SCHEDULES OF REINFORCEMENT AND DISCRIMINATION LEARNING¹

EDWARD J. GREEN, RICHARD M. SANDERS,
and
ROGER W. SQUIER, JR.²

DARTMOUTH COLLEGE

Previous investigations have been directed toward the analysis of discrimination learning under fixed-ratio schedules of reinforcement. Other parameters have been investigated, such as the length of the interval during which discrimination stimuli are presented, the degree of overlap between S^D and S^A subsets of stimulus elements, and the effects of using the same or different stimulus series upon the speed and variability of discrimination learning (Green, 1955, 1956). The aim of the present study is to evaluate the effects of a wide range of schedules of reinforcement upon discrimination learning under standard conditions, with apparatus adopted by other investigators and a method of stimulus presentation which lends itself easily to replication.

METHOD

Seventy adult male students from an introductory course in psychology served as S s. Their ages ranged from 18 to 20 years. Five S s were assigned to 13 major experimental groups and one additional control group. Except with three of these groups, each experimental session lasted for 1 hour, during which 60 stimulus patterns were presented visually to the subject. Discriminative stimuli consisted of combinations of lights presented upon a panel, where ten jewel lights were mounted in two rows of five lights each. Thirty of the presentations of stimuli were occasions for reinforcement, S^D ; and half, or 30, of the stimuli were occasions for nonreinforcement, S^A . A different sequence of stimulus presentation was used with each S in every group. After each S was run, leads to the stimulus panel were disconnected, randomly shuffled, and reattached. In this way, no S could communicate to another the particular light or combination of lights which occasioned reinforcement. During S^D , responding by pulling a Lindsley Manipulandum was reinforced under the appropriate schedule. Reinforcement consisted of awarding S a point on an electrical counter mounted on the panel before him. His instructions directed him to obtain as high a score as possible during the time available to him. In the first experimental group, reinforcement was programmed on a continuous schedule. Reinforcement was programmed on a fixed-interval schedule for three more groups, the intervals being 5, 15, and 30 seconds. In another set of three groups, reinforcement was programmed on the basis of variable-interval reinforcement, the intervals also being 5, 15, and 30 seconds. Two sets of ratio groups were also run. Under the fixed-ratio condition, FR 5, FR 10, and FR 20 schedules were used. An additional FR 15 group was also run.

Data from VR 5, VR 10, and VR 20 groups were obtained in a related study (Burdick, 1958) under conditions identical to the present, except that only 25 S^D 's and 25 S^A 's were presented during discrimination training.

¹This investigation was supported in part by Research Grant M-2200 from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service. Apparatus was provided by a Faculty Research Grant awarded by Dartmouth College.

²The authors are greatly indebted to Dr. D. C. Butler for assistance in the analysis of data, and to Dr. H. A. Burdick for use of VR data that he gathered in a related study.

The experimental chamber was dark, except for the lights comprising the stimulus patterns and a small 7-watt bulb illuminating the electrical counter. White noise was introduced for masking purposes both through headphones and through a loud-speaker placed in the room. To further rule out the possibility, especially under the interval conditions, of the *Ss* discriminating clicks from the programming apparatus, a timer was installed to generate spurious clicks in a random sequence. Initially, about half of the *Ss* did not respond at a high enough rate to meet the conditions of the schedule used. In those cases, *E* stopped the experiment after the first stimulus presentation, entered the experimental room, and instructed *S* that he would have to respond more frequently if he were to get any points at all. A record was kept of such prompts, and no systematic relationship was found between either the number of prompts required and the schedule used, or the number of prompts required and the subsequent behavior of the individual *S*. Having prompted, *E* would return to the recording room and begin presentation of the stimulus series from the beginning. The following instructions were read to each *S* individually before the experimental session was to begin:

