CORRELATED DELAY OF REINFORCEMENT[1]

GORDON H. BOWER[2]

Yale University

It is well established that variations in an operant response can be differentiated by arranging appropriate reinforcement conditions. For example, Skinner (1938) and Ferster and Skinner (1957) have selectively reinforced duration, force, and rate of lever pressing in rats. In each case, reinforcement was applied whenever the "correct" response variation occurred and was withheld for all other variations of the response. However, this all-or-none reinforcement is probably not a necessary condition for response differentiation; it may be sufficient that the specified response variation receive a slightly more favorable reward (e.g., larger amount, shorter delay, higher percentage of reward) than do alternative variations of the operant. Under such conditions, if the normal variability in details of the operant ensures some experience with the reward differential, then we would expect an increase in the relative frequency of the specified response variation.

The present experiments deal with the selective reinforcement of the temporal characteristics of a behavior chain. Rats were reinforced according to the time they took to traverse a straight runway. For convenience, we shall refer to the time the rat takes to reach the end box as "response time" and the reciprocal of this time as "response speed." Such measures do not imply anything about the form of the response; e.g., a long response time of 5 sec. (say) can be achieved by a large number of behavior sequences other than "running forward at a steady slow pace."

In the studies to be reported the delay of

reinforcement was made contingent upon response time, the time of delay being shorter for long response times. The contingency was two-valued; a cutoff time was selected, and if the rat took longer than the cutoff time to reach the goal box, it was rewarded immediately; if its response time was shorter than the cutoff time, then a fixed delay intervened between completion of the response and delivery of reward.

Logan (1960) has studied one condition of this sort. Rats were given reward immediately if they took longer than 5 sec. to reach the goal end of a 5-ft. runway, and were given the reward after a 10-sec. delay if they took less than 5 sec. For convenience, conditions like this will be referred to by the expression 5-0, 10D, where the first number represents the cutoff time and the last two numbers represent the delay of reward for response times longer and shorter than the cutoff time, respectively. Logan found that rats trained under this condition gradually learned to respond slowly, their mean speeds stabilizing at an asymptotic level near the cutoff speed (i.e., reciprocal of cutoff time). At this point about 40% of their responses were slow enough to receive reward immediately.

In conjunction with each of these 5-0, 10D Ss a matched control mate was trained. The matched control rat received the same sequence of delays as that earned by its experimental partner; however, these delays were given irrespective of the control S's response speed. At the end of training the matched control Ss were responding considerably faster than their experimental mates; indeed, none of their response times was greater than 5 sec. The difference in performance indicates that response speed depends not only on the delay of reward, but also upon the correlation between delay and response speed.

The purpose of the present research was to provide a more systematic description of the effects of correlating delay of reward with response time. For the two-valued contingencies considered here, the relevant variables

are the cutoff time, the delay for speeds below the cutoff, and the delay for speeds faster than the cutoff speed. In these studies the delay for speeds below the cutoff was always zero, whereas the cutoff and the delays for speeds faster than the cutoff differed for different groups of Ss.

Considering only how quickly the rat can get the reward, one would not expect some conditions of this type to favor slow responses. If the cutoff time is large relative to the delay of reward incurred by response speeds above the cutoff, then the rat may be able to get the reward sooner by responding fast. For example, in a 6-0, 2D condition, although response times greater than 6 sec. get the reward without delay, the rat can get the reward sooner if it responds in less than 4 sec. and takes the 2-sec. delay.

There are two ways to modify this condition so that the S would be more likely to be influenced by the reward differential. One is to lower the cutoff time (say, from 6 to 2 sec.) so that S need not spend so much time in the alley to get the reward immediately; the other is to lengthen the delay (say, from 2 to 30 sec.) for speeds faster than the cutoff. If the delay for fast responses is large relative to the cutoff time, then S can get the reward sooner by responding at speeds slower than the cutoff speed.

