

SUBSTITUTABILITY BETWEEN CONDITIONED AND PRIMARY REINFORCERS IN DISCRIMINATION ACQUISITION

BEN A. WILLIAMS AND ROGER DUNN

UNIVERSITY OF CALIFORNIA-SAN DIEGO AND
SAN DIEGO STATE UNIVERSITY

Rats and pigeons were trained on a series of reversals of a conditional simultaneous discrimination. The percentage of reinforcement for correct trials was varied across reversals. When nonreinforced correct trials produced the same feedback as incorrect trials, the number of errors to reach an acquisition criterion was greater for smaller percentages of reinforcement, but the number of reinforcers required was either approximately constant or smaller for the smaller percentages. When a stimulus paired with food (the conditioned reinforcer) was added on nonreinforced correct trials, both measures were substantially decreased. When the same stimulus was presented, but without a history of food pairing, learning rate was similar to when no stimulus was presented on nonreinforced trials. The results provide direct evidence that conditioned reinforcers may substitute, although imperfectly, for a primary reinforcer, and that pairing with the primary reinforcer is a necessary condition for such substitutability to occur.

Key words: conditioned reinforcement, percentage of reinforcement, invariance in reinforcements to acquisition, serial reversal learning, simultaneous discrimination, lever press, key peck, rats, pigeons

The concept of conditioned reinforcement has a venerable but disputed status. It has been commonly invoked as an explanation of the maintenance of behavior when responding is not temporally contiguous with reinforcement (e.g., Spence, 1947), but others have claimed that its effects, although clearly of major significance, occur for reasons other than reinforcement per se. Instead of an initially neutral stimulus gaining reinforcement properties of its own because of a history of pairing with the primary reinforcer, the claim of these alternative explanations is that such stimuli affect behavior via their discriminative properties quite apart from their ability to strengthen behavior (e.g., Longstreth, 1971; Schuster, 1969).

The great majority of studies that have addressed this issue have used free-operant procedures in which response rate has been the dependent variable. Although it is possible to demonstrate powerful effects of "conditioned reinforcement" contingencies with such procedures (e.g., Royalty, Williams, & Fantino, 1987), the problem of interpretation has been

to disentangle response-strengthening effects from those that occur for other reasons. For example, stimuli may provide discriminative cues for schedule transition or time to reinforcement, which modulate both the rate and pattern of responding, but without such effects being dependent on the stimulus having reinforcement properties in its own right, as shown by the similarity of the effects of stimuli with versus without a history of pairing with the reinforcer (e.g., Cohen, Calisto, & Lentz, 1979). Moreover, even in cases in which the stimulus effects do depend on the associative history of the stimulus, these may occur because the previous pairings with the reinforcer increase the salience of the stimulus rather than its value as a reinforcer (e.g., Stubbs, Vautin, Reid, & Delehanty, 1978).

An alternative approach to the study of conditioned reinforcement is to examine its possible effects on the acquisition of discrimination learning. Procedures such as discrete-trial simultaneous discrimination learning should be insensitive to the discriminative properties of the stimulus, because the rate of learning is the dependent variable and should be unaffected by any stimulus effects on the rate or pattern of behavior. Thus, to the extent that presentations of consequent stimuli facilitate such learning, and this facilitation depends upon the history of pairing of the stimuli with

This research was supported by NIMH Grant MH42797 to the University of California-San Diego and MH 40853 to San Diego State University. Correspondence and reprint requests should be addressed to the first author, Department of Psychology, C-009, University of California-San Diego, La Jolla, California 92093.

primary reinforcement, clear evidence would be provided for the view that stimuli may acquire conditioned reinforcement effectiveness because of association with primary reinforcers and thus may substitute for primary reinforcers.

Surprisingly few studies of conditioned reinforcement have used discrimination learning as a measure. Several studies of maintained discrimination on a matching-to-sample (or oddity) problem have varied the percentage of reinforcement for such behavior on fixed-ratio (FR) schedules, and these often included brief presentations on the nonreinforced correct trials of a stimulus also present during food delivery (e.g., Fujita, 1985; Thomas, 1979), but the effects of the presence or absence of the conditioned reinforcer generally were not investigated systematically. Clark and Sherman (1970) did provide such an investigation by training pigeons on a continuous version of a matching-to-sample problem in which food occurred on a fixed-interval (FI) 8-min schedule while conditioned reinforcers were presented on a variable-interval (VI) 1-min schedule contingent either on matching behavior or on mismatching behavior. Presentation of the brief stimuli contingent on matching behavior had little effect, perhaps because the behavior was very accurate even when the brief stimuli were never presented, whereas presentation of the brief stimuli contingent on mismatching behavior substantially decreased overall matching accuracy. Moreover, the effect of the contingency for mismatching was substantially greater when the stimuli were paired with food. However, some increase in mismatching behavior did occur even when the stimuli were unpaired with food.

The study that provides the strongest evidence for facilitation of discrimination by conditioned reinforcement was reported by Hursh (1977), who trained monkeys on repeated acquisition of different response chains. For a given session, the correct chain consisted of responses to three particular response locations (out of a possible six); the particular elements constituting the correct sequence changed over sessions. The role of conditioned reinforcement was studied by presenting a distinctive stimulus on a given response location following each correct response in the sequence, and this distinctive stimulus remained illuminated until the entire sequence was completed. The

effect of the distinctive stimulus was then evaluated by omitting its presentation following some or all of the individual members of the chain. In general, stimulus presentations alone sustained discrimination acquisition even when food was omitted at the end of the chain, and omission of the stimuli decreased discrimination performance in proportion to the number of links of the chain for which the omission occurred. Moreover, when the stimulus was omitted from only one member of the chain, the deterioration in performance was greater for the response preceding that omitted stimulus presentation than for the others, demonstrating a clear response contingency effect. Thus, Hursh's results provide strong evidence for a conditioned-reinforcement effect. It was unclear, however, whether the effects of the stimulus depended upon its history of pairing with the food reinforcer.

The present study provides an additional investigation of the role of conditioned reinforcement in discrimination acquisition using a different procedure. To provide within-subject comparisons, a serial reversal procedure of a two-choice simultaneous conditional discrimination was used. Previous work with the reversal procedure has shown that the acquisition of repeated reversals, after improvement on the problem has stabilized, provides a stable baseline for studying a variety of procedural features (e.g., Williams, 1971, 1976). Of greatest relevance to the present research are those previous studies which varied the percentage of reinforcement for responses to the S+ (Williams, 1981, 1989). As expected, the number of errors required to reach an acquisition criterion was increased by smaller percentages of reinforcement, but the number of reinforcers required was approximately constant across the different percentages, an effect that was labeled "invariance in reinforcements to acquisition" (see also Eckerman, 1969; Gibbon, Farrell, Locurto, Duncan, & Terrace, 1980). The effects of percentage of reinforcement on serial reversal learning suggest a simple method of evaluating the role of conditioned reinforcers: These stimuli may be presented on the nonreinforced correct trials, and the effect of their presentation on the number of food reinforcers required for discrimination acquisition can then be assessed. To the extent that the conditioned reinforcers effectively substitute for primary reinforcers, the number of

food presentations (and errors) required to reach the discrimination criterion should be reduced in direct proportion to the number of conditioned reinforcers presented.

EXPERIMENT 1

In Experiment 1, we first attempted to determine the generality of the invariance effect with a new procedure. Rats were trained on a two-choice conditional discrimination in which a light and a noise served as discriminative stimuli, with the light cueing one response lever for a given session and the noise cueing the other. After that problem was learned to a criterion, the relation between the light/noise and particular lever was reversed for the next session. For some reversals the percentage of reinforcement for a correct trial was 100%; for other reversals it was 50%.