In this experiment, we are investigating the formation of a discrimination. On the panel you see before you, I am going to present a series of light patterns. Some of the light patterns are right and some of them are wrong. There is a rule which makes a pattern right or wrong. When you think one of the patterns of the right group is being presented, I want you to pull the lever before you on the panel. You will have to guess at first, perhaps for several of the patterns, before you will be able to tell for certain that a right pattern is being shown to you. Try to find out the rule for the light patterns which are right. You will know when a right light pattern is being shown to you because you will always get at least one point for pulling the lever when one is before you. Do your best to get as high a score as you can; you never get points for a wrong light pattern. Remember, you will have to guess at first, perhaps for several patterns, before you will begin to get hunches as to what the rule is. Remember, too, that it is only by pulling the lever when a light pattern of the right group is before you that you can raise your score on the counter. Pull the lever when you think a right light pattern is up. You can get more than one point during the time that a right light pattern is before you, so you must pull the lever repeatedly while it is up in order to get all the points you have coming to you. After you have learned the rule, pull the lever only when you think a right light pattern is before you.

The particular discrimination problem selected was one which is identified in this laboratory as a 1-2-5 discrimination. In this problem, one light is always associated with the S^D condition. Five lights may appear only when S^A 's are occasioned. Two of the lights constitute intercept elements and appear equally often with S^D as with S^A conditions. This particular discrimination problem was selected because it had been found previously to be a difficult enough problem that *Ss* did not immediately solve it, and yet not so difficult that discrimination could not proceed within 1 hour. Obviously, all extraneous stimulation in the room which is not systematically associated either with S^D or S^A constitutes intercept, in addition to the two lights so designated by *E*. This specification of the lights in their functional relationship to the conditions of reinforcement and nonreinforcement has the advantage of permitting the possibility of an unambiguous analysis of the stimulus relevant to discrimination.

The stimulus interval of 1 minute was selected on the basis that it was not so short as to prevent the use of the longer schedules, and yet was short enough to allow a sizable number of stimulus presentations during an hour's session.

As in previous experiments with similar methods, the first one or two stimulus presentations were always S^D 's. This procedure was adopted to ensure that rates of responding following the contingent reinforcement would be high enough to sustain a reasonable per-

formance during the hour. Unless one is specifically interested in working with low rates of responding, it would seem pointless to initiate the discrimination processes with S^A . Aside from this limitation, the series were originally established by a procedure of randomization.

RESULTS

The primary analysis of data was based upon two nonparametric analyses of variance. A discrimination index, DI, was computed for each group, in which the ratio of S^D responding over total responding was taken to be a measure of the degree to which two curves diverged. This index varies between 0.0 and 1.0, where a more advanced discrimination is indicated by values approaching 1.0. Reversals in discrimination are indicated by values less than 0.5. The Kruskal-Wallis Test was performed on the combined ratio schedules and on the combined interval schedules to determine the effects of ratio and interval values upon the DI. Ratio and interval schedules were not grouped into a single analysis, since there is no basis for assuming the ordinal groupings of FR 5, FR 10, and FR 20 to be in any way comparable with groupings of FI 5 seconds, FI 15 seconds, and FI 30 seconds. Data from S s receiving continuous reinforcement were included in these analyses, since the continuous schedule logically represents a 1:1 ratio and/or a 0-second interval schedule.

In addition to the over-all ratio and interval analyses, the Kruskal-Wallis H was determined within both fixed and variable contingencies of these schedules. The results of the six analyses are shown in Table 1.

Table 1
Significance of Effects of Ratio and Interval Values
upon Discrimination Index

Schedule	H	df	P
Combined-ratio	22.70	3	<0.005
Variable-ratio	12.72	3	<0.025
Fixed-ratio	8.11	3	<0.01
Combined-interval	17.67	3	<0.005
Variable-interval	12.80	3	<0.005
Fixed-interval	13.68	3	<0.005

In all cases the different values of ratio and interval schedules are seen to have significant effects.

In the accompanying figures, data points are plotted for five S s across each of two stimulus presentations, so that each point represents ten observations. The S^D data are indicated by filled circles, and S^A by open circles. Clearly, ratio contingencies produce higher rates of responding in S^D than do interval contingencies. The S^D performances do not vary significantly between the two types of ratio schedule. As Fig. 1 shows, the over-all mean rates of S^D responding under variable-ratio reinforcement rise as a function of increases in ratio. The data on S^A performance under this condition, while suggestive of a reversed trend, are not so unambiguous.

