Reward sequence and reinforcement level as determinants of S— behavior in differential conditioning

STEVEN J. HAGGBLOOM

Arkansas State University, State University, Arkansas 72467

In two differential conditioning experiments, groups of 10 rats each differed with respect to average reward and schedule of reward received in S+. Nonreward (N) occurred on all S— trials. In both experiments, extinction of responding to S— (resistance to discrimination) was extensively regulated by reward sequence and was largely independent of average reward. In Experiment 1, resistance to discrimination was a function of transitions from N to rewarded (R) trials (N-R transitions). In Experiment 2, resistance to discrimination was increased by large reward on the R trial of N-R transitions and decreased by large reward on the R trial of R-N transitions. These schedule effects on resistance to discrimination parallel the effects of comparable schedules on resistance to extinction following partial reinforcement. The results are discussed in terms of sequential theory, reinforcement level theory, and their implications for various schedule manipulations that have previously shown S— behavior to be inversely related to average reward in S+.

Two factors have been shown to be particularly important in regulating behavior to the negative (S—) stimulus during differential conditioning, trial sequence (Capaldi, Berg, & Morris, 1975; Haggbloom, 1978, 1979, 1980; McHose & Blackwell, 1975) and the difference between the average incentive associated with the two discriminanda (McHose, 1970).

The incentive averaging view is illustrated by an experiment reported by McHose, Maxwell, and McHewitt (1971, Experiment 4). In that experiment, all groups received nonreward (N) in S—; one group received eight and another two 45-mg food pellets on all S+ trials; and a third group, crucial to the incentive averaging view, received a 50% partial reinforcement (PRF) schedule in S+, obtaining 16 pellets on rewarded (R) trials. Where the numerals indicate S+ and S— reward, these groups can be identified, respectively, as Groups 8-0, 2-0, and 16/0-0.

Using formulas suggested by McHose (1970) for computing average incentive, McHose et al. (1971) showed that the average incentive disparity between S+ and S— was equal in Groups 16/0-0 and 2-0 and less than that in Group 8-0. Consistent with the incentive averaging view, Groups 16/0-0 and 2-0 performed alike in S— and ran faster than Group 8-0.

These experiments were supported in part by a grant from Arkansas State University and by National Institute of Mental Health Grant 1 ROC MH33349-01 to the author. The author wishes to thank Professor John K. Beadles for the use of laboratory space in the Department of Biology. Requests for reprints should be addressed to Steven J. Haggbloom, Department of Psychology, Arkansas State University, P.O. Box 2127, State University, Arkansas 72467.

Recently, differences in S— behavior have been obtained among groups given the same average incentive but different trial sequences (Haggbloom, 1978, 1979, 1980; McHose & Blackwell, 1975). That fact is not in itself particularly damaging to the incentive averaging view; when there are no incentive differences among groups, sequential variables may assume greater importance in determining S— behavior. What is problematic for the incentive averaging view is that the same sequential effects provide an alternative explanation for the results of the McHose et al. (1971) investigation and similar experiments supporting the average incentive view (e.g., McHose, 1970; McHose, McHewitt, & Peters, 1972).

Recent experiments indicate that sequential variables known to regulate resistance to extinction following PRF have a highly similar effect on extinction of behavior in S— during differential conditioning (Haggbloom, 1978, 1979, 1980; McHose & Blackwell, 1975). Among the more powerful sequential effects in differential conditioning, as in PRF (Capaldi, 1966), is that of transitions from N to R trials (N-R transitions). Differential conditioning groups given a PRF schedule in S+ that contains N-R transitions run faster in S— and are more resistant to discrimination than groups given a PRF schedule that has no N-R transitions (Haggbloom, 1979, 1980; McHose & Blackwell, 1975).

The question raised by these findings is whether Group 16/0-0 in the McHose et al. (1971) experiment, and comparable groups in other tests of the incentive averaging view, ran faster in S— than Group 8-0 because of a smaller average incentive associated with S+ or because of N-R transitions experienced within S+. The purpose of the two experiments reported here was

to investigate the joint effects on S- behavior of average incentive and trial sequence. Experiment 1 employed three groups similar to those employed by McHose et al. (1971) and a fourth group given PRF in S+ on a schedule that did not contain N-R transitions. In Experiment 2, each of four groups received PRF in S+. Groups differed with regard to average incentive and/or the sequential variable of reward magnitude contained in transitions from rewarded to nonrewarded trials and nonrewarded to rewarded trials.

