# Enhancement of pigeons' conditional discrimination performance by expectancies of reinforcement and nonreinforcement

GAIL B. PETERSON
*Univeristy of Minnesota, Minneapolis, Minnesota 55455*

RUTH L. WHEELER
*Grinnell College, Grinnell, Iowa 50112*

and

MILTON A. TRAPOLD
*Memphis State University, Memphis, Tennessee 38152*

Prior work has shown that when the separate correct responses of a conditional discrimination are followed by different reinforcing outcomes, performance is enhanced relative to that obtained under the conventional, single-reinforcer procedure. Four experiments with pigeons yielded the analogous finding when the different outcomes were reinforcement and explicit nonreinforcement. Controls indicated that the results could not be attributed to the effects of intermittent reinforcement, to possible differences in cue duration, or to a variety of potential sources of conditioned reinforcement. An interpretation in terms of expectancy learning is proposed.

Trapold (1970) originally demonstrated what we will refer to in this paper as the *differential outcomes effect* (DOE). He found that rats learn a conditional discrimination (i.e., if stimulus $S_1$, then response $R_1$ is correct, but if $S_2$, then $R_2$ is correct) faster when correct responses to $S_1$ are consistently followed by reinforcing outcome $O_1$ (e.g., sucrose solution) and correct responses to $S_2$ are consistently followed by a different outcome $O_2$ (e.g., food pellets) than if correct responses to $S_1$ and $S_2$ are both reinforced by the same outcome. Trapold hypothesized that the DOE occurs because (1) when $S_1$ and $S_2$ are each followed by a unique outcome event, subjects learn an expectancy of $O_1$ to $S_1$ and an expectancy of $O_2$ to $S_2$, and (2) these different expectancies provide an additional source of differential stimulation that can assume stimulus control over the choice behavior. In contrast, subjects which receive the same outcome for correct responding to $S_1$ and $S_2$ learn the same expectancy to $S_1$ and $S_2$, and hence do not have available this additional source of potential differential stimulus control over $R_1$ and $R_2$.

Trapold provided additional support for this theo-

retical analysis by showing that subjects who are pretrained so as to preestablish the expectancies involved in a subsequent differential-outcomes conditional discrimination problem learn that problem faster than subjects pretrained to expect the wrong outcome.

Carlson and Wielkiewicz (1972, 1976) extended Trapold's findings by showing a DOE for food reward outcomes of different delays and different magnitudes in rats. Overmier, Bull, and Trapold (1971) demonstrated a DOE in dogs when the differential outcomes were avoidance of electric shocks to different loci on the body. Brodigan and Peterson (1976) and Peterson, Wheeler, and Armstrong (1978) found a DOE with pigeons using food and water as the different outcomes. These latter studies also provided considerable additional support for the assumption that the DOE is mediated by differential outcome expectancies which function as part of the discriminative stimulus complex to which the correct response becomes learned. Specifically, they showed that the DOE is much more profound when a delay is inserted between the offset of the conditional cue ($S_1$ or $S_2$) and the opportunity to choose between $R_1$ and $R_2$. With a standard procedure, insertion of such a delay of even a few seconds produces a marked deterioration of both initial learning and terminal performance on a conditional discrimination. Traditionally, these effects of a delay have been attributed to degraded stimulus control over $R_1$ and $R_2$ because of short-term memory losses of conditional cue

information during the delay. If expectancies persist better across a delay than the traces of the conditional cues, then we would expect that having differential expectancies conditioned to $S_1$ and $S_2$ would attenuate the normal decremental effect of a delay. Indeed, Peterson et al. (1978) found very good conditional discrimination performance in pigeons with delays that are much too long to support differential responding with a standard reinforcement procedure.

Thus, there is a growing body of evidence to support the empirical generalization that conditional discrimination performance is facilitated if different outcomes are made contingent upon the separate stimulus-response components of the problem. The evidence also supports the hypothesis that this DOE can be most parsimoniously interpreted in terms of the operation of expectancies that are relatively specific to particular outcomes, and which provide the subject with an additional source of stimulus control for the correct response.

This paper reports the results of several experiments directed at the question of whether nonreinforcement functions like other (reinforcing) outcomes that have so far been studied in the DOE paradigm. Nonreinforcement as a specifically expectable outcome is of special significance for at least two reasons. First, is the relatively straightforward question of whether a specific expectancy based upon nonreinforcement is learned and functions in the same manner as the other outcome events that have so far been studied. The concept of conditioned anticipatory frustration, which has played such a prominent role in theorizing regarding other learning phenomena (e.g., Amsel, 1962) is, in effect, an expectancy of nonreinforcement, and so provides one rationale for suspecting that a nonreinforcement expectancy may operate in the DOE paradigm. Second, and perhaps more important, is the fact that if subjects in a reinforcement/nonreinforcement differential-outcomes paradigm show a DOE relative to subjects who receive the same reinforcing outcome for both $S_1$ and $S_2$ correct responses, then we must conclude that the facilitative influence of receiving a different outcome on $S_1$ and $S_2$ trials more than outweighs the detrimental effect of being reinforced on only half the correct trials. In other words, the DOE is not minor or trivial when compared to the S-R association-strengthening function of the reinforcer.

