JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1969, 12, 551-559 NUMBER 4 (JULY)

PROBABILITY OF REINFORCEMENT AND THE DEVELOPMENT OF STIMULUS CONTROL[1]

CAROL O. ECKERMAN

UNIVERSITY OF NORTH CAROLINA

Pigeons were trained with a successive discrimination procedure in which responding during the negative stimulus was never reinforced and responding during the positive stimulus was reinforced according to one of four probability values. This discrimination training followed extensive training with a single, neutral stimulus and the same temporal distribution of reinforcements. The development of stimulus control was studied by tracing the difference in rate of responding between the positive and negative stimuli over the course of discrimination training. Response rate during the positive stimulus remained constant, while that during the negative stimulus decreased to zero. The probability of reinforcement associated with the positive stimulus affected both the total number of responses emitted during the negative stimulus and the number of negative stimulus presentations during which responding occurred. However, the number of reinforcements during the positive stimulus preceding the attainment of various degrees of stimulus control was similar for all probability values.

Stimulus control refers to a relationship between an antecedent stimulus and an organism's responding. If a change in some aspect of the antecedent stimulus results in a change in some measure of responding, then stimulus control is said to exist with respect to that aspect of the stimulus varied (*e.g.*, Terrace, 1966; Jenkins, 1965). The present experiment studied the development of the stimulus control of conditioned responding for various probabilities of reinforcement associated with an antecedent stimulus.

A successive operant discrimination procedure was developed, in which pigeons were trained to peck a response key. Two stimuli, red or green illumination of this key, were presented according to a random sequence, each for the same fixed time period. The probability of reinforcement associated with a stimulus was defined as the proportion of presentations of that stimulus that were terminated with the presentation of reinforcement for a key peck. The probability value associated with the negative stimulus was zero, while that associated with the positive stimulus was some non-zero value. Stimulus control was said to exist if there was a reliable difference between the rates of responding during the two stimuli. Since these rates were recorded repetitively throughout training, the manner in which stimulus control developed was studied; this development was compared for groups of subjects differing in the probability value associated with the positive stimulus.

Before discrimination training, all subjects were trained in the presence of a single neutral stimulus, white illumination of the key, with the same reinforcement conditions that would occur during discrimination training. During discrimination training, the positive and negative stimuli replaced white illumination of the key, while the temporal distribution of reinforcements remained constant for each subject. Thus, the development of stimulus control was traced in a situation where the only change was the manner of correlation of stimuli and reinforcement probabilities. This technique of inserting a stimulus into a constant schedule of reinforcement was used previously by Skinner (1938) to study the development of a discrimination and more recently by Farmer and Schoenfeld (1966).

Since the different reinforcement probabilities used might control such aspects of behavior as the time spent around the response keys, amount of pacing around the chamber, or

[1]These data are based on a dissertation submitted to the Department of Psychology, Columbia University. I wish to thank Drs. H. S. Terrace and W. W. Cumming for their advice and encouragement during the course of this research and Dr. M. B. Waller for the use of his experimental facilities. Reprints may be obtained from the author, Department of Psychology, University of North Carolina, Chapel Hill, North Carolina.

brooding, the presentation of the discrimination procedure stimuli was made contingent upon responding on another key. This requirement assured that all subjects were oriented toward the response keys and engaged in key pecking when the stimui were presented.

The effect of reinforcement conditions upon stimulus control has been analyzed in prior studies. Several of these studies varied the schedule of reinforcement during original training with a single stimulus and employed generalization procedures to determine the extent of stimulus control. If degree of stimulus control is equated with the steepness of the generalization gradient obtained, then Hearst, Koresko, and Poppen (1964) determined that stimulus control was greatest for subjects trained with variable-interval (VI) 30-sec or VI 1-min schedules and progressively less for subjects trained with longer mean interval lengths. Similar differences between generalization gradients have been obtained with VI 15-sec or VI 1-min and VI 4-min schedules (Haber and Kalish, 1963), variable-interval and variable-ratio schedules (Thomas and Switalski, 1966), variable-interval and differential-reinforcement-of-low-rate schedules (Hearst *et al.*, 1964) and with positive-reinforcement and avoidance schedules (Hearst, 1962).