After this initial training, a third condition was added in which the conditioned reinforcer was presented on the nonreinforced correct trials. This third condition was interspersed with sessions involving the initial two conditions of 100% and 50% without conditioned reinforcement. The issue was whether the 50% condition with conditioned reinforcement would be more similar to the 100% or the 50% conditions. In the final set of conditions, the percentage of reinforcement was changed to 30%, and the corresponding comparisons were again assessed.

METHOD

Subjects

Four Sprague-Dawley albino rats, approximately 3 months of age at the beginning of the study, were housed individually with a 14:10 hr light/dark cycle. All subjects had histories consisting of lever-press acquisition training in a different apparatus under different delayed reinforcement contingencies, followed by FR 1 training in that other apparatus for several hundred reinforcers. Food deprivation was maintained by 1-hr access to Purina® lab chow immediately following the experimental session. Water was continuously available in the home cage.

Apparatus

A standard two-lever rat chamber, with glass side walls, sheet metal ceiling and front and rear walls, and a grid floor, was housed inside

a sound-attenuating larger chamber equipped with a ventilating fan. The interior of the chamber was 30.5 cm wide by 20.3 cm high by 22.9 cm long. Two retractable levers (BRS/LVE model RRL-015), which protruded 1.5 cm into the chamber when in operation and which required a minimum force of 0.3 N for depression, were mounted on the front wall of the chamber, spaced 9 cm apart measured from side to side. The only feedback for a lever press was the sound of the microswitch inside the lever housing. Directly between and 6.5 cm below the levers was a food receptacle into which dropped 45-mg Noyes pellets (improved Formula A) that served as the reinforcer. Between and 3.3 cm above the levers was a 28-V pilot light, covered by a glass translucent cover. A sonalert, which delivered an 88 dB auditory signal, was mounted 4 cm above the left lever. In the center of the ceiling was mounted a speaker through which 77-dB white noise could be presented. Ambient noise level in the absence of the tone or white noise was 72 dB.

Procedure

For the discrimination contingency, either the left or right lever was designated correct on a given trial, depending on whether the conditional cue was the white noise or illumination of the pilot light. For some reversals the noise signaled that the left lever was correct and the right lever was incorrect, whereas the light signaled the opposite contingencies. For the succeeding reversal the contingencies for the two cues were reversed.

After a 10-s intertrial interval (ITI), a trial began with the onset of one or the other conditional cue and the presentation of the response levers. Responses during the first 4 s had no effect in order to ensure that the subjects were adequately exposed to the conditional stimulus before their choice response; the location of the first response after 4 s had elapsed determined whether the trial was correct or incorrect. The levers were then withdrawn, the conditional cue was terminated, and the trial outcome was delivered. For reinforced correct trials, the outcome consisted of a 0.5-s presentation of the sonalert, followed by a single pellet. For incorrect trials and for nonreinforced correct trials, no differential feedback was presented.

Training on a given reversal continued until the subject had reached a criterion of 10 con-

secutively correct trials within a session. Training within a session was terminated either when this criterion was reached or when a total of 100 reinforcers had been delivered. When the latter occurred, training on the same contingencies was resumed the next day.

During Phase 1, all subjects were assigned to the 100% condition for the initial acquisition of the discrimination and were then subdivided into the 100% and 50% conditions for the first and second reversals. Because the first two reversals of serial reversal learning typically produce many more errors than later training, they were omitted from the data analysis. Subsequently, the two different reinforcement percentages (100% vs. 50%) were presented alternately every two reversals. Thus, the 100% condition was presented for two reversals, followed by the 50% condition for two reversals, and so forth. Training continued for an additional 24 reversals, 12 in each condition.

During Phase 2, the same two conditions were continued with the addition of a third condition in which the tone stimulus was presented on nonreinforced correct trials. The contingencies for this third condition (50+T) were otherwise exactly the same as for the 50% condition. Each of the three conditions was presented for two successive reversals; these blocks of two reversals were randomly interspersed with the restriction that each condition occur for two reversals out of each block of six. A total of 36 reversals occurred during Phase 2, 12 in each condition.

During Phase 3, the contingencies remained the same except that the percentage of correct trials ending in reinforcement was changed to 30%. Thus, three different conditions were again interspersed: 100%, 30%, and 30+T. Training during Phase 3 continued for 36 reversals, 12 in each condition.

RESULTS

Figure 1 shows the average number of errors per reversal (left side) and average number of reinforcers per reversal (right side) for each subject over the last 24 reversals of training in Phase 1. Results from the initial acquisition of the discrimination and from the first two reversals are omitted because performance at that time was highly variable across subjects and learning was much slower than for subsequent reversals, thus preventing any meaningful comparison across experimental con-

ditions (i.e., whichever condition was randomly assigned to the first two reversals had a much slower rate of learning). Even with the first two reversals omitted, learning during the first block of training (Figure 1) was still quite slow, typically requiring two or three sessions before the acquisition criterion was attained. Nevertheless a clear pattern emerged for all subjects. With respect to the number of errors per reversal, more errors occurred during the 50% condition throughout training for all 4 subjects. This observation was tested with a two-factor ANOVA (Percentage \times Blocks). The Percentage factor was significant, $F(1, 3) = 42.2, p < .05$, as was the Blocks factor, indicating significant improvement in learning across reversals, $F(3, 9) = 24.2, p < .05$. The interaction term was not significant.

The results for the number of reinforcers per reversal were more complex. Early in training there were considerably more reinforcers per reversal for the 100% condition for 3 of the 4 subjects, but this difference gradually decreased over training, until by the end of training no consistent difference was apparent. Again these observations were tested with a two-factor ANOVA. The effect of percentage of reinforcement approached but did not attain significance, $F(1, 3) = 9.11, .06 > p > .05$; the effect of blocks was significant, $F(3, 9) = 15.4, p < .05$, as was the interaction between percentage and blocks, $F(3, 9) = 4.54, p < .05$. Because of the significant interaction term, a test of simple effects was conducted, which showed that the effect of percentage of reinforcement was significant only during the second block of training. Thus, by the end of training, there was no reliable effect of percentage of reinforcement on the number of reinforcers per reversal, despite there being a robust effect of percentage of reinforcement on the number of errors per reversal.

Figure 2 shows the results after the addition of the third condition, which included presentations of the tone stimulus (the conditioned reinforcer) on the nonreinforced correct trials. The bars represent the mean of the 12 reversals under each procedure, because there was no clear trend over the course of training. The results for the number of errors per reversal were consistent over subjects. Many more errors were required for the 50% condition without the conditioned reinforcers than for the 100% condition. The 50+T condition was

