Reward-punishment relationships[1]

ROBERT S. BEECROFT, University of North Dakota, Grand Forks, N. D. 58201, BRIAN M. KRUGER and BONNIE G. FISHER, University of Iowa, Iowa City, Iowa 52240

Response cessation was studied as a function of 16 reinforcement conditions (nine reward-punishment combinations and seven nonrewarded and/or nonpunished controls). Responding was quickly suppressed when every response or every fourth response was punished. When only every 16th response was punished, rats took more shocks and earned a substantial number of pellets. However, infrequent punishment coupled with extinction (no reward) resulted in prompt suppression. The generalization that intermittent punishment suppresses behavior less than continuous punishment may hold only when responding is also currently rewarded.

One effect of both reward and punishment is the cessation of responding. Satiation is the name given to reward-induced response cessation and suppression is the punishment-related term. In general one would expect punishment to stop responding sooner than reward, although this probably depends upon both the magnitude and frequency of the rewards and punishments. In this study we held reward and punishment magnitudes constant, although by no means equal, and varied the fixed ratios of rewards and punishments produced by responding.

Recent research tells us some things about punishment frequency and reward frequency in reward-punishment situations. For example, intermittent punishment suppresses responding less than continuous punishment (e.g., Azrin, Holz, & Hake, 1963) or, to put it another way, behavior continues longer in the presence of less frequent punishment. However, evidence showing that the reward schedule which maintains operant responding also modulates the suppressive effect of punishment is equivocal. Although Church & Raymond (1967), working with rats, found less suppression with a VI 0.2-min reward schedule than with a VI 5-min schedule, Holz (1968) reported a proportional reduction of response rates in pigeons as punishment intensity increased when reward was concurrently programmed on two keys at VI 1.9 min and VI 7.5 min.

Of possible reward-punishment combinations one would anticipate that the combination of frequent reward and infrequent punishment would most favor response perseveration. Accordingly, we chose to study small reward ratios (FR 1, FR 2, FR 4) in combination with larger punishment ratios (PR 1, PR 4, PR 16). Preliminary work convinced us that to get a moderate number of responses out of the rat before he ceased responding we would have to employ reward ratios smaller than the punishment ratios with which they were paired. So in five of the nine combinations reward was more frequent than punishment. Two equal-frequency combinations (FR 1/PR 1 and FR 4/PR 4) were studied and the other two cases paired continuous punishment (PR 1) with intermittent positive reinforcement.

Seven additional conditions were investigated. One of these was the simple extinction condition in which neither reward nor punishment was given. Three other nonpunishment conditions were satiation cells for the three reward ratios. The equivocal character of current behavior theory is well illustrated by the fact that one cannot predict whether satiation or extinction will occur the sooner. Finally, the three punishment ratios were applied to animals undergoing conventional rewardless extinction. Thus, 16 conditions in all were studied. The dependent variables were the number of responses made, the number of rewards earned, and the number of punishments received before response cessation occurred.

METHOD

The Ss were male hooded rats maintained at 80% ad lib weight. Their ages were between 90 and 150 days and their 80% weights were restricted to 275 ± 50 g. On the first day they received 50 reinforcements on FR 1 (including shaping) followed by 50 more reinforcements on VR 2. The VR 2 contained an equal number of FR 1, FR 2, and FR 3 components in a 30-step cycle. On the second and third days 100 reinforcements were earned on VR 2. Interresponse times for VR 2 training have been presented in another report (Beecroft & Kruger, 1967); the VR 2 responding was at a moderate but very steady rate. On the fourth (test) day Ss were allowed to earn 10 pellets on VR 2 before a particular reward-punishment combination was instituted. The response-cessation criterion was 1000 sec. This is a suppression-extinction-satiation criterion depending upon the reinforcement conditions. Bar pressing was done in a single-bar Grason-Stadler box located on a table in a soundproof room. All recording and control equipment was outside. Water was not available in the box. Reward was a 45-mg Noyes pellet; the punishment was a scrambled .40-mA shock of 1-sec duration. Rewards and punishments were locked in phase. For example, in the FR 4/PR 4 condition every fourth response was both rewarded and punished. Twenty-eight Ss were discarded for failure to satisfactorily complete the initial bar press training (including equipment failure). In addition, one rat extinguished on VR 2 during the second day and there were three equipment failures during testing on the fourth day. The analysis was based on 64 Ss, four per cell.

RESULTS

Means of the three performance measures are shown in Table 1.

The nonpunishment conditions (three satiation and one extinction cells) provide a base against which to evaluate the other conditions. They tell us how much bar pressing occurs when punishment is not a factor. Perhaps the most important fact is that all nonpunished Ss made more than 100 responses and, if rewarded, obtained at least 100 pellets. Satiation Ss met the response-cessation criterion after receiving 126 to 324 pellets; the mean was 221 pellets. Response means for the three reward ratios were 234, 397, and 920 for FR 1, FR 2,

Table 1
Mean Number of Shocks, Pellets, and Responses

	FR 1	FR 2	FR 4	No Reward
	4	5	4	4
PR 1	4	2	1	–
	4	5	4	4
	4	4	3	5
PR 4	14	7	3	–
	14	15	11	22
	6	10	11	3
PR 16	102	88	44	–
	102	168	177	48
	–	–	–	–
No	234	198	230	–
Punishment	234	397	920	234

The numbers from top to bottom within each cell are mean shocks received, mean pellets earned, and mean responses, all prior to meeting the 1000-sec response-cessation criterion.