Other investigators have varied the reinforcement conditions associated with the positive stimulus of a discrimination procedure, but only after discriminated responding was established. Nevin (1967) determined that the probability of reinforcement associated with the positive stimulus of a simultaneous discrimination procedure affected the probability of a response occurring on a trial, but not the probability that the response that occurred would occur to the positive stimulus. Cumming (1955) also found that the mean interval length of a variable-interval schedule did not affect the relative rates of responding during the positive and negative stimuli of a successive discrimination procedure. Thus, if degree of stimulus control is assessed by the relative rates of responding during the positive and negative stimuli, these reinforcement manipulations did not alter the existing level of stimulus control.

The present study extended these prior findings regarding reinforcement manipulations by determining the effect of reinforcement probability upon both the development and the final level of stimulus control.

METHOD

Subjects

Twelve male White Carneaux pigeons, seven months old at the start of the experiment, were housed in individual cages within a general vivarium and maintained at 75% of their free-feeding weights. Purina pigeon chow was fed immediately after the daily experimental session when needed. Water and grit were continuously available in the home cages. No subject had a prior experimental history.

Apparatus

A modified Grason-Stadler E3125A-300 pigeon chamber was used. The stimulus panel contained a grain feeder opening centered on the panel 2 in. above the grid floor and two 0.75-in. diameter Gerbrands pigeon keys centered at a height of 8 in. and separated by 4 in. (center to center).

The keys could be back-illuminated by red, green, or white jeweled pilot lights. The average luminance values recorded at the front surface of the left key with a Macbeth Illuminometer were 6, 4, and 14 foot lamberts for red, green, and white illumination of the key. A mask was fitted behind the right key (later called Key A) which allowed only three vertical bars of illumination to be projected onto the key. The bars were 0.13-in. wide and 0.13 in. apart. White illumination of the right key produced alternating bars with an average luminance at the key of 25 foot lamberts.

The force required to activate the keys was 23 g for the left key and 27 g for the right. White noise was provided in the chamber through a speaker mounted on the upper left of the stimulus panel. Scheduling and recording equipment were located in a separate room. Standard relay equipment was used.

Procedure

Basic schedule. The basic schedule of reinforcement was a two-component chain (See Fig. 1). Pecking of Key A when it was illuminated by Stimulus 1 (S_1) was followed, according to a VI 6.25-sec schedule, by the illumination of Key B with either Stimulus 2 (S_2) or Stimulus 3 (S_3).

There was a probability value x that after 12 sec of S_2 the next response on Key B would be followed by a 3-sec presentation of the food hopper filled with mixed grain. Likewise, a probability value y was associated with reinforcement for pecking on Key B after 12 sec of S_3. If reinforcement was available on a given stimulus presentation, S_2 or S_3 remained on until the delivery of the reinforcer. The delivery of the reinforcer was followed by a 12-sec intertrial interval (ITI) during which neither key was illuminated, the houselight remained on, and key pecks had no scheduled effect. If reinforcement was not available, the stimulus condition, S_2 or S_3, was removed after 12 sec, and a 15-sec ITI, rather than a 12-sec ITI, followed in order to maintain constant temporal relations between successive stimulus presentations. At the end of the ITI, S_1 was again presented on Key A.

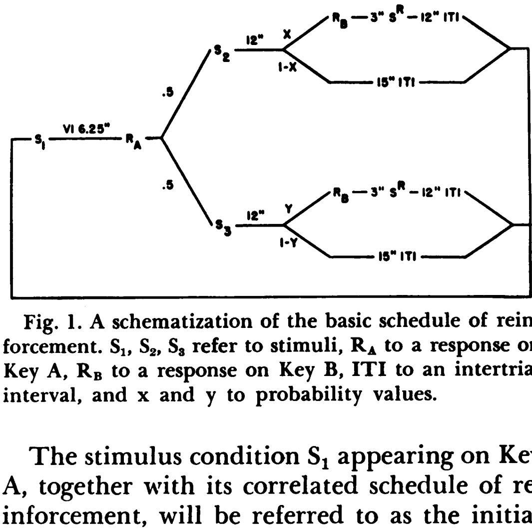

Fig. 1. A schematization of the basic schedule of reinforcement. S_1, S_2, S_3 refer to stimuli, R_A to a response on Key A, R_B to a response on Key B, ITI to an intertrial interval, and x and y to probability values.

