

# Infrequent reward facilitates observing responses in rats<sup>1</sup>

JAMES S. McMICHAEL, DENISON UNIVERSITY; JOHN T. LANZETTA, DARTMOUTH COLLEGE;  
JAMES M. DRISCOLL, UNIVERSITY OF LOUISVILLE

Rats had the option of pressing either of two stimulus levers to gain access to a third lever which produced food reinforcement on 20%, 50%, or 80% of the trials for different groups. One stimulus lever led to stimuli correlated with reward and no reward; the other lever always led to the same stimulus. Delay of reward facilitated choice of the correlated stimuli, while increasing deprivation time had equivocal effects. Ss rewarded on 20% and 50% of the trials increased their frequency of choice of the correlated stimuli, while the 80% group showed a decrease with respect to baseline responding.

Observing responses ( $R_O$ ) are studied in the situation where the outcome of an instrumental response is uncertain, e.g., a given response may lead to reward, punishment, or nothing; S may be on a mixed schedule of reinforcement, etc. A second response,  $R_O$ , provides exteroceptive cues which are correlated with the outcome, e.g., indicates the availability of food, warns of punishment, or converts the mixed schedule into a multiple one. A number of investigators (e.g., Kendall, 1965; Prokasy, 1956; Lockard, 1963; Driscoll, Lanzetta, & McMichael, in press) have shown that a variety of species will choose, via  $R_O$ , to be informed of the nature of the outcome. Wyckoff (1952, 1959) has provided a framework within which observing behavior may be interpreted by stating that cues which reliably predict reinforcement acquire greater secondary reinforcing properties than do unreliable cues. Other theorists have proposed that the reliable cues produced by  $R_O$  enable S to make preparatory responses which are appropriate to the outcome (Knapp, Kause, & Perkins, 1959; Prokasy, 1956) or that the stimuli correlated with the outcome reduce conflict (Berlyne, 1962). It also seems possible that  $R_O$  may be made to avoid any frustration attendant upon nonreward after the equivocal stimulus following not making  $R_O$ .

The present study was undertaken chiefly to investigate the parameter of reward probability. In the course of carrying out the experiment the additional variables of deprivation time and delay of reward were introduced.

## Method

The Ss were 24 naive, male, albino rats aged 100 days at the beginning of the experiment.

The apparatus was a 19 x 13 x 9 in. aluminum box with a floor of 3/32 in. stainless steel rods spaced 1/2 in. apart. On one end wall, 1-3/4 in.

from the floor and 2-1/4 in. apart were two retractable levers (stimulus levers), and centered on the other end wall 1-3/4 in. from the floor was a third retractable lever (food lever). To the right of the food lever was the food dish with a Gerbrands dispenser capable of delivering a 45-mg food pellet. Above the apparatus was a speaker for presentation of white noise with an intensity of 78 dB (re .0002 dyne/cm<sup>2</sup>) measured at the floor of the experimental chamber. The entire apparatus was enclosed in a sound-deadened compartment. Appropriate electro-mechanical programming and recording devices were located in an adjacent room.

All Ss were first shaped to press the food lever and then to press one of the two stimulus levers to gain access to the food lever. Following shaping, training trials were run wherein the stimulus levers were retracted and the food lever was extended until S responded, ending the trials. During this phase of the experiment reward was delivered on each trial following the response on the food lever. Trials were run with a 1 min. intertrial interval. If S failed to respond to the stimulus levers within 15 sec., the levers were retracted and the trial was rerun after a 4 min. interval. Trials were then run in blocks of six trials within which S was forced to respond three times to each stimulus lever.<sup>2</sup> Forced trials were implemented by failing to extend a lever once it had been responded to on three of the trials, and therefore, the number of forced trials was a measure of strength of choice of the other lever.

For the first 60 trials, neither the lights nor the noise were presented—the purpose of the trials being to determine S's initial side preference. Subsequently the non-preferred side was used for the stimulus lever which produced the correlated stimulus. Following these trials, the Ss were divided into three eight-S groups with groups receiving reward on 20%, 50%, and 80% of the trials, respectively. Prior to introducing the stimuli, 15 six-trial blocks (one block per day) were run to obtain base-line measures of forced trials to the non-OR lever for each of the experimental groups. Subsequently, pressing the OR lever produced the intense light on reward trials and produced the dim light on no reward trials, while presses on the non- $R_O$  lever produced noise on both reward and non-reward trials. Each S was run for 15 six-trial blocks one day a week for seven weeks. For the first four weeks, Ss were always run at two days of food deprivation, and for Weeks 5 through 7,

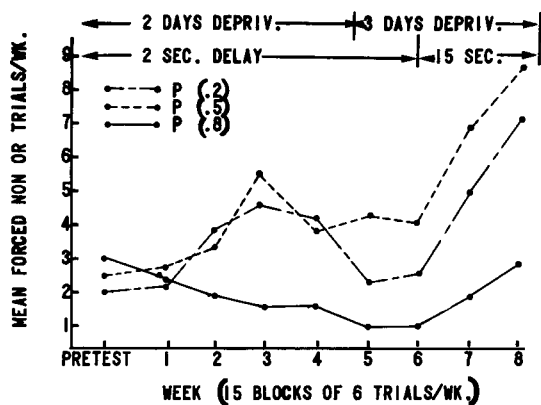


Fig. 1.

deprivation was increased to three days. In addition, for Weeks 6 and 7, a 15 sec. delay was introduced between the pressing of the stimulus lever (stimulus onset) and the extension of the food lever. The average delay prior to the introduction of this experimental condition was the average time for the animals to run from the stimulus to the food levers, on the order of 2 sec.