## EXPERIMENT 1

The purpose of Experiment 1 was to determine if a DOE would occur when nonreinforcement was one of the outcomes. Thus, after initial acquisition with food as the outcome for both components, one group of pigeons (Group D) was switched to a $S_1$-$R_1$-$O_1$/$S_2$-$R_2$-$O_2$ conditional discrimination in which $O_1$ was food + tone and $O_2$ was tone alone. One con-trol group (Group M) was switched to food + tone for half their correct trials to both $S_1$ and $S_2$, and tone alone for the remaining half of $S_1$ and $S_2$ trials. A second control group (Group S) continued to receive food + tone for all correct responses to both $S_1$ and $S_2$.

## Method

### Subjects

The subjects were 12 adult Roller pigeons obtained from local breeders. They were housed in individual cages in a continuously illuminated, temperature-controlled colony room. After having eaten Purina Pigeon Chow ad lib for at least 3 weeks, each bird was reduced to 80% of its free-feeding weight where it was then maintained.

### Apparatus

The apparatus consisted of two three-key pigeon operant conditioning chambers in individual sound enclosures. One-plane projectors mounted behind the response keys made possible the presentation of various visual stimuli. Overhead houselights provided approximately 2.7 lx of ambient light measured at the center of the experimental space, with the detection surface parallel with and pointed toward the key panel. Each chamber was equipped with a food hopper, mounted behind an access aperature in the middle of the front wall beneath the center of the response key. The hopper was illuminated with white light when raised.

Control and data recording were by conventional electromechanical equipment located in adjacent room.

### Procedure

The subjects were randomly assigned to one of the two chambers and then to one of the three groups such that two subjects from each chamber were assigned to each group. The groups were designated S for single outcome, D for differential outcomes, and M for mixed outcomes.

**Preliminary training.** After the pigeons were trained to eat from the hopper, they were autoshaped to peck the center key. The center key was transilluminated with white light on the average of once every 30 sec. In the absence of a keypeck, the white key light went off after 6 sec and was immediately followed by a 3-sec access to grain. If the pigeon pecked the key before the 6 sec had elapsed, the key light went out and the grain was presented immediately. Keypecks during the intertrial interval (ITI) had no effect.

**Phase 1: Standard training.** When the subjects were responding reliably to the onset of the center key light, all were trained on the conditional simultaneous discrimination task for 45 sessions. A trial began with the presentation of either a green or a red light on the center key. A single peck on the center key resulted in the termination of the color stimulus and onset of the choice stimuli, vertical lines and horizontal lines, located on the side keys. A single peck to one of the choice stimuli terminated both choice key lights. Vertical was correct on green trials, horizontal on red trials. Correct choices resulted in 3-sec access to grain, the first .75 sec of which was accompanied by a moderately loud 1-kHz tone. A 5-sec ITI followed. An incorrect choice produced a 10-sec blackout of the houselight, followed by the 5-sec ITI, after which the trial was repeated (i.e., a correction procedure was employed). This process continued until the trial was correctly completed. Responses during the ITI, when all keys were darkened, reset the ITI clock, thereby delaying onset of the next trial for 5 sec.

Each session consisted of 80 correctly completed trials. Order of trials was random, with the restrictions that each block of 16 trials contained 8 green and 8 red trials and that, within each of the 8-trial subsets, the correct choice stimulus occur equally often on the left and right side keys.

**Phase 2: Group treatments.** When performance had stabilized under the standard training conditions, the group treatments were begun. Group S (single outcome) continued to receive food + tone for correct responses on both red and green trials. Group D (dif-
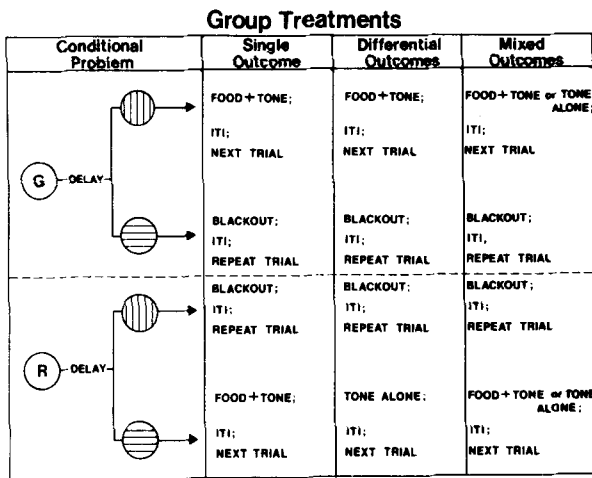
Figure 1. Schematic diagram of the procedures employed in Experiment 1.

ferential outcomes) continued to receive food + tone for correct responses of one kind of trial (e.g., green-vertical) but received the .75-sec tone alone (immediately at offset of the key light) for correct responses on the other kind of trial (e.g., red-horizontal). On tone-alone presentations, the 3-sec feeder timer operated as usual (but without operating the hopper) and initiated the ITI when it timed out. Thus, the time to the next trial was the same whether the trial ended in food + tone or tone alone; the only difference was the presence or absence of food and its accompanying hopper sound and light. One Group D subject in each of the two chambers received food + tone on green trials and tone alone on red trials, while the other two subjects received the reverse.