An interesting result of extending the ratio under fixed-ratio reinforcement appears in Fig. 2. Where the mean S^D curve rises from the FR 5 to the FR 10 condition, it is seen to

Figure 1. Mean response curves under variable-ratio schedules of reinforcement.

Figure 2. Mean response curves under fixed-ratio schedules of reinforcement.

decrease markedly from the FR 10 to the FR 20 condition; moreover, as indicated in Table 2, the FR 20 condition yields an S^D curve of greater variability than that produced by any other ratio condition. Again, S^A performances differ but little as a function of the ratio used here, although the ordinal positions of the four mean curves are in perfect inverse relation to the ratio used when high responding is ranked first and low responding ranked fourth.

As Fig. 3 shows, nothing systematic seems to emerge as a consequence of the variable-interval conditions, except for the narrowing of the distance between S^D and S^A responding as the intervals are increased to 30 seconds.

Table 2
Variances and Results of F Tests of Rates of Responding in S^D
over Last 10 Stimuli under Ratio Reinforcement

Schedule		F	p
FR 10	73.44	13.08	<0.005
FR 15	960.96		
FR 20	1243.02		
VR 10	542.37	11.16	<0.005
VR 20	48.60		

It was anticipated that in FI performance, the initial increase in rate of S^D responding should be followed by an inversion in the curve to a terminal decrease in rate of responding in S^D . Such an effect would be a consequence of the establishment of temporal discrimination. That is, reinforcement would initially result in increased frequency of S^D responding. With the development of temporal discrimination, there should be a subsequent reduction in frequency of responding to the appropriate extinction ratio. As Fig. 4 shows, some suggestion of this appears in the 15-second FI and, particularly, in the 30-second FI conditions. Since other factors are operative in these cases, reversals in S^D curves cannot be taken unambiguously to indicate the existence of temporal discrimination.

Figure 3. Mean response curves under variable-interval schedules of reinforcement.

Figure 4. Mean response curves under fixed-interval schedules of reinforcement.

DISCUSSION

Our previous attempts to bring the operant behavior of human adult subjects under the control of various schedules of reinforcement within short experimental periods have largely met with failure. This has been true particularly with a free operant, where the Ss simply performed on the standard apparatus for periods of time in S^D under various schedules. The general result was that humans under these circumstances tend to give high rates of responding which showed little change as a function of various schedules of reinforcement. Exploratory work suggested that responding might be more sensitive to the effects of schedules of reinforcement under conditions of discrimination than was true in the ordinary free-operant situation.

Certain of the present results confirm the effects of schedules of reinforcement in other organisms, especially on the effect of increasing the ratio under the fixed-ratio schedules. In the original FR groups the intertrial variability in the S^D curves of the FR 5, FR 10, and FR 20 conditions was increased. This increased variability possibly was an artifact of a particular sample of Ss used in the FR 20 condition. Consequently, the FR 15 group was run to interpolate between the FR 10 and FR 20 groups. As mentioned before, variability does increase consistently through the FR 15 schedule. A similar phenomenon does not occur under the VR schedules, and, in fact, as Table 2 shows, variability here is greater in the VR 10 group than in the VR 20.

This increased variability is taken to be evidence of "strain" of the ratio frequently observed under fixed-ratio schedules of reinforcement.

In 30-second interval performances, a distortion appears as the consequence of special experimental conditions. As noted before, the S^D and S^A curves show a very slight separation at these higher intervals. This separation is due to at least two factors. The first, and most obvious, is that as the interval between possibilities for reinforcement becomes longer, the S initially finds it more difficult to distinguish between S^D and S^A conditions. Indeed, the initial overlap between S^D and S^A curves is greater in the 30-second interval groups than in any other groups in the study. Second, from animal studies, the over-all rate of responding under longer intervals should be somewhat less. An additional complication in the present study lies in the limitation of the S^D period to 1 minute. The maximum number of reinforcers that could possibly be presented in the 30-second interval schedule is two per S^D presentation. Since programming of reinforcement usually was not made precisely contingent upon the onset of S^D , a smaller number of reinforcers is actually delivered. The effect is a bounding of the effective schedule. To illustrate an extreme, a 61-second interval schedule of reinforcement could not lead to much, if any, discrimination during a 1-hour period, since almost as many S^D presentations would pass without reinforcement as S^A presentations. While a systematic drop in rate of responding as a function of increased interval should occur even without the bounding by the stimulus interval, the 1-minute stimulus interval does accentuate the effect. This bounding also may be partially responsible for accentuating the ratio strain seen in the FR 20 condition.