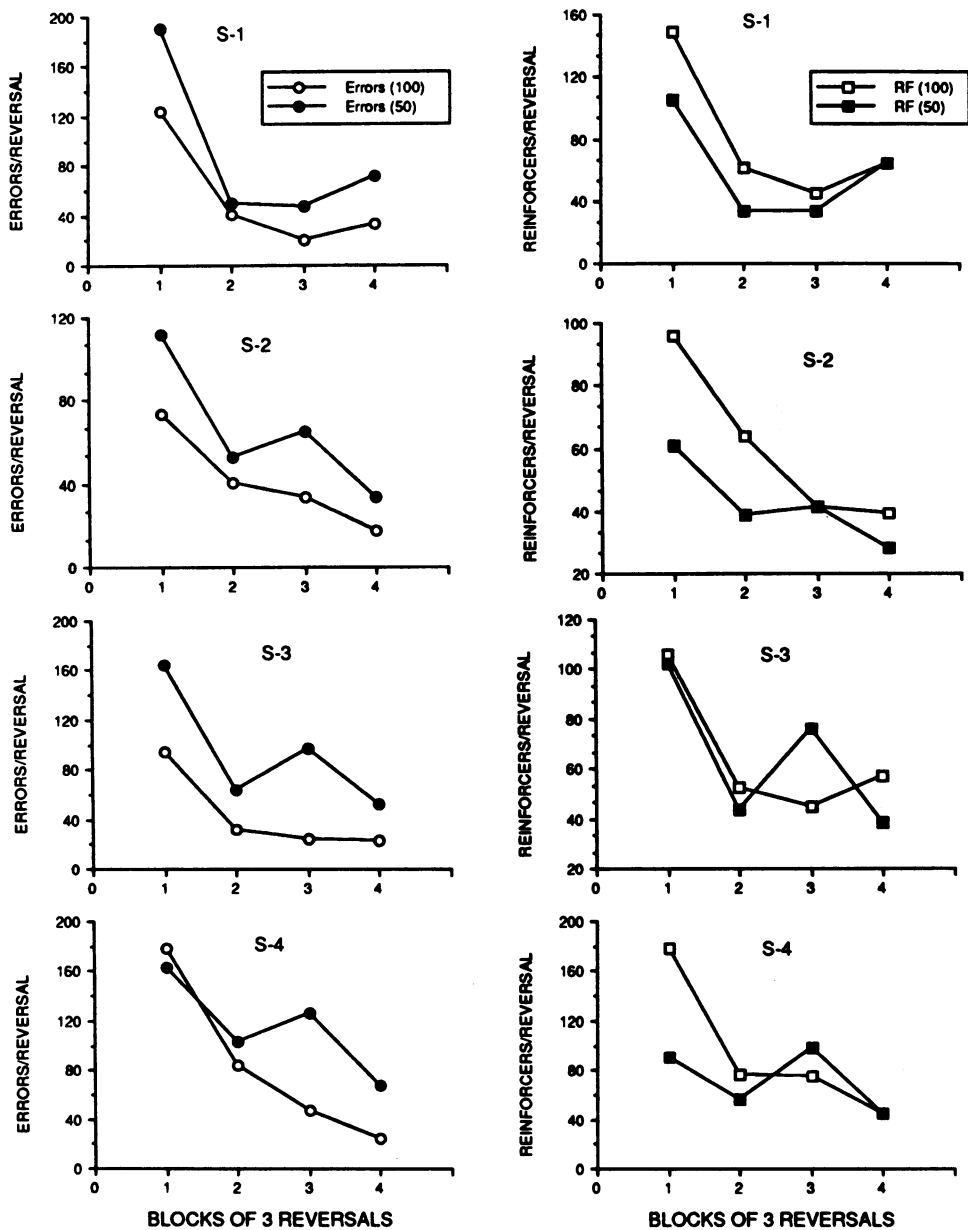


Fig. 1. Errors (left panels) and food reinforcers (right panels) required to reach the acquisition criterion across Phase 1 of Experiment 1. The abscissa refers to the number of reversals presented with each percentage of reinforcement. The correspondence of the 100% and 50% conditions to the different functions is shown in the legends.

much more similar to the 100% condition than to the 50% condition, although there were slightly fewer errors per reversal for the 100% than for the 50+T condition for 3 of the 4 subjects. These observations were tested with a one-way ANOVA. The effect of the different experimental conditions was significant, $F(2,$

$6) = 34.7, p < .05$. Paired comparisons between conditions using the Newman-Keuls test (with a .05 significance level for these and all subsequent tests) showed that the difference between the 100% and 50% conditions was significant, as was the difference between the 50% and 50+T conditions. The difference be-

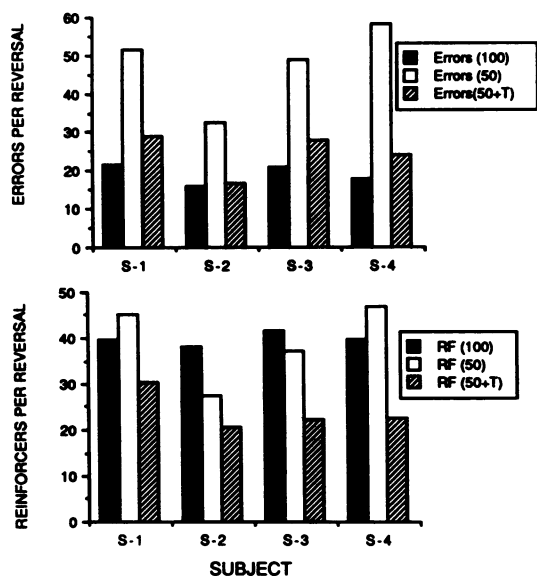


Fig. 2. Errors per reversal (top) and food reinforcers per reversal (bottom) averaged over all reversals presented in Phase 2 of Experiment 1.

tween the 100% and 50+T conditions was not significant.

The bottom portion of Figure 2 shows the results for the number of food reinforcers per reversal. The differences here were not as consistent across subjects. Considering first the comparison of the 100% versus 50% conditions, 2 of the subjects required a larger number of food presentations for the 100% condition, but 2 others exhibited the opposite pattern. Averaged over all subjects, the mean for the 50% condition was 39.1 and for the 100% condition was 39.8. Thus, the mean effect across subjects showed that the number of food reinforcers required for acquisition was approximately constant for the different percentages of reinforcement, consistent with the end of Phase 1 and previous studies. The number of food reinforcers required for the 50+T condition was smaller than for either of the other two conditions for all subjects, with an average of 24.0. These differences were tested statistically with a one-way ANOVA. The effect of conditions was significant, $F(2, 6) = 13.3$, $p < .05$. Comparisons of the differences between pairs of means with the Newman-Keuls test revealed that the difference between the 50% and 100% conditions was not significant, but the differences between the 50+T condition and each of the remaining conditions were.

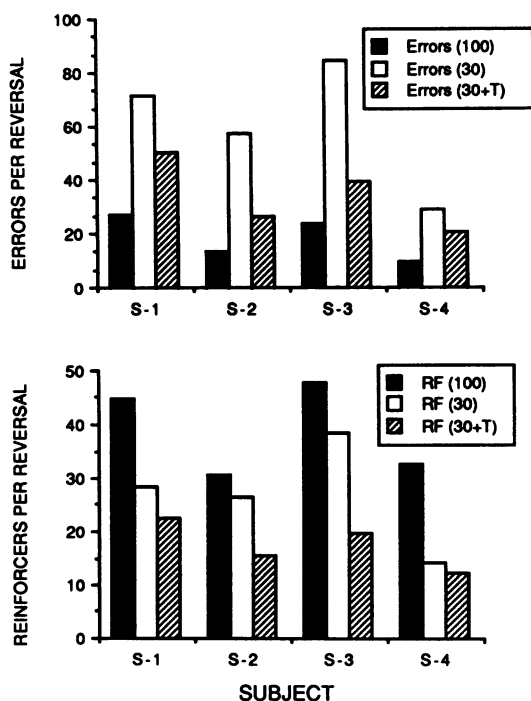


Fig. 3. Errors per reversal (top) and food reinforcers per reversal (bottom) averaged over all reversals presented in Phase 3 of Experiment 1.