The stimulus condition S_1 appearing on Key A, together with its correlated schedule of reinforcement, will be referred to as the initial component of the chain schedule; the presentation of S_2 or S_3 on Key B, together with the associated reinforcement conditions, will be referred to as the final component. A session consisted of 100 presentations of each component.

Throughout the experiment, S_1 was a pattern of alternating dark and light vertical bars obtained by back-illuminating Key A with white light. The VI 6.25-sec schedule associated with responding on Key A consisted of 2.5-, 5.0-, 7.5-, and 10.0-sec intervals, arranged in a random sequence of 40 intervals, with the restriction that each interval occur equally often within the two blocks of 20.

S_2 and S_3 were presented in random order with the restriction that S_2 and S_3 occur equally often in each block of 20 presentations. Three 100-presentation sequences of S_2 and S_3 presentations were constructed from a random number table, and these three sequences were used in random permutations of three.

Reinforcement probabilities (x or y) of 1.00, 0.60, 0.30, and 0.16 were used. For 0.60 and 0.30, the stimulus presentations ending with reinforcement were determined at random for each of the three S_2-S_3 sequences, with the restriction that six (for 0.60) or three (for 0.30) reinforcements were scheduled within each block of 10 presentations of a stimulus. For 0.16, the presentations ending with reinforcement were determined at random for each session, with the restriction that four reinforcements were scheduled within each block of 25 presentations.

Experimental design. The variables manipulated were the stimuli appearing on Key B, S_2 and S_3, and the probabilities, x and y, that S_2 and S_3 ended with the presentation of reinforcement. The probability value x is the major variable, and the sequence of values of x distinguishes four groups of subjects. Table 1 presents the sequence of values of S_2 and S_3 and of x and y for each of the groups.

An experimental stage is a series of sessions during which the values of x, y, S_2, and S_3 were constant. All stages of training were completed first for Groups 1 and 3 and later for Groups 2 and 4. All subjects received one session of preliminary training during which they were trained to peck Key A illuminated with the vertical bars of light and Key B illuminated with red, green, or white light. Each stimulus condition was presented three times according to a non-systematic order, and a total of 20 regular reinforcements was delivered during each condition.

During Stage 1, all subjects received the same training on the two-component chain schedule with a probability of reinforcement of 1.00 and both stimulus values, S_2 and S_3, as white illumination. Subjects were ranked according to their rates of responding during the final component of the chain schedule, and the ranks were non-systematically assigned to the various groups. Responding during Stage 1 also provided a baseline for evaluating the effects of the different probabilities of reinforcement presented during Stage 2.

Table 1

Sequence of Reinforcement Probabilities, x and y, and Associated Stimulus Conditions, S_2 and S_3

Stage	Number of Sessions	S_2	S_3	Group 1 x	Group 1 y	Group 2 x	Group 2 y	Group 3 x	Group 3 y	Group 4 x	Group 4 y
1	30	W	W	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	18	W	W	1.00	0.00	0.60	0.00	0.30	0.00	0.16	0.00
3	18	R (G)	G (R)	1.00	0.00	0.60	0.00	0.30	0.00	0.16	0.00
4	18	W	W	1.00	1.00	0.60	0.60	0.30	0.30	0.16	0.16
5	18	W	W	—	—	1.00	1.00	1.00	1.00	1.00	1.00

W = white.
R = red.
G = green.

From Stage 1 to Stage 2, the overall probability of reinforcement associated with the final component of the chain schedule (*i.e.*, the average of x and y) was reduced from 1.00 to either 0.50, 0.30, 0.15, or 0.08 for Groups 1 to 4, while S_2 and S_3 remained the same white illumination. The temporal distribution of reinforcements during a session was the same as during Stage 3.

