Conditioned reinforcement and complex discrimination: Can conditioned reinforcing value be specific to a complex of three cues?

SUZETTE L. ASTLEY
Cornell College, Mt. Vernon, Iowa

The present experiment is concerned with the nature of the cues that might acquire conditioned reinforcing value, and the ways in which such cues might interact with one another. Red and green colored keylights were differentially paired with food dependent upon the houselight context (A or B) and the trial type (training or choice/forced). The duration of the colored keylights was varied between groups in an attempt to manipulate the effectiveness of the short-term memory of trial-type cues at the trial's end. The red and green stimuli were of 30 sec duration for Group 30 and of 3 sec duration for Group 3. The results indicated that the choices of the pigeons in Group 30 were influenced by the houselight context present and by the keylight color. The choices of the pigeons in Group 3 seemed to be influenced by the houselight context present, the keylight color, and the memory of trial-type cues. Memory cues for trial antecedents were not overshadowed by presumably more salient external houselight stimuli for the pigeons in Group 3. Two alternative explanations for the results are discussed, and determined to be unlikely based on the results of an earlier experiment. The present results are related to a model of the conditioned reinforcing value of momentary stimuli and of transmission of conditioned reinforcing value.

Much recent interest has focused on complex discrimination learning (e.g., Fedorchak & Bolles, 1986; Zentall, Hogan, & Edwards, 1984). Unfortunately, until recently, little has been done to explore the ways in which complex discrimination might influence the acquisition and demonstration of conditioned reinforcement. Three recent experiments have directly addressed this issue (e.g., Astley & Perkins, 1985; Hancock, 1982; Winter & Perkins, 1982). In the Hancock (1982) experiment, conditioned reinforcing value was decreased markedly when a potential conditioned reinforcer was preceded by one stimulus on training trials and by another on test trials. In their experiments, Winter and Perkins (1982) and Astley and Perkins (1985) varied the duration of stimuli and demonstrated that the conditioned reinforcing value of stimuli could be restricted to contexts distinguished by the prevailing short-term memories (STMs) of antecedent events. In the present experiment, this general approach to discovering the cues that effectively acquire conditioned reinforcing value was followed. In the three studies cited above, pairs of cues (i.e., the component stimulus and the STM of antecedents) provided the basis for differential conditioned reinforcing value. The present experiment was designed to determine whether conditioned reinforcing value would be restricted to an even more complex set of available cues. The present study included both on-

going environmental stimuli and memory cues as potential discriminative stimuli to discover whether the discriminative behavior that has been achieved in the past on the basis of STM cues might be disrupted by a presumably more salient environmental stimulus.

METHOD

Subjects

Fifteen mature pigeons served as subjects. All subjects were maintained at 75% of their free-feeding weights in individual cages, where they had free access to water and grit. The room in which they were housed was continuously illuminated.

Apparatus

The experiment was conducted in four standard two-key pigeon chambers. The two response keys were 21.5 cm from the floor of the chamber and 5 cm apart, center to center. Stimuli were presented on the translucent keys by means of miniprojectors mounted behind each key. Pressure with a force of .06 N registered as a peck. An aperture providing access to the Lehigh Valley food magazine was located on the intelligence panel beneath and between the two keys. The aperture measured 5 × 4.5 cm, and the bottom edge was 3 cm above the floor of the chamber. This opening was illuminated by a 1.1-W light whenever food was available. Reward consisted of 3-sec access to a mixture of wheat and milo. All houselights were extinguished while reward was available.

Centered in the ceiling was a 6-W white houselight in a translucent housing. On the wall opposite the intelligence panel was a 6-W blue houselight which, when on, flickered with a cycle of 2 sec on and .5 sec off.

Procedure

All conditions were fully counterbalanced. For ease of description, however, only one of the counterbalanced conditions will be described. All trials throughout the experiment were preceded by

368

a 2-sec blackout and a 60-sec intertrial interval (ITI), during which only the houselight was illuminated. The houselight remained on during the trial. On half of the trials, the houselight was white; on the other half, it was blue and blinking. For purposes of description, the houselight conditions are referred to as houselight A and houselight B. Since all conditions were fully counterbalanced, houselight A was the white houselight for half of the pigeons and the blue blinking houselight for the remainder of the pigeons in each group.