Figure 3 shows the results when the percentage of reinforcement was reduced to 30% from 50%. Again, the bars represent the mean of all 12 reversals in each condition because there was no trend over training. The number of errors was substantially greater for the 30% condition than for the 100%; the results for the 30+T condition were intermediate. This intermediate pattern contrasts with the results shown in Figure 2, where the 100% and 50+T conditions were much more similar. The results shown in the top portion of Figure 3 were subjected to a one-way ANOVA. The overall effect of conditions was significant, $F(2, 6) = 18.99$, $p < .05$. Paired comparisons using the Newman-Keuls test showed that the difference between the 100% and 30% conditions was significant, as was the difference between the 30% and 30+T conditions. However, the difference between the 100% and 30+T conditions was not significant.

The bottom portion of Figure 3 shows the results for the food reinforcers per reversal. For all subjects, the 100% condition required the greatest number of reinforcers, and the 30+T condition required the fewest. Averaged

over subjects the mean values were 38.9, 26.9, and 17.5 for the 100%, 30%, and 30+T conditions, respectively. Thus, unlike the results in Phase 2, here the invariance effect did not occur. These results were also analyzed with a one-way ANOVA. The overall effect of conditions was significant, $F(2, 6) = 22.7$, $p < .05$. Paired comparisons with the Newman-Keuls test showed that the differences between all possible pairs of conditions were significant.

Comparison of Figures 2 and 3 suggests that the tone stimulus was more substitutable for food with the 50% reinforcement schedule than with the 30% schedule. To quantify this observation, an index of the trade-off between the number of nonreinforced tone presentations and the reduction in food reinforcers needed for acquisition produced by these tone presentations is necessary. Two methods of calculating this trade-off are available. First, comparing only the 100% and 50+T conditions, the difference in food reinforcers per reversal is divided by the number of nonreinforced tone presentations. The second method of calculation involves the comparison of the 50% and 50+T conditions. The difference in number of food reinforcers per reversal is again divided by the number of tone presentations. The two methods of calculation will yield exactly the same results as long as the reinforcers per reversal are exactly the same for the 100% and 50% conditions. Together the two measures offer an estimate of the substitutability of the tone for food.

Table 1 shows the results of these calculations for individual subjects under the 50% and 30% reinforcement schedules. Substantial variance is evident for both types of calculations, but in general the degree of substitutability was higher for the 50% schedule than for the 30% schedule.

DISCUSSION

The results of Experiment 1 clearly establish that a stimulus with a history of pairing with food facilitates discrimination acquisition when presented on correct trials without food reinforcement, both in terms of the number of errors required per reversal and in terms of the number of food reinforcers per reversal. Relative to the conditions in which the same percentage of food reinforcement was presented but without the stimulus presentations, the reduction was substantial. The results thus

Table 1

Substitutability of the tone for food in Experiment 1. Separate calculations are shown for the comparison of the 50+T condition with the 100% and 50% conditions, and for the comparison of the 30+T with the 100% and 30% conditions.

Subject	100	50	<i>M</i>	100	30	<i>M</i>
	vs. 50+T	vs. 50+T		vs. 30+T	vs. 30+T	
S-1	0.31	0.51	0.41	0.38	0.10	0.24
S-2	0.87	0.33	0.60	0.43	0.31	0.37
S-3	0.86	0.66	0.76	0.62	0.41	0.52
S-4	0.76	1.07	0.92	0.70	0.07	0.39
<i>M</i>	0.70	0.64	0.67	0.53	0.22	0.38

demonstrate that such stimulus presentations can effectively substitute for the food reinforcers. The degree of this substitutability varied with the percentage of food-reinforced trials and was quite high in the 50% condition but was substantially lower in the 30% condition.

The results of Experiment 1 also have implications for the generality of the invariance-in-reinforcements-to-acquisition effect. The effect failed to occur early in training, as significantly fewer reinforcers were required for the 50% condition than for the 100% condition. But by the end of Phase 1, the mean number of reinforcers per reversal was highly similar for the two conditions, indicating that the effect occurred here, as it has in other reversal learning situations (Williams, 1981, 1989). The similarity between the 100% and 50% conditions continued to occur during Phase 2, at least with respect to the means across subjects, but clearly did not occur in Phase 3 when the percentage of reinforcement was reduced to 30%. Thus, the invariance effect in the present study was not as robust as has been obtained in previous studies. It is noteworthy that the direction of the deviation from invariance that was obtained was consistent with previous work (Williams, 1989), in that fewer reinforcers per reversal were required with the smaller percentages of reinforcement. This effect is surprising, because the introduction of nonreinforced correct trials should diminish the information value of nonreinforced trials, or, from an alternative perspective, counteract the response-strengthening effects of the reinforced correct trials. Just why the reductions in percentage of reinforcement actually decrease the required number of reinforcers for

acquisition can only be determined by further investigation.

EXPERIMENT 2

The results of Experiment 1 clearly established the effectiveness of the tone stimulus in facilitating discrimination learning, but they did not address whether this facilitory effect depended upon the tone having a history of being paired with the food presentations. Experiment 2 thus repeated the basic comparisons of Experiment 1 except that the tone presentations were not paired with food. In the first phase, on food-reinforced trials a correct response was followed immediately by delivery of the food pellet, so that the tone occurred only on nonreinforced correct trials. At issue was whether these tone presentations would facilitate discrimination learning. It is plausible that such facilitation might occur if the tone is regarded as an informative stimulus, especially in view of other results showing strong effects of contingent stimuli without those stimuli being paired with reinforcement (e.g., Neuringer & Chung, 1967).

METHOD

Subjects and Apparatus

Four new subjects, similar in history to those used in Experiment 1, served. The apparatus was also the same as in Experiment 1.

Procedure

The general procedure was identical to that used in Experiment 1 except that the tone presentations were omitted on food-reinforced correct trials, which meant that a correct response was followed immediately by the delivery of the food pellet. The only other change in procedure was that trial outcome was determined by the first response after 2 s had elapsed, instead of after 4 s as in Experiment 1.

Initially, three experimental conditions were presented for two consecutive reversals within each block of six reversals. The 100% condition was like that of Experiment 1 except that correct trials were followed immediately by food reinforcement without the intervening tone presentation. The 50% condition was also like that of Experiment 1 except that food immediately followed the correct responses on a randomly determined 50% of the correct trials, whereas the feedback for a nonreinforced cor-

rect trial was the same as for incorrect trials (withdrawal of the levers and the onset of the ITI). The 50+T condition was like the 50% condition except that the 0.5-s presentation of the sonalert signal was presented only on the nonreinforced correct trials. At no time was the tone paired with food. Training continued during this phase for a total of 48 reversals, 16 in each condition.

Following training with the tone unpaired with food, the conditions of Experiment 1 were instituted during the second phase to investigate whether tone-food pairings would establish the tone as an effective conditioned reinforcer. Only the 50% and 50+T conditions of Experiment 1 were presented, each for alternate blocks of two reversals. A total of 32 reversals were presented, 16 in each condition.

RESULTS

The course of improvement across reversals is shown in Figure 4, averaged over subjects. Mean data are presented because there was substantial variability in the rate of learning from reversal to reversal, with the result that the functions for individual subjects were somewhat erratic across blocks. However, the mean data are representative of major trends obtained with the individual subjects. For both the errors per reversal and the reinforcers per reversal, there was a substantial increase in the rate of learning between the first and second block of reversals and relatively little change thereafter. Unlike Experiment 1, the differences between the conditions were not well defined, although there was a strong tendency for the 100% condition to have fewer errors per reversal throughout training.

