

EFFECT OF NUMBER AND LOCATION  
OF S- ON DISCRIMINATION

EFFECT OF NUMBER AND LOCATION OF NEGATIVE  
TRIALS ON A FEATURE POSITIVE DISCRIMINATION

by

Maria Catalina Filip de Duarte, B.A.

A Thesis

Submitted to the Faculty of Graduate Studies  
in Partial Fulfillment of the Requirements  
for the Degree  
of Master of Arts

McMaster University.  
April, 1977

MASTER OF ARTS (1977)  
(Psychology)

McMaster University

TITLE: Effect of number and location of negative trials on a feature  
positive discrimination

AUTHOR: Maria Catalina Filip de Duarte, B.A. (Universidad Iberoamericana)

SUPERVISOR: Dr. H.M. Jenkins

NUMBER OF PAGES: 46, vii

### ABSTRACT

Pigeons were trained to discriminate between two displays differentiated only by a distinctive feature on the positive display. The development of a simultaneous discrimination was measured by the percent of key pecks on the positive display that were directed to the distinctive feature. The effect of a) the sequence of positive and negative trials, b) the number of negative trials, and c) the shape of the distinctive feature on the development of discrimination were examined.

An alternating sequence of a block of positive trials and a block of negative trials was interrupted by session breaks either between positive trials, or between positive and negative trials. The location of the session break preserved or prevented within session transitions from negative trials to positive trials. Sequences that preserved the within session transition from negative to positive trials were highly effective in producing control of key pecking by the distinctive feature. Sequences that prevented within session transitions from negative to positive trials were less effective, particularly when the shape of the distinctive feature was less preferred. The more trials in the negative block the more rapidly discrimination occurred. However, not even the longest block was capable of producing control of key pecking by the less preferred distinctive feature when the session break prevented a within session transition from negative to positive trials.

The preparation employed in this experiment was particularly suitable for evaluating the effect of sequential variables on stimulus control. While differential responding between positive and negative displays can be controlled by the sequence of reinforcement and non-reinforcement itself, differential responding within the positive display cannot.

### ACKNOWLEDGEMENTS

I want to express my sincere gratitude to Dr. H.M. Jenkins for his invaluable guidance and patient assistance throughout all the course of this work.

I would also like to thank Dr. F. Barrera and Ms. Lila Krishnan for many practical suggestions.

I am grateful to Durval, my husband, for encouraging my scientific aspirations, and to my parents for their unconditional support.

## CONTENTS

	<u>Page</u>
Introduction	1
Method	18
Results	21
Discussion	30

## FIGURES

- Figure 1. Idealized shapes of generalization gradients obtained after simple conditioning and after discrimination training with a random sequence of S+ and S- trials.
- Figure 2. Top panel: Arrangements of S+ and S- trials to which different groups were exposed.  
Lower panel: Two pairs of display used as stimuli. The location of the features on the display was varied at random from trial to trial.
- Figure 3. Mean percent of responses on S+ trials that were directed to the DF (dot) during the first and second blocks of positive trials on each session of discrimination training.
- Figure 4. Mean percent of responses on S+ trials that were directed to the DF (star) during the first and second blocks of positive trials on each session of discrimination training.
- Figure 5. Mean number of response to S- displays for each session of discrimination training in experimental groups (dot as DF on positive displays).
- Figure 6. Mean number of responses to S- for each session of discrimination training in the experimental groups (star as DF on positive trials).

Figure 7. Top panel: Median number of S- trials presented before reaching the shift criterion for each experimental group.

Lower panel: Median number of active S- trials before reaching the shift criterion (dot is DF) for each experimental group. Individual data is shown by dark dots. Active S- trials are trials with at least one keypeck.

Figure 8. Left panel: Median number of S- trials presented before reaching the shift criterion for each of the PNP Groups.

Right panel: Median number of active S- trials before reaching the shift criterion for each of the PNP Groups. Individual data shown by dark dots (star is the DF).

Figure 9. Top panel: Mean percent of responses on S+ trials that were directed to the DF (star) during sessions 17 and 18 of discrimination training with nonrandom sequences and during each of the six training sessions with a random sequence.

Lower panel: Mean number of S- responses during sessions 17 and 18 of discrimination training with nonrandom sequences and during the six training sessions with a random sequence.



## Introduction

3. A common procedure for training a discrimination between successively presented stimuli is to present them in an irregular sequence, associating reinforcement with one stimulus and nonreinforcement with the other. Pavlov's so-called method of contrast explicates the procedure in classical conditioning:  $CS_1^+$  trials (reinforced) are presented in a mixed sequence with  $CS_2^-$  trials (nonreinforced). The parallel arrangement in operant conditioning involves periods of reinforcement of the response in the presence of  $S_1^+$  intermixed with other periods when  $S_2^-$  is presented without reinforcement. In each case the result of training, assuming appropriate parameters, is a go/no-go performance; the response to the positive stimulus is maintained while the response to the negative stimulus is eliminated or at least reduced.

It is well known that training with a mixture of  $S_1^+$  and  $S_2^-$  trials can produce a greater performance differential to  $S_1-S_2$  than is found after simple conditioning to  $S_1^+$  followed by a test of performance with  $S_1$  and  $S_2$ . Intermixed training selectively reduces the tendency to respond to  $S_2$ . For example, if  $S_1$  were a tone at 1000 Hz and  $S_2$  a tone at 3000 Hz the repeated reinforcement of a response to  $S_1$  would, in a typical case, produce a strong tendency to respond when  $S_2$  was presented. The  $S_2$  strength would have two sources: 1) stimulus generalization based on the similarity between  $S_1$  and  $S_2$ , and 2) conditioning to nonauditory features of the experimental situation that accompanied both  $S_1$  and  $S_2$  presentations,

or, in other words, conditioning to common features. When these sources of responding on  $S_2$  trials are substantial, performance will not be sharply controlled by the  $S_1$ - $S_2$  stimulus variable. In order to obtain sharp stimulus control these sources of strength must be reduced. The intermixture of  $S_1^+$  and  $S_2^-$  trials during training can accomplish the needed reduction and thereby produce a great increase in stimulus control.

Although it is known that interspersing  $S_2^-$  trials with  $S_1^+$  trials produces strong stimulus control, the process by which interspersing accomplishes its effect is not well understood. A better understanding of the process will require systematic data on the effects of different degrees of interspersing, or more generally, on the effects of varying the location of  $S_2^-$  trials within the training sequence. The principal purpose of the present experiment is to obtain such information.

The usual strategy of previous research on the location of  $S_2^-$  trials has been to vary the degree of interspersing of  $S_2^-$  with  $S_1^+$  trials during training and then to test stimulus control in a separate phase. For subjects trained with minimal interspersing of  $S_1^+$  and  $S_2^-$ , the trials are presented in uninterrupted blocks of one or the other trial type. This sequence, like every nonrandom sequence of trials, introduces the possibility that responding will come under the control of sequential cues for reinforcement and nonreinforcement. There is ample evidence that a partial go/no-go discrimination can develop on the basis of regularities in the sequence of reinforced and nonreinforced trials when no differential trial stimuli are present (e.g. Verplank, 1942; Jenkins, 1965; Capaldi, 1971). Moreover

as experiments reviewed later demonstrate, a discrimination based on the predictability of the sequence can develop even when reinforced and nonreinforced trials are also distinguished by discriminably different trial stimuli.

The possibility of effective sequential cues means that in order to obtain an unconfounded assessment of control by the discriminative stimuli it is necessary to discontinue training under the nonrandom sequence. A common procedure is to evaluate stimulus control by means of stimulus generalization tests in which reinforcement is discontinued and the formerly positive and negative stimuli are presented together with a range of intermediate stimulus values in a random series.

In the present experiment a new procedure is used to study the effect of the location of  $S_2^-$  trials on discriminative performance. The new procedure allows one to examine the development of a discrimination based on the  $S_1-S_2$  trial stimuli without an interruption of training by a switch from a nonrandom to a random sequence of trials. Before describing the new procedure it will be useful to review the principal results that have been obtained to date in experiments employing a separate test of stimulus control.

An evaluation of the effect of a particular sequence of  $S_1^+$  and  $S_2^-$  trials typically entails comparing gradients of generalization after each of three training conditions: 1) simple conditioning with  $S_1^+$  presentations only, 2) fully interspersed presentations of  $S_1^+$  and  $S_2^-$ , and 3) a particular nonrandom sequence that involves reduced interspersion obtained by a single or repeated alternation of

blocks of  $S_1^+$  and  $S_2^-$  trials. Comparisons among these gradients allows one to assess the effects on stimulus control of a given form of reduced interspersion in relation to the effects of fully interspersed presentations of  $S_2^-$ .

The stimulus generalization gradient obtained after fully interspersed training (sometimes called the post discrimination gradient) commonly differs from the one obtained after simple conditioning in several respects. The differences are brought out in Figure 1 which shows, in idealized form, typical gradients of generalization after each type of training.

---

Insert Figure 1 about here

---

The literature in which generalization gradients are compared has developed a nomenclature, not completely standardized, which includes the following terms. When compared with the gradient obtained after simple conditioning the post discrimination gradient is said to show sharpening of control, selective reduction of responding at  $S_2$ , elevation of the peak and peak shift. Sharpening of control refers to the steeper decline in responding as one moves away from the stimulus value at which maximal responding occurs. As a result, a larger percent of the total responses in the entire gradient occurs at the point of maximal responding. Selective reduction has been variously defined (and sometimes not defined) but for the purposes of this review of previous experiments we consider that selective reduction occurred at  $S_2^-$  if the percent reduction in responding from  $S_1$  to  $S_2$  was greater in the post discrimination gradient than in the gradient

