IMMEDIATE REINFORCEMENT IN DELAYED REWARD LEARNING IN PIGEONS

JANET WINTER AND CHARLES C. PERKINS

KANSAS STATE UNIVERSITY

Pigeons were trained on simultaneous red-green discrimination procedures with delayed reward and sequences of stimuli during the delay. In Experiment 1, three stimuli appeared during the 60-second intervals between the correct responses and reward, and the incorrect responses and nonreward. The stimulus that immediately followed a correct response also preceded nonreward, and the stimulus that followed an incorrect response preceded reward. These stimuli were 10 or .33 second in duration for different groups. Stimuli during the remainder of the delay interval differed following correct and incorrect responses. Group 10 initially persisted in the nonrewarded choice, but shifted to a preponderance of rewarded responses after further training. Group .33 rapidly acquired the correct response. Similar results were obtained in Experiment 2 where delay intervals consisted of opposite sequences of two stimuli of equal duration and total delays were 6, 20, or 60 seconds. Early in training, generalization of differential conditioned-reinforcing properties from the conditions preceding reward and nonreward to postchoice conditions had a greater effect relative to backchaining than it did later. It was concluded that delayed-reward learning is best analyzed in terms of the conditioned-reinforcing value of the patterns of cues that follow immediately after rewarded and nonrewarded responses.

Key words: delayed reward, conditioned reinforcement, discrimination learning, pattern of cues, key peck, pigeon

The extent to which delayed rewards have a reinforcing effect is greatly influenced by events during the delay interval. A consistent interpretation of these effects seems to follow if it is assumed that the effect of delayed rewards results from acquisition of differential conditioned reinforcing value by the cue patterns that follow immediately after correct and incorrect responses. In a discrimination learning procedure with delayed reward, the pattern of cues that immediately follows the correct response may acquire greater conditioned-reinforcing value than the cue pattern that follows the incorrect response by two different mechanisms. One is direct generalization of differential conditioned-reinforcing value from the cue patterns preceding reward and nonreward to the cue patterns following the choice responses. The other is backchaining of differential attractiveness through two different sequences of cue patterns.

These mechanisms can be better described if a distinction is made between what we will

refer to as "patterns of cues" and what we will refer to as "sets of stimuli." A set of stimuli refers to all the external environmental conditions present at the moment. A pattern of cues is more inclusive and consists not only of environmental stimuli but also includes internal or inferred cues such as traces and temporal cues (readings on internal clocks). Thus, two patterns of cues may be discriminated by the subject even though each includes the identical set of stimuli. That trace cues are discriminable is illustrated, for example, by delayed matching to sample. The importance of temporal cues has been directly demonstrated by temporal discriminations as well as by scalloped fixed-interval cumulative records and Pavlovian "inhibition of delay" (Gibbon, 1977; Killeen, 1975).

Backchaining is illustrated by Napalkov's (1959) experiment cited by Kelleher and Golub (1962) in which a pigeon was trained to make a sequence of seven different responses each of which produced a different stimulus. The entire sequence terminated in food from which the antecedent stimulus (S_1 , a white light) acquired conditioned reinforcing properties transmitted in turn to S_2 , etc. It is here

Reprints may be obtained from Charles C. Perkins, Department of Psychology, Kansas State University, Manhattan, Kansas 66506.

assumed that conditioned reinforcing value may also be transmitted back through a sequence of stimuli that occur in sequence even when no response is required to produce the next stimulus.