METHOD

Subjects

In both experiments, the subjects were 40 90-day-old male rats obtained from the Holtzman Company, Madison, Wisconsin.

Apparatus

The apparatus consisted of two parallel straight alleys, 105 cm long $\times$ 9 cm high and wide. The walls and floor of one alley were painted white, and the walls and floor of the other alley were painted black. The last 25 cm of each alley constituted a goalbox separated from the rest of the alley by manually operated guillotine doors. The doors and goalboxes were painted the same color as the alley in which they were located. Each goalbox contained an unpainted wooden goal cup. A gray startbox, 9 cm high and wide $\times$ 25 cm long, could be aligned to permit entry to one alley or the other. The startbox had a gray, manually operated guillotine door.

The alleys were divided into three sections over which times were recorded (respectively, start time, run time, and goal time). Start times were recorded from the opening of the startbox door, which triggered a 1/100-sec clock, to a point 32 cm into the alley. Run and goal times were recorded over the next 40 and 30 cm, respectively. The offset of the first clock and the operation of the run and goal clocks were controlled by photoelectric circuitry.

Procedure

In both experiments, the rats were placed on a 12-g-per-day food-deprivation schedule on Day 1. On Days 12-14, they were handled in squads of four for 4 min per squad and were fed 45-mg Noyes pellets in their home cages in an amount equal to that subsequently received on R trials during training. The amount of food received in reinforcement pellets was always subtracted from each rat's daily ration.

During discrimination training in both experiments, all groups received four S+ trials on each odd-numbered day and four S- trials on each even-numbered day. These procedures avoid the confounding of variation of reward sequence in S+ with variation in other sequential variables (Haggblom, 1980). S- trials were always nonreinforced with rats confined to the unbaited goalbox for 20 sec.

Experiment 1 was similar in design to the McHose et al. (1971, Experiment 1) experiment described above. Groups 8 and 2 received, respectively, eight and two Noyes pellets weighing 45 mg each on all S+ trials. Groups NR and RN both received 50% PRF in S+ and 16 pellets on both R trials. Like the Group 16/0-0 employed by McHose et al., Group NR received a PRF schedule containing N-R transitions within S+. Four sequences were used recursively over successive blocks of 4 S+ days: RNNR, NRNR, RNRN, NNRR (NR schedules). Group RN, on the other hand, received the sequence RRNN (RN schedule), which did not contain N-R transitions, on each S+ day. There were 20 days of discrimination training in Experiment 1.

In Experiment 2, groups differed with respect to the sequence of large reward (L, 16 pellets), small reward (S, 2 pellets), or

nonreward trials experienced in S+. The four groups and their respective PRF sequences were as follows: Group SNS (SSNS or SNSS), Group SNL (SSNL or SNLL), Group LNS (LLNS or LNSS), and Group LNL (LLNL or LNL). Each group received the first sequence on S+ Days 1, 5, 9, and 13 and the second sequence on S+ Days 3, 7, 11, and 15. There were 16 days of training in Experiment 2.

In both experiments, a trial was begun by placing a rat in the startbox and opening the startbox door after approximately 3 sec regardless of the rat's orientation. The rats were run in squads of four containing one rat from each group. The order in which successive squads were run was held constant across days, while the order of administering trials within a squad varied randomly from day to day. The intertrial interval was 3-4 min. A maximum time of 30 sec was allowed in each alley section. If 30 sec was exceeded in any alley section, the additional time was added to the latency score of the next section forward. If the rat did not enter the goalbox within 90 sec, it was placed in the goalbox.

RESULTS

All time scores were converted to speeds in centimeters per second. For both experiments, only total speeds, which were representative of responding in individual alley sections, are reported. Figure 1 shows the mean speed in S- in blocks of two S- days in Experiment 1 (left panel) and Experiment 2 (right panel).