Stage 3 was a successive discrimination procedure. One stimulus value, S_2 or the positive stimulus, was associated with a reinforcement probability of either 1.00, 0.60, 0.30, or 0.16, while the second stimulus value, S_3 or the negative stimulus, was associated with a zero probability value. For two subjects of each group, green illumination of Key B was the positive stimulus and red illumination the negative stimulus. For the remaining subject of each group, the stimulus values were reversed.

During Stage 4, S_2 and S_3 were again a single stimulus value, white illumination, and the probability of reinforcement associated with the white illumination was the same as that associated with the positive stimulus of Stage 3.

For Groups 2, 3, and 4, a fifth stage of training was presented, consisting of the same conditions as Stage 1. Stage 4 conditions were identical to Stage 1 conditions for Group 1.

Data recording. For each session, the total number of responses on Key A during S_1 and the total time in the presence of S_1 were recorded. Also, the number of responses on Key B occurring during the first 12 sec of a stimulus presentation was recorded for each S_2 or S_3 presentation during a session.

RESULTS

Chain Schedule Performance before Discrimination Training

Chain schedule performance was quickly established. Both responding on an inappropriate key during a stimulus presentation and responding during the intertrial interval were eliminated within the first three sessions of Stage 1.

Overall rates of responding were computed for each component of the chain schedule. For the overall rate during the final component, the total number of responses on Key B during the first 12 sec of a stimulus presentation was divided by the total time period over which the responses were summed. The overall rate during the initial component was the total number of responses on Key A during S_1 divided by the total duration of S_1.

The data of one of the subjects assigned to Group 4 are not reported because this subject was discarded from the experiment at the end of Stage 2 when sessions were not completed within 15 hr.

The overall rates of responding varied considerably among subjects. The rate values averaged over the last six sessions of Stage 1 ranged from 0.37 to 2.80 responses per second during the final component and from 0.04 to 1.29 responses per second during the initial component. Similarly, the average rates during the last six sessions of Stage 2 ranged from 0.48 to 4.52 responses per second for the final component and from 0.02 to 0.72 responses per second during the initial component.

The reduction in the probability of reinforcement from Stage 1 to Stage 2 was accompanied by a general increase in the rate of responding during the final component and a decrease in the rate during the initial component. Using the data of the last six sessions of Stages 1 and 2, the average percentage change in rate of responding was computed for each subject by subtracting the Stage 1 rate from that for Stage 2 and expressing this difference as a percentage of the Stage 1 rate. The average percentage change during the final component was +26, +21, +74, and +59% and during the initial component −64, −28, −32, and −94% for Groups 1 to 4. Of the 11 subjects, seven showed increases of greater than 10% in rate of responding during the final component, two showed no change, and two showed decreases; 10 showed decreases in rate of responding during the initial component and one showed no change. The four groups of subjects did not differ reliably in the number of subjects showing the general changes, and there was considerable overlap among the percentage change values for the subjects of the various groups. At the end of Stage 2, the four groups of subjects did not differ reliably in the overall rates of responding during either the initial or the final component of the chain schedule.

Development of Stimulus Control

Stimulus control was assessed by comparing the overall rates of responding during S_2, the positive stimulus, and S_3, the negative stimulus, during the discrimination training procedure of Stage 3.

Over the course of the discrimination training, the rate of responding during the positive stimulus remained approximately constant, while the rate of responding during the negative stimulus changed to zero from approximate equality with that for the positive stimulus. The manner in which stimulus control developed was assessed by tracing the difference between the rates of responding during the positive and negative stimuli as a function of the amount of training. Two measures of amount of training were analyzed: the number of presentations of a stimulus value, or the amount of exposure to a stimulus value, and the number of prior reinforcements during the positive stimulus.