These studies investigate these factors using different groups of Ss. Experiment 1 was a factorial design combining two cutoff conditions with three conditions of delay for speeds faster than the cutoff. In Experiment 2 the cutoff was varied over a larger range with the delay contingencies held constant. In both experiments response speeds slower than the cutoff speed received reward immediately.

METHOD

Subjects

Thirty-four male albino rats of the Sprague-Dawley strain, about 90 days old, served as Ss; 22 were in Experiment 1 and 12 in Experiment 2. They were housed in individual living cages with free access to water. They were put on a 12-gm. daily feeding cycle and handled daily for ten days before training began.

Apparatus

The apparatus was a 5-ft. straight alley with photobeams spaced at 1-ft. intervals along the runway.

TABLE 1

EXPERIMENT 1 CONDITIONS: CELL ENTRIES INDICATE THE NUMBER OF RATS IN EACH CONDITION

Delay (sec.) for Fast Responses	Cutoff Time (sec.)	
	4	6
2	2	5[a]
10	2	2
30	3	3

[a] Matched controls were run for this condition.

Details of construction are described elsewhere (Logan, Beier, & Ellis, 1955). A Hunter interval clock recorded the time from the opening of the start door to S's breaking a photobeam 2 in. in front of the food cup. One interval timer timed the cutoff period from the opening of the start door; a second timer timed the delay of reinforcement in case S's response time was shorter than the cutoff time. A solenoid, which made an audible click when operated, dropped 10 P. J. Noyes food pellets (45 mgm. each) into the food cup after the delay of reinforcement. Between trials Ss were placed in a detention box, 24 in. by 24 in. by 4 in., where water was continuously available.

Procedure

On Days 11 and 12 of Experiment 1 each S was placed in the alley and allowed to roam about for 2 min. On Days 13 to 15 each S was placed directly in the goal box with food present and allowed 5 min. to eat. Goal-box training continued at five trials a day for each S until it ate within 15 sec. on three consecutive placements in the goal box. This alley and goal-box habituation was eliminated in Experiment 2, the Ss beginning training on Day 11 with no special pretraining.

Experiment 1 was a 2 × 3 factorial design (see Table 1) with cutoff time (4 or 6 sec.) and delay for response times less than the cutoff time (2, 10, or 30 sec.) as the two factors. The number of Ss in each condition is indicated in the appropriate cell of Table 1. In those cells containing two Ss, a third S began the experiment but died during the prolonged course of training. Since it was expected that Ss in the 6-0,2D condition would eventually respond faster than their cutoff speed of .167, the question arose as to whether their terminal performance would be affected by the experience early in training with immediate reward for response speeds less than .167. To answer this question, a matched control group was trained along with the 6-0,2D Ss. Because of Logan's prior data it was not considered necessary to run matched control rats for all conditions just to show that performance was affected by the correlation between response time and delay time.