For purposes of statistical analysis, the data from the first block of sessions shown in Figure 4 were excluded, leaving the remaining 36 reversals (12 per condition) for which there was little consistent change with continued training. These results are shown in Figure 5 for individual subjects. Considering first the errors per reversal, all subjects had fewer errors in the 100% condition than in the 50% condition, although for S-7 this difference was very small. In general the errors in the 50+T condition were similar to the 50% condition; there was no consistent difference across subjects produced by the presence of the tone on nonreinforced correct trials. Averaged over subjects, the mean number of errors per reversal was

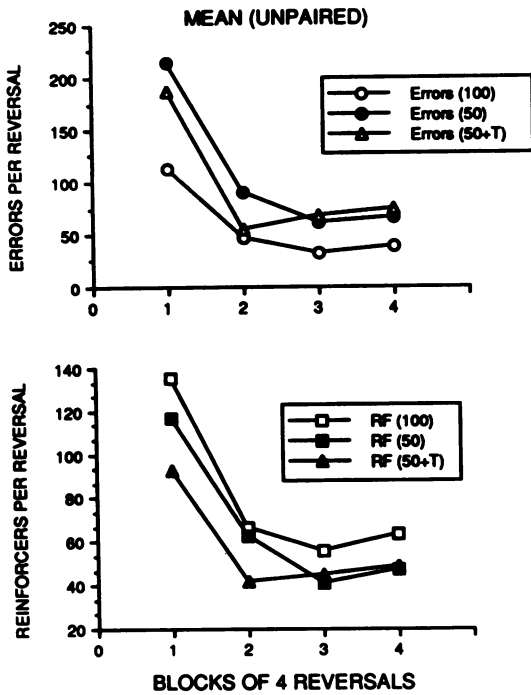


Fig. 4. Mean results across subjects with respect to the errors (top) and food reinforcers per reversal (bottom) for different blocks of reversals during Phase 1 of Experiment 2.

33.9, 64.8, and 65.6, for the 100%, 50%, and 50+T conditions, respectively. These differences were analyzed with a one-way ANOVA. The overall effect of conditions was significant, $F(2, 6) = 10.77, p < .05$. Paired comparisons with the Newman-Keuls test showed that differences between the 100% condition and both of the 50% conditions were significant, but the difference between the 50% and 50+T condition was not.

The bottom portion of Figure 5 reveals considerable variability across subjects in the number of food reinforcers required to reach the discrimination criterion. Averaged over subjects the mean number of reinforcers per reversal was 57.9, 45.3, and 45.0 for the 100%, 50%, and 50+T conditions, respectively. A one-way ANOVA showed that the effect of conditions was not significant ($F < 1$). Thus, neither for errors per reversal nor for reinforcers per reversal was there any significant effect of the tone presentations on nonreinforced correct trials.

The results shown in Figures 4 and 5 are notably less orderly than those shown in Ex-

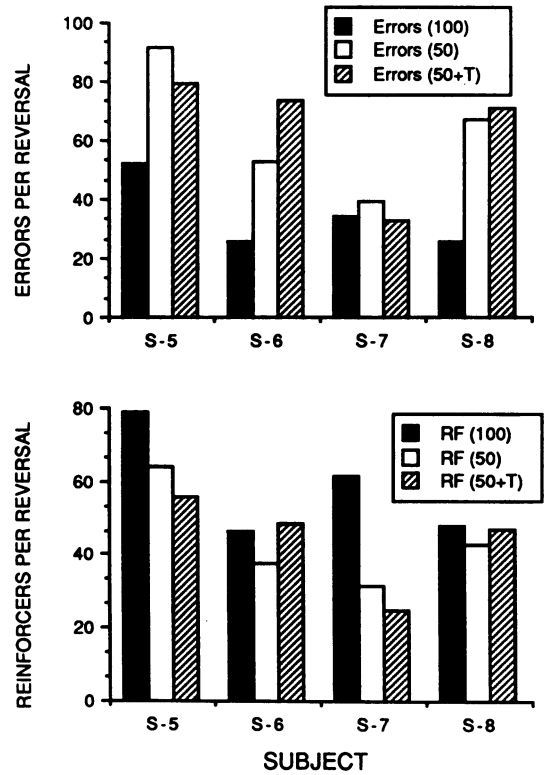


Fig. 5. Results for individual subjects in Phase 1 of Experiment 2 for the last 36 reversals of training.

periment 1. One possible reason is that a choice response was possible after only 2 s of exposure to the conditional stimulus, rather than the 4 s used in Experiment 1. Whatever the reason for the less orderly data in Experiment 2, it is possible that the failure to obtain any facilitatory effect of the tone was due to the increased amount of experimental noise rather than to the lack of reinforcement efficacy of an unpaired stimulus. It was thus essential to establish that the tone would facilitate discrimination learning under the present conditions after the tone had been paired with food.

Figure 6 shows the mean results across subjects after the tone was also presented on reinforced correct trials. Only the 50% and 50+T conditions were presented. During the first block of reversals there continued to be no consistent effect of the tone presentations, but with continued training both the errors per reversal and reinforcers per reversal showed a substantial reduction. Because Figure 6 shows no major changes in the rate of learning after the first block of four reversals per condition,

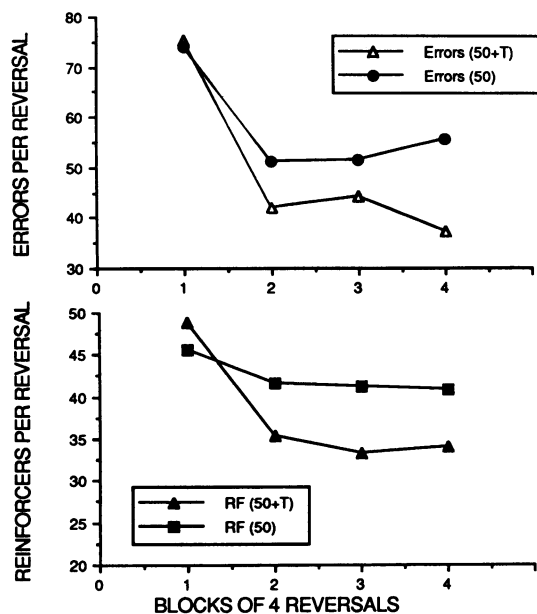


Fig. 6. Mean results across subjects for errors per reversal and food reinforcers per reversal for the successive blocks of training in Phase 2 of Experiment 2, after instituting tone-food pairings.

the data for the remaining reversals were averaged for each individual subject and are presented in Figure 7. For all subjects, fewer errors and reinforcers were required in the 50+T condition than in the 50% condition, indicating that presentation of the tone stimulus facilitated learning. For both measures these differences were statistically significant. For errors per reversal, the mean difference was 11.7, $t(3) = 11.7$, $p < .05$; for reinforcers per reversal, the mean difference was 7.0, $t(3) = 3.86$, $p < .05$.