As can be seen in Figure 1, Groups RN, 8, and 2 performed alike in S- and ran more slowly than Group NR in Experiment 1. In Experiment 2, Group SNL ran fastest in S-, Group LNS was the slowest, and Groups SNS and LNL performed at an intermediate level. These observations were evaluated by analyses of variance (ANOVA) applied to speeds on the last block of two S+ and two S- days in both experiments. In Experiment 1, the ANOVA had groups (4) and brightness (2) as between-subjects variables and discriminanda (2), days (2), and trials (4) as within-subject

Figure 1. Mean speed in S- over the total alley in blocks of 2 days for each of the four groups in Experiment 1 (left panel) and Experiment 2 (right panel).

variables. The ANOVA of Experiment 2 was identical except for the partitioning of the groups variable into two factors, two levels of reward magnitude contained in the R-N transition (R-N magnitude) and two levels of reward magnitude contained in the N-R transition (N-R magnitude).

In Experiment 1, there was a reliable Groups by Discriminanda interaction [$F(3,32) = 14.61$, $p < .001$], which was partitioned into simple effects of groups in S+ [$F(3,54) = 1.70$] and groups in S- [$F(3,54) = 6.39$, $p < .01$]. The interaction was also partitioned into simple effects of discriminanda for each group. Those tests indicated that all groups ran reliably faster in S+ than in S- [smallest $F(1,32) = 29.15$, $p < .001$, for Group NR]. The reliable groups effect in S- was evaluated using the Newman-Keuls procedure, which showed that only the differences between Group NR and each of the other three groups were reliable (all $ps < .05$).

In Experiment 2, there was a reliable interaction of N-R Magnitude by Discriminanda [$F(1,32) = 5.86$, $p < .05$], but no R-N Magnitude by Discriminanda interaction [$F(1,32) = 3.30$] or N-R Magnitude by R-N Magnitude by Discriminanda interaction [$F(1,32) = 1.39$]. Comparisons between S+ and S- speeds within each group showed that each group ran reliably more slowly in S- than in S+ [smallest $F(1,32) = 5.59$, $p < .05$, for Group SNL]. Comparisons among groups within S+ and within S-, using the Newman-Keuls procedure, revealed no group differences in S+, while, in S-, Group SNL ran reliably faster than each of the other three groups ($ps < .05$) and Groups SNS and LNL both ran reliably faster than Group LNS ($ps < .05$).

Under conditions in which S+ and S- trials occur on separate days, as in both of the present experiments, behavior is regulated at least as extensively by reward-produced cues as by brightness cues (Haggbloom & Tillman, *in press*), although reward-produced cues may not exercise differential control over behavior on the first trial of the day (Haggbloom, 1979). Thus, behavior on the first trial of each day in the present experiments may have been more under the control of brightness cues than was behavior on subsequent trials. Nevertheless, an examination of behavior on the first trial of each day revealed exactly the same pattern of results in both experiments, as described above for means collapsed across all four trials of each day.

DISCUSSION

In both experiments reported here, behavior in S- was extensively regulated by sequence of reward outcomes experienced in S+ and was largely independent of average incentive associated with S+. In Experiment 1, average reward was equal in Groups NR, RN, and 2 and was less than that in Group 8. According to the incentive averaging view, Groups NR, RN, and 2 should have performed alike and run faster in S- than Group 8.

Instead, Groups 2, 8, and RN performed alike in S- and ran more slowly than Group NR. Only Group NR received N-R transitions. N-R transitions in S+ have previously been shown to retard extinction of responding to S- (Haggbloom, 1979, 1980).

In Experiment 2, according to the incentive averaging view, Groups SNL and LNS, given the same average reward in S+, should have responded alike in S- and at an intermediate level, that is, more slowly than Group SNS, which received a smaller average reward, and faster than Group LNL, which received a larger average reward. Instead, S- speeds were faster in Group SNL than in Group LNS despite equal average reward in those groups, and S- speeds were equal in Groups LNL and SNS despite a large average reward disparity between those groups. Group SNL ran faster than Group SNS, although average reward was greater in Group SNL; Group LNL ran faster than Group LNS, although average reward was greater in Group LNL.