Autopeck pretraining. The pigeons were first habituated to eating from the food magazine. They were then given autoshape pretraining to peck the horizontally and vertically striped stimuli, which were later used as discriminative stimuli for the choice responses. The pigeons were given single and two-key autoshape pretraining with the striped stimuli. There were an equal number of autoshape pretraining trials in each of the houselight contexts. Pretraining sessions were administered daily.

There were 80 trials in each single-key pretraining session. The horizontal and vertical stimuli were presented equally often following the ITI and equally often on the left and right side keys. In the absence of a peck, the single-key pretraining trials were of 10 sec duration and ended with a food presentation. A peck to the lit key ended the trial immediately with a food presentation. Training continued until a pigeon pecked on at least half of the trials in a daily session.

Two-key pretraining began on the following day. In two-key pretraining, the horizontally striped stimulus was presented on one of the keys and the vertically striped stimulus on the other following each ITI. To discourage a preference for one side key over the other, only one of the side keys was effective in ending the trial and producing food, and the side of the effective key changed from trial to trial. The side of the effective key was determined in a pseudorandom order, such that the left and right keys were equally often effective. To further discourage side preferences, the alternative keylight was extinguished if the designated key was not pecked within 10 sec. If the designated key was not pecked within an additional 20 sec, it was extinguished and the ITI was initiated. The daily two-key pretraining sessions consisted of 40 trials, and continued until the number of pecks to the left key was within 10% of the number of pecks to the right key.

Conditioned reinforcement training and testing. During Phase 1 of conditioned reinforcement training, trials consisting of red or green stimuli on both keys followed the ITI and were differentially associated with food. In the presence of houselight A, the red keylights were followed by food and the green keylights were not. In the presence of houselight B, the contingencies were reversed; the green keylights were followed by food and the red keylights were not. On these trials, the appropriate stimulus appeared on both keys. In each Phase 1 session, there were 10 presentations of each key color in each context, for a total of 40 trials. In Phase 1, all stimuli (including food) were presented independently of the behavior of the pigeon. Phase 1 continued for 2 days.

The pigeons were randomly assigned to one of two groups prior to Phase 1. The red and green keylights were of 3 sec duration for Group 3 and of 30 sec duration for Group 30. Eight pigeons were originally assigned to each of the groups. A death near the end of the experiment, however, left 7 pigeons in Group 3.

Figure 1 is a schematic diagram of the procedure for Phase 2 of training. In Phase 2, five pairs of test trials, each consisting of one choice trial and one forced trial, were interspersed among the training trials on alternate days. These pairs of trials were presented equally often in each of the two houselight contexts. As shown in Figure 1, choice trials began with presentation of one of the choice stimuli on each key. For a given pigeon, the horizontal stimulus always appeared on one key (e.g., right) and the vertical stimulus always appeared on the other key (e.g., left). Transition from the choice stimuli to the colored keylights was arranged by means of a fixed-interval 3-sec (FI-3) schedule. An effective peck to the right key produced the red keylights on both keys and an effective peck to the left produced the green keylights on both keys. (Recall that there was full counterbalancing across pigeons, and that the procedure described is for only one of the counterbalanced conditions.)

