The reinforcing signal as a conditioned stimulus in human operant discrimination training

HOWARD B. ORENSTEIN
Western Maryland College, Westminster, Maryland

DONALD A. SCHUMSKY
University of Cincinnati, Cincinnati, Ohio

THOMAS ROTH
Henry Ford Hospital, Detroit, Michigan

and

JOHN TRINDER
University of Melbourne, Parkville, Victoria, Australia

We demonstrated that subjects are capable of forming an on-off discrimination based on the fact that they are presented with periods of response-contingent reinforcement discriminable from periods of no such reinforcement. Four groups of 16 college students were run in a double-blind free-operant eyeblink conditioning experiment. Group designation was determined by whether the subject was an experimental or yoked control and whether he was provided with an external signal (S+, a red light) for on periods. Results indicated a significant degree of discrimination in both experimental groups when compared with operant periods and/or yoked controls. Contrary to expectation, no statistically significant advantage accrued to the subjects provided the light signal. Discussion considers the implication of this type of discrimination as an experimental control and conditions under which the provision of S+ should lead to greater advantage.

In operant conditioning, to establish stimulus control in "true discrimination" training one might employ a procedure that alternates presentation of the reinforced conditioned stimulus (S+) with the presentation of a nonreinforced stimulus (S−). Yet, if one examines the acquisition data and the subject reliably responds to S+ and not to S−, can one be certain that the only stimulus controlling responding is S+? To properly demonstrate discrimination, that is, control over responding by S+, one must include the following two tactics: (1) a condition in which the administration of S+ (and S−) is uncorrelated with the administration of reinforcement (often referred to as pseudodiscrimination; see Bresnahan, 1970; Honig, 1969) and (2) a test for the effects of both true discrimination and pseudodiscrimination training in which S+ is presented without reinforcement during extinction.

Even if, as a result of these two tactics, true discrimination training relative to pseudodiscrimination training leads to: (1) high levels of responding to S+ during acquisition and (2) high (initial) levels of responding to S+ during extinction, something other than S+ may be responsible for these outcomes. Over 50 years ago, Skinner (1938) showed that the reinforcer itself might be considered a stimulus that has achieved control over the learned response (see Reid, 1957). More recently, Jenkins (1965) presented a critical analysis of discriminative conditioning in which he faulted the use of these two tactics because extinction interrupted conditioning and "the effects of antecedent reinforcement and nonreinforcement on response probability ... [are] confounded with the effects of the discriminative stimuli on response probability" (p. 365). To avoid such confounding, he devised a "discriminated trial procedure" in which a correct trial (response to S+ or no response to S−) was followed by a new random trial, and an incorrect trial (a response to S−) was repeated. The results of his study indicated that antecedent reinforcement increased, while antecedent nonreinforcement decreased, the probability of responding to a subsequent stimulus, but only early in discrimination training.

A closer analysis of Jenkins's (1965) procedure raises some important issues. The antecedent nonreinforcement effect may have resulted from extensive training on S+ alone prior to discrimination training, a procedure precluding the potential influence of a second type of incorrect trial, that is, no response to S+. Thus, it is unclear whether discrimination training following only minimal training on S+ would result in an antecedent nonreinforcement effect (on S+ and S−) identical to that following maximal training. Despite Jenkins's objections to the use of extinction, to more fully explore how antecedent reinforcement and nonreinforcement affected acquisition

---

Correspondence should be addressed to D. A. Schumsky, Department of Psychology, University of Cincinnati, Dyer Hall (ML 376), Cincinnati, OH 45221-0376.

stages differentially, it might have been instructive to compare performance on S+ during both the early and late segments of discrimination training by inserting a series of nonreinforced trials with S+.

Discriminative stimulus control in Jenkins's (1965) study may have resulted from a high level of "directedness" of the response (e.g., Thomas & Curran, 1988), given that the stimuli were localized and contiguous with the response. Consider, in contrast, the development of a simple on–off discrimination of the human eyeblink in which the response has a (presumed) low level of directedness. Schumsky, Richman, and Trinder (1967) and Schumsky, Trinder, and Richman (1971), for example, using a double-blind yoked control procedure (analogous to pseudodiscrimination training), have shown conditioning of the human eyeblink as a free operant. Using this paradigm, suppose that S+ (a light) and a reinforcing signal (a tone in the subject's ear) are presented during an on interval that is randomly alternated with an off interval, that is, no S+ and no reinforcer. The nature of the free operant is such that the response at issue is one that has a fixed probability of spontaneous occurrence. In addition, the probability of the response occurring during an on period and being reinforced with a tone is usually set to be very close to 1. Thus, it seems likely that the first tone occurring in this way can act as the signal for the start of a period of high likely reinforcement. In short, the reinforcing signal itself rather than (or in addition to) S+ may become a cue that the subject is in an on period (see Jenkins, 1965).