Group M (mixed outcome) received an unpredictable sequence of the food + tone and tone-alone outcomes for correct responses on both kinds of trials such that both kinds of outcomes occurred equally often on both kinds of trials within each successive eight-trial block. The group treatments are schematized in Figure 1.

After 12 80-trial sessions under these conditions, additional sessions were given during which a .5-sec delay was imposed between the offset of the conditional cue and onset of the choice stimuli. During the delay interval, all keys were darkened and responses, no matter on which key they occurred, had no effect. Following 15 sessions of delay training, the delay was removed for 12 additional sessions.

**Phase 3: Standard retraining.** Following Phase 2, the subjects were all returned to the single-outcome procedure (at zero delay) for 12 sessions. Then the .05-sec delay was again imposed for 12 sessions.

## Results and Discussion

Figure 2 presents the proportion of trials on which the correct choice response was made on the first presentation of the trial (percent first correct) across three-session blocks of trials for the three groups for all phases of this experiment.

The leftmost panel of Figure 2 shows that all groups initially acquired the conditional discrimination under the standard training conditions at comparable rates, reaching a terminal performance level of 85%-90% first correct.

The next panel shows the effect of introducing the three outcome patterns. Upon introduction of the differential outcomes for correct $S_1$ and $S_2$ responses, Group D showed an abrupt and significant [t(3) =

8.56, p < .05] decrement in performance, followed by a sharp increase to a level higher than (though not significantly so) the level attained at the end of standard training.

Introducing a .5-sec delay between the offset of the conditional cue and the onset of the choice stimuli produced a sharp separation of the three groups. Group D continued to perform at the 95% level achieved under 0-sec delay, whereas Groups S and M showed an abrupt and significant [t(3) = 3.30 and 3.24, respectively, p < .05] decrease to 78% and 69%, respectively. Groups S and M were not significantly different from one another, but were significantly different from Group D [F(1,6) = 14.69, p < .01]. When the delay was subsequently removed, Groups S and M quickly returned to higher levels of performance. However, Group M's performance continued to be significantly [F(1,6) = 13.55, p < .025] below that of Group D.

When the subjects were returned to the standard training condition in Phase 3, the differences among Groups S, M, and D vanished and remained absent when a .5-sec delay was added to that condition.

These results, then, confirm those obtained previously with other species and different reinforcing outcomes. Differential outcomes, even if one of those involves no primary reinforcement, facilitate conditional discrimination performance and, moreover, do so to a much more pronounced degree when a delay is inserted between the conditional cue and the choice stimuli. Furthermore, at .5-sec delay, the DOE was large enough to completely overcome the
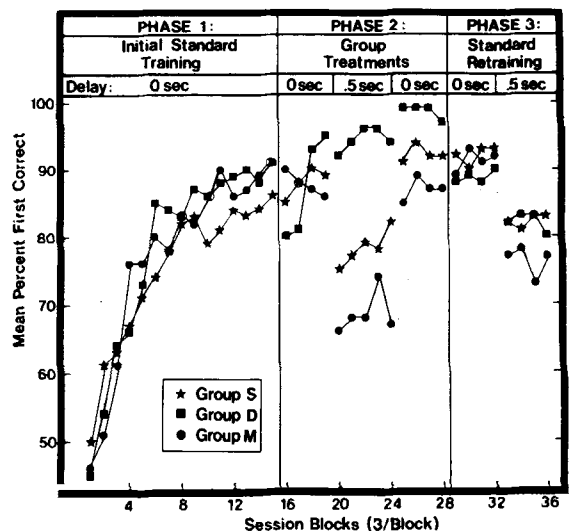


Figure 2. The results of Experiment 1 in terms of mean percent first correct for the three groups across the successive stages of the experiment. During initial acquisition (left panel), all groups were trained under the standard single-outcome procedure. The center panel presents the effects of the different group treatments under 0- and .5-sec prechoice delay. The right panel presents data from 0-sec delay and .5-sec delay conditions when all groups were trained once again under the standard procedure.

decremental effect of reinforcing only one of the two types of trials. That is, providing primary reinforcement for an unsystematic half of the correct responses (Group M) produced a decrease in performance (albeit not significant) from the level maintained by giving primary reinforcement for every correct response (Group S). Providing reinforcement for a systematic half of the correct responses, not only canceled out the decremental effect of eliminating half the primary reinforcements, but actually produced a performance level significantly better than that produced by 100% primary reinforcement.

## EXPERIMENT 2

Using a differential outcome paradigm with food and water reinforcement in pigeons, Peterson et al. (1978) showed that the DOE becomes much more profound as the delay between conditional cue offset and choice stimuli onset is lengthened, and that differential outcomes will support very good discrimination performance across delays that are too long to support above-chance correct responding with a standard or a mixed-outcome procedure. The purpose of Experiment 2 was to determine if the same would hold true for a differential outcome procedure based upon reinforcement and nonreinforcement.