Cronin's (1980) recent findings illustrate the importance of immediate differential conditioned reinforcement for delayed-reward learning. All her pigeons were trained on a simultaneous red-green discrimination procedure with a 60-sec delay of primary reward (food). For her "nondifferential group," environmental conditions during the 60-sec interval between the correct choice response (e.g., pecking red) and reward were the same as during the corresponding interval following a nonrewarded response (e.g., pecking green). This group gave no evidence of learning the discrimination. Apparently the delayed rewards and nonrewards had no differential reinforcing effect on the choice responses. Cronin's "differential group" had one stimulus throughout the delay interval following correct responses and another following incorrect responses. Her "reinstatement group" received one houselight condition (e.g., yellow blinking houselight) only during the 10 sec immediately after a correct choice and the 10 sec immediately before rewards. Another houselight condition (e.g., constant blue houselight) prevailed only for the 10 sec after the incorrect response and the 10 sec before nonreward. The stimuli during the 40-sec middelay interval were the same on all trials. The birds in the differential and reinstatement groups rapidly acquired the correct choice response. This result was interpreted as indicating that the patterns of cues that immediately preceded reward and nonreward acquired differential conditioned-reinforcing properties that made them effective as reinforcers when they were presented immediately after choice. This interpretation was supported by the results for a group that was exposed to reversed cues. This group was trained on a procedure like that for the reinstatement group except that responses to the correct stimulus (S^D) were followed immediately by 10 sec of the stimulus that also preceded nonreward and responses to the incorrect stimulus (S^A) were followed immediately by 10 sec of the stimulus that preceded reward. The birds in the reversed-cue group acquired the incorrect

choice as rapidly as the birds in the differential and the reinstatement groups acquired the correct choice. For the reversed-cue group too, the prereward stimuli apparently acquired conditioned-reinforcing properties that were effective when those stimuli immediately followed choice responses.

Cronin's (1980) results may seem to contradict Lawrence and Hommel's (1961) frequently cited finding that differential conditioned-reinforcing value presumed to have been acquired by brightness stimuli in the goal box during a delayed-reward procedure transferred little or no differential immediate conditioned reinforcement to the choice compartment. However if conditioned-reinforcing value is situation specific, i.e., the conditioned reinforcing value of a component stimulus does not generalize to that same stimulus in a quite different context, then Lawrence and Hommel's results are quite consistent with Cronin's.

In Cronin's reinstatement and reversed-cue groups the 40-sec middelay interval (during which environmental conditions were identical following correct and incorrect responses) seemed to provide an effective barrier preventing backchaining of differential conditioned-reinforcing value from prereward stimuli (S^+) and prenonreward stimuli (S^-) to postresponse stimuli. Although the 10-sec prereward stimuli (S^+) presumably acquired greater reinforcing value than the 10-sec prenonrewarded stimuli (S^-), there was no indication that their differential conditioned-reinforcing value was backchained to events antecedent to the middelay interval.

EXPERIMENT 1

Experiment 1 was designed in part to provide further information about the effects of sequences of stimuli in delayed reward discrimination learning procedures with reversed cues. The procedure for one group was virtually identical to that for Cronin's reversed-cue group except that stimuli during the 40-sec middelay interval differed following the two choice responses. We expected the differential conditioned-reinforcing value acquired by S^+ and S^- to be transmitted through the 40-sec middelay stimuli to the postresponse stimuli. Thus, backchaining should transmit differential conditioned-reinforcing value through the

entire delay interval. The effect of this back-chaining would then oppose the effects that led Cronin's birds to select S^A.

If the postresponse stimuli and the stimuli preceding reward and nonreward are all very brief, then the cue patterns during identical stimulus situations, one of which precedes reward (or nonreward) and the other immediately follows a choice response, would include quite different traces of antecedent conditions. In this case the differential conditioned-reinforcing value acquired by the cue patterns preceding reward and nonreward might not generalize to the postresponse cue patterns. A second group was therefore included to assess whether the marked preferences acquired by Cronin's reinstatement and reversal-cue groups depended upon the lengths of the post-response and prereward stimuli. This second group received the same treatment as the first except that stimuli preceding reward and nonreward, and the postresponse stimuli lasted only .33 sec and the middelay stimuli were 59.33 sec in duration.

METHOD

Subjects

Eight experimentally naive pigeons served. They were maintained at 75% of free-feeding weights by grain obtained during experimental sessions and, when necessary, by supplemental feeding shortly after completion of sessions. Birds were individually housed in a continuously illuminated colony room where they had free access to grit and water.