The present experiment was designed to evaluate the extent to which a subject can, in fact, form a discrimination when supplied no external discriminative cues other than the reinforcing signal itself. It would be interesting to compare both acquisition and extinction data from such a No-S+ group with those from a group provided a light (S+) signaling on periods in an on–off discrimination procedure. If such an analysis is correct, the only group showing discriminative behavior during extinction should be the one receiving S+ (light signal) both during acquisition and extinction. If the reinforcing signal serves as the discriminative signal for No-S+ subjects, then no discrimination behavior related to the S+ should or could be discernible for these subjects during extinction. On the other hand, if some uncontrolled source of discriminative cue is available to them, the elimination of only the reinforcing signal during extinction should not immediately affect discrimination behavior.

METHOD

Subjects
Sixty-four male undergraduates at the University of Cincinnati participated to fulfill a course requirement.

Apparatus
The subjects were run in pairs, and both members of the pair were seated in front of a one-way-vision mirror screen. A 2-m-high plywood barrier prevented the pair of subjects from seeing each other during the course of the experiment. Each subject wore a pair of earphones through which the tone that represented his reinforcing signal could be presented.

To one side of the vision screen for each subject was a red light (to serve as S+) placed so that it could be easily seen by the subject but could not be seen by the experimenter counting the subject's responses.

On the experimenter's side of the apparatus were three areas, each separated by a plywood screen. On the outside were seated the experimenters, each counting one subject's eyeblink responses. These experimenters were only able to see their own subjects. In the center of the experimenter side of the apparatus sat the experimenter, who controlled the administration of the reinforcing signal via a button switch. Also available in this central compartment was a switch that turned on and off the red signal light (S+) on the subject's side of the apparatus. The central compartment was arranged so that both subjects could be clearly viewed. However, the experimenters could not see each other.

Procedure
Each pair of subjects was composed of a conditioned subject and a yoked control subject. Upon entry into the experimental room, they were seated on both sides of the barrier in front of the one-way-vision mirror. Each subject was fitted with a pair of earphones.

Each subject was told that this was an experiment involving competition. The competition involved getting more points than his opponent. The problem before him was to determine what it was he had to do to get points and do it more often than his opponent. Points were signaled by means of a tone through the earphones. Each subject was told that the tone would be administered to him by an experimenter on the other side of the mirror who also would be counting points for him.

Each subject was randomly assigned to one of four experimental conditions: S+ or No-S+, factorially combined with the experimental or yoked designation. On and off periods were each 15 sec long and were programmed from a randomization of such periods. The randomization of on–off periods was such that the subject was presented with one on–off cycle per 30-sec period, with the order of periods varied. Two such randomizations of periods were each employed for half of the subjects in the experiment. Thus, the two experimenters counting the subjects' eyeblinks had no foreknowledge as to which subject was experimental or yoked, whether S+ was being employed on this set of trials, or whether a given period was an on or off period.

Experimental Design
The experiment was run in three separate phases, each following one another continuously. The first minute, Trial Periods 1–4, of the experiment was employed for a determination of operant level. No differential treatments were applied to the subjects during this period. The next 5 min was the acquisition period. The final 2 min was the extinction period.

The experimental design during the acquisition phase of the experiment was a $2 \times 2 \times 2 \times 10$ factorial design with repeated measures on the last two variables. That is, four groups of 16 subjects were defined by whether the subject was in a pair to receive S+ or not and whether he was an experimental or yoked control subject. Each subject was administered 10 on–off cycles. During the 15-sec on period, both subjects were administered a tone each time the experimental subject blinked his eye. During the off period, no tones were administered. For both experimental and yoked control subjects in groups receiving S+, the red light was turned on during on periods and off during off periods.

During the extinction phase, the groups were subdivided. Half of each acquisition group ($n=8$) was administered S+ during extinction, and half were not. Thus, half of the subjects trained with S+ received S+ during extinction, and half did not. The No-S+ groups were similarly subdivided. The on–off trial-period designation during extinction for the subjects not receiving S+ was merely for purposes of scoring and statistical analysis. All the subjects during this phase were administered four such on–off cycles, although previous research with similar procedures indicated that extinction is usually complete in two trial periods (Schumsky et al., 1967).