DISCUSSION

In contrast to the results of Experiment 1, the presentation of the tone stimulus on non-reinforced correct trials had no consistent effect on the acquisition of the repeated reversals of the conditional discrimination when the tone was itself not paired with food. But when the tone-food pairings were instituted in Phase 2, facilitation did occur. Thus, the facilitation seems clearly to depend on the associative status of the tone and not on perceptual marking, "information," or any of a number of other conceivable functions that a neutral stimulus might possess. The results thus strongly support the view that the tone's effect depended

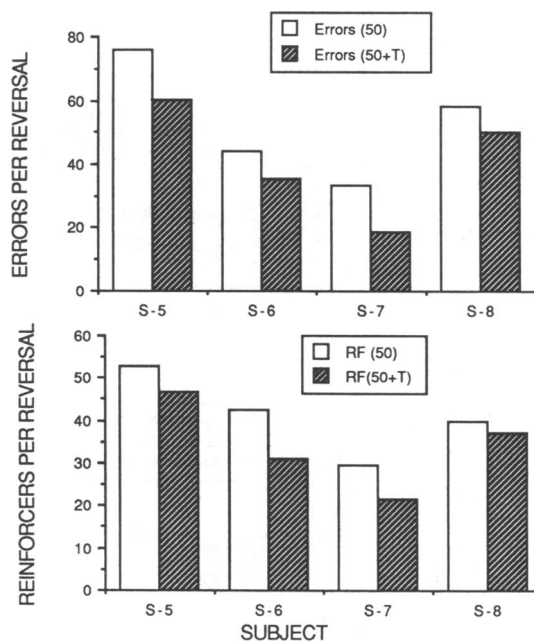


Fig. 7. Results for individual subjects in Phase 2 of Experiment 2. Data are from the last 24 reversals of training (corresponding to Blocks 2-4 of Figure 6).

on it being a conditioned reinforcer. This conclusion is supported by additional studies with a variety of different parameters (e.g., duration of the tone, length of the ITI) that we do not report here, all of which failed to find any facilitatory effect of the tone when it was not paired with food.

EXPERIMENT 3

Experiment 2 conflicts with several previous studies that have suggested that unpaired stimuli significantly affect the behavior on which they are contingent (e.g., Neuringer & Chung, 1967). The great majority of such studies have used pigeons as subjects, so it is of some interest to determine whether the same pattern of results reported in Experiments 1 and 2 can be obtained with a different preparation. Experiment 3 was thus a conceptual replication of Experiments 1 and 2 but with pigeons as subjects. The major change in design was that the comparison between the paired and unpaired conditions was accomplished by using two different brief stimuli, one paired with food and one unpaired, both of which were presented to the same subjects.

METHOD

Subjects

Three male White King pigeons, maintained at 85% of their free-feeding weights, served as subjects. When necessary, supplemental feedings were given approximately 4 hr after the experimental sessions. Water and grit were freely available in the home cages. All subjects had prior experience on a conditional discrimination task.

Apparatus

The experimental chambers were cubes, 32 cm on a side. One side panel was a Plexiglas door; the remaining sides and ceiling were aluminum. The chambers were housed in wooden enclosures. Three translucent response keys were mounted on the front panel, each 2.5 cm in diameter and evenly separated 24 cm above the grid floor. The keys could be transilluminated with various colors and required a minimum force of approximately 0.16 N for operation. The food hopper opening was located 9 cm beneath the center key. When activated, the solenoid-operated hopper was illuminated by white light and allowed access to mixed grain. A houselight mounted in the center of the ceiling provided general chamber illumination except during operation of the hopper. Stimuli, contingencies, and data collection were controlled by an XT-compatible computer with Turbo-Pascal® software.

Procedure

The conditional discriminations alternated between matching-to-sample and oddity tasks. After a 30-s ITI, a trial began with presentation of either red or green as the sample on the center key. Upon completion of an FR 10 schedule, the center key was darkened and the side keys were illuminated with the comparison stimuli. One side key was red and the other was green, with their respective locations varied randomly over trials. On the matching task, a single response to the comparison key of the same color as the sample was correct. On the oddity task, a response on the comparison key with the color that differed from the sample was correct. When reinforced, correct responses were followed by a 1-s presentation of yellow on the correct side key and 3-s access to food. Incorrect responses were fol-

lowed by a return to the ITI with 4 s added to the duration of the ITI.

Nonreinforced correct trials were followed by one of three outcomes. In the no-feedback (NF) conditions, nonreinforced correct responses were followed by a return to the ITI with 4 s added (the same contingencies as for incorrect responses). In paired (P) conditions, nonreinforced correct responses were followed by the 1-s presentation of yellow on the side key and a return to the ITI with 3 s added. In unpaired (UP) conditions, nonreinforced correct responses were followed by a 1-s presentation of blue on the side key and a return to the ITI with 3 s added. For 1 subject, Bird 13, the yellow and blue color assignments were reversed.

Each stimulus condition was presented with three percentage reinforcement schedules: 75%, 50%, and 25%. In addition, a condition with a 100% reinforcement schedule was presented. The parameters for each condition (task type, stimulus contingency, and percentage reinforcement) were selected randomly without replacement until all possible conditions had been completed four times. Sessions continued for 50 reinforcers or until the acquisition criterion of 10 consecutively correct trials. Two sessions were presented each day, separated by a 6-hr intersession interval.

RESULTS

Figure 8 shows the results for the NF condition for individual subjects across the four different percentages of reinforcement. Each data point represents the mean of the four reversals presented with each percentage. The top panel shows the data for errors per reversal; the bottom panel shows the number of reinforcers per reversal. The pattern of results is exactly opposite for the two measures. The largest number of errors occurred during the 25% reinforcement schedule, and the smallest number occurred for the 100% condition. The results for the 50% and 75% conditions were intermediate and similar to each other. Averaged over subjects, the mean numbers of errors were 254, 371, 362, and 715, for the 100%, 75%, 50%, and 25% conditions, respectively. For reinforcers per reversal, the differences were generally smaller but also were monotonically related to percentage of reinforcement, with the exception of one reversal in trend for Bird 4. Averaged over subjects, the

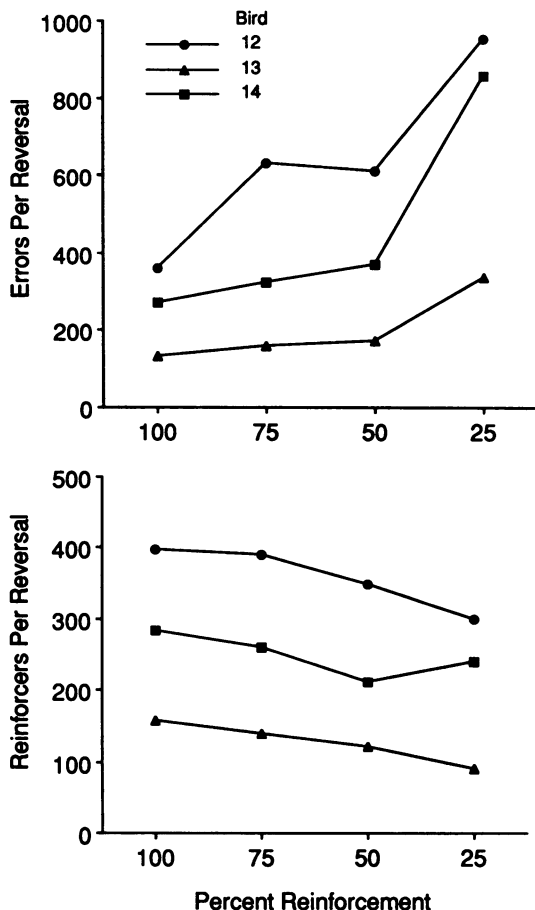


Fig. 8. Results for individual subjects of the no-feedback conditions in Experiment 3 for the different percentages of reinforcement. The top panel shows the errors per reversal. The bottom panel shows the food reinforcers per reversal.

mean numbers of reinforcers were 280, 263, 214, and 210 for the 100%, 75%, 50%, and 25% conditions, respectively. To evaluate these effects statistically, separate ANOVAs were conducted for the error and reinforcer measures. For both, the effect of reinforcement percentage was significant: errors, $F(3, 6) = 9.31, p < .05$; reinforcers, $F(3, 6) = 5.12, p < .05$.