Apparatus

The apparatus was the same as that used by Cronin except that two houselights were added. The two 33 by 28 by 31 cm two-key experimental chambers were painted flat gray throughout. Two translucent Gerbrands pigeon response keys were mounted 5 cm apart (center to center) on the steel intelligence panel, 21 cm above the floor. The keys were positioned symmetrically on either side of the vertical midline of the panel. Industrial Electronics Engineers multiple stimulus projectors behind the panel could transilluminate the response keys with red, green, or white light. A 5 by 4 cm opening 13 cm below and midway between the response keys provided access to a Lehigh Valley grain feeder contain-

ing a mixture of 50% wheat and 50% milo. During grain presentation the feeder tray was illuminated by a 1.1-W incandescent lamp and all other lights in the chamber were extinguished. Chamber illumination was provided by one of five 6-W lamps. Three of these houselights were located above a round translucent plastic window in the ceiling of the chamber. They were constant blue, constant white, and blinking yellow. The other two houselights were white. They were mounted on upper left and lower right corners of the intelligence panel and were shielded by white styrofoam drinking cups.

The response keys could be operated by a minimum force of approximately .1 N. When a response key was illuminated, each response to that key produced an audible relay click. No feedback click was provided for responses to dark keys. White noise delivered through a loudspeaker located behind the response panel and noise from ventilating fans masked extraneous sounds. Programming and recording equipment were located in an adjacent room.

Procedure

Birds were randomly assigned to two groups of four. The pretraining procedure for all birds was the same as that used by Cronin (1980) except for the number of sessions. Birds were first taken through a sequence of phases from magazine training to the final pretraining procedure. At the end of pretraining, sessions consisted of 24 trials on which red and green colors were presented simultaneously, one on each key. Each color was on right and left keys equally often. The first peck to a preselected one of the keys, which was more than 4 sec after presentation of the stimuli, changed that key to white for 60 sec and darkened the other key. Each color-key combination was preselected an equal number of times (e.g., red on the right). Half the effective pecks (pecks that produced white) to each color-key combination were followed by food 60 sec later, at termination of the white keylight.

During training, choice trials for all birds began with simultaneous presentation of a red stimulus on one of the response keys and a green stimulus on the other. Within each block of four trials, red was on the right for a random half of the trials. After four sec

with the keylights on, the first key peck initiated a sequence of events terminating in reward if the effective peck had been to one color (S^D) and nonreward if the effective peck had been to the other (S^A). The events following effective pecks to S^D or S^A included 60 sec of white keylight on the key pecked. For half the birds in each group, the sequence of houselight illumination following one choice was blue, then lower-right, then yellow; the sequence following the other choice was yellow, then upper-left, then blue. The duration of the blue and yellow houselights was 10 sec and .33 sec for Groups 10 and .33, respectively. The duration of the lower right and upper left houselights was 40 and 59.33 sec for Groups 10 and .33, respectively. The color of S^D and S^A , the order of yellow and blue houselights following correct and incorrect responses, and the sequence in which the lower right and upper left houselights occurred were counterbalanced across subjects.

All birds received a maximum of 20 trials per daily session. Trials were separated by variable intertrial intervals (ITIs) constructed from Fleshler and Hoffman (1962) schedules, with a mean of 120 sec. During ITIs the overhead white houselight was illuminated and the keys were dark and ineffective. A forced-trials procedure was employed in order to equate the number of exposures to rewarded and nonrewarded sequences occurring in each session. After the bird had chosen either red or green on 10 trials, the remainder of the 20 trials in that session were single-key trials in which only the discriminative stimulus that had not been pecked 10 times was presented.

Several criteria were employed to determine session termination. A session terminated immediately if the bird completed all 20 trials within a 2-hour time limit. If the bird failed to complete all 20 trials within the 2-hour limit, the session terminated immediately if an ITI was in effect, at the beginning of the next ITI, or as soon as S^D or S^A had been on a key for at least 45 sec.

Phase 1 training continued until a bird reached the criterion of choice of one of the stimuli (either the positive or the negative) on at least 9 of the first 10 trials for five consecutive sessions or choice of that stimulus on all of the first 10 trials for three consecutive sessions.

At the completion of Phase 1 reversal train-

ing was initiated. There were two kinds of reversal training. For Birds 74 and 719 in Group 10 and Birds 3274 and 73 in Group .33, the sequences of events following selection of red and green stimuli were reversed (red-green reversal). For the other birds food terminated the nonreward sequence of Phase 1 and no food followed the previous reward sequence (food-food reversal). As described in the results, one Group 10 bird (74) received a second reversal.

RESULTS

With the exception of Pigeon 73 in Group .33 during reversal (see below), only a very few nonrewarded and no rewarded trials were missed by failure to peck with short enough latencies to complete the sessions within 120 min.