The number of stimulus presentations. In Fig. 2, the total number of responses on Key B during the first 12 sec of each stimulus presentation is plotted cumulatively as a function of the number of presentations of that stimulus value over the first two sessions of Stage 3. A third session is presented for P61 since the major change in the curvature of the cumulative response curve for S_3 occurred during the third session for this subject. For the remaining subjects, the form of the S_3 cumulative curve is adequately described by the data of the first two sessions. The cumulative response curves for S_2 are essentially linear, although there is an initial acceleration for seven subjects. The cumulative response curves for S_3 are negatively accelerated, approaching a slope of zero. With the exception of P58 and P61, the number of negative stimulus presentations during which responding occurred increased as the probability value associated with the positive stimulus decreased.

The relationship between the total number of responses during S_3 and the probability

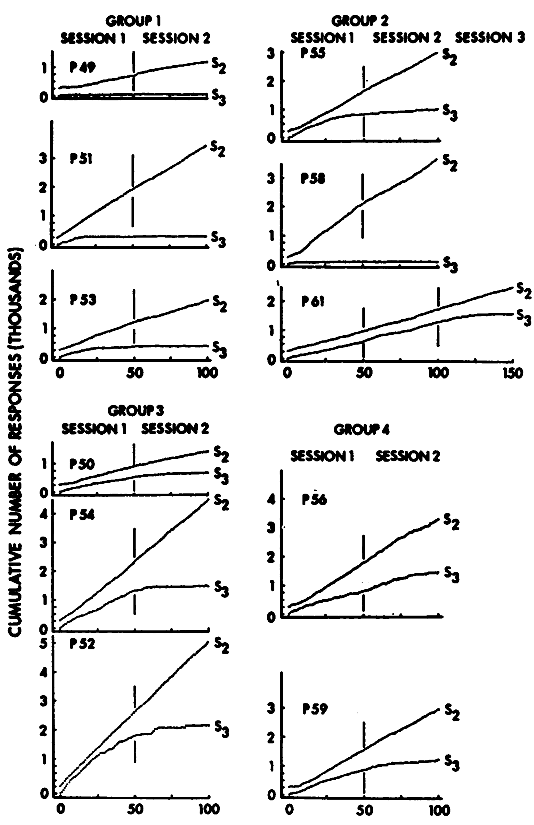

Fig. 2. The cumulative number of responses during the positive (S_2) and negative (S_3) stimuli as a function of exposure time to the stimuli. The curve for S_2 is displaced upwards along the response axis by 250 responses.

value associated with S_2 is summarized in Fig. 3. Since individual subjects differed greatly in their overall rate of responding during all stages, a relative measure has been employed. The total number of responses emitted during the S_3 presentations of the 18 sessions of Stage 3 was divided by the total number of responses emitted during the S_2 presentations of these sessions. This ratio is plotted as a function of the probability value associated with S_2. For probability values 1.00, 0.30, and 0.16 (*i.e.*, Groups 1, 3, and 4) this ratio increased as the probability value decreased, and there was no overlap among the groups. The ratios for subjects of Group 2, however, varied widely and overlapped with those of the other groups. The very high ratio and the very low ratio for Group 2 represent P61 and P58.

The number of preceding reinforcements. The difference in rate of responding between the positive and negative stimuli was also determined for various numbers of preceding reinforcements. The mean number of responses during the stimulus presentations between the delivery of the n^{th} and the $(n + 1)^{th}$