After the blackout and illumination of the houselight that was present before and during the choice trial, the forced trial began. It began with the illumination of the striped choice key which had
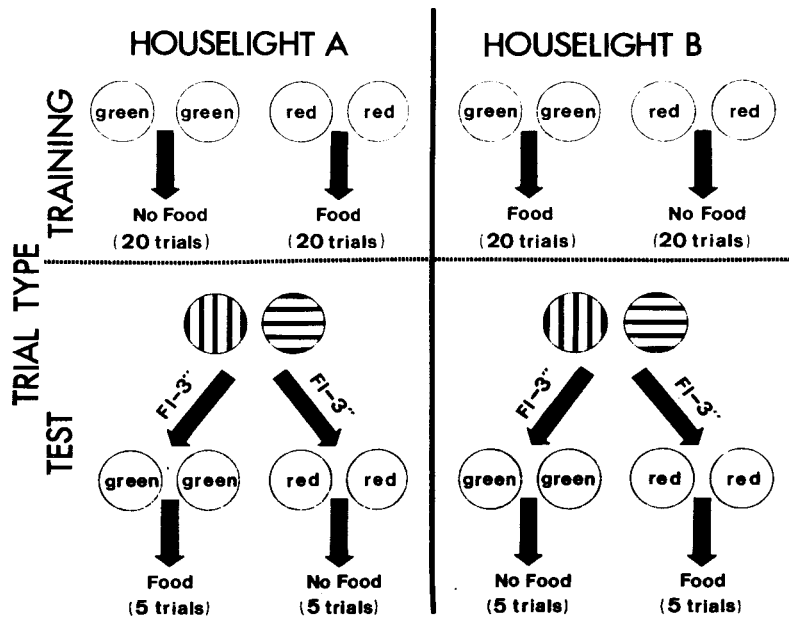

Figure 1. A schematic representation of the relative probabilities of reward following Sr+ and Sr− on the two trial types in the two contexts. The numbers of each type of trial shown are for pairs of sessions in Phase 2. (Context A was the white houselight for 4 of the 8 pigeons in Group 30 and for 3 of the 7 pigeons in Group 3.)

not been pecked on the preceding choice trial. After the FI-3 interval timed out, a peck to the illuminated key changed both keys to red if it was on the left and to green if it was on the right.

Reward conditions on the choice/forced (test) trials were the reverse of those on the training trials. If reward followed red on training trials in a given context, reward did not follow red on test trials in that context. During houselight A, choice and forced presentations of the green keylights were followed by food, and those of the red keylights were not. During houselight B, choice and forced presentations of the green keylights were not followed by food and presentations of the red keylights were.

Phase 2 continued for 90 days. Two of the pigeons in Group 30 (B2 and B3) received an additional 130 sessions of Phase 2 training.

RESULTS

The colored-key stimulus rewarded on the greater proportion of all trials within a houselight context will be referred to as Sr+. As there was one test trial for every four training trials in Phase 2, Sr+ was the stimulus rewarded most often on training trials in a given context (houselight). In the counterbalanced condition described in the method section, the red keylight was Sr+ in house-

light A and the green keylight was Sr+ in houselight B. The measure of conditioned reinforcement used in the following analyses is the percentage of the choice trials (averaged over the two houselight contexts) in which the first peck after the FI-3 requirement produced Sr+. This measure is referred to as percent choice of Sr+.