### Method

#### Subjects and Apparatus

The subjects and group assignments were the same as in Experiment 1.

#### Procedure

Experiment 2 was a continuation of Experiment 1. All birds first received 24 sessions of additional training under their respective group treatment procedures at a .5-sec delay. This was followed by 24 sessions with a 1-sec delay, then 12 sessions with a 2-sec delay, 12 sessions with a 3-sec delay, and finally, 12 sessions with a 5-sec delay.

### Results and Discussion

Figure 3 presents the results of this experiment. The more extended training at .5-sec delay resulted in a diminution of the group differences seen in Experiment 1, with Group S finally achieving the same performance level as Group D, and with Group M at a lower (but not significantly so) level.

Extending the delay to 1 sec produced a significant (p < .05 in all cases) drop in performance level of all three groups, and caused the three groups' performance levels to separate slightly, but not significantly.

Extending the delay beyond 1 sec, however, produced a very large separation among the three groups. In general, Group D's performance stayed at about the 90% level out to the maximum 5-sec delay employed here. The performance of Group S and M, on the other hand, dropped off very sharply
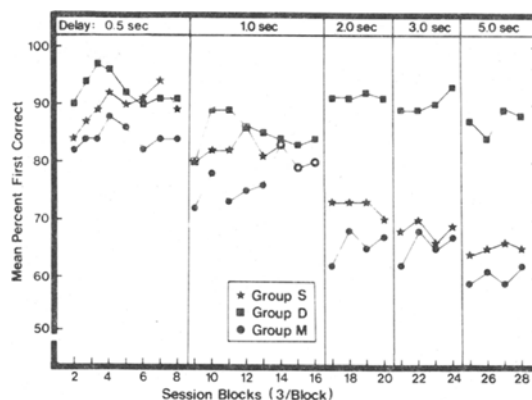


Figure 3. The results of Experiment 2 in terms of mean percent first correct for the three treatment groups as a function of successively longer prechoice delays.

at a 2-sec delay, to a level significantly below that of Group D [F(1,6) = 7.38, p < .05], and continued to get progressively lower as the delay got longer. Consistent with the results of Experiment 1, and with what one would expect on the basis of the S-R association-strengthening nature of reinforcement, Group S generally performed at a somewhat (but not significantly) higher level than Group M.

These results, then, effectively replicate those of Experiment 1, and also demonstrate that, like differential outcomes that are both reinforcers, differential reinforcement/nonreinforcement outcomes produce a much more profound facilitation of conditional discrimination performance when there is a delay interval between conditional cue and choice opportunity.

## EXPERIMENT 3

There were two features of the procedures employed in Experiments 1 and 2 that may have permitted possible confounding differences to arise among the three treatment groups. First, Grant (1976) has shown, with a standard reinforcement procedure, that the longer the duration of the conditional cue, the better the long-delay conditional discrimination performance of pigeons. Thus, if the differential outcome procedure somehow produced slower responding to the conditional cue than either the S or the M procedure, that could have been responsible for the observed superiority of Group D in Experiments 1 and 2. Although casual observation revealed no obvious differences of this sort, we deemed it wise to expressly control conditional cue exposure duration in Experiment 3. This was done by leaving the conditional cue on for a 4-sec duration regardless of whether the bird pecked it or not. To ensure that the birds were attending to the conditional cue when it was pre-

sented this way, the center key always lit up white at the onset of a trial and a peck at the white key was required to produce the conditional cue.

A second possible source of confounding involves side-key pecking during the delay period. As is customary in experiments of this type, responses during the delay period had no effect. However, we noticed, in Experiment 2, that several birds in the S and M conditions engaged in frequent side-key pecking during the delay interval, whereas no birds from the D condition were observed to do this. These delay responses often "ran over" into the choice component of the task, resulting in the scoring of correct and incorrect choices that were probably not under the control of any of the stimuli related to choice behavior. Mixing this presumably random choice determination in with the stimulus-controlled choice responding of Groups S and M, but not of Group D, would serve to reduce the overall level of first-correct responding in the former groups relative to the latter. This possible confounding was eliminated by use of a procedure in Experiment 3 which aborted a trial if any side-key responses occurred during the delay interval.

This experiment also examined the effects of switching subjects from the D to the S condition, and vice versa.

## Method

### Subjects and Apparatus
Subjects, apparatus, and group designations were the same as in Experiments 1 and 2.

### Procedure
The basic procedure was the same as in the previous experiments, except for the two changes mentioned above. All trials now began with the onset of white light on the center key. A single peck to this white-lighted key resulted in the offset of the white light and the onset of a color cue, red or green. No keypecks were required to the color cues. If any occurred, they had no effect. After 4 sec, the color cue went out automatically and the delay interval began. Center-key pecks during the delay interval continued to have no effect, but side-key pecks during the delay now resulted in an immediate abortion of that particular trial and an unsignalled return to the ITI. Thus, the subject could proceed from the conditional cue to the choice stimuli only if it did *not* emit pecks to side keys during the delay interval. All other aspects of the procedure remained as they were in Experiments 1 and 2.