Figure 9 shows the effects on errors per reversal of the presentations of the additional stimuli following correct trials not followed by food. Separate panels correspond to the different percentages of reinforcement, and separate measures are presented for each of the three stimulus conditions (NF, P, and UP). Comparing first the results from the NF con-

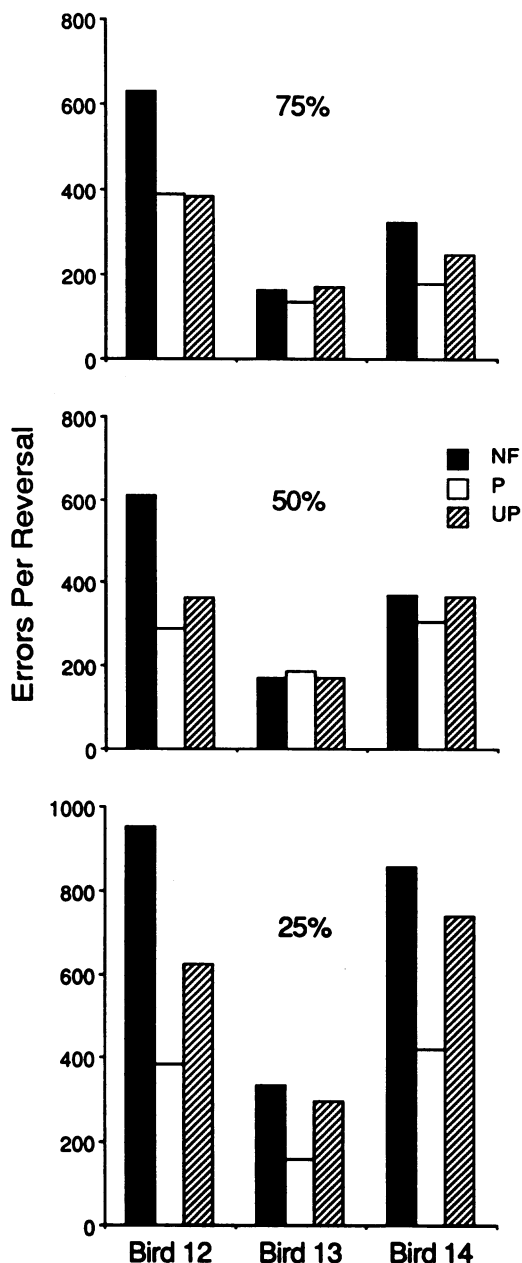


Fig. 9. Errors per reversal for individual subjects in the three different stimulus conditions in Experiment 3. Different rows correspond to the different percentages of reinforcement. Each bar corresponds to the mean of four reversals per condition.

ditions with those from the P conditions, in all but one case the presentation of the paired stimulus reduced the number of errors. Thus, as in Experiments 1 and 2, the paired stimulus

substantially substituted for food in facilitating the discrimination. In all but two cases, the numbers of errors during UP conditions were intermediate between those in the NF and P conditions. The difference between the number of errors in the P and UP conditions was greatest in conditions with 25% reinforcement. Because the individual subjects differed greatly in the overall number of errors, for purposes of statistical analysis the data for individual subjects were normalized by converting the error measures for the different conditions into percentages of those that occurred in the NF conditions. Averaged over subjects and the three percentage reinforcement conditions, these values were 100%, 59.6%, and 81.8% for the NF, P, and UP conditions, respectively. A simple one-way ANOVA showed that there was a significant effect of conditions, $F(2, 4) = 14.8$, $p < .05$. Comparison between individual conditions using the Newman-Keuls test showed that the P condition was significantly different from both the NF and UP conditions, but the NF and UP conditions were not significantly different.

Figure 10 shows the corresponding results with the measure of food reinforcers per reversal. As with the errors measure, the comparison of the NF and P conditions showed that paired stimuli reduced the number of reinforcers required to attain the discrimination criterion. Again the results for the conditions with the unpaired stimulus presentations were intermediate. The numbers of reinforcers per reversal in the UP conditions were lower than those in the NF condition in all but one comparison and higher than those in the P condition, again with only one exception. The difference between the P and UP conditions was most pronounced in the conditions with 25% reinforcement. Again for statistical analysis the results were normalized with respect to the number of reinforcers that occurred during the NF conditions. Averaged over the different percentages the means across subjects were 100%, 65.5%, and 83.0% for the NF, P, and UP conditions, respectively. A one-way ANOVA showed that the effect of conditions was significant, $F(2, 4) = 13.5$, $p < .05$. Comparisons between individual conditions using the Newman-Keuls test showed that the P condition was significantly different from the NF conditions, but none of the remaining comparisons reached significance.

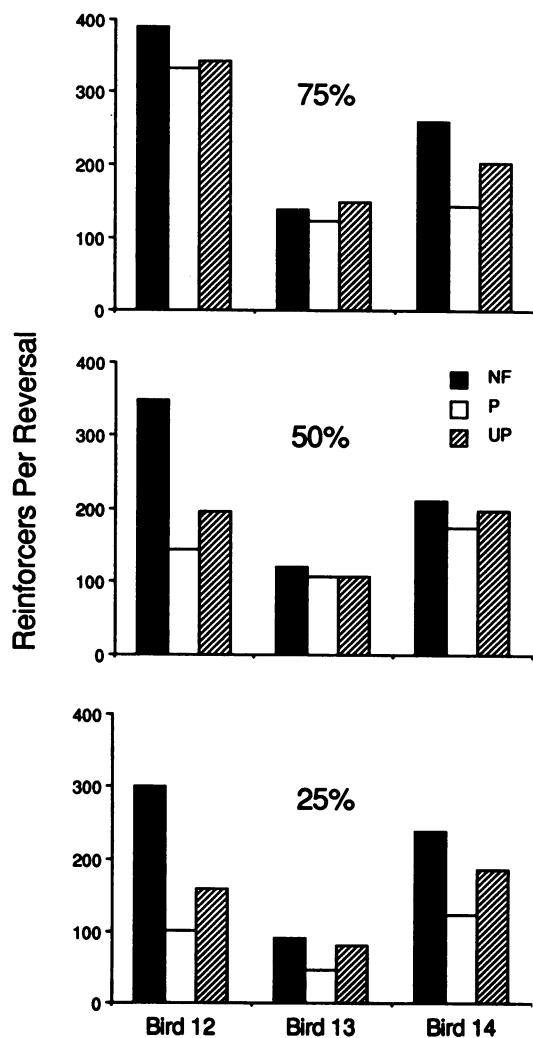


Fig. 10. Food reinforcers per reversal for individual subjects in the three different stimulus conditions in Experiment 3. Different rows correspond to the different percentages of reinforcement. Each bar corresponds to the mean of four reversals per condition.

DISCUSSION

The results of Experiment 3 were similar to those of Experiments 1 and 2 despite the change in subjects and discrimination task. Presentations of a stimulus on nonreinforced correct trials that previously had been paired with food facilitated discrimination acquisition in all conditions, indicating that the stimulus was a conditioned reinforcer that effectively substituted for the food. On the other hand, presentation of a stimulus not paired with food on nonreinforced correct trials also appeared

to facilitate the discrimination, but to a lesser extent. It is likely that this facilitation was due to stimulus generalization between the two different stimuli. That is, both the paired stimulus, which occurred on all food reinforced trials regardless of the condition, and the unpaired stimulus were colored lights presented on the response keys. It is thus plausible that the unpaired stimulus received some degree of "reinforcement value" because of its similarity to the paired stimulus.