Figure 1 summarizes the major findings for Group 10 during Phase 1 of training. The percentage of the first 10 trials on which S^D was pecked on successive pairs of sessions is plotted for individual birds. As depicted in Figure 1, all four birds initially increased choice of S^A .

Fig. 1. Percentage of the first 10 trials of a session in which individual birds in Group 10 of Experiment 1 made a choice response to S^D on successive pairs of sessions during original learning (Phase 1).

Fig. 2. Percentage of choice responses to S^D during the first 10 trials of a session for successive pairs of sessions during reversal (Phase 2). Each curve represents the performance of a single pigeon in Group 10 of Experiment 1.

Fig. 3. Percentage of the first 10 choice responses that were to S^D on successive pairs of sessions for individual birds in Group .33 of Experiment 1. The left hand portion of the figure represents original learning (Phase 1) and the right hand section represents reversal (Phase 2).

Each bird pecked S^A more than S^D during Sessions 7 to 10. Subsequently three of the birds increased pecking on S^D and eventually met the criterion of all 10 pecks on S^D on three consecutive sessions or pecks on S^D on at least 9 of the first 10 trials for five successive sessions. The fourth pigeon (74) met the criterion of preference for S^A .

As shown in Figure 2 during Phase 2 all four Group 10 birds came to peck S^D . Three of them met the criterion. Pigeon 15 which had a food $\rightarrow$ food reversal was terminated after he responded on S^D on over 80% of the first 10 choices during Sessions 22 to 36. Pigeon 74 met the criterion of preference by pecking mainly S^A in Phase 1 and S^D in Phase 2. He did not switch his red-green preference between Phases 1 and 2. Therefore this bird was again reversed and met the criterion of preference with pecking on the new S^D after 23 sessions. This indicates that like the others, Pigeon 74 eventually learned to discriminate postresponse from prereward conditions.

As shown in the left portion of Figure 3 all four birds in Group .33 reached the criterion of S^D preference after between 6 and 10 sessions (including criterion sessions). As depicted in the right-hand portion of Figure 3, three of the four pigeons also met the S^D criterion after reversal. The fourth bird (73) never pecked S^D (formerly S^A) during Phase 2.

There is no evidence that rate of reversal differed for red-green and food-food reversal procedures.

DISCUSSION

The S^A preferences of each 10-sec bird early in training may be ascribed to the greater reinforcing value of S^+ , which followed responses

to S^A as well as preceding food, than of S^- , which both followed S^D and preceded nonreward. The failure to obtain any initial S^A preference by the .33-sec birds may be ascribed either to a failure of such brief stimuli to acquire differential conditioned-reinforcing value when S^+ precedes food and S^- does not or to a failure of the differential reinforcing value of these cue patterns to generalize from the prereward to the postresponse patterns of cues. We favor the latter explanation because trace cues from antecedent stimuli were the basis for an eventual discrimination between postresponse and prereward cue patterns for Group 10. Such trace cues should be much more salient during the briefer postresponse and prereward stimuli of Group .33. For this group the immediate postresponse cue patterns appear to have been so different from the patterns of cues during the same sets of stimuli during the last .33 sec of the delay interval that the differential conditioned-reinforcing value of the latter pair of cue patterns did not generalize to the postresponse cue patterns.

The eventual choice of S^D by the 10-sec birds indicates that they came to discriminate between postresponse and prereward conditions and that differential conditioned-reinforcing value of the prereward cue patterns backchained through the entire delay sequence.

Cronin's reversed-cue pigeons continued to peck the S^A key throughout 15 sessions of training. Only one of the Group 10 birds reliably pecked S^D after 15 training sessions. It seems likely that if birds continued key pecking with further training on Cronin's reversed-cue procedure, they would either lose their S^A preference or this preference would become less marked. Our results indicate that late in training, backchaining had a greater effect than direct generalization, but they do not demonstrate that all direct generalization from prereward to postresponse cues had been eliminated.

