5-sec intervals unless a lever press occurred, in which case shock was postponed for 20 sec. Houselights were turned on at the beginning of the session and, coincident with a lever press, were momentarily extinguished.

After 39 daily 2-h training sessions under a free-operant-avoidance schedule, a signal (CS), followed within 60 sec by a shock (US), was superimposed on FOA responding. The CS-US was presented on a variable-interval schedule, with a mean rate of 4 min. Shock was the same intensity as that used under the FOA procedure. Thirty CS-US pairings were presented on each session.

After 20 sessions in which the CS-US was given, the experimental conditions were changed, and the CS was presented alone without being followed by US. Ten sessions were given under this extinction procedure.

RESPONSE MEASURES

The effects of the CS-US procedure were investigated in terms of number of FOA shocks occurring in signaled and nonsignaled periods of the session. A distribution of responding was obtained in the following manner: The 60-sec pre-CS, CS, and post-CS periods were divided into five 12-sec intervals; responses occurring during each interval were separately recorded.