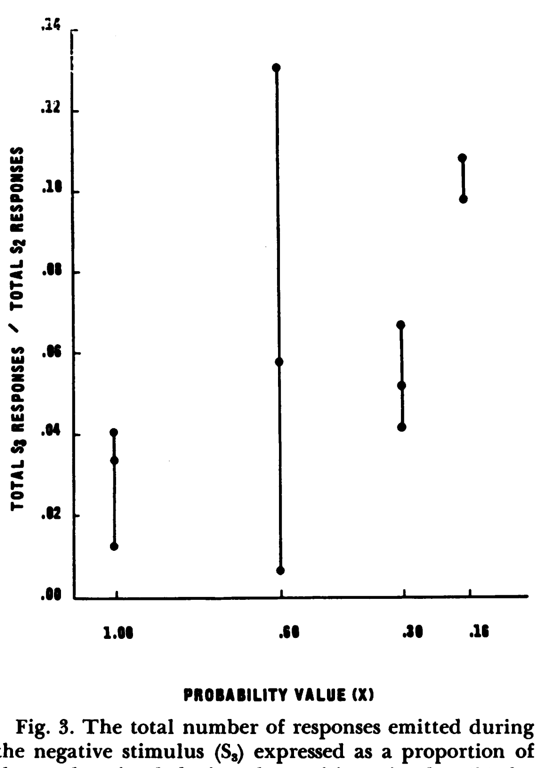

Fig. 3. The total number of responses emitted during the negative stimulus (S_3) expressed as a proportion of the total emitted during the positive stimulus (S_2) for the various probabilities of reinforcement associated with S_2.

reinforcement was calculated separately for S_2 and S_3.

With respect to this measure of training time, there were no consistent differences among the groups. There was no consistent change for any group in the mean number of responses during the positive stimulus. Figure 4 presents the mean number of responses during negative stimulus presentations as a function of the number of preceding reinforcements. For all groups, the mean number of responses during the negative stimulus decreased to a near-zero value within 30 reinforcements. P61 is the only exception to this relationship. To characterize the decrease in the number of S_3 responses, a straight line was fitted by the method of least squares to the data points preceding either 31 reinforcements or the criterion of three successive zero values for S_3 responding, whichever occurred first. There is considerable variability of the data points about the straight lines, but the deviations do not appear to be systematic.

Characteristics of these fitted lines may be summarized for the four groups of subjects. The value at which the fitted line intersects the abscissa, *i.e.*, the number of prior reinforcements yielding a zero response rate during S_3, was very similar for probability values 1.00, 0.30, and 0.16, and there was considerable overlap among the values for the various groups. The average number of preceding reinforcements was 24.3, 25.9, and 24.3 for probability values 1.00, 0.30, and 0.16. Again the intercept values for the 0.60 probability value varied greatly with the two discrepant values representing P61 and P58. There was no systematic variation in the slope values for the various probability values, although the slopes varied from −0.38 to −2.29 for the subjects of Groups 1, 3, and 4.

Determinants of Responding during the Positive Stimulus

The determinants of the overall rate of responding during the positive stimulus may be studied by within-subject comparisons of S_2 responding across Stages 2, 3, and 4. The experiment was designed to allow a comparison between the overall rate of responding during S_2 in Stage 3, where S_2, the positive stimulus, was associated with a probability value x and presented in the context of an equal number of presentations of a negative stimulus, with:

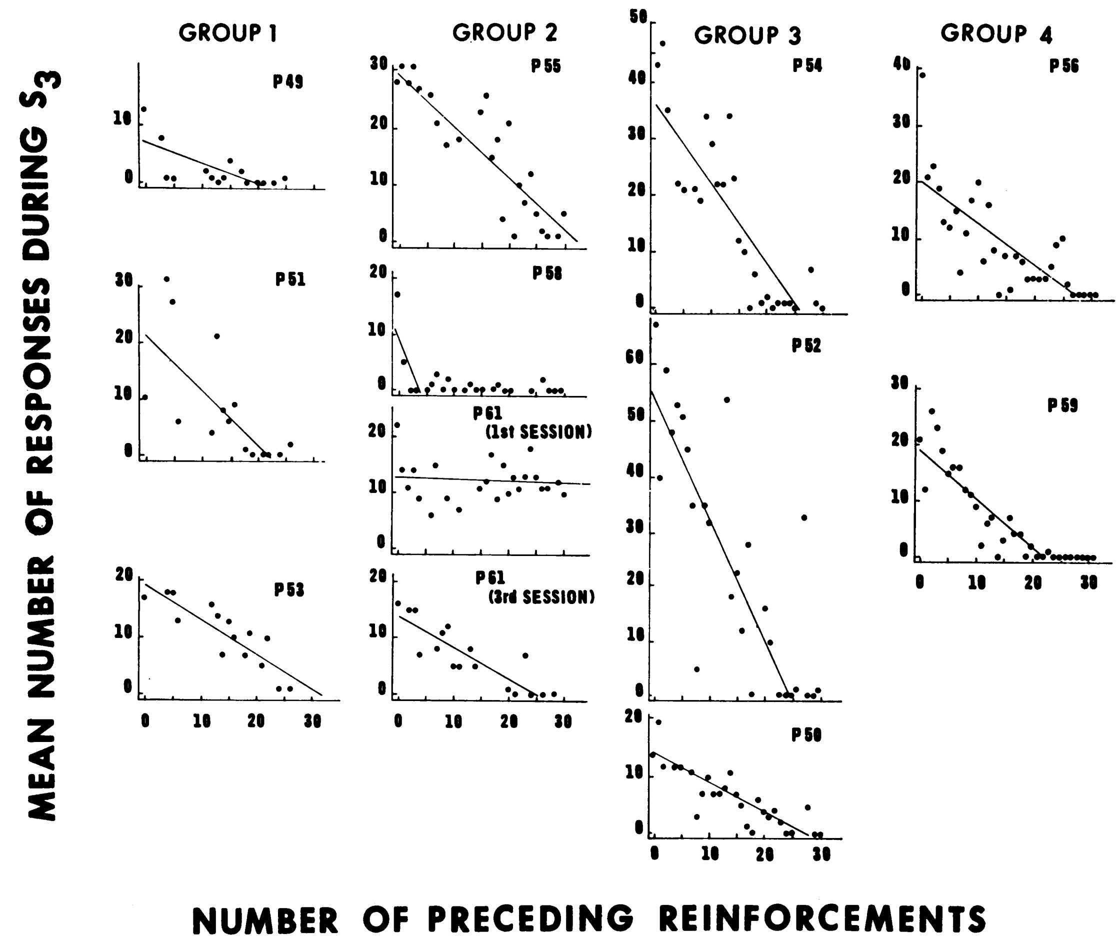

Fig. 4. The average number of responses per negative stimulus presentation between the n^{th} and the $(n + 1)^{th}$ reinforcement as a function of n, the number of preceding reinforcements. Straight lines have been fitted by the method of least squares.

(1) the rate of responding during a single stimulus condition associated with the same overall probability of reinforcement, x/2, (Stage 2), and (2) the rate during a single stimulus condition associated with the same probability value x as the positive stimulus itself (Stage 4).

The rate of responding during the positive stimulus of Stage 3 did not bear a systematic relation to either the rate during Stage 2 or that during Stage 4. For six subjects, the rate during S_2 for Stage 3 was greater than that for Stage 4, while for five subjects the opposite relation held. Likewise, for six subjects the S_2 rate during Stage 3 was greater than that for Stage 2, while for five subjects, the Stage 3 rate was less than that for Stage 2.

The data indicate certain limitations upon these comparisons across stages of training. A comparison of the response measures during Stage 5 for Groups 2, 3, and 4 and during Stage 4 for Group 1 with those during Stage 1 demonstrated that the response measures generally were not recoverable for a given set of reinforcement conditions. The overall rate of responding during S_2 on the second determination was greater than that for the initial determination for four subjects, less than that for the initital determination for six subjects, and the same for one subject.

DISCUSSION

Clear relationships between antecedent stimuli and behavior were established for all subjects. Repetitive measurement of this relationship yielded orderly changes in the degree of stimulus control with both increasing numbers of stimulus presentations and increasing numbers of reinforcements for responding during the positive stimulus.

In the present experimental design, the antecedent stimuli to be studied were correlated with different distributions of reinforcement availability after the response rate in question had stabilized. Further, these correlations between antecedent stimuli and reinforcement availability were effected without altering the overall distribution of reinforcement availability for responding in the experimental situation. Under these conditions, the rate of responding during the positive stimulus did not differ systematically from the rate of responding already established during prior training, nor did it change systematically over the course of discrimination training. The rate of responding during the negative stimulus, however, decreased to zero, and the probability of reinforcement associated with the positive stimulus affected the manner in which this decrease occurred.