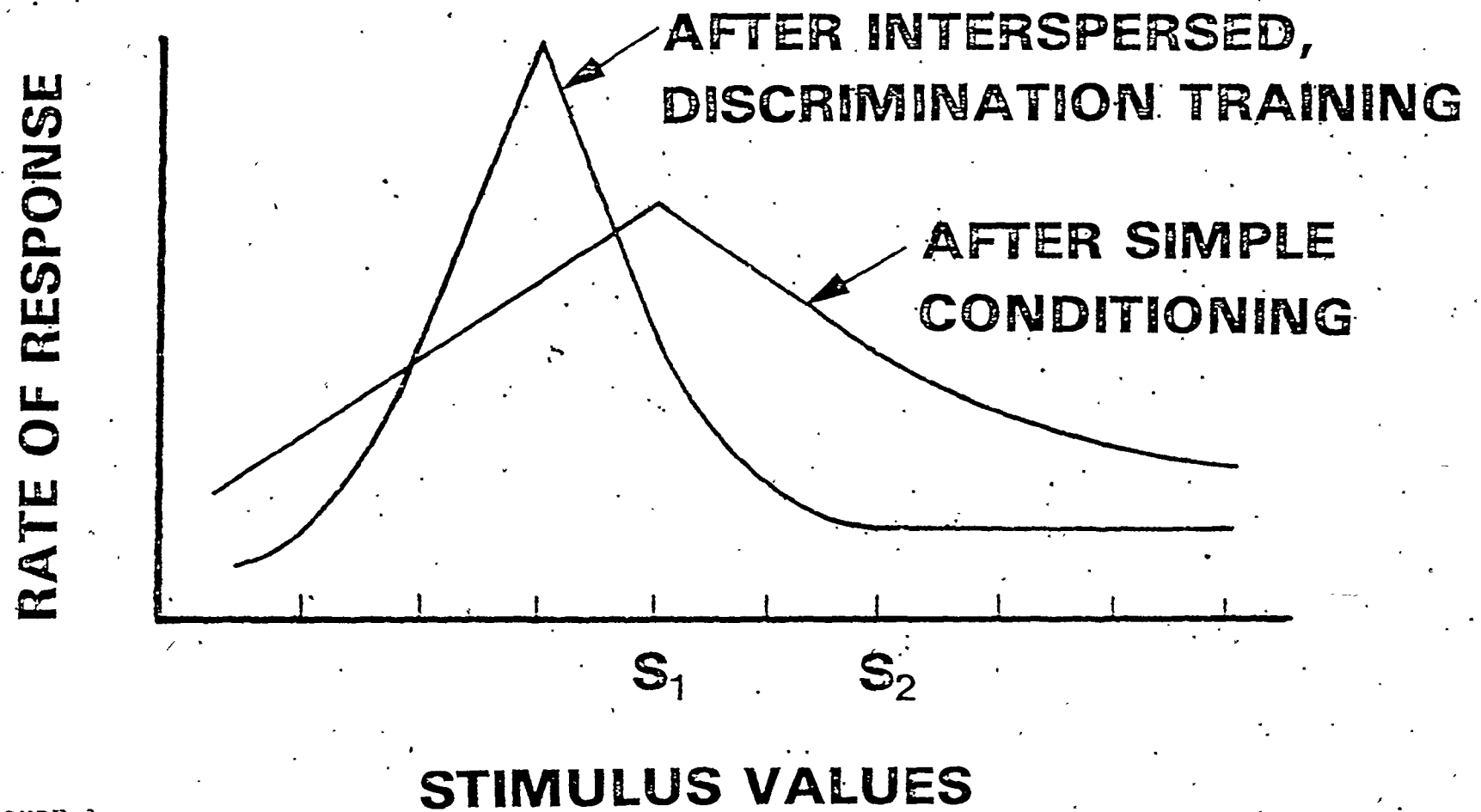


FIGURE 1

obtained after simple conditioning. Peak shift occurs when the point of maximal responding is shifted from  $S_1^-$  in a direction away from  $S_2^-$ . Elevation of the peak occurs when the total amount of responding at the point of maximal responding is greater in the post discrimination gradient than in gradient after simple conditioning.

It is convenient to review experiments on the effect of the sequencing of  $S_2^-$  trials under three categories: 1) those in which only one block of  $S_2^-$  trials is presented (massed extinction), 2) those in which interspersed discrimination involving one  $S^-$  (say  $S_x^-$ ) precedes a massed extinction on  $S_2^-$  (massed extinction preceded by interspersed training) and 3) those in which blocks of  $S_1^+$  trials are repeatedly alternated with blocks of  $S_2^-$  trials (repeated blocks).

Massed extinction. Honig, Thomas and Guttman (1959) were the first to evaluate the effects of massed extinction on the generalization gradient. They reinforced pigeons on a variable interval schedule for pecking a key illuminated by a narrow band of wavelengths centered at 550 nm ( $S_1^+$ ). After 10 sessions of training (simple conditioning) half of the subjects were exposed to a single extinction session with the keylight at 570 nm ( $S_2^-$ ). Generalization gradients along the wavelength dimension were measured the next day. All subjects then received further training with  $S_1^+$  and  $S_2^-$  interspersed followed by a second test of generalization. When compared with the gradient obtained after simple conditioning the gradient after interspersed discrimination training showed all of the effects illustrated in Figure 1. The post discrimination gradient was sharper, showed peak shift, an elevation of responding at the peak, and a selective reduction of responding at  $S_2$ . None of these byproducts of interspersed

discrimination training appeared reliably in the gradients obtained after massed extinction on  $S_2$ . Compared with simple conditioning, massed extinction resulted in a lower tendency to respond at all stimulus values but the percent reduction was just as large at  $S_1$  as at  $S_2$ . The experiment provided a clear demonstration of what had been suspected: a single block of extinction on  $S_2^-$  is less effective in increasing stimulus control than is interspersed training with  $S_2^-$ . It raised the stronger possibility that without interspersion (i.e. without a return to  $S_1^+$ ) the  $S_2^-$  trials had no selective effects on the profile of responding to stimulus values. The weaker conclusion has held up in subsequent research but the stronger one has not.

Honig et al. stated a general conclusion in a way that was later challenged. They concluded that the effect of discrimination training with interspersed  $S_1^+$  and  $S_2^-$  presentations was more than could be accounted for by a combination of excitatory effects at  $S_1$  and inhibitory effects at  $S_2$ . That conclusion assumed that extinction at  $S_2^-$  was sufficient to make  $S_x$  inhibitory, but that is an assumption that requires a separate test.

Weisman and Palmer (1969) examined the question of whether massed extinction produces inhibition with a procedure for testing inhibition that was described by Jenkins and Harrison (1962). In this procedure the nonreinforced stimulus,  $S_2$ , consists of  $S_1$  plus a second stimulus. In the Weisman and Palmer experiment  $S_1$  was a blank white pigeon key and  $S_2$  was a vertical black line on the same white background. The added stimulus is chosen so that it can be varied during a generalization test without making the  $S_2$  presentation as a whole any more, or less, like  $S_1$  (the added stimulus is said to

be orthogonal to  $S_1$ ). If a test of generalization over line orientations were to show a U-shaped gradient with a minimum at  $S_2$  (the vertical orientation of the line in the present example), one would conclude that responding was inhibited by  $S_2$ . Since the increase in responding is not due to more excitation from  $S_1$  it must be due to less inhibition from  $S_2$ . Weisman and Palmer gave one group of pigeons discrimination training with interspersed  $S_1^+$  and  $S_2^-$  periods and obtained a U-shaped gradient of inhibition over line orientation. A second group received massed extinction on  $S_2$  and it showed only a zero level of responding along the entire stimulus dimension. They concluded that massed extinction was not sufficient to make the nonreinforced stimulus inhibitory. Hearst (1972) has, correctly, questioned this conclusion on the grounds that floor effects may mask possible inhibitory effects. The fact remains, however, that the same test yielded evidence of inhibition after interspersed training and none after massed extinction. One could conclude that the effects of nonreinforced  $S_2$  trials on inhibitory control by  $S_2$  depends on whether the trials occur in a single block (massed extinction) or are interspersed with  $S_1^+$  trials and it is therefore not safe to assume that massed extinction is sufficient to produce inhibition.

Massed extinction preceded by interspersed training. Friedman and Guttman (1965) showed that some of the effects on the generalization gradient produced by interspersed training with  $S_2^-$  can also be produced by a massed extinction on  $S_2^-$  if the massed extinction is preceded by interspersed discrimination training on an unrelated  $S_2^-$ , say  $S_x^-$ . In one experiment, two groups were first trained to peck at a 550 nm keylight for food reinforcement on a variable



interval schedule. They then received interspersed discrimination training between the 550 nm keylight ( $S_1^+$ ) and the same light on which a dark cross ( $S_x^-$ ) was superimposed. The experimental group then received a single 40-min extinction session with a 570 nm keylight ( $S_2^-$ ) while the control group received no treatment. Generalization gradients were obtained during extinction in a separate test session. Gradients from the experimental groups were sharper than for the control and showed a peak shift whereas those in the control group did not. The authors concluded that the initial interspersed discrimination training on the unrelated stimulus had produced a state of susceptibility to specific extinctive effects on  $S_2^-$  during massed extinction.

Yarczower (1974) also examined the effect of prior discrimination training in an experiment with pigeons and used stimuli similar to those used by Friedman and Guttman. There were four groups. The first was trained only on  $S_1^+$  (simple conditioning). The second was trained on  $S_1^+$  and then given a single massed extinction on  $S_2^-$ . The third and fourth groups first received interspersed presentations of  $S_1$  and  $S_x$  and then a massed extinction on  $S_2^-$ . The third group received reinforcement on  $S_1^+$  and nonreinforcement on  $S_x^-$ , as did the corresponding group in the Friedman and Guttman experiment (differential training). In the fourth group, however,  $S_1$  and  $S_x$  were equally reinforced (nondifferential training). The results were surprising on two counts. First, simple conditioning followed by massed extinction produced gradients that were clearly sharper than those obtained after simple conditioning alone. Yarczower suggested

that he obtained sharpening as the result of massed extinction whereas Honig et al. did not because he subjected his pigeons to a more complete extinction on  $S_2^-$ . The second surprising result was that both groups exposed to  $S_1$  and  $S_x$  prior to massed extinction showed sharper gradients than the other groups. The absolute level of responding was greater in the group differentially trained ( $S_1^+, S_x^-$ ) but in terms of relative generalization gradients (percent of total responses at each stimulus test value) it made no difference whether  $S_x$  had been reinforced or nonreinforced. It now seems possible that the effect of prior, interspersed discrimination training on the effectiveness of a subsequent massed extinction is simply due to the reinforced exposure to different stimuli rather than to differential, go/no-go, discrimination training between those stimuli.

Repeated blocks. We now review two experiments that provide a comparison of the effects of training on interspersed  $S_1^+, S_2^-$  presentations with training on repeated alternating blocks of  $S_1^+$  and  $S_2^-$  presentations.

Yarczower and Switalski (1969) trained goldfish to nose a key illuminated with light at 550 nm,  $S_1^+$ , for a food reward. In each of 25 sessions, one group received 20  $S_1^+$  trials only (simple conditioning), another received the same block of  $S_1^+$  trials followed by an equal block of  $S_2^-$  trials (590 nm), and a third group received 20  $S_2^-$  interspersed with the 20  $S_1^+$  trials. A partial go/no-go discrimination developed in both groups that received  $S_2^-$  trials. However, when  $S_2^-$  trials occurred in a single block at the end of each session the discrimination appeared to be based largely on the

occurrence of nonreinforcement, which predicted further nonreinforcement, rather than on the change in the visual stimulus. This was suggested by the fact that a high percentage of all  $S^-$  responses occurred on the first few nonreinforced trials (62% in the first 5) whereas when the  $S^-$  trials were interspersed, the percentage of all  $S^-$  trials responded to was not greater for the first 5 trials than for later trials (25% in the first 5). Generalization gradients on wavelength showed a selective reduction in responding in the vicinity of  $S_2^-$ , when compared with simple conditioning, for both the blocked and interspersed presentations of  $S_2^-$ . However, the selective reduction was significantly greater after interspersed training than after blocked training. Only interspersed training produced a significant overall sharpening of the gradient when compared with simple conditioning.

The most instructive experiment to date on the effects of trial sequencing on stimulus control was reported by Ellis (1970). The experiment provides comparisons among simple conditioning, interspersed discrimination training, and training with repeated alternation of blocks of positive and negative trials in four different arrangements. In one group, labelled AE (acquisition-extinction), each session consisted of a block of positive trials followed by a block of negative trials so that the only within-session transition was from  $S_1^+$  to  $S_2^-$ . A second group (EA) received only the transition from  $S_2^-$  to  $S_1^+$  in each session. The remaining groups (AEA and EAE) received both types of transition in each session. In order to examine the development of stimulus control, four separate generalization tests were given after

different amounts of training. In the last phase of the experiment, subjects were placed on a random sequence of positive and negative trials. The initial go/no-go performance on the random sequence provides another assessment of control by the trial stimuli since sequential cues are removed. The experiment involved the pigeon's key peck to lights of different brightness. A go/no-go discrimination developed in all experimental groups. The development was more rapid in the group with interspersed  $S_2^-$  trials than in other groups, but by the end of 14 session of training all groups were showing approximately the same overall difference in responding to  $S_1^+$  and  $S_2^-$ . The generalization test data showed, however, that the go/no-go performance was more under the control of the visual stimulus in groups that had been exposed to the transition from negative to positive trials within sessions (EA, AEA, EAE) than in the group that received the positive to negative transition only between sessions (AE). When compared with the gradient after simple conditioning the groups with the EA transition showed sharpening, peak shift, elevation of the peak, and selective reduction of responding at  $S_2^-$ . The AE group, on the other hand, showed neither sharpening, peak shift, nor elevation of the peak although by the end of testing there was selective reduction of responding at  $S_2^-$ . Performance on the random sequence was consistent with generalization test data: the AE group responded indiscriminately to positive and negative trials during the first session on the random sequence whereas the groups trained with a within-session EA transition maintained their differential performance. The results show in a convincing way that the negative to positive transition within a

session makes a very substantial difference to the degree of stimulus control that is developed from repeated exposure to positive and negative trials.