The results of Experiment 3 are also relevant to the generality of the invariance in reinforcements to acquisition. Whereas Experiments 1 and 2 showed an approximation of the effect under some conditions (Phases 1 and 2 of Experiment 1) but not under others (the 30% conditions of Experiment 1), here the invariance effect was clearly violated for all 3 subjects. In contrast to the previous findings showing that the number of reinforcers required to produce an acquisition criterion was approximately constant regardless of the percentage of correct trials that are reinforced (cf. Williams, 1981, 1989), here substantially fewer reinforcers were required with the smaller percentages. Why the effect occurs under some circumstances and not others is unclear. One significant difference between the present study and previous work showing the effect was the difficulty of the discrimination. In previous work, the discriminations were substantially easier than those used here. For example, Williams (1981) presented pigeons with successive reversals of a simple simultaneous color discrimination, in contrast to the matching/odddity problem used here in Experiment 3. Why difficult discriminations should cause smaller percentages of reinforcement to become relatively more effective is uncertain, but one possibility involves the role of responding to irrelevant cues such as spatial position. It is possible that larger percentages of reinforcement are correlated with stronger position habits; these position habits then persist longer with the larger percentages, thus requiring more reinforcers for the discrimination criterion to be attained. At present, however, such an explanation of the present deviation from the invariance effect can only be speculative (but see Weaver & Michels, 1961, and Williams, 1977, for related effects). Regardless of their explanation, the counterintuitive nature of these results bears reemphasis: Smaller per-

centages of reinforcement of correct responses result in fewer reinforcers required for discrimination to occur. These results pose a substantial challenge for any general explanation of discrimination learning.

GENERAL DISCUSSION

The results of all three experiments taken together provide strong support for the necessity of the concept of conditioned reinforcement. When a stimulus was presented on nonreinforced correct trials, learning was facilitated, but only if that stimulus was paired with food on other trials or was similar to the paired stimulus. Moreover, the degree of substitutability of the stimulus for food was dependent on the percentage of stimulus presentations paired with food (Phases 2 vs. 3 of Experiment 1), as would be expected by the view that the conditioned reinforcement properties of the stimulus were due to its Pavlovian relationship to food delivery.

Unlike many previous results using free-operant procedures in which response rate or pattern has been the dependent variable, the present effects cannot be explained by the discriminative properties of the stimulus. Here the dependent measure was the rate of learning, and there was no cue relationship between the stimulus following choice on one trial and the correct choice on the next trial. According to accounts of conditioned reinforcement effects that ascribe the stimulus effects to the discriminative effects of the stimulus (e.g., Schuster, 1969), the nonreinforced presentation of the stimulus should have been "frustrating" in its effects, because it presumably provided a cue that reinforcement was available when in fact it was not delivered. There is no reason, according to such accounts, to suppose that these effects should facilitate learning. But in fact the presentation of the stimulus did produce a strong enhancement of the response on which it was contingent.

The present results are consistent with a variety of others showing that conditioned reinforcement contingencies affect choice behavior, both in terms of acquisition of simultaneous discriminations (Hursh, 1977) and in the sustained maintenance of preference for the choice alternative associated with the conditioned reinforcer contingency (Clark & Sherman, 1970; Nevin & Mandell, 1978).

These results, combined with other previous findings (Dunn, Williams, & Royalty, 1987; Royalty et al., 1987), leave little doubt that the concept of conditioned reinforcement is essential for the proper analysis of many different behavioral situations and cannot be displaced by alternative theoretical concepts, as suggested by some previous investigators (Baum, 1973; Longstreth, 1971; Schuster, 1969). Stimuli paired with primary reinforcers such as food acquire the properties of reinforcers themselves.

REFERENCES

- Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, *20*, 137-153.
- Clark, H. B., & Sherman, J. A. (1970). Effects of a conditioned reinforcer upon accuracy of match-to-sample behavior in pigeons. *Journal of the Experimental Analysis of Behavior*, *13*, 375-384.
- Cohen, S. L., Calisto, G., & Lentz, B. E. (1979). Separating the reinforcing and discriminative properties of brief-stimulus presentations in second-order schedules. *Journal of the Experimental Analysis of Behavior*, *32*, 149-156.
- Dunn, R., Williams, B., & Royalty, P. (1987). Devaluation of stimuli contingent on choice: Evidence for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, *48*, 117-131.
- Eckerman, C. O. (1969). Probability of reinforcement and the development of stimulus control. *Journal of the Experimental Analysis of Behavior*, *12*, 551-559.
- Fujita, K. (1985). Effects of ratio reinforcement schedules on discrimination performance by Japanese monkeys. *Journal of the Experimental Analysis of Behavior*, *43*, 225-234.
- Gibbon, J., Farrell, L., Locurto, C. M., Duncan, H. J., & Terrace, H. S. (1980). Partial reinforcement in autoshaping with pigeons. *Animal Learning & Behavior*, *8*, 45-59.
- Hursh, S. R. (1977). The conditioned reinforcement of repeated acquisition. *Journal of the Experimental Analysis of Behavior*, *27*, 315-326.
- Longstreth, L. E. (1971). A cognitive interpretation of secondary reinforcement. In J. K. Cole (Ed.), *Nebraska symposium on motivation* (Vol. 19, pp. 33-80). Lincoln: University of Nebraska Press.
- Neuringer, A. J., & Chung, S. H. (1967). Quasi-reinforcement: Control of responding by a percentage-reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, *10*, 45-54.
- Nevin, J. A., & Mandell, C. (1978). Conditioned reinforcement and choice. *Journal of the Experimental Analysis of Behavior*, *29*, 135-148.
- Royalty, P., Williams, B. A., & Fantino, E. (1987). Effects of delayed conditioned reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior*, *47*, 41-56.
- Schuster, R. H. (1969). A functional analysis of conditioned reinforcement. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 192-234). Homewood, IL: Dorsey Press.
- Spence, K. W. (1947). The role of secondary reinforcement in delayed reward learning. *Psychological Review*, *54*, 1-8.
- Stubbs, D. A., Vautin, S. J., Reid, H. M., & Delehanty, D. L. (1978). Discriminative functions of schedule stimuli and memory: A combination of schedule and choice procedures. *Journal of the Experimental Analysis of Behavior*, *29*, 167-180.
- Thomas, J. R. (1979). Matching-to-sample accuracy on fixed-ratio schedules. *Journal of the Experimental Analysis of Behavior*, *32*, 183-189.
- Weaver, L. A., Jr., & Michels, K. M. (1961). Methodological factors affecting the formation of learning sets in rats. *Animal Behavior*, *9*, 4-7.
- Williams, B. A. (1971). The effects of intertrial interval on discrimination reversal learning in the pigeon. *Psychonomic Science*, *23*, 241-243.
- Williams, B. A. (1976). Short-term retention of response outcome as a determinant of serial reversal learning. *Learning and Motivation*, *7*, 418-430.
- Williams, B. A. (1977). Contrast effects in simultaneous discrimination learning. *Animal Learning & Behavior*, *5*, 47-50.
- Williams, B. A. (1981). Invariance in reinforcements to acquisition, with implications for the theory of inhibition. *Behaviour Analysis Letters*, *1*, 73-80.
- Williams, B. A. (1989). Partial reinforcement effects on discrimination learning. *Animal Learning & Behavior*, *17*, 418-432.

Received May 10, 1990
Final acceptance August 24, 1990