The 12 Ss in Experiment 2 were equally divided into three conditions differing only in the cutoff time (2.5, 5, or 10 sec.). For each of these Ss a 10-sec. delay of reward was imposed for response times less than S's
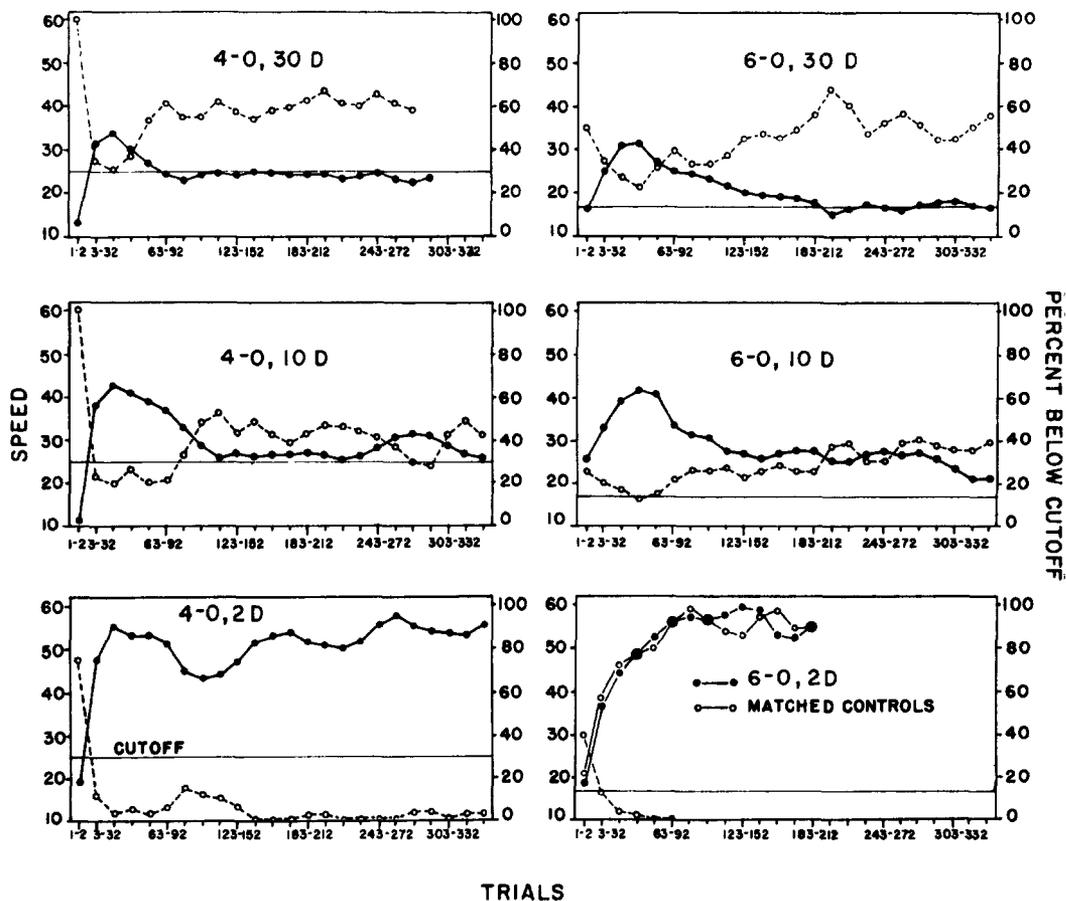
TRIALS

FIG. 1. Learning curves for the groups in Experiment 1, plotted in sliding blocks of 30 trials. Solid curves are mean speeds, referred to left ordinate; dashed curves are percentages of speeds below cutoff, referred to right ordinate. The horizontal line on each graph is the cutoff speed.

appropriate cutoff time, whereas reward was given immediately for response times longer than its cutoff time.

With the exception of the first four days when Ss were run 1, 1, 2, and 3 trials, respectively, the Ss had 5 trials a day with an average intertrial interval of 35 sec. during which time S was in the detention cage with water available. The running order of Ss was randomly determined but remained fixed throughout the experiment. Ten minutes after its last trial of the day each S received its daily ration of 12 gm. minus the amount of food received in the runway. The Ss in Experiment 2 received 300 trials and the Ss in Experiment 1 received 367 trials with the exception of the 4-0,30D group (312 trials) and the 6-0,2D group (212 trials). The last 100 trials for a particular group were used in statistical analyses of terminal performance. At the end of the trials specified above the groups were shifted to other cutoff and delay conditions of interest; the results of these shifts are reported in the dissertation (Bower, 1959) but not in this article.

RESULTS

The learning curves for the groups in Experiments 1 and 2 are shown in Figures 1 and 2, respectively. The solid curves are group mean speeds and the dashed curves are the percentages of speeds which were below the cutoff speed and, hence, received immediate reward. The horizontal line on each graph is the cutoff speed.

Inspection of Figure 1 reveals several significant features. First, with the short 2-sec. delays, at the bottom of Figure 1, the Ss learned to respond at speeds much faster than the cutoff and the percentage of speeds below the cutoff dropped to an asymptote of zero. In the lower right panel it may be seen that the 6-0, 2D Ss and their matched controls did

not differ in terminal performance; at the end of training both groups were receiving the 2-sec. delay every trial.