That the transition from negative to positive periods has a powerful influence on stimulus control was also demonstrated in an experiment by Rosen and Terrace (1975). This experiment in which pigeons were trained with certain wavelengths or line orientations on the key, was concerned with the minimally adequate conditions for obtaining inhibitory gradients, a selective reduction in responding in the vicinity of  $S_2^-$ , and peak shift. After extensive training on  $S_1^+$  some groups received only extinction on  $S_2^-$  (massed extinction) before the generalization test. Other groups received various treatments immediately following extinction on  $S_2^-$ . Among these treatments was a return to reinforced responding to  $S_1^+$ . In this case, the training was comparable to a single cycle of AEA training in the experiment by Ellis. When compared with simple conditioning, massed extinction produced some selective reduction in responding in the vicinity of  $S_2^-$  but it produced neither peak shift nor inhibitory gradients. The return to  $S_1^+$  at the end of extinction, even though the exposure was limited to three one-minute periods, resulted in a greater selective reduction at  $S_2$ , peak shift, and inhibitory gradients. Surprisingly, similar effects were obtained when the post extinction treatment was a return to  $S_1$  but without further reinforcement, or the free delivery of unsignalled food. Leaving aside this puzzling fact, the results, like those of Ellis, point to the within-session transition from nonreinforcement to reinforcement as especially

effective in the development of stimulus control.

The review of previous literature on the sequencing of positive and negative trials leads to the following summary.


When compared with simple conditioning, discrimination training with positive and negative trials interspersed is commonly found to increase stimulus control, to produce a shift in the point of maximal responding away from the negative stimulus (peak shift), and to elevate the level of responding at the peak.

Interspersed discrimination training may also make the negative stimulus inhibitory. Massed extinction on the negative stimulus (no return to positive trials prior to testing) is not equivalent to interspersed training although under some conditions it produces in an attenuated form certain of the effects produced by interspersed training. Honig et al. (1959) found no increase in control to result from massed extinction but increases have been found in some subsequent experiments. The effectiveness of massed extinction may be enhanced by prior training involving an unrelated negative stimulus (Friedman and Guttman, 1965) but that does not appear to be a necessary condition for obtaining an increase in control from massed extinction (Yarczower, 1974). Peak shift sometimes results from massed extinction (Yarczower, 1974) but it is less reliable and less marked than after interspersed discrimination training. Elevation of the rate of responding at the maximal point on the generalization gradient has not been found after massed extinction, which tends to lower responding at all points when compared with simple conditioning (Honig et al: 1959; Yarczower, 1974; Ellis, 1970). Under conditions

in which interspersed discrimination training produces U-shaped gradients, massed extinction may fail to do so (Weisman and Palmer, 1969).

The repetition of within session transitions from positive to negative trials does not increase stimulus control to the extent that interspersed training does. The go/no-go discrimination that develops when each session consists solely of a block of reinforced trials followed by a block of nonreinforced trials may be based largely on the cue value of nonreinforcement for subsequent nonreinforcement (Yarczower and Switalski, 1969; Ellis, 1970). The missing ingredient is the within-session transition from negative to positive trials which has been shown to be powerful in increasing control by the discriminative stimuli presented on the trial (Ellis, 1970; Rosen and Terrace, 1975).

Any comparison of the effects of interspersed training with those of massed extinction or of training with repeated alternation of blocks of trials, brings into play several possible contributing factors. There will be differences in the number and types of transition, in the lengths of runs of positive and negative trials, and in the amount of responding on negative trials. Complexity is reduced when the training sequences to be compared have the same ordering of positive and negative trials and differ only in how the sequence is segmented or broken up by the stopping and starting of sessions. The experiment by Ellis (1970) illustrates this strategy since, with minor exceptions, the various treatments differ only in terms of where the alternating sequence of positive and negative



trials is interrupted by the session breaks. Although this variable itself introduces more than one kind of difference it is analytically more manageable than altering the composition of the trial sequences. Further, segmentation is important: interrupting the sequence in a way that prevents within-session transitions from negative to positive trials greatly reduces the development of stimulus control when compared with interruptions that allow within-session transitions from positive to negative trials.

The present experiment examines the effect of interrupting the sequence of trials by the between-session break in such a way as to allow or to prevent within-session transitions from negative trials to positive trials. The course of discrimination training is compared when, in each session, a block of  $S_2^-$  trials occurs between two blocks of  $S_1^+$  or only after all  $S_1^+$  trials have been presented. One can consider the variable to be the location of the block of negative trials within the session or, equivalently, the location of the session break in the sequence of trials. The design of the experiment also provides information on how the development of the discrimination is affected by the number of  $S_2^-$  trials in the block. As suggested by Yarczower, the effects of continuous extinction might be stronger after prolonged exposure to  $S_2^-$ . Finally, the effect of the location and number of  $S_2^-$  trials per session was examined for a readily formed discrimination and one that forms more slowly.

An important objective of the present research is methodological. We wish to examine the possibility of obtaining a continuous assessment of the development of discrimination when the positive and negative



trials are presented in nonrandom sequences; specifically, in alternating blocks of positive and negative trials. A method that allows continuous measurement is likely to be more sensitive and more economical to use than the previous method of separate training and testing, but the method must overcome the difficulty that in nonrandom sequences the presence or absence of reinforcement predicts, although imperfectly, reinforcement or nonreinforcement on the next trial. Therefore, the development of a go/no-go discrimination cannot be unequivocally attributed to control by the trial stimuli.

The present experiment attempts to overcome this difficulty by the use of a feature-positive discrimination task (Jenkins and Sainsbury, 1969, 1970; Wasserman, 1974). In this task, a pair of displays sharing common features is distinguished by a distinctive feature on the positive display. For example, both displays may contain the common feature of dots while the distinctive feature, say a star, appears only on the positive display. In the course of learning a discrimination between such displays presented successively, the location of the peck within the positive display switches from the common feature, to which the pigeon was initially trained to respond, to the distinctive feature. Although reinforcement is not made to depend on the location of the peck, the response tracks the distinctive feature as it is moved from place to place on the display over a series of trials. The shift from the common to the distinctive feature typically occurs before responses to the negative display show a substantial decline. In other words, a simultaneous discrimination within the positive display precedes the go/no-go discrimination between the positive and negative displays. Although

the go/no-go discrimination can be based in part on regularities in the sequence of positive and negative trials, the simultaneous discrimination must be based solely on the display stimuli. The development of the simultaneous discrimination is therefore expected to provide a continuous measure of discriminative control by the distinctive feature under blocked sequences of positive and negative trials. In order to examine the relation between the simultaneous discrimination and control by the display stimuli in the go/no-go discrimination, the present experiment includes ~~certain~~ groups that were placed on random sequences of positive and negative trials after prolonged training on blocked sequences.

## Method

### Subjects

The subjects were 72 experimentally naive adult White King pigeons obtained from the Palmetto Pigeon Plant, Sumter, South Carolina. They were maintained at approximately 80-85% of their free feeding weight with Purina Chow Checkers.

### Apparatus and stimuli

Four standard pigeon chambers obtained from Lehigh Valley Electronics were used. The response key was a translucent square of plastic 3.5 cm on a side, divided into four square sectors. Each sector operated as a separate key. Black metal strips 2 mm wide separated the sectors and reduced the possibility that a single key peck would operate more than one key. A force of 15 gm operated the keys. The displays to be discriminated were back projected on the key by a Kodak Ektographic RA-960 projector with a 120-volt, 500-watt light bulb. It was run, however, at approximately 25 volts. Exposure was controlled by an external shutter. The displays (see Figure 2) consisted of a brightly illuminated geometric figure at the center of three of the four sectors. The background remained dark. The figures

---

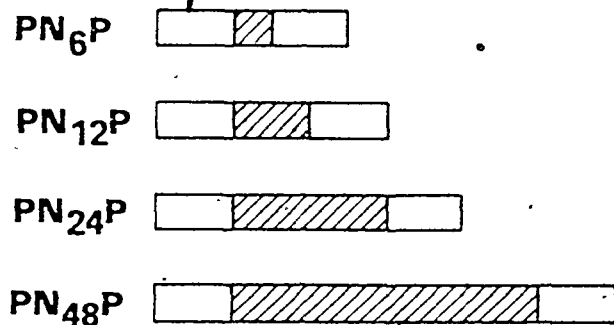
Insert Figure 2 about here

---

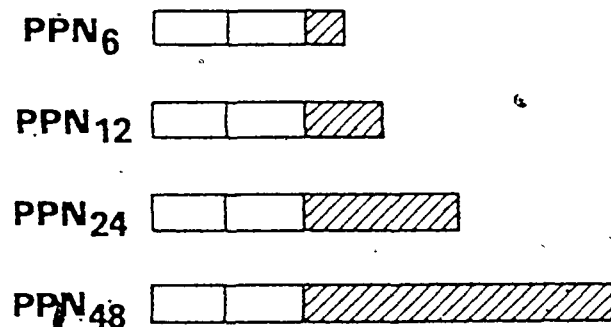
were a circle 0.5 cm in diameter or a five-pointed star of the same area. There were two types of displays: those consisting only of common features and those consisting of common features plus a distinctive feature. When the common feature was a circle, for example, the common-features-only display showed three circles.

## EXPERIMENTAL DESIGN

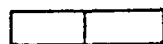
### PNP



### PPN



### CONTROL



## DISPLAY STIMULI

### DOT AS DF AND STARS AS CF



### STAR AS DF AND DOTS AS CF



FIGURE 2

The common-feature-plus-distinctive-feature display showed two circles and a star. The position of the distinctive feature and the position of the blank on the key varied randomly from trial to trial. A masking noise of approximately 80 dB was supplied to each chamber by a speaker mounted on the panel containing the response key. The working compartment was diffusely illuminated at all times by a number 1820 bulb housed in a reflector and mounted 5.5 cm above the upper edge of the key. A PDP-8E computer scheduled the presentation of the displays and the delivery of the reinforcer, and stored the number and location of responses on each trial.

#### Design

Essentials of the design are shown in Figure 2. Each session of training in experimental groups consisted of two blocks of 12 positive trials and a block of 6, 12, 24 or 48 negative trials. On positive trials the display consisted of common features plus the distinctive feature; on negative trials it consisted only of common features. The block of negative trials was placed between two blocks of positive trials (PNP: positive, negative, positive), or after two blocks of positive trials (PPN). In PPN sequences, the second block of positive trials followed the first without interruption. The dot was the distinctive feature (star common) for half the subjects; the roles of dot and star were interchanged for the remaining subjects. The combinations of two locations of the block of negative trials, four lengths of negative blocks, and the two representations of the distinctive feature, makes 16 experimental groups. Two control groups, one with the dot as the distinctive feature and one with the

star, received no negative trials. Four subjects were assigned at random to each of the 18 groups.

#### Procedure

Birds were trained to approach quickly and to eat from the food tray which contained Purina Chow Checkers. The tray opening was lighted whenever the tray was in the raised position. The birds were first allowed to eat from the tray which was heaped with grain and held in the raised position. The tray was then operated intermittently for durations that were reduced gradually to a final value of 5 sec. The time between reinforcers was varied. The mean time was gradually increased to the final value of 60 sec. One or two sessions were required to complete tray training. A discrete trial procedure was used for the initial training of the key peck and in all subsequent training. Trials were marked by the lighting of the display for a maximum of 8 sec. Intertrial intervals were variable with a mean of 64 sec and a range of 8 to 360 sec. The compartment was lighted throughout the session. Responses between trials had no effect. On positive trials, throughout the experiment, there was both a response-dependent and a response-independent reinforcement contingency. Positive trials were terminated and immediately reinforced by the completion of a response requirement or they were terminated and reinforced at the end of 8 sec, whichever came first. The response-independent contingency allows the initial acquisition to be accomplished through autoshaping. Negative trials remained on for 8 sec. regardless of responding and were terminated without reinforcement.