(Continued on page 232)

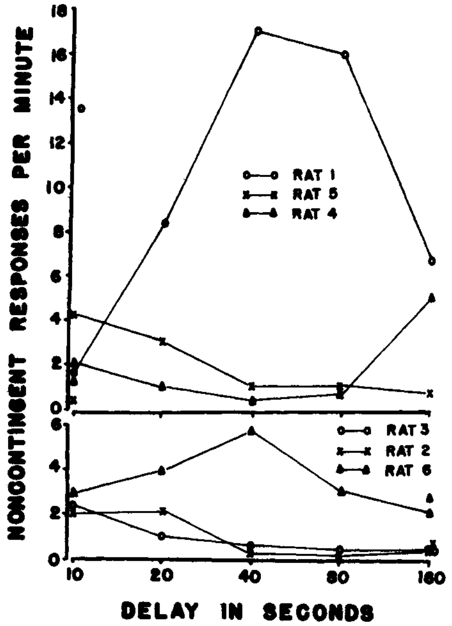

Fig. 2. Noncontingent response rates as a function of delay in seconds. Ss in upper graph received short to long delay sequence; Ss in lower graph received the reverse. Unconnected data points are recovery data.

delay value. Numbers of responses were recorded in each of several subintervals of the delay value. Such an analysis would reveal particular patterns of responding, such as a fixed-interval "scallop." These data showed no systematic trends, either within or between Ss, and hence are not shown.

The lack of order in the noncontingent data might be predicted from Ferster's (1953) analysis of light acting as both a conditioned reinforcer for contingent responding, and as an S^D for other behaviors which might adventitiously be reinforced by pellet delivery. On the other hand, this analysis would not be supported by the subinterval data if it were predicted that lever pressing, a prepotent response, would show consistency during the delay interval. Therefore, the data presented here cannot, at the present time, be viewed as supporting any particular interpretation.

REFERENCES

AZZI, R., FIX, D. S. R., KELLER, F. S., & ROCHA E SILVA, M. I. Extereoceptive control of response under delayed reinforcement. Journal of the Experimental Analysis of Behavior, 1964, 7, 159-162.

DEWS, P. B. Free-operant behavior under conditions of delayed reinforcement. I. CRF-type schedules. Journal of the Experimental Analysis of Behavior, 1960, 3, 221-234.

FERSTER, C. B. Sustained behavior under delayed reinforcement. Journal of Experimental Psychology, 1953, 45, 218-224.

FERSTER, C. B., & HAMMER, C. Variables determining the effects of delay in reinforcement. Journal of the Experimental Analysis of Behavior, 1965, 8, 243-254.

SKINNER, B. F. *The behavior of organisms.* New York: Appleton Century Company, 1938.

(Continued from page 230)

and FR 4, respectively. The Ss given simple extinction without punishment averaged 234 responses, the same number as FR 1 satiation Ss.

Punishment, on either PR 1 or PR 4 schedules, was highly effective in suppressing responding at all three reward ratios. No PR 1 rat made more than seven responses or received more than five pellets; no PR 4 animal made more than 24 responses or received more than 24 pellets. The PR 16 animals showed much less suppression. They made from 32 to 373 responses; seven of the 12 Ss made more than 100 responses. Pellets received ranged from 8 to 241, although only two Ss earned over 100 pellets. Thus, there was at least some overlap of PR 16 performance with that of nonpunished Ss. At the most favorable reward ratio, the PR 16 rats averaged 42% of the pellets received by satiation Ss. This implies that at a higher punishment ratio, perhaps around PR 40, satiation would occur before suppression.

In one sense it is a foregone conclusion that Ss punished less frequently will make more responses. For example, a PR 16 animal who quits responding after three shocks will have made at least 48 responses while a PR 1 rat who also stopped responding after three shocks would only have three responses to his credit. However, the pertinent fact is that the PR 16 Ss took more shocks than either PR 1 or PR 4 Ss. If one looks at the distribution of shocks received in the PR 1 and PR 4 conditions, the number of shocks range from one to seven with a mode (and median) of three. Six of the 12 PR 16 Ss took more than seven shocks.

Punishment was also applied at three different ratios when no reward was given. The extinction-with-punishment animals quit responding after a few shocks, even with the PR 16 condition.

DISCUSSION

Theoretically, the effects of punishment should depend upon both punishment parameters (e.g., punishment intensity) and factors related to the strength of the behavior being punished (e.g., reward magnitude). However, reward variables often count for little in punishment studies. In the present experiment punishment quickly stopped responding under all conditions save those involving reward coupled with infrequent punishment. The sustaining function of reward was only revealed from the effect of its absence in the case of punished extinction. When reward is discontinued punishment frequency seems to lose its differential control of behavior. Thus, intermittent punishment may suppress behavior less than continuous punishment or frequent punishment only when some reward is available.

REFERENCES

AZRIN, N. H., HOLZ, W. C., & HAKE, D. F. Fixed-ratio punishment. Journal of the Experimental Analysis of Behavior, 1963, 6, 141-148.

BEECROFT, R. S., & KRUGER, B. M. VR 2 interresponse times. Psychonomic Science, 1967, 9, 383-384.

CHURCH, R. M., & RAYMOND, G. A. Influence of the schedule of positive reinforcement on punished behavior. Journal of Comparative & Physiological Psychology, 1967, 63, 329-332.

HOLZ, W. C. Punishment and rate of positive reinforcement. Journal of the Experimental Analysis of Behavior, 1968, 11, 285-293.

NOTE

1. This research was supported by Grant MH 11734-02 from the National Institutes of Health to Judson S. Brown.