The finding here, in Experiment 1, that resistance to discrimination is increased by N-R transitions in S+ is in agreement with previous findings (Haggbloom, 1980; McHose & Blackwell, 1975) and with the effects of reward schedule on resistance to extinction following PRF. The schedule effects obtained in Experiment 2 also parallel those of comparable schedules on resistance to extinction. In conventional PRF, as here, resistance to extinction is increased by L in the N-R transition and decreased by L in the R-N transition (Leonard, 1969a).

The present experiments emphasize the role of sequential variables in discrimination learning at the expense of the average incentive view. Recently, Capaldi (*in press*) proposed a model of reinforcement, reinforcement level theory, which, like the incentive averaging view, predicts that S- behavior will be inversely related to average reward in S+ (hereafter called the inverse relationship). Thus, the present results would also be inconsistent with reinforcement level theory, except that the reinforcement level principle has been incorporated into sequential theory. According to sequential theory, predictions regarding S- behavior must take into account both sequential variables and reinforcement level. When sequential variables are equated between groups that differ in reinforcement level associated with S+, the inverse relationship should obtain.

The inverse relationship occurs regularly between groups given different magnitudes of consistent reinforcement (CRF) in S+ and the same sequence of S+ and S- trials (e.g., MacKinnon, 1967; Matsumoto, 1969; McHose & Howard, 1973; McHose et al., 1971; McHose & Peters, 1973). In Experiment 1 here, sequential variables were equated in Groups 8 and 2, and those groups differed in reinforcement level (and average reward) associated with S+, yet, the inverse relationship did not occur. That finding is not in agreement with reinforcement level theory or with most previous experiments.

One possible explanation of the present results is that neither group experienced the N-R transitions that usually occur in differential conditioning from S- to S+ and which, like N-R transitions within S+, retard extinction of responding to S- (Haggbloom, 1978, 1980). In the absence of N-R transitions, the discrimination may have been learned too rapidly here so that a "floor effect" precluded the observation of differences in S- behavior between Groups 2 and 8.

Experiment 2 provides a good illustration of results that are consistent with predictions derived simultaneously from sequential theory and reinforcement level theory. According to the sequential view, behavior in S- under the present conditions was extensively regulated by SN in addition to the S- cue. While all four groups received N-R transitions, the four groups would be ordered in strength of conditioning to SN: SNL > LNL > SNS > LNS. The latter prediction follows from the reinforcement level principle that the strength of conditioning is an increasing function of the size of the discrepancy between expected reward and obtained reward.

Expectancy of reward is learned and is acquired slowly. Thus, at the outset of training, the discrepancy between expected and obtained reward on trials on which responding would be conditioned to SN would be greater in Groups SNL and LNL than in Groups SNS and LNS. With continued training and growth of expectancy, conditioning to SN would become stronger in Group SNL than in Group LNL, since expectancy of reward would be less in Group SNL, and stronger in Group SNS than in Group LNS, since expectancy of reward would be less in Group SNS. This is the predicted order identified above.

The predicted order does not agree with how the groups were ordered in resistance to discrimination at the end of training, that is, SNL > LNL = SNS > LNS. Here, it is necessary to turn to the role of expectancy in determining inhibition and the notion that running speed is a function of habit strength minus inhibition.

According to reinforcement level theory, inhibition is acquired on trials on which expected reward exceeds obtained reward. The amount of inhibition acquired is thought to be an increasing function of the size of this discrepancy. Thus, Group LNL would acquire considerably more inhibition than Group SNS. Moreover, since expectancy of reward is acquired slowly, inhibition is acquired later in training than habit strength. Note (see right panel of Figure 1) that, early in training, Group LNL ran faster in S- than Group SNS (greater habit strength in Group LNL and little inhibition) but that, later in training, speeds in Group LNL were identical to speeds in Group SNS (due to the acquisition of inhibition). Of course, inhibition would also contribute to the differences in S- speed between Groups SNL and LNL and between Groups SNS and LNS.