EXPERIMENT 2

Experiment 2 was designed to determine more about how the relative durations of stimuli intervening between choice and reward (or nonreward) affect the establishment of discriminative responding based on delayed

reward. The procedure was simplified by eliminating the middelay interval and by following one choice with a blue-yellow stimulus sequence during the delay, and by following the other choice with a yellow-blue sequence. The procedure for the three groups differed only in the durations of blue and yellow stimuli and thus also of the total delay interval. The delays were 6, 20, and 60 sec for Groups 6, 20 and 60. Blue and yellow houselights were each on for half the delay interval.

METHOD

Subjects

Twelve experimentally naive pigeons obtained locally served. They were housed and maintained as described for Experiment 1.

Apparatus

The apparatus was the same as that used in Experiment 1 except that the upper left and lower right houselights had been removed.

Procedure

Birds were randomly assigned to three groups of four. Pretraining was similar to that employed in Experiment 1. The last phase of pretraining was the same as in Experiment 1 except that the delay interval was 6, 20, and 60 sec for Groups 6, 20, and 60, respectively. At this time pecks on both S^D and S^A were followed by a delay during which the effective key and houselight were white. For each stimulus, delayed food presentation followed effective pecks half the time.

During training, a choice trial began with the illumination of one of the response keys by a red stimulus and the other by a green one. Within blocks of four trials, red was on the right for two randomly selected trials. After four sec with the keylights on, the first peck on either key changed the stimulus on that key from red or green to white for the duration of the delay and darkened and deadened the other key.

A peck on one key was also followed by a blinking yellow houselight during the first half of the delay period and a constant blue houselight during the final half of the delay. A peck on the other key was followed by a constant blue houselight in the first half and a blinking yellow houselight in the second half of the delay period. Initial and final houselight cues following pecks on the red

and green stimuli were counterbalanced across subjects. Delayed food reward followed pecks on the red key for half the subjects. Food followed pecks on the green key for the other half.

As in Experiment 1, following choice of red or green on 10 trials, the other stimulus was presented alone on one key and pecks on that key had the same effects as on choice trials. Criteria for termination of sessions were the same as described for Experiment 1.

The three experimental groups differed only in the length of the delay period. Initial and final delay cues were each 3 sec for Group 6, 10 sec for Group 20, and 30 sec for Group 60. Phase 1 training was completed when either criterion of choice used in Experiment 1 had been met on the S^D , or after 30 sessions. When the choice criterion was met, Phase 2 training was initiated on the following session. There was one exception: Pigeon 75 in Group 60 was not reversed until meeting the criterion after 41 Phase 1 sessions.

Phase 2 was a reversal of the S^D and S^A stimuli of Phase 1. The reversal was the type described as a food-food reversal in Experiment 1. The sequence of houselight stimuli following each choice was the same as in Phase 1, but food followed the choice and delay sequence which had not been followed by food in Phase 1. Birds were run to the same criterion as in Phase 1. Those birds in Groups 6 and 20 reaching criterion within 30 sessions received Phase 3 training.

In Phase 3, the delay was gradually extended to 60 sec for each bird that met the criterion of preference (three successive 100% choice sessions or five successive sessions of pecking S^D on 9 of 10 trials). If they met criteria, birds in the 6-sec group were trained successively with 20-sec, 35-sec, and 60-sec delays. Birds in the 20-sec group were trained on 35-sec and 60-sec delays. Training was terminated after 30 sessions of any stage in which a bird failed to reach criterion. Note that no reversal of stimulus conditions occurred between Phase 2 and Phase 3 or between different delays within Phase 3. In other words, for each bird, both the S^D color and the sequence of the delay cues following each choice remained the same throughout Phases 2 and 3.

Sessions were run daily with the exception that there was a 17-day break between the 46th and 47th session for each bird.

Fig. 4. Percentage of the first 10 trials of a session during which the choice response was to S^D on successive pairs of sessions for individual birds in Experiment 2 during original learning. Each graph represents the results for a different delay condition. The arrow on the curve for Bird 75 indicates the point at which there was a 17-day interruption of training.

RESULTS

The point at which the 17-day break came for individual birds is indicated by arrows in Figures 4, 5, 6, and 7. Examination of these figures indicates this break had no clear effect on the measures obtained on subsequent sessions.

