Maximization of pellet density by rats in a two-choice operant situation*

GRANT COULSON, KENNETH KOFFER, and VALERIE COULSON
York University, Downsview, Ont., Canada

Four food-deprived experimentally naive male albino rats were given a choice between two levers. Pressing either lever after a food pellet delivery started a time period of variable length which ended with a pellet delivery, providing neither lever had been pressed in the immediately preceding 5 sec. Pressing one of the levers delivered a pellet more often than did pressing the other. Responses were generally allocated to the lever producing pellets more frequently. We conclude that rats can distinguish between pellet densities and will maximize pellet densities even when no response is immediately reinforced by being followed by a pellet.

Many classical two-choice situations (e.g., mazes, jumping stands, etc.) require the S to make a correct response before receiving the reinforcer associated with that response; an incorrect choice may have no reinforcer correlated with it or may even be punished. However, in operant two-response choice situations, both responses are usually continuously available and are distinguishable by being correlated with different reinforcement schedules.

When the choice is between two different aperiodic reinforcement schedules, the distribution of responses is generally directly proportional to the reinforcement frequency associated with the two responses (Herrnstein, 1970). This pattern of matching response frequency to reinforcement frequency tends to maximize reinforcement density and probably occurs because the programmers for the two schedules run independently and because neither stops except while a reinforcer is available to its correlated response. This means that if S switches from one response to another frequently enough, he will obtain all the reinforcers almost as soon as they are available.

A different set of contingencies associated with a two-choice situation might produce a different result. In the present experiment, if either of two levers was pressed after a pellet delivery, one of two timers started and a food pellet was delivered after a variable length of time. Pressing one lever delivered pellets sooner than did pressing the other. Presses made when either of the timers was running had no effect.

*This research was completed while the first two authors held predoctoral fellowships from the Addiction Research Foundation of Ontario.

At least two patterns of responding might be expected under these contingencies. The Ss could, as in the experiments reviewed by Herrnstein (1970), match relative response rates to relative pellet densities or maximize pellet density by exclusively pressing the lever associated with the higher density of pellet delivery. Both patterns, of course, require discrimination between two different rates of pellet delivery. The present experiment was designed to determine which, if either, of these two patterns would occur.

SUBJECTS

The Ss were four experimentally naive male Sprague-Dawley rats, 90 days old at the beginning of food deprivation. Ss were housed in individual cages and had free access to water.

APPARATUS

The operant chamber was a 30-cm cube constructed of aluminum, except for a Plexiglas door and stainless steel grid floor. Two Lehigh Valley rat levers (Model 1532) were mounted on one end wall (8 cm from the floor, with their centers 8.5 cm from either side wall). Each lever projected 4.2 cm into the chamber and required a downward force of 14 g to actuate it. A foodwell was set into the wall between the levers, centered 6.5 cm above the floor. Regular formula 45-mg Noyes pellets were delivered by a Gerbrands pellet dispenser. The chamber was illuminated by a 7-W bulb in a light socket mounted above the foodwell. Both the operant chamber and feeder were housed in a sound-attenuating shell, and electric fans provided ventilation and masking noise. Electromechanical programming equipment was housed in an adjacent room, and the Ss were monitored via closed-circuit television.

PROCEDURE

The Ss were deprived to 80% of

their free-feeding weights during the 7 days prior to magazine training and held at 80% (with corrections for normal growth) throughout the experiment. On the first magazine session, after 10 min of exploring the experimental space, pellets were delivered on a variable interval (VI) 30-sec schedule for 1 h. During the next session, pellets were delivered on a VI 120-sec schedule for 2 h. Neither lever was present during magazine training. Sessions subsequent to magazine training lasted 2 h or until 200 pellets were delivered, whichever came first. Consecutive sessions were run on consecutive days.

On the third session, both levers were made available and the following contingencies were introduced: pressing one of the levers started a tape programmer which ran for 10 sec on the average; a fixed-duration timer then started and ran for 5 sec, during which a response on either lever reset the fixed-duration timer. A pellet was delivered when this timer timed out. This schedule is labeled: tandem variable interval of delay 10 sec, differential reinforcement of other behavior 5 sec (tand VID 10 DRO 5).

During the first phase, the other lever delivered pellets on a tand VID 25 DRO 5 schedule for all the Ss. When the DRO timer was running, presses on either lever always reset the timer; responding when either of the tape programmers was running had no effect. The first response after pellet delivery was the *effective* response in determining which schedule delivered the next pellet. For Ss 1 and 2, the left lever and for Ss 3 and 4 the right lever was associated with the tand VID 10 DRO 5 schedule during the first phase of the experiment. After a number of sessions, the schedules associated with the levers were switched. Since Ss 3 and 4 persisted in pressing the lever producing the lower pellet density, the density of this lever was changed from 1 pellet/30 sec (tand VID 25 DRO 5) to 1 pellet/60 sec (tand VID 55 DRO 5) and then to 1 pellet/120 sec (tand VID 115 DRO 5), and finally, for No. 3, no pellets were delivered for pressing the less dense lever.