The present experiments indicate that experiments in which average reward is varied between groups by employing PRF in S+ (e.g., McHose et al., 1971) do not provide unequivocal support for the role of average incentive in determining S- behavior. A similar problem exists with regard to the effects of a number of other reward manipulations in S+ on S- behavior. For example, average reward increases, and behavior in S- decreases, as percentage of reward in S+ increases (e.g., McHose et al., 1972; McHose & Peters, 1973). However, unless sequential variables are specifically equated between groups, the number of N-R transitions usually decreases as reward percentage increases (see Capaldi, 1966). Thus, differences in S- behavior attributed to percentage of reward in S+ might, instead, be due to differences in the number of N-R transitions. Similarly, behavior in S- is inversely related to the proportion of S+ to S- trials (McHose et al., 1972; McHose & Howard, 1973). However, as the proportion of S+ to S- trials increases, the number of transitions from S- to S+, an important variable regulating resistance to discrimination (Haggbloom, 1978), is likely to decrease. Finally, the use of varied small (S) and large (L) reward trials in S+ to manipulate average incentive (e.g., McHose, 1970) introduces S-L transitions into the situation. S-L transitions, like N-R transitions, increase resistance to extinction (e.g., Leonard, 1969) and may increase resistance to discrimination. Additional experiments are needed to identify the joint effects of reinforcement level (or average incentive) and sequential variables on S- behavior when S+ reward is manipulated in the ways just described.

REFERENCES

- CAPALDI, E. J. Partial reinforcement: An hypothesis of sequential effects. *Psychological Review*, 1966, **73**, 459-477.
- CAPALDI, E. J. Reinforcement level: An expectancy-associative approach to relative reinforcement and nonreinforcement effects. In G. McCain & J. Bearwalt (Eds.), *The Arlington Symposium on Learning*. Stamford, Conn: Greylock, in press.
- CAPALDI, E. J., BERG, R. F., & MORRIS, M. D. Stimulus control of responding in the early trials of differential conditioning. *Learning and Motivation*, 1975, **6**, 217-229.
- HAGGBLOOM, S. J. Intertrial interval effects on internal stimulus control of behavior in brightness differential conditioning. *Learning and Motivation*, 1978, **9**, 347-358.
- HAGGBLOOM, S. J. The differential reinforcement of reward-produced and response-produced stimuli. *Learning and Motivation*, 1979, **10**, 364-381.
- HAGGBLOOM, S. J. Resistance to discrimination and subsequent resistance to extinction as a function of the sequence of partial S+ reward in differential conditioning. *Animal Learning & Behavior*, 1980, **8**, 441-446.
- HAGGBLOOM, S. J., & TILLMAN, D. J. Sequential effects on discrimination reversal. *Learning and Motivation*, in press.
- LEONARD, D. W. Amount and sequence of reward in partial and continuous reinforcement. *Journal of Comparative and Physiological Psychology*, 1969, **67**, 204-211. (a)
- LEONARD, D. W. Varied reward sequence and resistance to extinction in the straight alley. *Psychonomic Science*, 1969, **16**, 129-130. (b)

- MACKINNON, J. R. Interactive effects of the two rewards in a differential magnitude of reward discrimination. *Journal of Experimental Psychology*, 1967, **75**, 329-338.
- MATSUMOTO, R. T. Relative reward effects in differential conditioning. *Journal of Comparative and Physiological Psychology*, 1969, **68**, 589-592.
- McHose, J. H. Relative reinforcement effects: S_1/S_2 and S_1/S_3 paradigms in instrumental conditioning. *Psychological Review*, 1970, **77**, 135-146.
- McHose, J. H., & BLACKWELL, D. R. Performance in differential instrumental conditioning as a function of the pattern of partial S+ reward. *Animal Learning & Behavior*, 1975, **3**, 63-66.
- McHose, J. H., & HOWARD, G. S. Performance in differential instrumental conditioning with infrequent S+ presentations. *Bulletin of the Psychonomic Society*, 1973, **1**, 132-134.
- McHose, J. H., MAXWELL, F. R., & McHEWITT, E. R. Effects of nonreward in S+ and S- on performance in differential conditioning. *Journal of Experimental Psychology*, 1971, **88**, 282-284.
- McHose, J. H., McHEWITT, E. R., & PETERS, D. P. Average reward as a determinant of S- performance in differential conditioning. *Psychonomic Science*, 1972, **29**, 129-132.
- McHose, J. H., & PETERS, D. P. Differential instrumental conditioning as a function of percentage and amount of positive stimulus reward. *Journal of Experimental Psychology*, 1973, **100**, 413-415.

(Received for publication February 14, 1980;
revision accepted April 21, 1980.)