With the longer delays, in the middle and top panels of Figure 1, the Ss learned to respond with speeds near the cutoff and speeds below the cutoff occurred more frequently. By testing the significance of differences among the group means, it was found that the longer the delay for fast speeds, the slower the terminal speed ($p < .001$ by an F test, $df = 2, 8$). Also, the longer the delay, the greater the frequency of speeds below the cutoff speed ($p < .001$). When only the 10- and 30-sec.
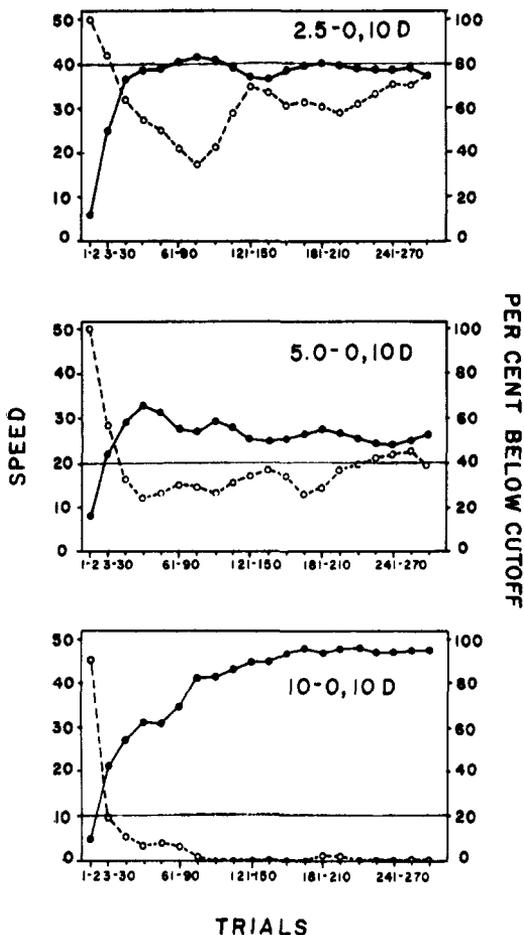

FIG. 2. Learning curves for the groups in Experiment 2, plotted in sliding blocks of 30 trials. Solid curves are mean speeds, referred to left ordinate; dashed curves are percentages of speeds below cutoff, referred to right ordinate. The horizontal line on each graph is the cutoff speed.

delay groups were considered, the effects of delay were still reliable ($p < .001$ for speed; $p < .05$ for percentage below cutoff).

In the longer delay conditions, the lower cutoff speed (right side of Fig. 1) produced slower mean speeds ($p < .001$). Although the percentage of speeds below the cutoff was higher for the 4-sec. than for the 6-sec. cutoff, this difference was not statistically significant.

Figure 2 (for Experiment 2) shows the result when the cutoff was varied over a wider range, holding constant the delay for fast speeds. The percentage of speeds below the cutoff dropped proportionately as the cutoff speed was lowered ($p < .001$). Comparison of the 2.5- and 5-sec. cutoffs reveals that mean speed was slower for the lower cutoff speed ($p < .001$) as in Experiment 1; however, when the cutoff speed was still lower (10 sec.), the Ss failed to meet it and responded faster than those Ss on the intermediate 5-sec. cutoff. Thus, here was a quadratic relationship between terminal speed and the length of cutoff ($p < .001$).

In those conditions that produced slow responding the mean-speed learning curves had similar shapes. They began at slow speeds, rose quickly to a maximum above the cutoff speed, and then fell progressively to slower speeds near the cutoff. The percentage of speeds below the cutoff reflected these changes in mean speeds; the percentage measures began at a high value, dropped to a minimum, then rose gradually to a stable level. The speed curve for the 2.5-sec. cutoff group (top panel of Fig. 2) did not appreciably overshoot the cutoff speed; however, improvement in their performance may be seen in the increasing percentage of speeds below the cutoff.