Five to six sessions of preliminary training were given.

Each session of preliminary training consisted of 36 reinforced trials on the common-feature display (later, during training, to be the negative display). A single response was required for early termination of the trial during the first two sessions. Beginning with the third session, and thereafter, four responses were required for early termination. After preliminary training, subjects with the dot as the distinctive feature received 10 sessions of training. Those with the star received 18 sessions of training followed by 6 training sessions in which 24 positive and 24 negative trials were presented in random sequence.

## Results

### Blocked sequences

By the last pretraining session almost all trials (mean of 99 percent overall) were terminated by the completion of four pecks before the maximum trial duration of 8 sec elapsed. Responses were predominantly directed to a feature rather than to a blank sector, on which only 2 percent of the total occurred. During the intertrial periods, when the entire key was dark, responses were very infrequent and they remained infrequent throughout the experiment.

Responses to distinctive feature. The percent of responses to the distinctive feature is shown by sessions of training in Figures 3 and 4. When the dot was the distinctive feature, (Figure 3) the

---

Insert Figures 3 and 4 about here

---

following trends appeared. In every experimental group responses came to be directed almost exclusively to the distinctive feature. The

MEAN PERCENT OF RESPONSES ON S+ TRIALS MADE TO THE DF (DOT)

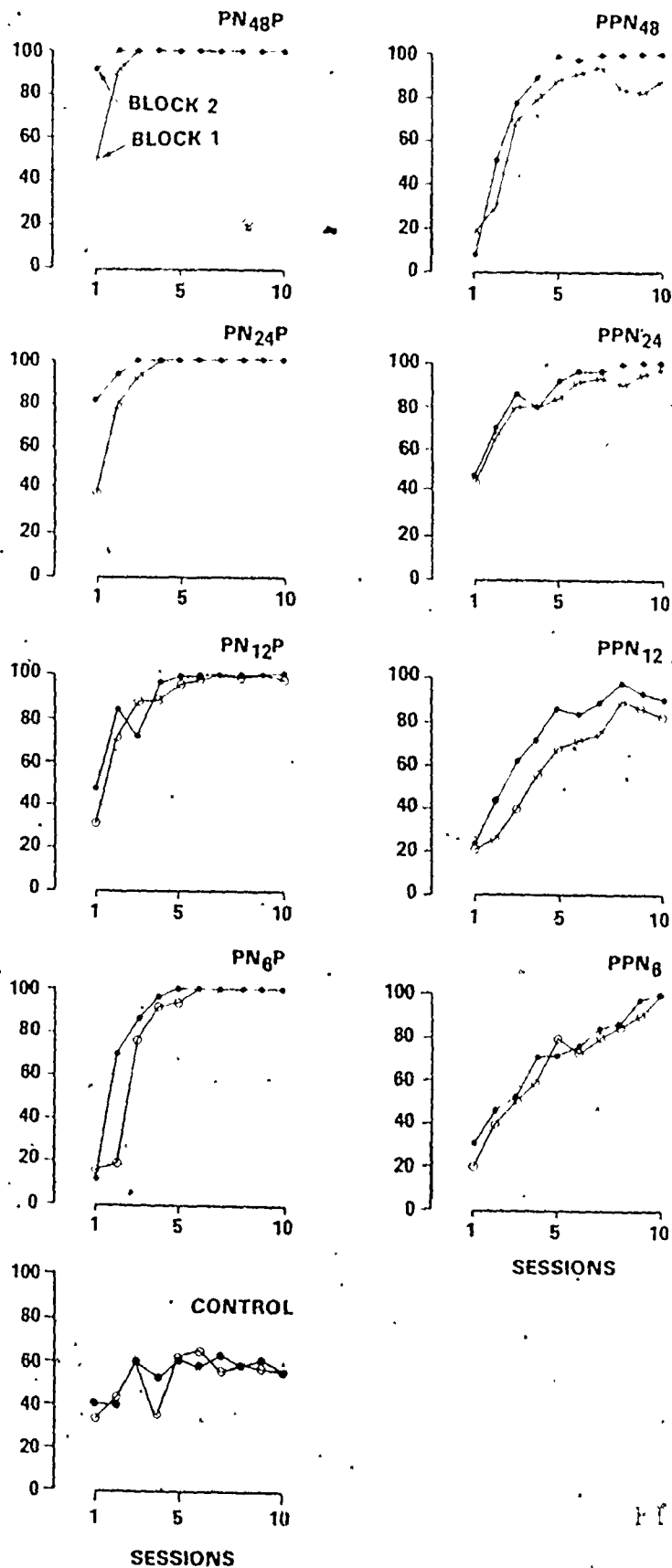


FIGURE 3



MEAN PERCENT OF RESPONSES ON S+ TRIALS MADE TO THE DF (STAR)

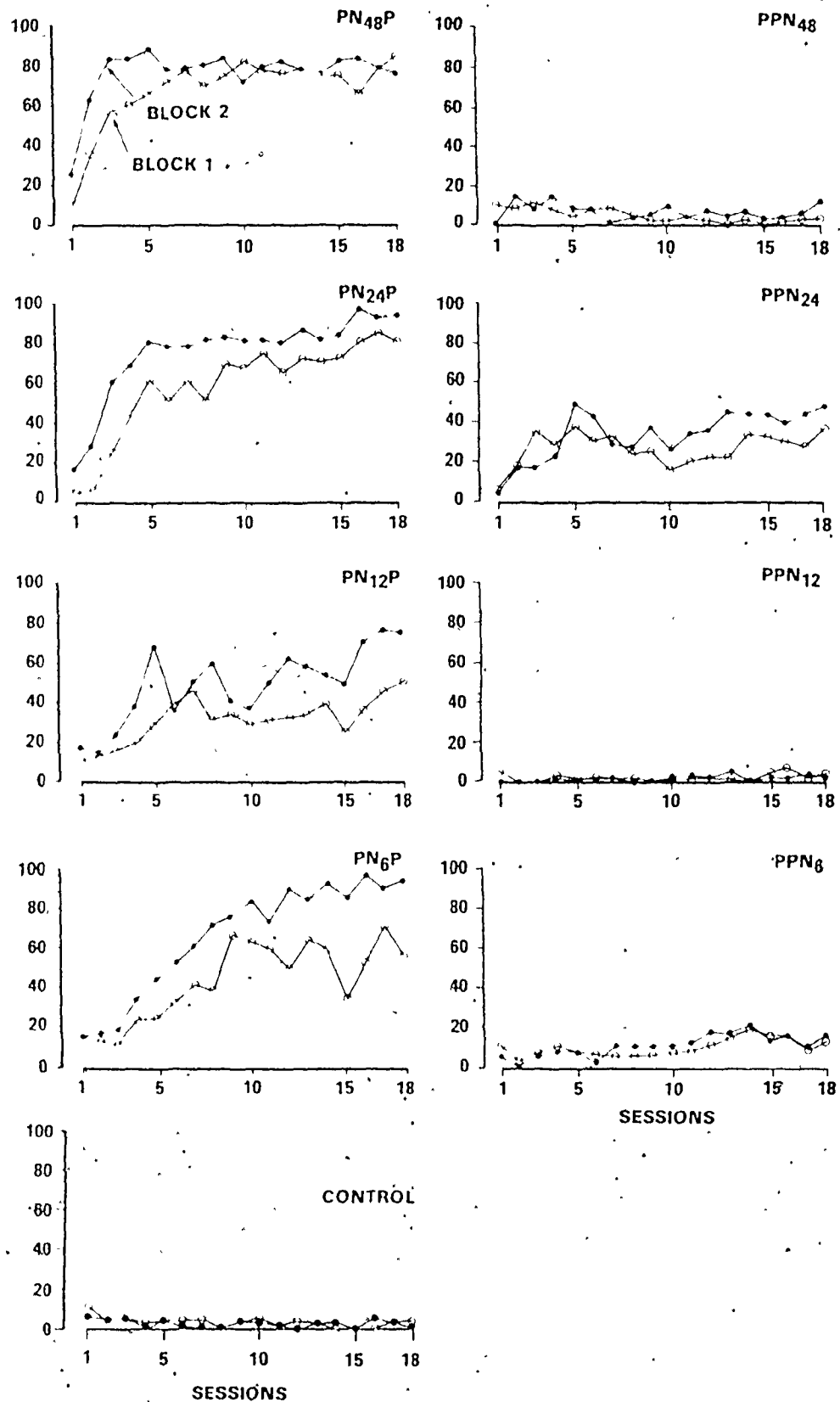


FIGURE 4

control group, which received only positive trials, also showed an increase with training in the percent of responses directed to the distinctive feature, but the increase was less than in any experimental group. Preference for the distinctive feature developed more rapidly, measured by sessions, the greater the number of negative trials in the block. The location of the negative block produced a clear effect. For each pair of groups with the same number of trials in the negative block, the group with the negative block between positive blocks showed a more rapid development of preference for the distinctive feature than did the group with the negative block at the end of the session. Finally, there was a tendency for the preference to develop more quickly in the second block of positive trials than in the first. This was especially marked early in training for groups which received a large block of negative trials between positive blocks; i.e. group  $PN_{48}P$  and group  $PN_{24}P$ .

When the distinctive feature was represented by the star (figure 4) a strong preference for the distinctive feature developed only under the most favourable conditions of training. The star was avoided in favor of the dot in the absence of any negative trials. This is shown by the very low percent of responses to the star in the initial block of positive trials and throughout training in the control group. The initial preference for the dot was converted to a very strong preference for the star only when the block of negative trials was long and when it was located between the positive blocks. It is particularly striking that the shift did not occur, or occurred only weakly, when the block of negative trials was presented after

all positive trials. With negative trials between two blocks of positive trials, preference for the distinctive feature developed more quickly in the second block of positive trials than in the first. Preference remained stronger in the second block throughout training except in the most favorable condition (48 negative trials) where the preference eventually became very strong, and equally strong, in the first and second positive blocks.

The mean percent of positive-trial responses that were directed to the distinctive feature, averaged over ten sessions, is given for each experimental group along with individual mean scores, in Table 1. A summary of the analysis of variance of these means is given in Table 2. Highly significant main effects were obtained for

---

Insert Tables 1 and 2 about here

---

the number of negative trials in the block, the location of the block and the representation (dot and star) of common and distinctive features. The significant interaction of location with feature reflects the greater effect of location when the distinctive feature was the star than when it was the dot.

Responses to negative trials. The mean number of responses in the entire block of negative trials is shown by sessions in Figure 5 (dot as distinctive feature) and Figure 6 (star). Consider first the results when the distinctive feature was the dot. A general decline

---

Insert Figures 5 and 6 about here

---

Table 1. Summary of the results of the analysis of variance for the data in Table 1.

Group	Subgroups			Mean	Group	Subgroups			Mean
	1	2	3			1	2	3	
$P_{1,8}P$	96.9	98.4	99.5	97.1	$PP_{1,8}$	60.1	61.3	61.5	60.5
$P_{2,2}P$	96.5	97.2	93.8	90.1	$PP_{2,2}$	92.3	70.2	80.4	82.4
$P_{3,1}P$	79.2	95.6	87.7	86.7	$PP_{3,1}$	82.7	81.2	82.7	81.9
$P_{4,5}P$	74.6	85.2	87.3	85.5	$PP_{4,5}$	28.5	11.0	92.3	60.4

Table 2. Summary of the results of the analysis of variance for the data in Table 2.