Experiment 3 began by retraining all subjects under their respective group treatment conditions at 0-sec delay for 24 sessions, then at 2-sec delay for 24 sessions and at 5-sec for 12 sessions. With the delay at 5 sec, the subjects of Group D were switched to the single-outcome procedure while Groups S and M were switched to the differential-outcomes procedure. Following 12 sessions of training under these conditions, Group D was returned to the differential-outcomes procedure, so that for the final 12 sessions of the experiment all subjects in all groups were trained under the differential-outcomes procedure.

## Results and Discussion

Figure 4 presents the results of Experiment 3. All groups began at the performance level last attained
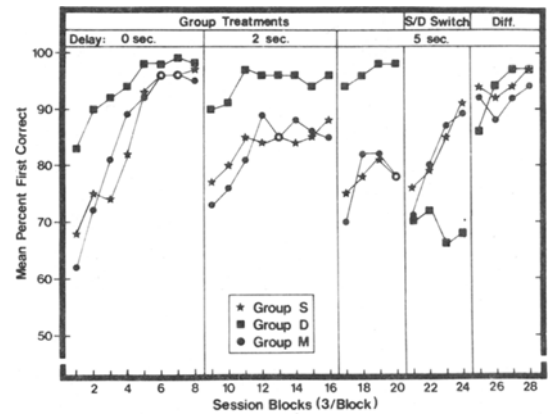


Figure 4. The results of Experiment 3 in terms of mean percent first correct for the three treatment groups as a function of successively longer prechoice delays and successive changes in the outcome contingency for correct choices. The panel labeled "S/D Switch" depicts the results of switching Group/D to the single-outcome procedure and Groups S and M to the differential-outcomes procedure. The rightmost panel presents the data obtained when all three groups were run under the differential-outcomes data obtained when all three groups were run under the differential-outcomes procedure.

under the 5-sec delay in Experiment 2, and gradually increased to a common terminal-performance level of approximately 95% first correct.

Reinstitution of the 2-sec delay resulted in significant (p < .05 in all cases) decrements in the performances of all three groups. Groups S and M performed consistently at levels significantly [F(1,6) = 6.50, p < .05] lower than that of Group D throughout the 2-sec delay phase and also throughout the next phase when the delay was extended to 5 sec [F(1,6) = 10.01, p < .05]. For reasons that are not clear, Group S did not show consistently better performance than Group M, as had been true in Experiment 2.

When Group D was switched to the S condition, its performance abruptly dropped from about 95% to about 70% [t(3) = 14.64, p < .001], and then abruptly returned to 90% or better [t(3) = 15.64, p < .001] when it was returned to the D condition. When Group S and M were switched to the D condition, their performance gradually approached the 95% level previously attained by Group D.

This experiment therefore demonstrates that neither the possible differences in conditional cue duration nor the differences in side-key pecking were likely responsible for the DOE seen in Experiments 1 and 2, since both sources of possible confounding differences among groups were controlled in this experiment.

These results also show that even after a very good level of delayed conditional discrimination performance was achieved and maintained for many sessions using the reinforcement/nonreinforcement differential-outcomes procedure, that same high level

of performance could not then be maintained by shifting to reinforcement for all correct responses. They also show that when subjects of Groups S and M, for whom a 5-sec delay was too long to maintain strong discrimination performance, were shifted to a differential-outcomes procedure, their performance improved. This was true even though, for Group S, this shift entailed a substantial decrease in the overall number of reinforcements received.

## EXPERIMENT 4

The procedures employed in Experiments 1, 2, and 3 were such that, even though we have consistently described one of the outcomes as "nonreinforcement," that outcome probably involved conditioned reinforcement. On "reinforced" trials, the birds received 3-sec access to food plus a tone. On "nonreinforced" trials, they received the tone alone. According to what is known about the conditions which foster a neutral stimulus becoming a conditioned reinforcer (cf. Mackintosh, 1974), the reinforced trials could have made the tone a conditioned reinforcer, so that in actuality the subjects were receiving reinforcement on "nonreinforced" trials.

Also, it is conceivable that, since "houselight off" was consistently correlated with no reinforcement whereas all primary reinforcement was delivered with the houselight on, "sustained lights on" following a response might have become a conditioned reinforcer. Similarly, the relatively short 5-sec ITI may have permitted other stimulus events associated with progress through the program to become conditioned reinforcers.

Accordingly, Experiment 4 was an attempt to evaluate the importance of conditioned reinforcement to the DOE based upon reinforcement and "nonreinforcement." Two strategies were employed to achieve this end. First, the procedures of Experiments 1, 2, and 3 were modified so as to minimize the likelihood that incidental stimulus changes accompanying the nonreinforcement outcome event would become conditioned reinforcers. The blackout following incorrect responses was eliminated; incorrect responses now resulted only in termination of the choice stimuli and the unsignalled initiation of the ITI. Furthermore, the ITI was extended to 15 sec.