The percentage of the time S^D was chosen by individual pigeons on the first 10 trials during successive pairs of sessions in Phase 1 is shown in Figure 4. Results for each group are presented in a separate graph. As can be seen in Figure 4, the 6-sec delay birds soon pecked S^D on most trials. Two of the 20-sec delay birds (6 and 7) showed an initial S^A preference and then increased pecking of S^D until they met the criterion for completion of Phase 1. The other two 20-sec birds exhibited

Fig. 5. Percentage of the first 10 trials of a session during which the choice response was to S^D on successive pairs of sessions during reversal (Phase 2). As the labeling indicates, each graph is for a different group, and each curve is for an individual bird. Arrows indicate the location of the 17-day interruption of training.

no clear preference during the first five sessions and then pecking of S^D increased until they met the criterion. All of the 60-sec delay pigeons acquired a clear initial preference for S^A within 10 sessions and maintained it until Session 20 or longer. Two of the birds later switched to an S^D preference and met criterion after 26 (Bird 4) or 41 (Bird 75) sessions. The other two birds maintained their S^A preference through Session 30.

Performance during Phase 2 (reversal) is depicted in Figure 5 in the same manner as Phase 1 results are represented in Figure 4. Note that the only 60-sec birds which received Phase 2 training were the two which met the criterion in Phase 1 and that Bird 75 had received extended Phase 1 training.

All birds receiving reversal training eventually acquired a preference for the S^D in Phase

Fig. 6. Choice of S^D on the first 10 trials of a session for pairs of sessions in Phase 3 during delays of 20 and 35 sec. Each curve is for an individual bird. Arrows indicate the location of the 17-day interruption of training.

2. There was a marked tendency for reversal to be more rapid, the shorter the delay. Three of the four 6-sec birds met the criterion for reversal before any birds with longer delays. Neither 60-sec pigeon met a criterion for reversal within 30 sessions, within which time all 6-sec birds and three of four 20-sec birds had met a criterion.

Figure 6 shows the performances of the birds originally trained with 6-sec delays when the delays were increased to 20 and 35 sec. All maintained their S^D preferences during 20-sec delays but only one (Bird 9) maintained a strong preference for S^D during 35-sec delays. Another (71) may have been recovering an S^D preference before his training was terminated after 30 sessions with a 35-sec delay. Pigeons 2, 6, and 7, which were originally trained with 20-sec delays, all maintained a strong S^D preference when the delay was increased to 35 sec and met a criterion on Sessions 14, 5, and 13, respectively. The other bird in Group 20 did not receive Phase 3 training.

The percentage of the time S^D was chosen after the delay had been increased from 35 sec to 60 sec for the four birds (one from Group 6 and three from Group 20) which received this treatment is shown in Figure 7. As can be seen there, none of the birds maintained a very strong preference for S^D during this training. The percentage of the choices of S^D by individual birds on Sessions 27 to 30 ranged between 30 and 77.5.

A comparison of the terminal choice performance for birds in different groups when they were exposed to the same procedure, e.g., 60-sec delays, provides no indication that earlier training with shorter delays improved subse-

Fig. 7. Choice during 60-sec delay in Phase 3 for those birds from Groups 6 and 20 that received this training. The axes are the same as in Figures 1 to 6.

quent steady-state performance with longer delays.

DISCUSSION

The results of Experiment 2 are in line with those obtained in Experiment 1. It appears that when the stimulus situation immediately following responses on S^D is the same as the stimulus situation preceding delayed nonreward and when environmental stimuli following responses on S^A are the same as those preceding delayed reward, then there is a positive relationship between the duration of these stimulus situations and responding on S^A early in training. With 3-sec stimulus situations no S^A preference was acquired; with 10-sec stimuli, two of the four birds acquired a brief S^A preference; and with 30-sec stimuli, all four birds acquired an S^A preference. Two of these birds maintained their S^A preference throughout 30 sessions.

Although it is not clear what would have happened had birds been given more extensive training with 30-sec stimuli (and 60-sec delays), results for the 60-sec group and for other birds during Phase 3 suggest that with 30-sec stimuli some pigeons might never acquire a clear S^D preference.

Differential responding on S^D may be taken as evidence that the birds discriminated post-response from prereward stimuli and that because of this conditional discrimination based on trace and/or temporal cues, the differential conditioned-reinforcing value of the prereward patterns of cues did not continue to generalize enough directly to the postresponse cue patterns to overcome backchaining. The failure of some of the birds to acquire an S^D preference with a 60-sec delay indicates that

direct generalization continued when the stimulus situation lasted 30 sec and/or that there was little or no backchaining of differential conditioned-reinforcing value through the full delay interval.