Group	Subgroups			Mean	Group	Subgroups			Mean
	1	2	3			1	2	3	
$P_{1,8}P$	80.9	40.2	72.6	70.4	$PP_{1,8}$	3.3	25.2	0	1.6
$P_{2,2}P$	20.5	61.8	71.6	63.7	$PP_{2,2}$	18.2	18.4	30.5	21.2
$P_{3,1}P$	41.4	30.5	20.5	36.5	$PP_{3,1}$	2.0	2.4	1.0	1.6
$P_{4,5}P$	75.5	29.2	18.6	37.5	$PP_{4,5}$	2.7	1.0	0	5.7

# TABLE 1

Mean Percent of Recruits or Adults per Female in the First Year of Life

Control	Control	Control	Control
1	1	1	1
3.2	3.2	3.2	3.2
Control Data as of	Control Data as of	Control Data as of	Control Data as of
5.0	5.0	5.0	5.0
Control Star as of	Control Star as of	Control Star as of	Control Star as of
1.0	1.0	1.0	1.0
2.0	2.0	2.0	2.0

Table 2

Analysis of Variance: Mean Percent of Responses to the Distinctive Feature

Source	DF	MS	F	P
Number (A)	3	121,635.0	5.15	.005
Location (B)	1	1,121,750.0	47.48	.001
Feature (C)	1	4,503,420.0	190.63	.001
A X B	3	32,416.0	1.37	
A X C	3	13,037.3	0.55	
B X C	1	191,076.0	8.09	.01
A X B X C	3	16,688.0	.71	
Error	48	23,624.3		

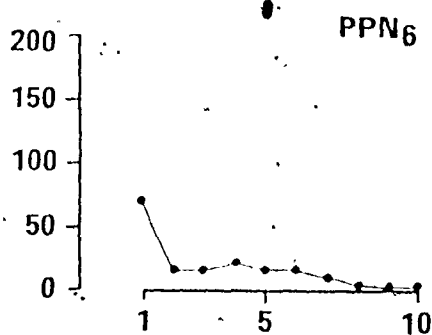
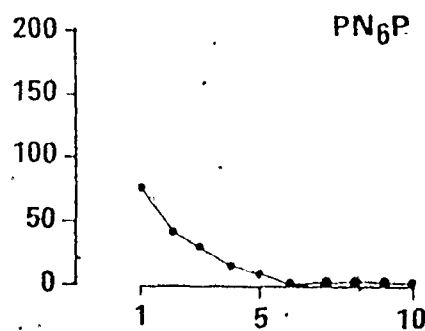
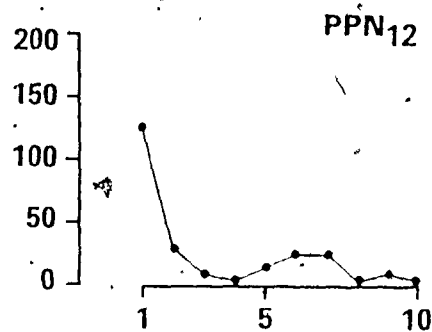
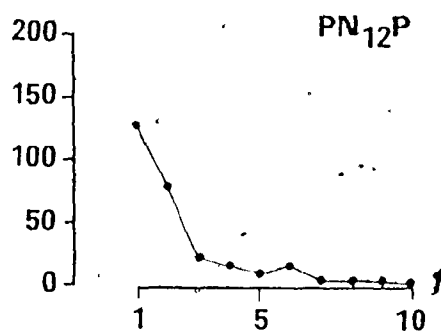
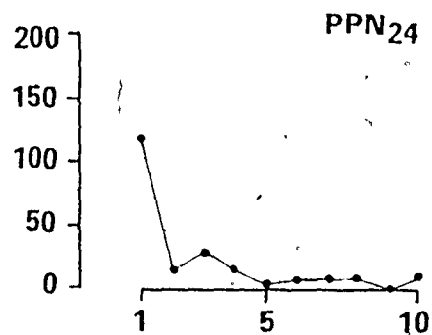
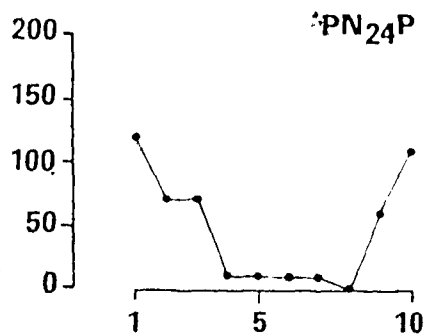
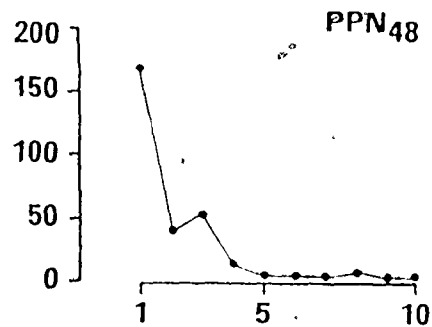
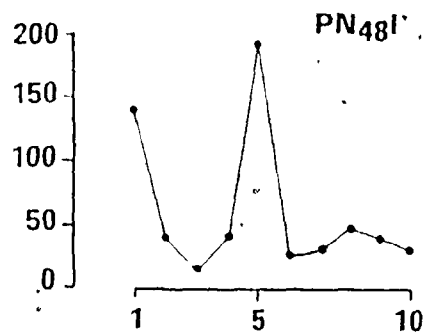
Number: number of negative trials presented each session (48, 24, 12 or 6).

Location: position of the session break: PNP or PPN.

Feature: shape of the distinctive feature, dot (star as the common feature) or star (dot as the common feature).

NOTE: An analysis of variance performed on the arc sine transformation of the data shows the same main effects and the same interaction.

MEAN NUMBER OF RESPONSES ON S- TRIALS



SESSIONS

SESSIONS

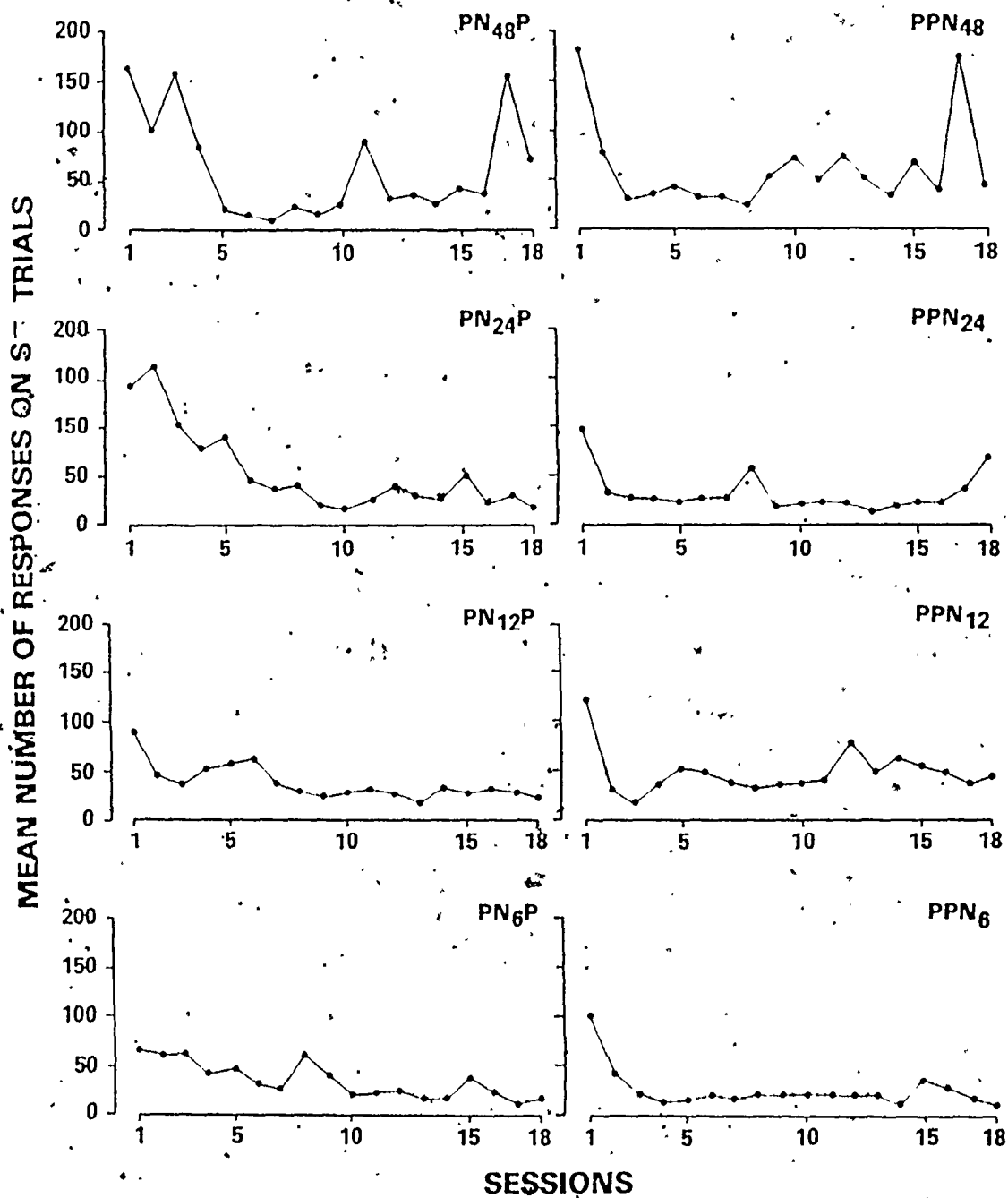


FIGURE 6



in negative-trial responses occurred over sessions although erratic, upward jumps in responding sometimes occurred even after very low levels had been reached. The decline in responding was more rapid when the negative block was presented after positive blocks than when it was presented between positive blocks.

The mean levels of responding on negative trials when the star was the distinctive feature (Figure 6) was generally higher than when the dot was the distinctive feature. Otherwise, the trends were similar; in particular, the decline was more rapid for negative blocks in the after position than in the between position.

These results show that a go/no-go discrimination is formed in all groups, but they do not, of course, establish that the discrimination is controlled by the same cues in different conditions of training.

The mean total number of negative responses over 10 sessions of training is shown, for each group, along with individual totals, in Table 3. The one bird in Group PPN<sub>48</sub> with an extraordinary total of

---

Insert Table 3 about here

---

almost 7,000 responses was not included in Figure 5 nor in statistical analyses of negative-trial responding. An analysis of variance of mean negative-trial responses is given in Table 4. Significant main effects

---

Insert Table 4 about here

---

of the number of negative trials in a block (the more trials per block the greater the total number of responses), of the location of the

TABLE 3

Mean and Standard Deviations of the Scores on the Test of the Ability to Read and Understand

Test of the Ability to Read and Understand

Group	Subjects			Mean	Group	Subjects			Mean
	1	2	3			1	2	3	
PN <sub>48</sub> P	18.7	37.8	128.5	56.5	60.1	49.1	16.0	37.9	24.1
PN <sub>24</sub> P	105.8	15.1	23.6	46.0	47.6	7.1	29.7	24.2	23.2
PN <sub>12</sub> P	12.4	42.4	36.3	22.0	28.3	26.1	19.1	36.1	23.6
PN <sub>6</sub> P	20.1	22.6	11.6	22.1	19.2	36.1	8.2	11.2	16.6

Star as Discriminative Feature

Group	Subjects			Mean	Group	Subjects			Mean
	1	2	3			1	2	3	
PN <sub>48</sub> P	60.8	104.8	49.3	34.7	62.1	52.1	29.5	62.2	25.1
PN <sub>24</sub> P	137.3	46.8	25.3	84.6	73.5	40.1	49.6	23.2	33.5
PN <sub>12</sub> P	43.2	72.3	17.6	48.4	45.4	61.8	28.9	36.8	43.0
PN <sub>6</sub> P	47.3	57.3	55.7	22.9	45.8	21.0	39.3	23.6	29.1

Table 4

Analysis of Variance: Mean Number of Responses to S- Trials

Source	DF	MF	F	P
Number (A)	3	183,519.0	2.89	.05
Location (B)	1	392,502.0	6.19	.02
Feature (C)	1	476,100.0	7.51	.01
A X B	3	65,610.0	1.03	
A X C	3	2,264.0	0.04	
B X C	1	690.0	0.01	
A X B X C	3	33,346.7	0.53	
Error	47	63.425.1		

NOTE: The extreme score in group PPN<sub>48</sub> was substituted by the group mean. The degrees of freedom corresponding to Error would be equal to  $abc(n-1) - 1$ .

block, and of the representation of the distinctive feature were found. The same analysis was carried out on the mean total number of negative trials on which at least one response occurred. The results, which are shown in Table 5, show the same pattern of significance but at higher levels.