Those Ss that performed near the cutoff responded slower on later trials of the day. This within-day effect was significant in both experiments ($p < .001$ in Experiment 1; $p < .05$ in Experiment 2). When the delay was 10 sec., those Ss with the intermediate cutoffs (4, 5, or 6 sec.) ran exceptionally fast on their first trial of the day, slowing down over later trials of the day; in contrast, Ss in the other conditions showed only a slight drop over trials within the day. This differential within-day effect on speed was reflected in a reliable trials-by-delay interaction in Experiment 1
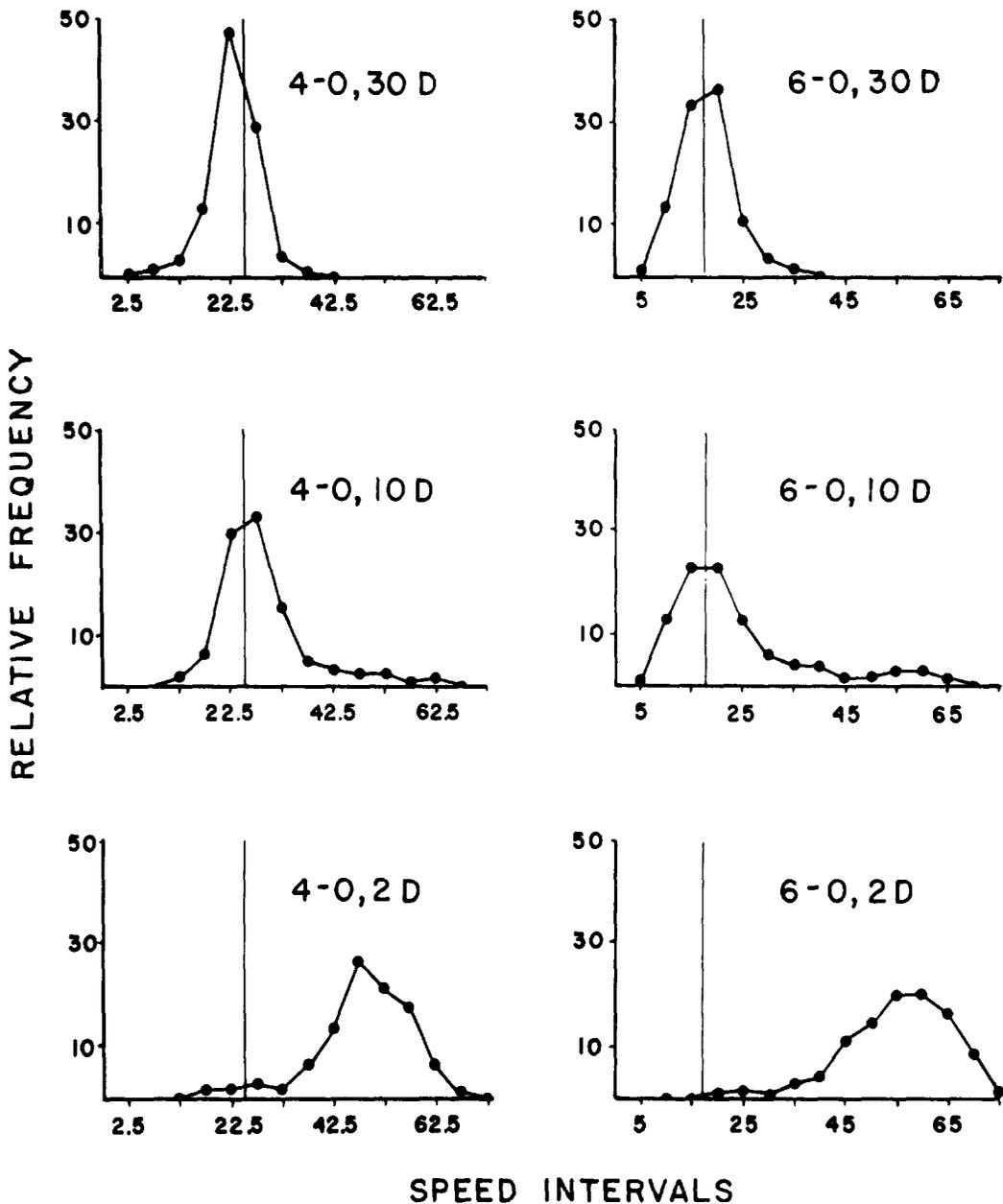
FIG. 3. Speed distributions for the groups in Experiment 1, plotted in intervals of .05 speed points. The vertical line on each graph is the cutoff speed.