In Fig. 4 the rise and subsequent fall in frequency of responding in S^D under fixed-interval reinforcement indicates a discrimination of the temporal interval. The inversion in the curve occurs earlier in the series, where the interval is short, and it is delayed with longer intervals.

In the earlier studies, the characteristic S^D and S^A curves obtained in conditioning of discrimination primarily constitute discriminations of the schedule of reinforcement. Those studies showed that where stimulus series were presented in which no discrimination was possible but where the contingencies of reinforcement were the same as in the normal experimental procedures, S^D and S^A curves differed in no respect from ordinary discrimination curves. The difference appeared in extinction, however, where the control groups could not possibly sustain discriminative performance, and the two curves collapsed into one. But extinction of the regular discrimination resulted in the usual separation between S^D and S^A responding. Data gathered under variable-ratio conditions have verified this finding (Burdick, 1958).

Although a discrimination of the visual-stimulus subsets occurs in all experimental groups, the response-frequency curves reported here primarily represent discriminations of schedules; and probably the conclusion of greatest significance to be drawn from these studies is that the definition of discrimination must be limited operationally to the precise contingencies used. The effects of schedules in these studies constitute a type of response differentiation; and these data emphasize again that there is no "true" set of discrimination curves.

A final comment is in order about the DI used here. In this study, as in others on discrimination learning, both S^D and S^A rates cannot realistically be included in a single analysis. The distributions of response frequencies under these contingencies are found to be skewed in opposite directions when data are taken over longer periods of time, and a trans-

formation which normalizes one of these distributions exaggerates the skewness of the other. Two possibilities remain for including data from both aspects of the discrimination process. One is to take the differences in response frequencies between the two as a measure of the extent to which discrimination has proceeded. The other is the present procedure of deriving the DI. The simple difference has the disadvantage that extraneous factors may produce wide differences in rate of response, thus augmenting error variance. Differences in the absolute rate of responding are washed out by the DI. Nevertheless, as much information on the separation between S^D and S^A responding can be derived from this index as would be possible with the simple difference. In one respect, more information can be gained from the DI, since it is a measure of the percentage of the total behavior appearing in S^D . This information is not present in a simple difference.

SUMMARY

In a series of experiments on the formation of a visual operant discrimination in human subjects, it was found that during an hour's experimental session a discrimination index defined as the ratio of S^D responding over total responding differed significantly as a function of the value of ratios used and as a function of the value of the intervals used. Discrimination, as defined by the separation between S^D and S^A response curves, is more advanced with low ratios and short intervals, and is retarded by longer ratios and longer intervals.

Supporting evidence varifying the effects of schedules of reinforcement upon the behavior of lower organisms was gathered in these studies, but the generality of this evidence is limited by the special bounding conditions imposed by the interval of stimulus presentation. Evidence both of temporal discrimination and of ratio strain were observed in the fixed contingencies, and the general conclusion is drawn that the characteristic effects of schedules of reinforcement upon operant responding represent special cases of response differentiation.

REFERENCES

- Burdick, H. A. Personal communication. October, 1958.
Ferster, C. B., and Skinner, B. F. *Schedules of reinforcement*. Appleton-Century-Crofts, New York, 1957.
Green, E. J. Concept formation: a problem in human operant conditioning. *J. exp. Psychol.*, 1955, **49**, 175-180.
Green, E. J. Stimulus variability and operant discrimination in human subjects. *Amer. J. Psychol.*, 1956, **69**, 269-273.

Received August 3, 1959