RESULTS

The statistical outcomes reported here were evaluated with a rejection level of $p < .01$.
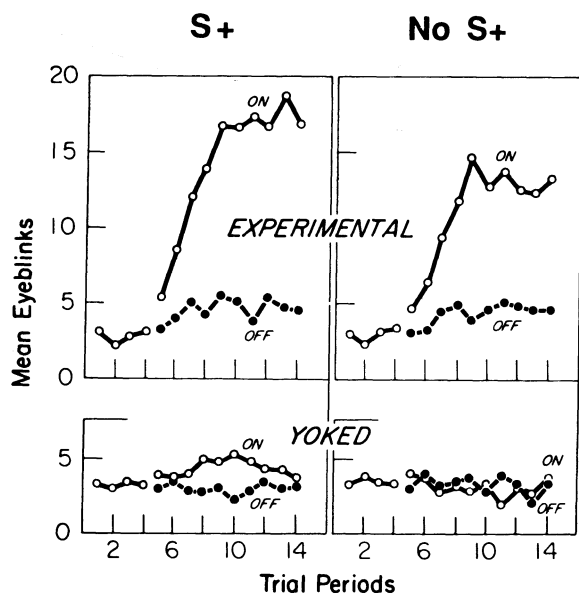
Figure 1. Mean number of eyeblinks per 15-sec trial period during operant periods (1–4) and discrimination acquisition periods (5–14). Separate results are presented for experimental subjects (top) and their yoked controls (bottom) and for S+ (left) and No-S+ (right) subjects.

Operant Period

The main dependent variable of this experiment was the number of eyeblinks emitted by each subject during each 15-sec period of the experiment. Since the four main experimental groups received no differential treatment during the operant period of the experiment, a 4 (group) × 4 (trial period) analysis of variance (ANOVA) was carried out to analyze operant period responding. As expected, the Fs for groups, periods, and their interaction were all nonsignificant (all Fs < 1.00). Therefore, the hypothesis of common conditioning origin, as assessed in the operant period, is not rejectable. Figure 1 presents the mean number of eyeblinks per 15-sec trial period for all groups during the operant period (Trial Periods 1–4). As can be seen from this figure, the four groups were quite homogeneous with regard to mean number of eyeblinks during this period.

Acquisition Period

The data representing the acquisition period are presented in Figure 1 as Trial Periods 5–14. During the acquisition period, separate response functions are plotted representing responses from on and off periods for each experimental group. A comparison of the data from the experimental subjects (top) and their yoked controls (bottom) reveals rather strong evidence for the development of conditioning in the experimental subjects, for whom the administration of reward was response contingent. This is further supported by the results of the ANOVA, where significant mean squares are associated with the experimental versus yoked designation [$F(1,60) = 20.95$], trial periods [$F(9,540) = 9.04$], and their interaction

[$F(9,540) = 9.47$]. Also, as can be seen from Figure 1, good evidence of differential responding seems apparent. The difference between on periods, when reward was forthcoming, and off periods, when it was not, is apparent from the figure. The outcome of the statistical analysis supports this. The ANOVA revealed significant mean squares associated with the on versus off period designation [$F(1,60) = 27.38$] and the on-off × trial period interaction [$F(9,540) = 6.89$]. Clearly, the development of the on–off discrimination was based on whether the subject was experimental or a yoked control [$F(1,60) = 27.13$] and changes in these effects over trial periods [$F(9,540) = 6.68$].

A primary concern of the present experiment was with the effect of the presence or absence of a red light (S+) signaling on periods. In comparing the two experimental groups across the top of Figure 1, there seems to be a noticeable advantage accruing to those subjects receiving the light signal (left) when compared with the response of those not receiving this signal (right). Surprisingly, this difference was not statistically significant. No mean square in the ANOVA associated with the S+ versus No-S+ designation as a main effect or in interaction even approached statistical significance. To provide a more sensitive test of this difference, an additional analysis was employed. In this analysis, only the on periods for the two experimental groups were compared as a function of trial period in a 2×10 ANOVA. It was hoped that this analysis, isolating the crucial effect in question, would provide a more sensitive test, since possible obscuring of effects due to averaging of groups could not occur. The results of this more sensitive test were in accord with the results of the overall analysis. Neither the S+ versus No-S+ designation [$F(1,30) = 1.00$] nor the interaction mean square across trial periods [$F(9,270) = 0.43$] was significant in this analysis. Not even when evaluating these mean squares against error terms, based on larger dfs, from the overall analysis was statistical significance obtained [S+ vs. No-S+, $F(1,60) = 2.21$; S+/No-S+ × trial period, $F(9,540) = 1.15$]. Thus, despite the apparent superiority of performance of S+ subjects when compared with No-S+ subjects in Figure 1, no statistically significant advantage accrued to subjects in this experiment because of the presence of S+ during acquisition.