The probability of reinforcement during the positive stimulus determined both the total number of responses occurring during the negative stimulus and the rapidity with which responding during the negative stimulus approached a zero rate. The lower the probability of reinforcement, the greater the number of responses emitted and the longer the period of responding during the negative stimulus. The functions relating responding during the negative stimulus to the number of reinforcements accumulated during the positive stimulus were essentially linear, reaching a zero response rate at approximately 25 reinforcements regardless of the number of interspersed positive stimulus presentations not ending with reinforcement.

The exact number of reinforcements preceding cessation of responding during the negative stimulus probably reflects the various parameters of the discrimination procedure: the stimuli used, the duration of a stimulus presentation, the sequencing of stimulus presentations, prior training, *etc.* For the present study, the important finding is not the exact number of reinforcements, but rather the constancy of this number. This constancy suggests that the interspersion of positive stimulus presentations not ending in reinforcement does not detract from the effectiveness of those presentations that do. However, in the present design, both the number of negative stimulus presentations and the number of positive stimulus presentations without reinforcement varied together. Research varying both the probability of reinforcement during the positive stimulus and the amount of exposure to the negative stimulus would help to determine the effects of the probability of reinforcement itself.

The present conclusions regarding the effects of probability of reinforcement are based on the behavior of nine of the 11 subjects. The close similarity for these nine subjects of the functions relating responding during the negative stimulus to number of prior reinforcements warrants such conclusions. Although the determinants of the remaining subjects' behavior are unknown, possible determinants may be suggested.

The discrimination performance of P58 may have resulted from an interaction between the effects of the reinforcement schedule and an "aversion" for green. For this subject, the rate of responding during the negative stimulus, green, was substantially less than that for the positive stimulus before any differential reinforcement occurred. The discrimination performance of the other subject, P61, however, resembles that described by a two-stage model of discrimination learning (see MacKintosh, 1965, for a review). For this subject, there was a sizable period of training, during which there was no evidence of a discrimination, followed by the rapid development of discriminative responding. Further, this development, once started, closely resembled that for the other subjects (see Fig. 2 and 4).

All subjects, however, showed similar final levels of stimulus control. Whether key pecking was reinforced every time a stimulus was presented or on the average of every sixth time, subjects essentially never responded during the negative stimulus and they responded during the positive stimulus at the same rate as they did before discrimination training (Stage 2). The probability of reinforcement affected the transitional behavior, not the final level of performance.

REFERENCES

Cumming, W. W. *Stimulus disparity and variable-interval reinforcement schedule as related to a behavioral measure of similarity.* Unpublished doctoral dissertation, Columbia University, 1955.

Farmer, J. and Schoenfeld, W. N. Varying temporal placement of an added stimulus in a fixed-interval schedule. *Journal of the Experimental Analysis of Behavior*, 1966, **9**, 369-375.

Haber, A. and Kalish, H. I. Prediction of discrimination from generalization after variations in schedules of reinforcement. *Science*, 1963, **142**, 412-413.

Hearst, E. Concurrent generalization gradients for food-controlled and shock-controlled behavior. *Journal of the Experimental Analysis of Behavior*, 1962, **5**, 19-31.

Hearst, E., Koresko, Minnie B., and Poppen, R. Stimulus generalization and the response-reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 1964, **7**, 369-380.

Jenkins, H. M. Measurement of stimulus control during discriminative operant conditioning. *Psychological Bulletin*, 1965, **64**, 365-376.

MacKintosh, N. J. Selective attention in animal discrimination learning. *Psychological Bulletin*, 1965, **64**, 124-150.

Nevin, J. A. Effects of reinforcement scheduling on simultaneous discrimination performance. *Journal of the Experimental Analysis of Behavior*, 1967, **10**, 251-260.

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

Terrace, H. S. Stimulus control. In W. K. Honig (Ed.), *Operant behavior: areas of research and application*. New York: Appleton-Century-Crofts, 1966. Pp. 271-344.

Thomas, D. R. and Switalski, R. W. Comparison of stimulus generalization following variable-ratio and variable-interval training. *Journal of Experimental Psychology*, 1966, **71**, 236-240.

Received 9 May 1968.