Second, Experiment 4 arranged an explicit comparison between one group of subjects, which received a stimulus (tone) on "nonreinforced" trials that should have been a good conditioned reinforcer, and a second group, for whom the tone should not have been as good a conditioned reinforcer. This was accomplished by having the tone in the former group also accompany food delivery on reinforced trials, whereas in the latter group it did not.

## Method

### Subjects and Apparatus

A new set of 12 naive Roller pigeons served as the subjects. The apparatus and the deprivation and housing conditions were the same as in Experiments 1, 2, and 3.

### Procedure

The birds were randomly assigned to one of three groups and two chambers, and then were given preliminary autoshaping to the white center key according to the procedure outlined in Experiment 1. When all subjects were responding reliably to the onset of the white center-key light, they were transferred to training on the conditional discrimination task according to their respective group procedures. The same fixed-cue duration procedure was employed as in Experiment 3. The houselight was on continuously throughout the experiment. All choice responses resulted in offset of the choice stimuli and a 15-sec ITI. If the choice was correct, the program then advanced to the next trial. If the choice was incorrect, the trial was repeated. Trial sequences were constructed as in previous experiments. Side-key pecks during the delay period aborted the trial. Pecks on any key during the ITI reset the ITI timer.

For Group S, correct choice responses of both kinds yielded the .75-sec tone presentation together with 3-sec access to grain. For Group D-P, correct choices of one kind yielded food + tone while correct responses of the other kind yielded the tone alone. Group D-U subjects received food alone for one kind of correct response and tone alone for the other.

Phase 1 training began with 20 sessions at a 0-sec delay between conditional cue offset and choice stimuli onset, followed by 20 sessions with a 2-sec delay and 20 sessions with a 10-sec delay. After 10 of these 10-sec delay sessions, Group S was switched to the D-P condition. Then all subjects received a further 10 sessions at a 15-sec delay.

In Phase 2 of this experiment, we made a further attempt to evaluate the role played by the tone on nonreinforced trials. After 10 sessions under the D-P or D-U conditions as described above, and at a 7.5-sec delay, the tone was eliminated on all trials for all subjects for an additional 10 sessions and then was returned for yet a further 10 sessions.

Phase 3 of this experiment was an attempt to replicate the results of Experiment 3, which showed a massive performance drop when consistent food reinforcement for all correct responses was substituted for the reinforcement/nonreinforcement differential-outcomes procedure. To that end, all subjects were switched to the S condition for a final 10 sessions.

## Results and Discussion

Figure 5 presents the results through Phase 1. In initial acquisition, the D-P and D-U groups learned the problem faster than Group S. An ANOVA of the data of the first 10 blocks yielded a significant Groups by Trials interaction [$F(18,81) = 3.52$, $p < .001$], as well as a significant Group main effect, [$F(2,9) = 5.46$, $p < .05$]. Moreover, Groups D-P and D-U did not differ throughout initial acquisition. When the delay was introduced, the D groups showed only a small transient decrement in performance whereas the S group showed an abrupt and dramatic drop in performance [$t(3) = 4.66$, $p < .05$], and then gradually recovered over the 10 sessions. When the delay was extended to 10 sec, the D groups showed an initial significant [$t(7) = 4.65$, $p < .01$] decrease in performance, followed by recovery to about 90% correct, whereas Group S dropped to
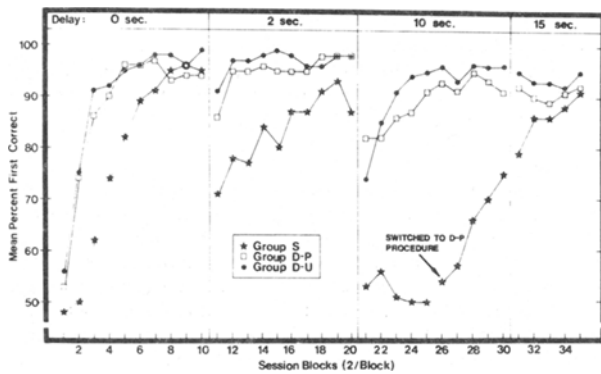
Figure 5. The results of Phase 1 of Experiment 4 in terms of mean percent first correct for the three treatment groups during initial acquisition at 0-sec delay and during subsequent stages at successively longer delays. Group S was switched ot the D-P procedure at the beginning of Block 26 and remained on that procedure for the rest of the experiment.

chance performance $[t(3) = 12.37, p < .01]$ and remained there. When Group S was switched to the D procedure, while remaining at a 10-sec delay, its performance immediately started to improve, and continued to improve as the delay was extended to 15 sec, until it reached the level of performance of the D groups (approximately 90% first correct).

Throughout Phase 1, no striking differences between Groups D-P and D-U were apparent.