Extinction

Results of previous research using the present conditioning techniques predicted that resistance to extinction would be quite short lived. Indeed, expectations were that any carryover effects into extinction should not be apparent beyond the first or second 15-sec trial period. Furthermore, the analysis presented in the introduction suggested that the only group predicted to have such carryover effects was that composed of experimental subjects who were administered S+ both in acquisition and extinction. This effect should be reflected in their responding during the first on period. This was indeed the case. The mean number ($M = 9.00$) of eyeblinks by this group in this period was the only one significantly different from the over-

all mean level of response during extinction [*M* = 4.27, *t*(15) = 11.48]. (For purposes of comparison, the next most deviant mean was equal to 6.50.)

DISCUSSION

The results of the present experiment provide evidential support for the contention that subjects can and do form an on–off discrimination on the basis of the differential administration of reinforcement. The subjects in the No-S+ experimental group gave clearcut, statistically significant evidence of the formation of a discrimination between on and off periods in the absence of any additional signal beyond the reinforcing tone. Comparison of the acquisition data of the No-S+ experimental group with that of their yoked controls argues that on versus off period differences were not merely a facilitation effect due to the high frequency of reinforcing signals during the on period. The experimental versus yoked control comparison further suggests that the administration of reinforcement must be made contingent upon the subject's making the appropriate instrumental response, in this case an eyeblink. Explanations based on temporal conditioning would also seem to have been ruled out by the use of a randomization of on and off periods. This procedure effectively provided for variable intervals between successive on periods.

The results of the present experiment cast doubt on studies purporting to demonstrate discrimination on the basis of an on–off difference when only acquisition data have been analyzed. Also, the inclusion of a No-S+ group would provide rather precise estimates of the degree of effect associated with S+, or relative effects of different S+s. Historically, books on design of such experiments (e.g., Sidman, 1960) warn against the problem of a conditioned stimulus acting as a reinforcer but offer no mention of the problem of concern here, that the reinforcer may act as a conditioned stimulus.

Statistical comparisons between the experimental groups receiving and not receiving S+ revealed no statistically significant advantage accruing to those subjects receiving the red-light signal. This was an unexpected finding. Graphical presentation of the data indicated that the groups receiving S+ did show a higher level of response than their No-S+ counterparts during on periods. If the present analysis is correct and the No-S+ subjects are discriminating the on periods because of the administration of the reinforcing signal, then the advantage afforded the S+ group in acquisition is a temporal one based on the earlier onset of the light for the S+ group as compared with the first tone reinforcement for the No-S+ group. The magnitude of the difference between the groups should be approximated by computing an average latency of the first eyeblink from the off periods of the No-S+ group and multiplying that by the

rate during on periods. (Computation led to an estimated on-off difference of 3.63 as compared with an actual difference of 4.47.) One would predict, then, that shortening the on-off periods would provide a relatively greater advantage for the S+ group and lengthening the periods would reduce the advantage. Choosing operant responses with varying probabilities of occurrence should have a similar effect.

A possible limitation upon the present findings seems worthy of mention. The present experiment employed a 1:1 ratio of reinforcement during on periods. One would presume that such a procedure would make on versus off periods most easy to discriminate. Consequently, as the ratio of reinforcement is increased, such a discrimination would be more difficult to make and acquisition would be slowed. The latter is the case in the simple administration of ratios of reinforcement during acquisition. Thus, it would seem necessary for further experimental investigation to define boundary conditions.

REFERENCES

BRESNAHAN, E. L. (1970). Effects of extradimensional pseudo-discrimination and discrimination training upon stimulus control. *Journal of Experimental Psychology*, **85**, 155-156.

HONIG, W. K. (1969). Attentional factors governing the slope of the generalization gradient. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 35-62). New York: Academic Press.

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

REID, R. L. (1957). The role of the reinforcer as a stimulus. *British Journal of Psychology*, **49**, 202-209.

SCHUMSKY, D. A., RICHMAN, C. L., & TRINDER, J. (1967). Free operant conditioning of the human eyelid response. *Psychonomic Science*, **8**, 237-238.

SCHUMSKY, D. A., TRINDER, J., & RICHMAN, C. L. (1971). Human operant eyelid conditioning: Auditory versus visual reinforcing signal. *Psychonomic Science*, **22**, 359-360.

SIDMAN, M. (1960). *Tactics of scientific research*. New York: Basic Books.

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

THOMAS, D. R., & CURRAN, P. J. (1988). The role of response directedness in discriminative and conditional stimulus control. *Bulletin of the Psychonomic Society*, **26**, 378-380.