---

Insert Table 5 about here

---

Shift in relation to negative-trial responses. The results already presented show that conditions favoring the shift to the distinctive feature tended to produce higher levels of responding on negative trials. This suggests the hypothesis that the shift was directly dependent on the amount of negative trial responding. The variables of location of the block of negative trials and the number of trials in the block would, on this hypothesis, exert their effect on the shift through their effect on the amount of negative trial responding. The hypothesis would be strongly supported if in the several groups the shift occurred after the same total number of negative trials on which at least one response had occurred. Such trials will be referred to as active trials. The upper part of Figure 7 shows the total number of negative trials (active plus inactive)

---

Insert Figure 7 about here

---

that were presented before the shift to the distinctive feature reached a criterion. The criterion was 80 percent or more of the responses in a block of 12 positive trials directed to the distinctive

Table 5

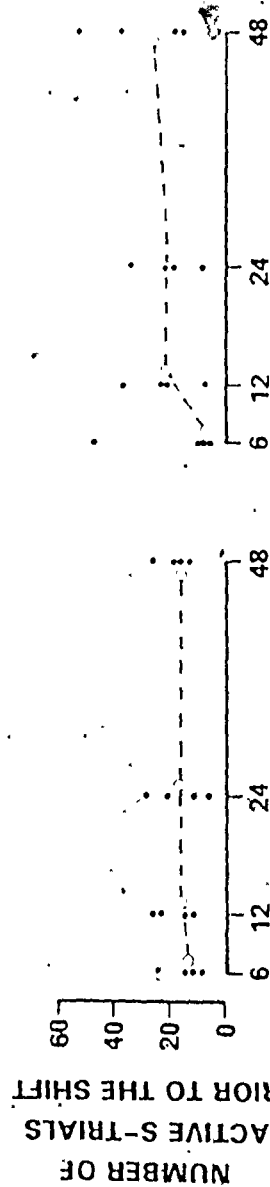
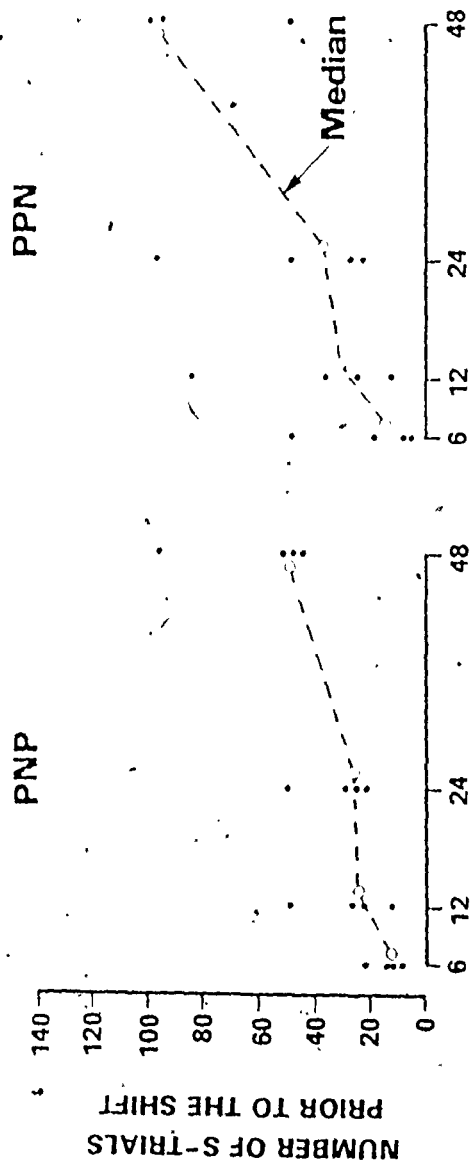
Analysis of Variance: Mean Number of Negative Trials with At least One Peck

Source	DF	MS	F	P
Number (A)	3	8,835.3	8.54	.001
Location (B)	1	7,119.2	6.88	.002
Feature (C)	1	11,209.5	10.84	.005
A X B	3	788.8	0.76	
A X C	3	184.1	0.18	
B X C	1	4.5	0.00	
A X B X C	3	746.1	0.72	
Error	47	1,034.1		

NOTE: The extreme score in group PPN<sub>48</sub> was substituted by the group mean. The degrees of freedom corresponding to Error would be equal to  $abc(n-1) - 1$ .

# DOT AS DF

(.192)



NUMBER OF S-TRIALS PRESENTED PER SESSION

feature. As the length of the negative block was increased, the number of negative trials required to produce the shift tended to increase. More negative trials were required, on the average, when they presented after positive trials than when they were presented between positive trials.

The number of active trials prior to the shift is shown in the lower half of Figure 7. The average number of active trials remained roughly constant across different lengths of negative trial blocks and across the two locations of the block. With the negative block between positive blocks the median number of active trials to the shift was 13.5; with the negative block after the positive block it was 19.5. In view of the variability evident in Figure 7, this relatively small difference would not be grounds for rejecting the hypothesis that the number of active trials required to produce the shift is constant.

The hypothesis can be rejected, however, on the results obtained when the distinctive feature was represented by the star. Of the 16 birds trained with the negative block after positive blocks only 2 met the shift criterion. Despite the failure to shift, the number of active negative trials in each of these groups eventually exceeded the number required to produce the shift in the comparable group with the negative block between positive blocks. The number of active negative trials is not the sole determiner of the shift and the hypothesis must therefore be rejected.

There is a much weaker hypothesis that might still prove acceptable. The length, although not the position, of the block

of negative trials might affect the shift entirely through the affect of length on the number of active trials per session. Figure 8 shows further data on this hypothesis from the groups trained with the star as the distinctive feature and the negative block located between positive blocks. The trends are roughly similar to those obtained

---

) Insert Figure 8 about here

---

with the dot as the distinctive feature. The results presented in the left panel show that more trials are required to produce the shift when the number of trials per block was 12 or more than when it was 6. The results in the right panel show that when active trials only are counted, the effect of block length was reduced and might be nil. Although the data were too variable to provide strong support, the hypothesis remains tenable that the length of the block of negative trials affects the rapidity of the shift entirely through its effect on the number of active negative trials per session.

Random sequence

Response to distinctive feature. The upper half of Figure 9

---

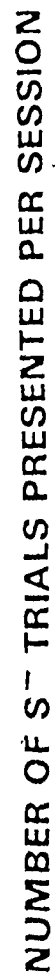
Insert Figure 9 about here

---

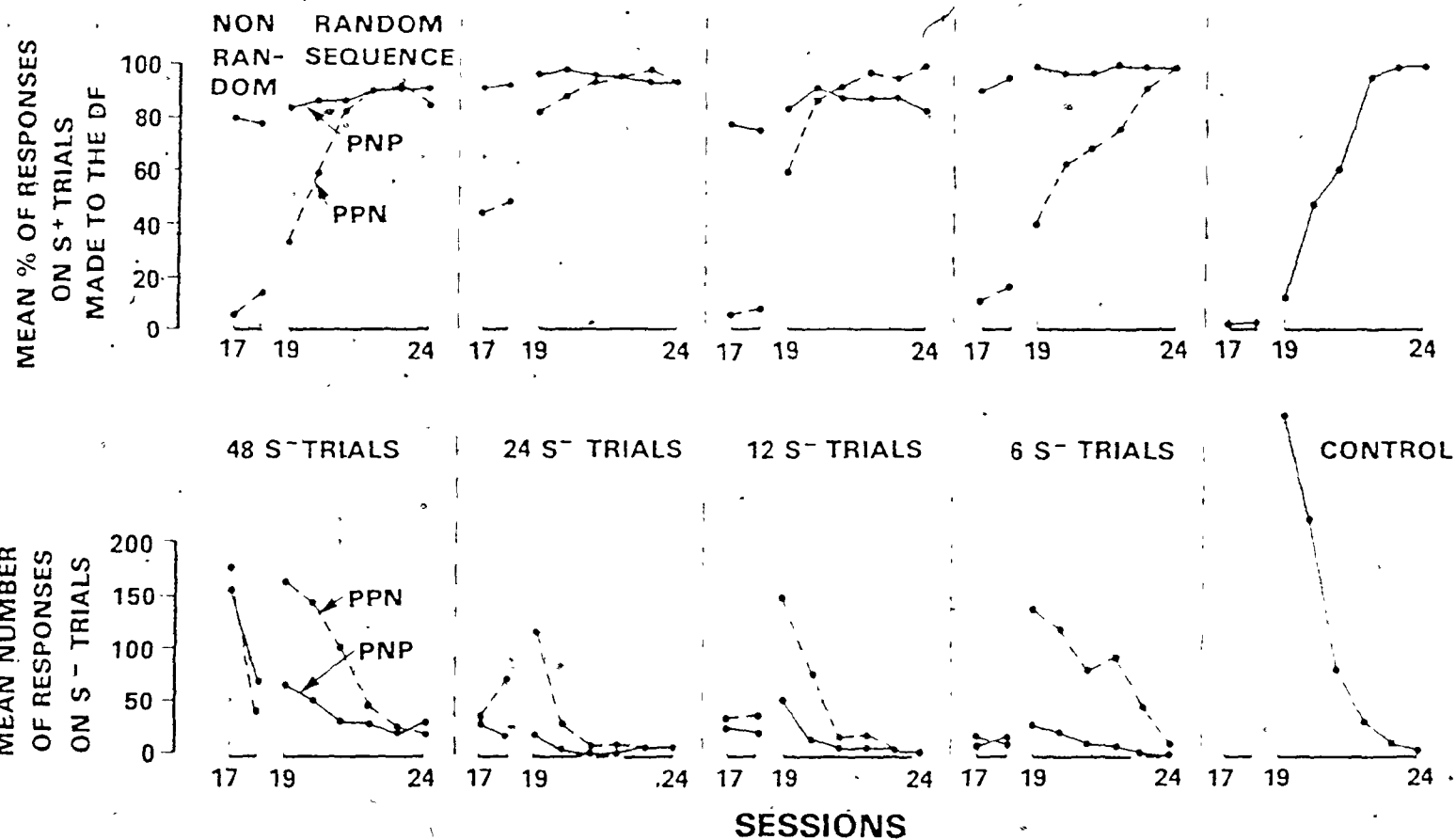
shows the percentage of responses to the distinctive feature in the last two sessions of training under the blocked sequence (sessions 17 and 18) and in the last 6 sessions of training with the random sequence (sessions 19-24). Groups that had been trained with the negative block of trials between positive blocks maintained their



(.271)