#### Results

Figure 1 shows that for the group being rewarded on 20% of the trials, the mean number of forced, non- $R_0$  trials per day (15 blocks of six trials) rose steadily with the introduction of the stimuli (Weeks 1 through 4), with increasing deprivation to three days (Week 5), and with the introduction of the 15 sec. delay of reward (Weeks 6 and 7). A similar trend is seen in the data of the group rewarded on 50% of the trials with the exception that increasing deprivation slightly reduced the tendency to make observing responses. For the Ss being rewarded on 80% of the trials,  $R_0$  frequency declined from the base-line. A comparison for the last five blocks between the means for the 80% and 20% groups showed significantly more forced non- $R_0$  trials in the 20% group ( $t=3.119$ ,  $df=13$ ,  $p<.01$ , two-tailed).

#### Discussion

The results indicate that probability and possibly delay of reward are important determinants of the OR. If one were to assume that the function of OR were simply to reduce uncertainty, one would predict that the 50% group would make more ORs, since for them the correlated stimuli were more informative, and one would expect the 20% and 80% groups to be equal in their frequency of ORs. The present findings indicate that factors other than sheer information content govern the choice of correlated stimuli. The tendency of the 20% reward group to make  $R_0$  most often among the groups tested could indicate that aversive consequences (frustration) of making non- $R_0$  and not being rewarded are respon-

sible for the maintenance of observing behavior. Alternatively, the difference between the 20% and 80% groups may reflect a greater importance of reducing uncertainty when reward is rather improbable than when it is probable (see Berlyne, 1962).

The finding that delaying reinforcement increases the frequency of  $R_0$  seems contrary to what one might expect from a preparatory response framework (Prokasy, 1956), since lengthening the delay might vitiate S's ability to time his preparatory response. Furthermore, the approximate 2 sec. interval S required to cross the apparatus after pressing a stimulus lever and to press the goal lever added to the time required to pick up and ingest the food pellet would seem sufficiently long for a preparatory response. A conflict resolution formulation could predict that  $R_0$  would occur only if S is permitted to respond differentially during the interval between the  $R_0$ /non- $R_0$  choice response and the outcome (i.e., resolve response uncertainty). In the present case where no delay of reward was imposed by the conditions of the experiment, all Ss, by their prior training, would be crossing the apparatus and pressing the goal bar during the interval in question, and such responding could preclude resolution of response uncertainty. The imposed 15 sec. interval would permit differential responding and would thereby maintain observing behavior at a higher level according to a response-uncertainty-resolution interpretation.

#### References

- Berlyne, D. E. Note on food deprivation and extrinsic exploratory responses. *Psychol. Rep.*, 1962, 11, 162.
- Driscoll, J. M., Lanzetta, J. T., & McMichael, J. S. Preference for information under varying conditions of outcome uncertainty, intensity, and delay. Unpublished.
- Kendall, S. B. An observing response analysis of fixed ratio discrimination. *Psychon. Sci.*, 1965, 3, 281-282.
- Knapp, R. K., Kause, R. H., & Perkins, C. C., Jr. Immediate vs. delayed shock in T-maze performance. *J. exp. Psychol.*, 1959, 58, 357-362.
- Lockard, Joan S. Choice of a warning signal or no warning signal in an unavoidable shock situation. *J. comp. physiol. Psychol.*, 1963, 56, 526-530.
- Prokasy, W. F. The acquisition of observing responses in the absence of differential external reinforcement. *J. comp. physiol. Psychol.*, 1956, 49, 131-134.
- Wyckoff, L. B. The role of observing responses in discrimination learning: Part I. *Psychol. Rev.*, 1952, 59, 431-442.
- Wyckoff, L. B. Toward a quantitative theory of secondary reinforcement. *Psychol. Rev.*, 1959, 66, 68-78.

#### Notes

1. This study was supported in part by the Air Force Office of Aerospace Research under contract No. AF 49(638)-1614. Data were gathered while the authors were at the University of Delaware.
2. A forced choice procedure was used to approximately equalize the number of rewards for  $R_0$  and non- $R_0$  responses. Though the absolute number of rewards differed for the 20%, 50% and 80% reward groups, the proportion of rewards following  $R_0$  and non- $R_0$  responses was approximately equal for each experimental group at each stage of the experiment.