Figure 6 shows the results of Phase 2 of Experiment 4. Removing the tone produced no significant
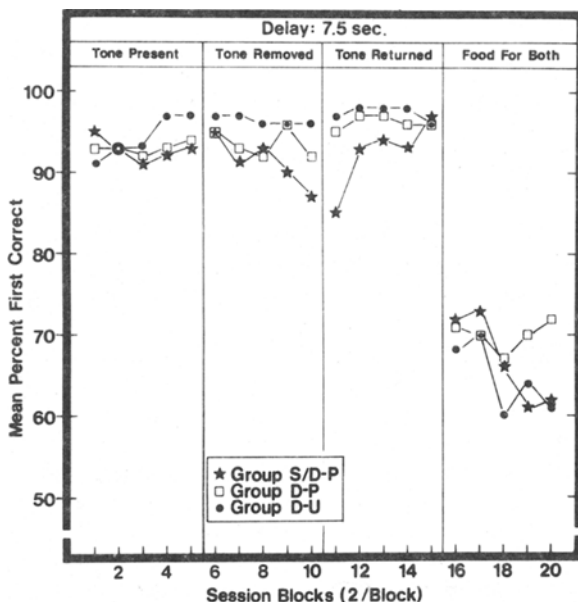


Figure 6. The results of Phase 2 of Experiment 4 in terms of percent first correct across successive stages in which the role of the tone feedback stimulus was assessed. The rightmost panel shows the effect of switching all groups from the differential-outcomes procedure to the single-outcome procedure. The pre-choice delay was 7.5 sec throughout Phase 2.

changes in the performance of any group. There was a slight indication that Group S/D-P's performance was somewhat disrupted by deletion of the tone, suggesting perhaps that the tone may play some role early in differential outcome training. However, the absense of any notable performance changes in Groups D-P and D-U when the tone was removed suggests that an explicit feedback stimulus on nonreinforced trials is not essential to maintenance of the DOE. Whether such a feedback stimulus is important to the superiority of the DOE procedure during initial acquisition is still an open empirical question.

The result of shifting all subjects to the S condition (rightmost panel of Figure 6) replicated nicely the results of Experiment 3. This shift produced a very dramatic loss of discriminative performance in all subjects $[t(11) = 11.64, p < .001]$.

## GENERAL DISCUSSION

The main empirical conclusion to be derived from these experiments is that differential outcomes based upon reinforcement/nonreinforcement function in much the same way in conditional discrimination learning as differential outcomes based upon qualitatively or quantitatively different reinforcers. Relative to either the standard procedure (the S condition here) or to the mixed procedure, differential reinforcement/nonreinforcement outcomes facilitate initial acquisition of a conditional discrimination, and lead to higher levels of performance on conditional discriminations, especially when a delay between conditional cue and choice stimuli produces poor performance under the S and M conditions. The effect of differential reinforcement/nonreinforcement outcomes under delayed conditions is, in fact, substantial enough to maintain 90% or better correct responding under delays which are capable of supporting no differential responding with a nondifferential outcome procedure. Moreover, the facilitative effect of differential reinforcement/nonreinforcement outcomes is substantial enough to be readily observable despite the fact that the nonreinforcement operation by definition means that differential-outcomes subjects receive fewer primary reinforcements for correct performance. Finally, the ability of "nonreinforcement" to function as a specific outcome in the differential-outcomes paradigm does not appear to be sensitive to manipulations which should produce marked differences in the conditioned reinforcing capacity of stimuli that comprise the nonreinforcement outcome. However, we do believe that use of a correction procedure may be crucial to obtaining a DOE involving nonreinforcement.

How are these results to be interpreted? Our preference is to employ the concept of an event-specific learned expectancy. An expectancy is a hypothetical

learned response to stimulus A, that is acquired as the result of the organism repeatedly experiencing A followed by some other stimulus, B. The expectancy is relatively specific or unique to the particular qualitative and quantitative nature of B, and has its own unique stimulus component which, in the same fashion as stimuli from other sources, can gain learned control over particular behaviors of the subjects.

In the standard conditional discrimination paradigm, the conditional cues both come to be regularly followed by the same outcome (say, food), and consequently the expectancy elicited by both conditional cues is the same ($E_{food}$). Thus, the subject is asked to learn to respond differently to two stimulus complexes, $S_1 + E_{food}$ and $S_2 + E_{food}$.

In the differential-outcomes paradigm, one conditional cue ($S_1$) is regularly followed by one outcome ($O_1$) and the other conditional cue ($S_2$) is regularly followed by the other outcome ($O_2$). Consequently, the expectancies that come to be elicited by the two conditional cues are different ($E_1$ and $E_2$). Thus, in this case, the subject learns to respond differently to the two stimulus complexes, $S_1 + E_1$ and $S_2 + E_2$. Other things being equal, this latter discrimination should be easier to learn than the former.

This simple analysis can account nicely for the superiority of the differential-outcomes procedure, and the extension of it to the reinforcement/nonreinforcement case involves only the additional assumption that a specific expectancy for nonreinforcement is also learned. This assumption, of course, is the heart of frustration theory (Amsel, 1962) and is quite familiar.