# TRANSFER TO A RANDOM SEQUENCE (STAR AS DF)



well developed preference for the distinctive feature. Those trained with the negative trial block after positive blocks entered the random sequence with a much lower percentage of distinctive feature responses. The random sequence brought about a rapid development of preference for the distinctive feature. The control group was exposed to negative trials for the first time in the random sequence. It also rapidly acquired a strong preference for the distinctive feature. Comparison of the results for the control group with the average results for all of the experimental groups previously trained with the negative block after positive blocks suggests that exposure to negative trials after positive trials contributed to a preference for the distinctive feature. A greater number of distinctive feature responses were made during the 6 sessions of random training in these experimental groups than in the control group. The percent of responses to the distinctive feature in the first session on the random sequence is shown in Table 6. An analysis of variance of the mean number of responses to the

---

Insert Table 6 about here

---

distinctive feature in the first session for experimental groups showed the location of the negative-trial block in previous training to be significant:  $F(1,24) = 23.11, p < .001$ . The length of the negative-trial block was not significant, nor was its interaction with the position of the block. The mean number of distinctive feature responses in the control group was significantly less than the mean number for the experimental groups that were previously trained with the negative

Table 6

Percent of Responses on S+ Trials that were directed to the  
Distinctive Feature in the First Session of Exposure to a  
Random Sequence of S+ and S- Trials

Group	Subjects				Mean
	1	2	3	4	
PN <sub>48</sub> P	100	40	100	94	83.5
PN <sub>24</sub> P	93	98	97	100	97.0
PN <sub>12</sub> P	93	80	87	73	83.3
PN <sub>6</sub> P	100	100	96	100	99.0
PPN <sub>48</sub>	1	50	52	29	33.0
PPN <sub>24</sub>	66	89	82	92	82.3
PPN <sub>12</sub>	78	75	25	58	59.0
PPN <sub>6</sub>	87	58	0	14	39.8
Control	22	0	12	19	13.3

block after positive trials ( $t(18) = 2.51$ ,  $p = .025$ ) as it obviously was for the experimental groups pretrained with the negative block between positive trials.

The lower half of Figure 9 shows responses on negative trials. When placed on the random sequence, subjects previously trained with the negative block between positive blocks maintained their go/no-go discrimination with little or no disruption. In contrast, those trained with the negative-trial block after positive trials showed a large increase in responses to negative trials on the first session, followed by a rapid reduction. The reduction in responses to negative trials paralleled the shift to the distinctive feature within the positive trials. The control group showed by far the highest initial level of responding on negative trials. Table 7 gives the data for responses

---

Insert Table 7 about here

---

on negative trials in the first session of training on the random sequence. An analysis of variance of the means for the experimental groups showed that only the position of the negative-trial block in previous training was significant:  $F(1,24) = 11.35$ ,  $p = .005$ . The mean number of negative trial responses in the control group was obviously significantly greater than in the experimental groups trained with the negative block between positive blocks. It was also significantly greater than the mean for all the experimental groups trained with the negative block after positive blocks ( $t(18) = 3.52$ ,  $p < .005$ ).

Table 7

Number of Responses on S-Trials in the First Session of Exposure  
to a Random Sequence of S+ and S- Trials

Group	Subjects				Mean
	1	2	3	4	
PN <sub>48</sub> P	13	227	0	21	65.3
PN <sub>24</sub> P	35	6	2	7	12.5
PN <sub>12</sub> P	37	100	37	18	48.0
PN <sub>6</sub> P	21	56	34	1	28.0
PPN <sub>48</sub>	93	52	124	397	166.5
PPN <sub>24</sub>	236	141	44	38	114.8
PPN <sub>12</sub>	86	103	171	227	146.8
PPN <sub>6</sub>	15	200	208	120	135.8
Control	313	369	290	301	318.3

## Discussion

The principal results may be summarized as follows. The shift to the distinctive feature was affected by the stimulus characteristics of the distinctive and common features, by the location of a block of negative trials within the session, and by the number of negative trials in the block. A much more rapid shift to the distinctive feature was produced by training under the combination of dot as distinctive feature and star as common feature than under the reverse combination. The dot, probably because of its more rounded shape, was preferred to the star as a target of pecks. A block of negative trials placed between two blocks of positive trials was more effective in producing the shift than was the same block placed after the blocks of positive trials. The shift to the less preferred star as a distinctive feature almost always failed to occur when the block of negative trials came after the positive blocks whereas it occurred regularly when the block of negative trials came between the block of positive trials. The shift to the more preferred dot as a distinctive feature was brought about by a negative block in either position but the shift occurred more rapidly when the negative block came between the positive blocks than when it came after the positive blocks. The more trials in the negative block the more rapidly, measured by sessions, the shift occurred. However, not even the longest block, which was 48 trials, was capable of producing the shift to the less preferred star as a distinctive feature when the block occurred after all positive trials.

During training with blocked sequences all groups showed a

progressive reduction of responding on negative trials. The negative block in the between location generated a larger total number of negative responses than did the negative block in the after location. However, the greater effectiveness of the PNP sequence in producing the shift cannot be solely due to the greater total number of negative responses which it generates. The shift to the star failed to occur in the PPN groups even though the total number of negative responses eventually exceeded the total reached when a clear shift did occur in subjects that received the negative trials in the between position. We may conclude that negative-trial responses in the PNP sequence contribute more to the shift than do negative trial responses in the PPN sequence.

Preference for the distinctive feature within the positive display during the exposure to block sequences was closely related to performance on the go/no-go discrimination when subjects were later exposed to the random sequence. If the response had shifted to the distinctive feature during blocked sequences the go/no-go performance was maintained on the random sequence. If the shift had not occurred the go/no-go performance was disrupted by a sharp increase in responding on negative trials. However, even when the negative trials failed to produce a shift during training with the blocked sequence they did facilitate the subsequent development of the shift under the random sequence and the eventual elimination of negative trial responses.

The present results on the effect of trial sequence on the development of stimulus control agree with the general trend of previously reported results. We find, as others have found with



different procedures; that when a session of training consists of a block of positive trials followed by a block of negative trials some increase in stimulus control is brought about by the negative block. In agreement with Ellis (1970), we find a much greater increase in stimulus control when the negative block is placed between positive blocks so that a transition from negative to positive trials occurs within a session.

The sequencing of positive and negative trials has closely related effects on the shift to the distinctive feature and on control by the display stimuli over the go/no-go performance. That makes the shift especially useful for studying the development of control by display stimuli when the trial sequence is not random and, as a consequence provides a potential source of stimulus control over the go/no-go performance. There is, however, a limitation to the use of the shift as a measure of control by display stimuli. The absence of a shift does not mean that the display stimuli have acquired no control. Subjects that failed to shift when exposed to a block of negative trials after all the positive trials, were quicker than controls to show stimulus control under random training; they were quicker to shift and to stop responding on negative trials. Something was learned as the result of exposure to negative trials even though the learning was insufficient to produce a shift.

A theory for the special effectiveness of a within session transition from negative to positive trials has not been offered in previous work. We turn now to the bearing of the present results on theory. It is important to recognize that, with unimportant exceptions

in the first and last session of training, the sequence of positive and negative trials was exactly the same in PPN and PNP groups.

The groups differed in where the sequence was interrupted by the between-session break. A theory of the effect must therefore explain why the location of the session break in the training sequence is important. Four consequences of the location of the session break are identified and their possible bearing on performance is considered.

1. In PPN groups the session start signals the resumption of positive trials after negative trials. In PPN groups the resumption of positive trials after negative trials is always marked by the start of a new session. To predict the resumption of reinforcement it is only necessary to notice that a session has begun and the bright figures are on the key. The start of the session, considered to be a stimulus, makes the distinctive feature redundant as a signal of reinforcement. It could, therefore, interfere with the acquisition of control by the distinctive feature. In contrast, to predict the resumption of reinforcement in PNP groups it is necessary to notice that a distinctive feature now appears on the display. Only the distinctive feature marks the beginning of the second block of positive trials. The theory formulated by Wagner and Rescorla (Rescorla and Wagner, 1972; Wagner and Rescorla, 1972) for the development of signal value in classical conditioning with multiple signal sources is capable of generating the needed prediction that a stimulus which makes another one redundant will interfere. The acquisition of signal value by the second stimulus is blocked by a stimulus that has already acquired signal value for the same reinforcing event.

Because the session start has, from pretraining onwards, signalled the beginning of reinforced trials, it could serve to block acquisition by the distinctive feature which is not introduced until the training sessions.

The more rapid development of stimulus control when the dot, as compared with the star, served as distinctive feature would be represented in the Wagner-Rescorla model by assigning the dot a higher salience. The theory would then generate the required prediction that the session start as a signal exerts a stronger blocking effect on the less salient star than on the more salient dot (1).

The principal trends in the present experiment could be encompassed by the Wagner-Rescorla theory but there is some question about the appropriateness of the coordination of the experimental events to the theory. The theory is usually applied when the blocking stimulus is concurrent with the blocked stimulus and that was not the case here. Neely and Wagner (1974) have, however, treated schedule of reinforcement as a stimulus in an application of the Wagner-Rescorla theory to blocking. Their extension of the concept

---

(1) An overshadowing account also compatible with the Rescorla-Wagner formulation is possible. In nonrandom sequences the regularities of reinforcement and nonreinforcement generate cues that can be conditioned and can control responding as efficiently as the display stimuli. It is possible then, that the stimuli produced in a nonrandom sequence of  $S_1$  and  $S_2$  trials compete with the display stimuli for the control of responding. Although in the context of the present experiment the blocking account and the overshadowing account predict roughly similar behavioural outcomes, the two alternatives can be confronted. Consider a single alternation arrangement. The session start is no longer a reliable predictor for reinitiation of reinforcement so blocking should be attenuated. According to the overshadowing hypothesis however, the pattern of reinforcement and nonreinforcement is still regular and predictable and would interfere with conditioning to the display cue.

of stimulus is similar to the one proposed here. One might also question the appropriateness of attributing the effect of interchanging the representation of common and distinctive features to the greater salience of the dot. Salience is commonly related to discriminability governed by the contrast between a stimulus and the background. There is no reason to think the dot more visible than the star. Rather, the dot is more attractive as a target for the peck than is the star (Fantz, 1957). The relevant attribute is preference rather than salience, but the Wagner-Rescorla theory has no place for preference because its experimental base is classical conditioning with diffuse, nonlocalizable stimuli.

2. In the PPN groups the session break leads to more rapid extinction of responses on the negative trials than in PNP groups.

In experiments on repeated acquisition and extinction the time between extinction trials and the resumption of reinforcement has been shown to influence resistance to extinction. A short wait between nonreinforced trials and the resumption of reinforcement generates greater resistance to extinction than does a long wait. According to Capaldi (Capaldi, Leonard and Ksir, 1968; Leonard and Capaldi, 1971) extinction rate is directly related to the difference between two intertrial intervals (ITI): a) the ITI between the last nonreinforced trial and the next reinforced trial (NP-ITI) and b) the ITI between two successive nonreinforced trials (NN-ITI). Capaldi assumes that the length of the reinforced ITI, the NP-ITI, acquires associative strength which generalizes to the length of the NN-ITIs. Resistance to extinction develops when the NP-ITI length is similar

to the NN-ITI lengths, as in PNP sequences, but not when there is a great discrepancy between two ITI lengths, as in PPN sequences. It remains to argue that greater resistance to extinction in PNP groups is favorable to the development of stimulus control by the distinctive feature. The argument might be that the bird is more effectively exposed to the common features on trials when it is responding to the display. We cannot entirely discount the possibility that this accounts for the greater effectiveness of the PNP arrangement but one finding militates against it. We have seen that in the case where the shift does not develop (PPN groups with star as distinctive feature) the total number of negative trials with responses eventually exceeds the number required to produce a shift in PNP groups. It is still possible that other parameters of negative trial responding (number per block or distribution over sessions) are controlling the shift but then the relation becomes more complex and less attractive as an explanation of the superiority of the PNP arrangement.