GENERAL DISCUSSION

An alternative way of referring to the context specificity of discriminative and reinforcing properties of cues is to treat these properties not as of component stimuli but rather as attributes or properties of the total pattern of cues. Because temporal cues (internal clocks) and traces of antecedent events are constantly changing, one might then expect that conditioned-reinforcing properties would be backchained not only from one set of stimuli to another but also from one pattern of cues to another within the same set of stimuli, i.e., within the same environmental conditions. Powell and Perkins (1957) provided direct experimental evidence that this is sometimes the case. They found that in rats the reinforcing value of food in a specific context on Trial *N* depended upon the length of time food had remained available in that context on antecedent trials and was not measurably affected by the length of time the food remained available on Trial *N*. Their rats learned to select the arm of a T-maze leading to a goal box in which they had been exposed to 150 sec of food on past occasions rather than a goal box in which they had been exposed to 10 sec of food on past occasions even though on each maze trial food was available for 20 sec in either goal box. In another experiment their rats failed to acquire a preference for the response leading to 150 sec of food in one goal box over the response leading to 10 sec of food in a different goal box when the average duration of availability of food was kept the same in each goal box. Straight runway trials interspersed among T-maze trials kept the average duration of availability of food the same in each goal box. Apparently in the Powell and Perkins experiments the reinforcing value of food-in-context was determined by the conditioned-reinforcing value of food onset and was unaffected by any delayed reinforcement from presence of food more than 10 sec later. Presence of food from 10 to 150 sec after food onset had no greater reinforcing effect than absence of food during that interval.

Under certain conditions a discrimination-learning problem with delayed-reward and nondifferential-delay stimuli will not lead to predominance of the rewarded choice when it would in otherwise identical conditions but with different stimuli throughout the delay following correct and incorrect trials (e.g., Cronin, 1980). Such failures to acquire the discrimination seem to occur if, and only if, the cue patterns at some point in the delay interval are the same following correct and incorrect responses. Inclusion of identical cue patterns in each sequence apparently blocks backchaining of differential conditioned-reinforcing value through the delay interval. Note, however, that in such cases if half the trials are correct, then correct and incorrect choice responses may both be maintained by the conditioned-reinforcing effect of the nondifferential-delay stimuli that are followed by a primary reward half the time.

In the case of Group .33 in Experiment 1 and Group 6 of Experiment 2, the trace cues during the postresponse stimuli apparently were so different from the trace cues during the prereward and prenonreward stimuli that the differential conditioned-reinforcing value acquired by the latter pair of cue patterns did not generalize measurably to the former. However, rapid backchaining through the delay sequence might have masked slight effects of direct generalization of conditioned-reinforcing value from prereward to postresponse stimuli. Presumably the differences in trace cues during prereward and postresponse cue patterns would prevent direct generalization of differential conditioned-reinforcing value from immediate prereward to immediate post-response patterns of cues for all birds in both experiments. How is it then that for all birds in Group 10 of Experiment 1 and Group 60 of Experiment 2 and two of the birds in Group 20 of Experiment 2, the differential reinforcing value of the prereward and prenonreward cue patterns initially "generalized" to the post-response cue patterns? For these birds the difference between the traces just before the termination of the 10-sec prereward stimuli and the traces just before the termination of the 10-sec postresponse-stimuli would have become increasingly similar during the 10-sec interval. Apparently 10 sec after stimulus onset, these traces were not always different enough to prevent initial generalization of differential condi-

tioned-reinforcing value from the former pair of cue patterns to the latter. Backchaining would then occur within the 10-sec postresponse stimuli so that the patterns of cues at onset of these stimuli would differ in reinforcing value. Later in training, cue patterns just before termination of the postresponse stimuli might come to be discriminated from the cue patterns at the termination of the pre-reward stimuli. Then the differential conditioned-reinforcing properties of the cue patterns preceding reward and nonreward might no longer generalize to the cue patterns at the termination of the postresponse stimuli. At least no such generalization would still dominate the opposing effects of backchaining throughout the entire sequence.