RESULTS

Figure 1 shows response rates per minute, averaged over two sessions, for each of the four Ss. The filled triangles or circles always represent the rate on the lever associated with the tand VID 10 DRO 5 lever. The open circles or triangles represent the rate on the lower density lever. For example, the label "60 sec" above the rates for a phase indicates that a tand VID 55 DRO 5 schedule was associated with the less dense lever. If no designation appears above the rate for a phase, the

Fig. 1. Responses per minute for individual Ss averaged over two sessions as a function of experimental session and contingencies in effect. Contingencies changed after each vertical dashed line. See text for further details.

schedule for the less dense lever was tand VID 25 DRO 5.

By the end of the first phase, Ss 1 and 2 were pressing the left (higher density) lever almost exclusively, although the shift to this lever took 14 sessions for S 2. When the associated contingencies were switched and responding on the right lever was reinforced on the tand VID 10 DRO 5 schedule, S 1 took 28 sessions and S 2, 8 sessions before most of their responses were made on the higher density lever.

Ss 3 and 4 began pressing the right (higher density) lever immediately, and both Ss made essentially zero responses on the less dense left lever. When schedules were switched for S 3 and the interpellet interval was increased on the less dense lever, he continued to allocate more responses to the right lever until no pellets were delivered for responding on the right lever. After changing contingencies for S 4, more responses were made on the left lever only when a tand VID 115 DRO 5 schedule was in effect on the right lever. Response rates for S 4 had been converging up to this point and may have crossed in any case.

In general, session lengths were inversely proportional to response rates on the denser lever; sessions with low rates, such as those of Phases 2, 3, and 4 for S 3, lasted 120 min. Sessions

with higher rates, such as those in Phase 1 for Ss 3 and 4, lasted an average of 73 and 85 min for the last five sessions, respectively. During the DRO component, responses on each lever were usually directly proportional to the overall response rate for each lever.

During the delay period, no noticeable within-delay behavior chains were observed, probably because the delays were variable rather than fixed.

DISCUSSION

The present experiment clearly indicates that when the ostensible reinforcer, the pellet, was removed by at least 5 sec from any response, and by an average of at least 15 sec from the effective response, Ss responded almost exclusively on the lever that produced the greatest change in pellet density, especially during the initial acquisition sessions. In other words, they maximized pellet density rather than distributing responses directly proportional to the frequency of reinforcement.

When the contingencies associated with the levers were switched, two of the Ss, Nos. 3 and 4, did not switch responding until the difference in pellet density produced by the two levers was increased. This potentiation, plus the initial nonzero rate of responding on the right lever for Ss 1

and 2, probably reflects a preference for the right lever.

The data indicate that rats can discriminate between pellet densities of 1 per 15 sec and 1 per 30 sec. More generally, they demonstrate that rats can "integrate over time [Sidman, 1966, p. 483]," a finding with important implications for the concept of reinforcement. Keehn & Coulson (1970) have shown that rats will learn to drink an initially nonpreferred solution of ethanol instead of distilled water when drinking ethanol increases pellet delivery from 1 per 120 sec to 1 per 15 sec. More pellets were delivered for responding than for not responding, and responding increased. This is another example of integration over time.

Similarly, discrete-trial experiments using the T-maze have shown that rats exhibit "temporal integration," referring to "... the process whereby rewards and punishments which are distributed over a time interval following a response are combined into a net value [Renner, 1967, p. 400]."

Logan (1965), for example, using a T-maze, found that delay of reinforcement and magnitude of reinforcement were, in essence, combined into a "net value." An increase in delay, up to a certain point, could be offset by an increase in reinforcer magnitude in determining choice behavior.

The usual notion of reinforcement is that reinforcers should immediately follow at least some responses, especially during response acquisition. This paradigm was violated in the present experiment, in which a temporally heterogeneous sequence of pellet delivery functioned as reinforcement during both response acquisition and maintenance even when no response was immediately followed by a pellet. The usual way in which reinforcers are delivered, immediately after some responses, may be a special case of an increase in reinforcer density.

REFERENCES

- HERRNSTEIN, R. J. On the law of effect. *Journal of the Experimental Analysis of Behavior*, 1970, 13, 243-266.
- KEEHN, J. D., & COULSON, G. E. Ethanol consumption by rats on a differential probability of reinforcement schedule. *Psychonomic Science*, 1970, 19, 283-284.
- LOGAN, F. A. Decision making by rats: Delay versus amount of reward. *Journal of Comparative & Physiological Psychology*, 1965, 59, 1-12.
- RENNER, K. E. Temporal integration: Modification of the incentive value of food reward with early experience with deprivation. *Journal of Experimental Psychology*, 1967, 75, 400-407.
- SIDMAN, M. Avoidance behavior. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application*. New York: Appleton-Century-Crofts, 1966. Pp. 448-498.