($p < .001$) and a reliable trials-by-cutoff interaction in Experiment 2 ($p < .05$). The same effects were reliable considering the percentage below cutoff measures.

In Figures 3 and 4 the pooled speed distributions over the last 100 trials are shown for the groups in Experiments 1 and 2, re-

spectively. The pooled distributions are representative of the distributions of individual Ss since differences between Ss within a given condition were very small. As an indication of the small variance between Ss in the slow conditions, the average terminal speeds of the three rats in the 4-0, 30D condition were .25,

.23, and .25, while the average speeds of the four rats in the most variable 5-0, 10D condition were .27, .26, .28, and .23. In the slow conditions (middle and top panels of Fig. 3 and 4) the distributions were unimodal with most of the response speeds clustered around the cutoff (vertical line), and the relative frequency declined to either side of the cutoff. The distributions for the 2.5-0, 10D and 4-0, 30D groups showed a distinct mode at that speed interval just below the cutoff speed,

FIG. 5. Hypothetical curves showing average speed as a function of cutoff, each curve corresponding to a given delay of reward when speed exceeds the cutoff speed. The curves were smoothed through data points, using the constraints explained in the text. The heavy line represents the cutoff speed.

whereas the 4-, 5-, and 6-0, 10D groups showed a slight mode at that speed interval just above the cutoff.

The variability of an S's performance was related to its condition of reinforcement. In those conditions that produced a high degree of successful slow responding, the S's performance from trial to trial was remarkably consistent; however, when the conditions produced only moderately successful slow responding, then the S was more variable in its performance. Thus, in Experiment 1 (see Fig. 3), the within-S variability of performance was higher for Ss on the 10-sec. delay than for Ss on the 30-sec. delay ($p <$ .001). Also, as Figure 4 clearly shows, the lower the cutoff speed, the more variable was S's performance ($p < .01$ in both experiments).

DISCUSSION

These studies have shown that if the reward is delayed when S responds faster than a cutoff speed, the resulting performance is a joint function of the delay and the cutoff involved. In an attempt to characterize these relationships, a family of hypothetical curves are drawn through the data points and are shown in Figure 5. These curves show average speed as a function of cutoff time, each curve corresponding to a given delay of reward when
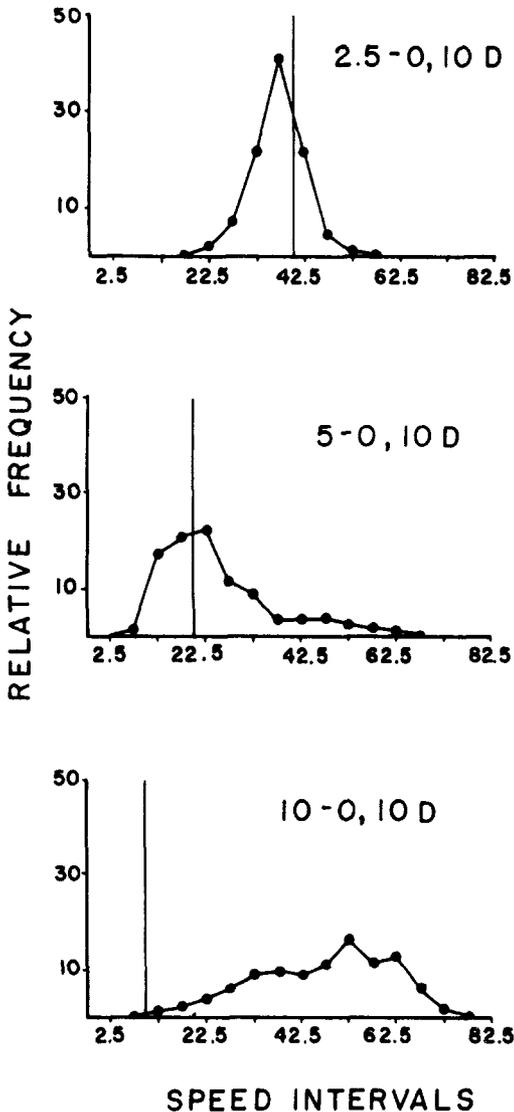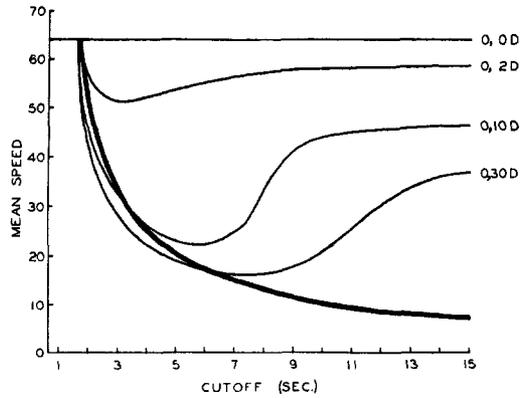