The fact that the differential-outcomes procedure becomes much more advantageous as the cue-choice delay increases requires a further additional assumption that the expectancies elicited by the conditional cue persist better across the delay period than does the subject's short-term memory for the conditional cue itself. Precisely why this is the case is a matter for future research. For now, it will have to suffice to state the assumption as explicitly as possible, and to note that there do exist some data from other contexts which support its plausibility (e.g., Cohen, Looney, Brady, & Aucella, 1976; Eckerman, 1970; Farthing, Wagner, Gilmour, & Waxman, 1977).

Two questions about the expectancy concept often arise: (1) Where is the expectancy located? Is it relatively peripheral and hence potentially identifiable and studiable, or is it buried inaccessibly somewhere in the organism's head? (2) How is the expectancy learned? Is it instrumental or is it Pavlovian? It has been argued elsewhere (Trapold & Overmier, 1972) that these questions are fundamentally unanswerable in our present state of ignorance. We simply do not possess sufficiently definitive criteria to permit us to unequivocally sort *observable* behaviors into those

that were learned instrumentally and those that were learned Pavlovianly, let alone to sort responses that are theoretical entities.

Further, since the concept of expectancy is a theoretical entity, invented precisely because it permits us to better summarize and integrate a set of empirical facts, the only criteria we possess for deciding whether any particular response, no matter how central or peripheral it may be, is "really the expectancy of x" is whether that response behaves the way expectancy theory says it should. If some particular response is examined on the suspicion that it may be a particular expectancy, and it proves not to act the way the theory says it should, we will decide that that particular response isn't it, but we certainly won't discard the theory.

So, the only truthful answer to the question "Is it central or peripheral" is "we don't know because we haven't found it yet." Moreover, since the concept is not very richly developed yet—that is, we have not been forced to make a very rich array of assumptions about it yet—to search for it at this point in time would probably be premature. We wouldn't really know what we were looking for

These logical considerations aside, however, one aspect of the data from these experiments severely strains the notion that the expectancy is instrumentally learned. That is the result of shifting subjects from a differential-outcomes to a single-outcome procedure. If we wish to argue that the expectancy is an instrumentally learned response, then that response must have been in place when subjects were performing well on the differential-outcomes procedure. Shifting those subjects to the S condition, then, should, if anything, increase the reinforcement the subject receives for making the instrumental expectancies, hence maintaning the expectancies and, therefore, the correct choice behavior. However, that is not what happened. Quite the contrary, in both Experiment 3 and Experiment 4, as well as in previous work with other outcomes (Peterson et al., 1978), shifting subjects from the D to the S condition devastated performance. This marked decrease in performance in response to improved reinforcement conditions is much more compatible with the notion that specific event expectancies are classically conditioned on the basis of the regular temporal relationships between conditional cues and outcomes. This reasoning has been elaborated in greater detail elsewhere (Peterson et al., 1978).

## REFERENCES

AMSEL, A. Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychological Review*, 1962, **69**, 306-328.

BRODIGAN, D. L., & PETERSON, G. B. Two-choice discrimination performance of pigeons as a function of reward expectancy, prechoice delay, and domesticity. *Animal Learning & Behavior*, 1976, **4**, 121-124.

CARLSON, J. G., & WIELKIEWICZ, R. M. Delay of reinforcement in instrumental discrimination learning of rats. *Journal of Comparative and Physiological Psychology*, 1972, **81**, 365-370.

CARLSON, J. G., & WIELKIEWICZ, R. M. Mediators of the effects of the magnitude of reinforcement. *Learning and Motivation*, 1976, **7**, 184-196.

COHEN, L. R., LOONEY, R. A., BRADY, J. H., & AUCELLA, A. F. Differential sample response schedules in the acquisition of conditional discriminations by pigeons. *Journal of the Experimental Analysis of Behavior*, 1976, **26**, 301-314.

ECKERMAN, D. A. Generalization and response mediation of a conditional discrimination. *Journal of the Experimental Analysis of Behavior*, 1970, **13**, 301-316.

FARTHING, G. W., WAGNER, J. M., GILMOUR, S., & WAXMAN, H. M. Short-term memory and information processing in pigeons. *Learning and Motivation*, 1977, **8**, 520-532.

GRANT, D. S. Effect of sample presentation time on long-delay matching in the pigeon. *Learning and Motivation*, 1976, **7**, 580-590.

MACKINTOSH, N. J. *The psychology of animal learning.* New York: Academic Press, 1974.

OVERMIER, J. B., BULL III, J. A., & TRAPOLD, M. A. Discriminative cue properties of different fears and their role in response selection in dogs. *Journal of Comparative and Physiological Psychology*, 1971, **76**, 478-482.

PETERSON, G. B., WHEELER, R. L., & ARMSTRONG, G. D. Expectancies as mediators in the differential-reward conditional discrimination performance of pigeons. *Animal Learning & Behavior*, 1978, **6**, 279-285.

TRAPOLD, M. A. Are expectancies based upon difference positive reinforcing events discriminably different? *Learning and Motivation*, 1970, **1**, 129-140.

TRAPOLD, M. A., & OVERMIER, J. B. The second learning process in instrumental learning. In A. H. Black and W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory.* New York: Appleton-Century-Crofts, 1972.