3. Because of the location of the session break the memory of the last N-trial when the first P-trial occurs will be older in PPN groups than in the PNP groups. A plausible hypothesis is that the process of discrimination requires a comparison of the contents of the reinforced and nonreinforced displays. Since the presentations are successive, the comparison must involve memory. The temporal conditions for the comparison of a previously reinforced display

with current nonreinforced display are the same in different groups since that comparison is based on the PN transition, but the temporal conditions for the comparison of a previously nonreinforced display with a current reinforced display are not the same in all groups. This comparison would be based on the NP transition and it would be favored only in the PNP groups. There are no well developed accounts of animal discrimination learning based on a comparison process involving memory so the matter is left with the mere recognition that such a process might explain the superiority of the PNP arrangement for the development of stimulus control.

4. Because of the location of the session break, the tendency to respond to the common features during the second block of P-trials will be stronger in the PPN-groups than in the PNP groups. The tendency to respond to the common feature is reduced by the N-block but it undoubtedly recovers from the end of one session to the next. The recovery could be due to the passage of time or to the start of a session as a signal for positive trials, or both. In any case, there will be a strong tendency to respond to the common features when positive trials begin and in the PPN groups reinforcement will keep it strong throughout both blocks of positive trials. In the PNP groups the situation during the second block of positive trials is different. Because the second block directly follows N-trials the tendency to respond to the common features will be lower than it is in the PPN groups. If it is further assumed that the distinctive feature gains more excitatory strength the weaker the common feature, then one can see why the PNP arrangement is more favorable to the

shift than the PPN arrangement.

There are reasons to believe this further assumption. For one, if the excitatory value of the common feature is lower, the total excitatory value of the positive trials will be reduced. The Wagner-Rescorla theory provides the required prediction that the lower the total excitatory value of the display as a whole, the greater the gain for the distinctive feature on a positive trial. Wagner, Saavedra and Leherman have demonstrated an effect of this kind in recent experiments on classical conditioning to compounds (Wagner, 1971b). A consideration more specific to the type of display used here leads to the same conclusion. Spatial separation allows the response to be directed at one feature. The lower the strength of the common features, the more likely it is that the response will be directed to the distinctive feature. The effect of the reinforcer might be channeled to the feature that captured the response, an assumption commonly made in simultaneous discrimination learning. The so-called choice scheme described by Jenkins and Sainsbury (1969) for feature positive and feature negative discriminations was based on this assumption of selective channeling. Although Jenkins (1973) later showed that the effect of a reinforcer is not exclusively channeled to the feature that was the target of the response, it remains plausible that the effect is concentrated on the target feature.

For present purposes it is not necessary to be specific about the channeling process. It is only necessary to assume that the greater the excitatory value of the common features the more effectively it

blocks acquisition of excitatory value by the distinctive feature. This account, centering on the common features as a blocking stimulus, is capable of integrating the effects of the PPN and PNP arrangements on the development of control by ~~the~~ distinctive feature. There is direct evidence that a recovery in excitatory value of the common features occur between sessions. We have noted that in PNP groups the common feature is responded to less often in the last block of P-trials, which followed N-trials directly, than in the first block of P-trials, which comes after the session break. The account also identifies the conditions that should favor the appearance of a shift in the very first block of positive trials following a single exposure to the negative block. The shift would of course be favored by a distinctive feature with a preferred shape (dot vs. star). It would be favored by longer series of negative trials in the block followed by an immediate presentation of positive trials, before common features recover strength. Results bearing on these expectations are presented in Table 8 which shows for each experimental group the mean percent

---

Insert Table 8 about here.

---

of distinctive feature responses in the first positive block of 12 trials after the first negative block. Although the results are not smooth, the expected trends are clear. An analysis of variance of the mean percents showed significant effects of the shape of the distinctive feature, ( $p < .001$ ), the number of trials in the negative block, ( $p < .005$ ), and the location of the session break, ( $p < .05$ ).



Table 8

Mean Percent Responses to Distinctive Feature on First Block  
of Positive Trials Following First Block of Negative Trials

Dot as Distinctive Feature			
Group	%	Group	%
PN <sub>48</sub> P	93.8	PPN <sub>48</sub>	30.3
PN <sub>24</sub> P	82.8	PPN <sub>24</sub>	66.8
PN <sub>12</sub> P	47.0	PPN <sub>12</sub>	26.0
PN <sub>6</sub> P	11.3	PPN <sub>6</sub>	41.0

Star as Distinctive Feature			
Group	%	Group	%
PN <sub>48</sub> P	24.0	PPN <sub>48</sub>	8.7
PN <sub>24</sub> P	14.5	PPN <sub>24</sub>	19.3
PN <sub>12</sub> P	15.5	PPN <sub>12</sub>	0.0
PN <sub>6</sub> P	14.5	PPN <sub>6</sub>	3.3

These factors would be expected to exert their effect throughout training. More than that, selective channelling of the reinforcer implies a positive feedback process. If the distinctive feature gains more, the common features gain less. On the next exposure to positive trials the common features are even lower in value, the distinctive feature gains even more, and so the process builds on itself. The other side of the coin is that selectivity makes possible a persistent failure to shift to the distinctive feature despite continued training. Unless some of the reinforcing effect is channeled to the distinctive feature, and this gain is carried over sessions, the process of building the excitatory value of the distinctive feature at the expense of the common feature will not take place.

Under the least favorable training conditions the shift failed to develop despite repeated training in the present experiment. Perhaps it would occur eventually, but even if the shift failed indefinitely this would not be hard to understand taken alone. It would be puzzling, however, when taken together with the observation that substantial savings were apparent on transfer to a random sequence. The savings suggest that something was learned about the status of the display features. If the learning process begins, why doesn't it build with training until the shift is complete? One possibility is that the distinctive feature is completely blocked and the transfer is due to a change in the status of the common features brought about by repeated conditioning and extinction. Yet blocked stimuli have also resulted in savings in a different

preparation (CER). When following compound training a blocked stimulus was reinforced alone, conditioning developed faster than in another group for which the same stimulus was a novel one (Kamin, 1968). Too little is known about what is responsible for the savings to warrant an extended discussion. The matter is worth noting, however, because persistent intermediate states of discrimination learning pose a serious problem for current conceptions of the discrimination learning process.

We have completed our discussion of how the location of the session break might have its effects. The fourth account, which has the location of the break affecting the strength of the common feature and the strength of the common features determining the extent to which it blocks the distinctive feature, is the most attractive one. This account provides an understanding of how the experimental variables affect the shift to the distinctive feature in the very first block of positive trials following a single exposure to negative trials. Leaving aside the least developed account in terms of memory loss, the remaining accounts predict effects only after repeated sessions training allow some aspect of the training regime to operate as a competing signal for reinforcement or nonreinforcement.

If the account centered on the common features is correct the within session transition from negative to positive trials is not necessary for effective discrimination training. It is only one of

several ways to keep the strength of the common features at a relatively low level during positive trials. It will be of interest to learn whether other ways of manipulating the strength of the common

features will produce the same results.

## REFERENCES

- Capaldi, E.J. Memory and learning: a sequential viewpoint. In W.K. Honig and P.H.R. James (Eds.), Animal Memory. New York, Academic Press, 1971. Pp. 111-154.
- Capaldi, E.J., Leonard, D.W. and Ksir, C. A re-examination of extinction rate in successive acquisitions and extinctions. Journal of Comparative and Physiological Psychology, 1968, 66, 128-132.
- Ellis, W.R. Role of stimulus sequences in stimulus discrimination and stimulus generalization. Journal of Experimental Psychology, 1970, 83, 155-163.
- Fantz, R.L. Formpreferences in newly hatched chicks. Journal of Comparative and Physiological Psychology, 1975, 50, 422-450.
- Friedman, H. and Guttman. Further analysis of the various effects of discrimination training on stimulus generalization gradients. In D. Mostofsky (Ed.), Stimulus Generalization. Stanford, California: Stanford University Press, 1965. Pp. 255-267.
- Hearst, E. Some persistent problems in the analysis of conditioned inhibition. In R.A. Boakes and M.S. Halliday (Eds.), Inhibition and Learning. London: Academic Press, 1972. Pp. 5-39.
- Honig, W.K., Thomas, D.R. and Guttman, N. Differential effects of continuous extinction and discrimination training on the generalization gradient. Journal of Experimental Psychology, 1959, 58, 145-152.

Kamin, L.J. 'Attention-like' processes in classical conditioning.

In M.R. Jones (Ed.), Miami Symposium on the Prediction of Behavior: Aversive Stimulation. Miami: University of Miami Press, 1968. Pp. 9-33.

Leonard, D.W. and Capaldi, E.J. Successive acquisitions and extinctions in the rat as a function of number of nonreward in each extinction session. Journal of Comparative and Physiological Psychology, 1971, 74, 102-107.

Jenkins, H.M. Measurement of stimulus control during discrimination operant conditioning. Psychological Bulletin, 1965, 5, 365-376.

Jenkins, H.M. Noticing and responding in a discrimination based on a distinguishing element. Learning and Motivation, 1973, 4, 115-137.

Jenkins, H.M. and Harrison, R.H. Generalization gradients of inhibition following auditory discrimination training. Journal of Experimental Analysis of Behavior, 1962, 5, 435-441.

Jenkins, H.M. and Sainsbury, R.S. The development of stimulus control through differential reinforcement. In N.J. Mackintosh and W.K. Honig (Eds.), Fundamental Issues in Associative Learning. Halifax: Dalhousie University Press, 1969. Pp. 123-161.

Jenkins, H., and Sainsbury, R.S. Discrimination learning with the distinctive feature on positive or negative trials. In D. Mostofsky (Ed.), Attention: Contemporary Theory and Analysis. New York: Appleton-Century-Crofts, 1970. Pp. 239-273.

Neely, J.H. and Wagner, A.R. Attenuation of blocking with shifts in reward: The involvement of schedule generated contextual cues. Journal of Experimental Psychology, 1974, 102, 751-763.

Rescorla, R.A. and Wagner, A.R. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black and W.F. Prokasy (Eds.), Classical Conditioning II. New York: Appleton-Century-Crofts, 1972. Pp. 64-99.

Rosen, A.P. and Terrace, H.S. On the minimal conditions for the development of peak shift and inhibitory stimulus control. Journal of Experimental Analysis of Behavior, 1975, 23, 385-414.

Verplank, W.S. The development of discrimination in a simple locomotor habit. Journal of Experimental Psychology, 1942, 31, 441-464.

Wagner, A.R. Elementary Associations. In H.H. Kendler and J.T. Spence (Eds.), Essays in Neobehaviorisms: A Memorial Volume to Kenneth W. Spence. New York: Appleton-Century-Crofts, 1971. Pp. 187-213.

Wagner, A.R. and Rescorla, R.A. Inhibition in Pavlovian conditioning: application of a theory. In R.A. Boakes and M.S. Halliday (Eds.), Inhibition and Learning. London: Academic Press, 1972. Pp. 301-336.

Wasserman, E.A. Stimulus-reinforcer predictiveness and selective discrimination learning in pigeons. Journal of Experimental Psychology, 1974, 102, 284-297.

Weisman, R.G. and Palmer, J.A. Factors influencing inhibitory stimulus control: discrimination training and prior non-differential reinforcement. Journal of Experimental Analysis of Behavior, 1969, 12, 229-237.

Yarczower, M. Discriminative effects of massed extinction. Journal of Experimental Analysis of Behavior, 1974, 22, 161-168.

Yarczower, M. and Switalsky, R. Stimulus control in the goldfish after massed extinction. Journal of Experimental Analysis Behavior, 1969, 12, 565-570.