SPEED INTERVALS

FIG. 4. Speed distributions for the groups in Experiment 2, plotted in intervals of .05 speed points. The vertical line on each graph is the cutoff speed.

speed exceeds the cutoff speed. The heavy line represents the cutoff speed.

Two logical constraints were used in generating these curves. First, if the cutoff speed is above the rat's physiological limit of speed, then the delay for speeds faster than the cutoff speed is irrelevant; thus, all curves must start on the left at that speed produced by a constant 0-sec. delay of reward. Secondly, there is some minimum cutoff speed (e.g., .05 or 20 sec.) that Ss practically always exceed if reinforced at all. At such a cutoff time and beyond, S would always receive the long delay, and hence the curves should be arranged according to the decreasing function relating speed to constant delay of reward (Logan, 1960).

These two constraints determine approximately where the curves start and where they end. The present results provide information about the paths they follow. The results for Experiment 1 show that the curve for the 0, 2D condition is considerably above that for the 0, 10D condition, and that the latter is slightly above that for the 0, 30D condition. The results of Experiments 1 and 2 permit a fairly detailed tracking of the 0, 10D curve showing speed as a U-shaped function of cutoff time. Presumably the other curves follow a similar course.

There is some evidence that performance on a given condition depicted in Figure 5 is not appreciably affected by S's preliminary training. After the initial 300 training trials the Ss in Experiment 2 were shifted to the next longer cutoff (e.g., 2.5 group to 5 sec., the 5 group to 10 sec.) and their performance appeared to stabilize eventually at the same level as that for Ss trained initially under the longer cutoff (Bower, 1959).

The rising limb of the 0, 10D curve disconfirms a prediction from Logan's equilibrium model (1960), which was designed to analyze conditions of correlated reinforcement. The prediction from the equilibrium model is that by correlating delay with slow speeds, one would be unable to induce the S to respond slower than it would were the long delay being received on every trial. Clearly, however, the 10-0, 10D Ss were responding faster than the 5-0, 10D Ss, even though the former Ss received the long delay on every trial.

All the results are consistent with the micromolar theory proposed by Logan (1960). This approach assumes that different speeds are different responses with separate response strengths, and that the likelihood of any particular speed depends on the reward for that speed relative to the rewards for alternative speeds. In the present studies the relevant dimension of the reward is how quickly it is received. This is called the interval of reinforcement and involves two components: the duration of the response itself and the delay of reward after the response.