The Egger-Miller information hypothesis—and the experimental findings that following presentation of an S1-S2-reward sequence S1 has a greater reinforcing effect than S2—might seem to indicate that the reinforcing value of S1 could not have been acquired entirely by backchaining through S2 (Egger & Miller, 1962). Hancock (1982) has recently demonstrated, however, that the greater reinforcing effect of S1 than S2 is a result of a greater generalization decrement from training to test conditions for S2 than S1. Furthermore, Hancock obtained a greater reinforcing effect of S2 than S1 when he controlled for the differential generalization decrement.

The quite striking effects of delay stimuli observed in some experiments including those reported here indicate that delay of reward effects often cannot be adequately described in terms of simple delay of reward gradients. Similarly, as Fantino (1977) has noted, his delay-reduction hypothesis does not always apply at least without modification or elaboration. The present interpretation of our results seems consistent with an elaboration of the delay-reduction hypothesis that also emphasizes the role of conditioned reinforcement.

Our interpretation of these results depends upon an analysis in terms of patterns of cues that emphasizes the role of mediating events such as traces and temporal cues, which are inferred but not directly observed. It seems to us that the results of our experiments and related findings may best be systematically described by postulating the existence of such processes and assuming that all causation is contiguous causation. If, instead, investigators

restrict their consideration to those events that can be directly observed at the moment in question, no such systematic description seems to be possible. It appears that we now know enough about some inferred events for reference to them to have considerable integrative value.

The emphasis on conditioned reinforcement as an attribute of cue patterns rather than of component stimuli is a great enough shift in emphasis from early formulations of conditioned reinforcement (Hull, 1943; Skinner, 1938) that it may be appropriate to employ a somewhat different set of terms than those which may still carry connotations that can now be seen as inappropriate. It is for this reason that a more comprehensive analysis of recent findings including those described here has been stated in somewhat different language elsewhere (Perkins, in press). In any case, it no longer seems appropriate to refer to delayed reinforcement rather than delayed reward because delayed rewards seem to be effective only when the cue patterns that immediately follow correct and incorrect responses acquire differential conditioned-reinforcing value.

Our findings and the results of related experiments could also be systematically described in a more cognitive language. For example, rather than referring to the acquisition of conditioned-reinforcing value, one could speak of the acquisition of expectations of the primary reward. If the conditions under which such expectations are acquired are specified appropriately, the implications of such an analysis may well be identical to the implications of an analysis in terms of the acquisition of conditioned-reinforcing properties by patterns of cues.

REFERENCES

- Cronin, P. B. Reinstatement of postresponse stimuli prior to reward in delayed-reward discrimination learning by pigeons. *Animal Learning and Behavior*, 1980, 8, 352-358.
- Egger, M. D., & Miller, N. E. Secondary reinforcement in rats as a function of information value and reliability of the stimulus. *Journal of Experimental Psychology*, 1962, 64, 97-104.
- Fantino, E. Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior*. Englewood Cliffs, N.J.: Prentice-Hall, 1977.
- Fleshler, M., & Hoffman, H. S. A progression for gen-

- erating variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 529-530.
- Gibbon, J. Scalar expectancy theory and Weber's Law in animal timing. *Psychological Review*, 1977, 84, 279-325.
- Hancock, R. A., Jr. Tests of the conditioned reinforcement value of sequential stimuli in pigeons. *Animal Learning and Behavior*, 1982, 10, 46-54.
- Hull, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
- Kelleher, R. T., & Gollub, L. R. A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 543-597.
- Killeen, P. On the temporal control of behavior. *Psychological Review*, 1975, 82, 89-115.
- Lawrence, D., & Hommel, L. The influence of differential goal boxes on discrimination learning involving delay of reinforcement. *Journal of Comparative and Physiological Psychology*, 1961, 54, 552-555.
- Perkins, C. C. The analysis of performance, In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behaviour* (Vol. 3). *Biological factors in learning*. Chichester, U. K. and New York: Wiley & Sons, in press.
- Powell, D. R., & Perkins, C. C. Strength of secondary reinforcement as a determiner of the effects of duration of goal response on learning. *Journal of Experimental Psychology*, 1957, 53, 106-112.
- Skinner, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.

Received June 24, 1981

Final acceptance March 19, 1982