Under conventional procedures in which delay is not correlated with speed, the rat can get the reward sooner by running quickly. This is not true in the present conditions if the cutoff time is short and a long delay is imposed when speed exceeds the cutoff. Thus, among conditions with the same cutoff, the longer the delay for fast speeds, the greater the relative advantage of responding slowly since the additional time taken by the response is compensated for by a large reduction in delay. Similarly, among conditions with the same delay for fast speeds, the longer the cutoff time, the less the relative advantage of responding slowly since the required response time offsets the reduction in delay time. Accordingly, the difference in the interval of reinforcement and hence in the differential incentive for speeds above and below the cutoff depends jointly on the cutoff and delay for speeds above the cutoff.

This theory also accounts for the observed differences in variability of S's speeds under the several conditions. In general, the predicted variance of the speed distribution is a decreasing function of the size of the incentive differential established by the reinforcement conditions. The variance will be minimal when a large incentive differential favors one particular speed; in contrast, the variance will be maximal when all response speeds have equal strengths. Lowering the cutoff speed or shortening the delay for fast responses is expected to result in less differential incentive for slow speeds; accordingly, these two operations will also increase the variance of an S's speed distribution.

One final point requires mention. It was observed that Ss learned to solve the problem

of taking time in the alley by performing "ritualistic" acts which filled the time interval. The specific rituals varied between Ss; a common and striking topography was for S to run down to the edge of the goal box, turn around and retrace to the startbox, and then dash down into the goal box. That Ss learned to take time by performing such rituals was not an unexpected result. It is clear that when we define and differentially reinforce response classes according to some minimal quantitative characteristic (e.g., taking 5 sec. to break the last photobeam), we aggregate a large number of topographically distinct behaviors, like the retracing mentioned above. Because of the interdependence of reinforcement and the probability of occurrence of particular behaviors of this large class, a "trapping" or "positive feedback" principle would be expected to apply. In such cases, we expect the S to be trapped into some topographically unique way of responding; however, because the reinforcement conditions are nondifferential with respect to the members of this class, they provide no basis for predicting which unique response will be learned. To make predictions of this later type would require further information concerning the initial likelihood of occurrence of the several behaviors constituting the quantitatively defined response class. Such information was clearly unavailable in the present experiment.

SUMMARY

Thirty-four rats were trained on a runway response under conditions of negatively correlated delay of reward. The delay of reward was zero if response speed was below a cutoff speed, and the delay for speeds faster than the cutoff was either 2, 10, or 30 sec. for different groups. The cutoff time was either 2.5, 4, 5, 6, or 10 sec.

Terminal speed was a decreasing function of delay of reward for fast speeds and a U-shaped function of cutoff time. The percentage of speeds below the cutoff was an increasing function of delay and a decreasing function of cutoff time. The within-subject variance of speed was an increasing function of cutoff time and a decreasing function of delay for fast speeds. Differential within-day effects were observed: Ss in the 0, 10D condition having either the 4-, 5-, or 6-sec. cutoff ran exceptionally fast on the first trial of the day but slowed down to near-cutoff speeds on later trials; by contrast, Ss in the other conditions showed relatively stable performance over trials within the day.

The results add support to a micromolar theory in which the incentive for a particular speed is a joint function of the within-chain delay and the nonchaining delay of reward associated with that speed. Subjects cannot be viewed as simply minimizing their delay of reward. Instead, it is the cutoff time relative to the delay of reward imposed for fast speeds which determines whether S will respond near the cutoff speed.

REFERENCES

BOWER, G. H. Correlated delay of reinforcement. Unpublished doctoral dissertation, Yale University, 1959.

FERSTER, C. B., & SKINNER, B. F. *Schedules of reinforcement*. New York: Appleton-Century-Crofts, 1957.

LOGAN, F. A. *Incentive: How the conditions of reinforcement affect performance*. New Haven: Yale Univer. Press, 1960.

LOGAN, F. A., BEIER, E. M., & ELLIS, R. A. The effect of varied reinforcement on speed of locomotion. *J. exp. Psychol.*, 1955, **49**, 260–266.

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