Figure 2 shows the mean percent choice of Sr+ during Phase 2 for the pigeons in each group. Note that the data were organized into blocks encompassing 10 successive days, only 5 of which included choice trials. Figure 2 indicates that the average percent choice of Sr+ for Group 30 increased through the first four blocks of Phase 2 and decreased somewhat in Blocks 5-9. The average percent choice of Sr+ for Group 3 declined over the nine blocks. These two groups were compared using a 2 (duration) × 9 (blocks) repeated measures unweighted means analysis of variance. This analysis revealed a significant main effect of duration [$F(1,13) = 26.45, p < .001$], a significant main effect of blocks [$F(8,104) = 8.00, p < .001$], and a significant duration × blocks interaction [$F(8,104) = 9.00, p < .001$]. A series of t tests
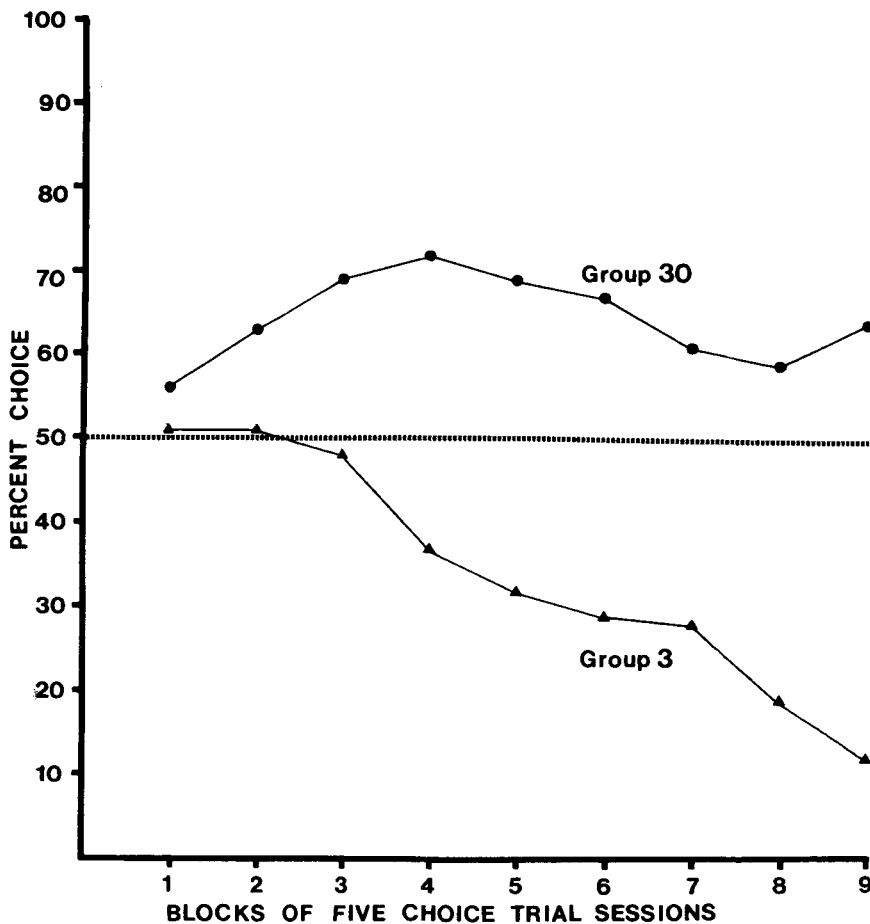

Figure 2. Mean percent choice of Sr+ for birds in each of Groups 3 and 30 over nine successive blocks of 10 sessions each during Phase 2.

Table 1
The *t* Values for Comparisons of Percent Choice for Blocks
1-9 Against 50% for Groups 3 and 30

Group	Blocks								
	1	2	3	4	5	6	7	8	9
3	.33	.25	.67	2.17	2.57*	2.62*	2.75*	4.42†	4.12†
30	.86	1.40	3.17*	3.67†	3.17*	3.40*	2.75*	2.25*	2.80*

*$p < .05$.   †$p < .01$.

compared the percent choice for each of Groups 3 and
30 against 50% in each of Blocks 1-9. These compari-
sons showed that the percent choice for Group 3 did not
differ significantly from 50% in Blocks 1-4, whereas the
percent choice for this group was significantly less than
50% in Blocks 5-9. The percent choice for Group 30 was
not significantly different from 50% in Blocks 1 and 2,
whereas the percent choice of this group was significantly
greater than 50% in Blocks 3-9. (The *t* values for each
of these comparisons may be found in Table 1.)

The percent choice data for individual subjects in
Group 3 are depicted in Figure 3. Note that the percent

choice of Sr+ for 5 of the 7 pigeons in Group 3 declined
to well below 50% over the nine blocks of trials, and those
of two of them changed little over Phase 2 and remained
close to 50%.

The percent choice data for individuals in Group 30 are
depicted in Figure 4. Subjects in Group 30 showed a
marked individual variation in performance. Two of the
subjects (B2 and B3) in Group 30 showed a percent choice
that increased over blocks of trials initially to well above
90% and then declined to nearly 50% by Block 9. Two
other subjects (C1 and C2) showed a similar but less
marked trend. Of the percent choice of the other four sub-
jects, that of C3 never varied far from 50%; that of E1
and E4 tended to increase over blocks throughout; and
that of C4 was initially high, decreased to about 50%,
and then increased again. Because the percent choice of
Subjects B2 and B3 of Group 30 rose initially to above
90% but had decreased to 50% by Day 90 of Phase 2,
they were given an additional 130 days of training to see
if the percent choice would decrease yet further. During
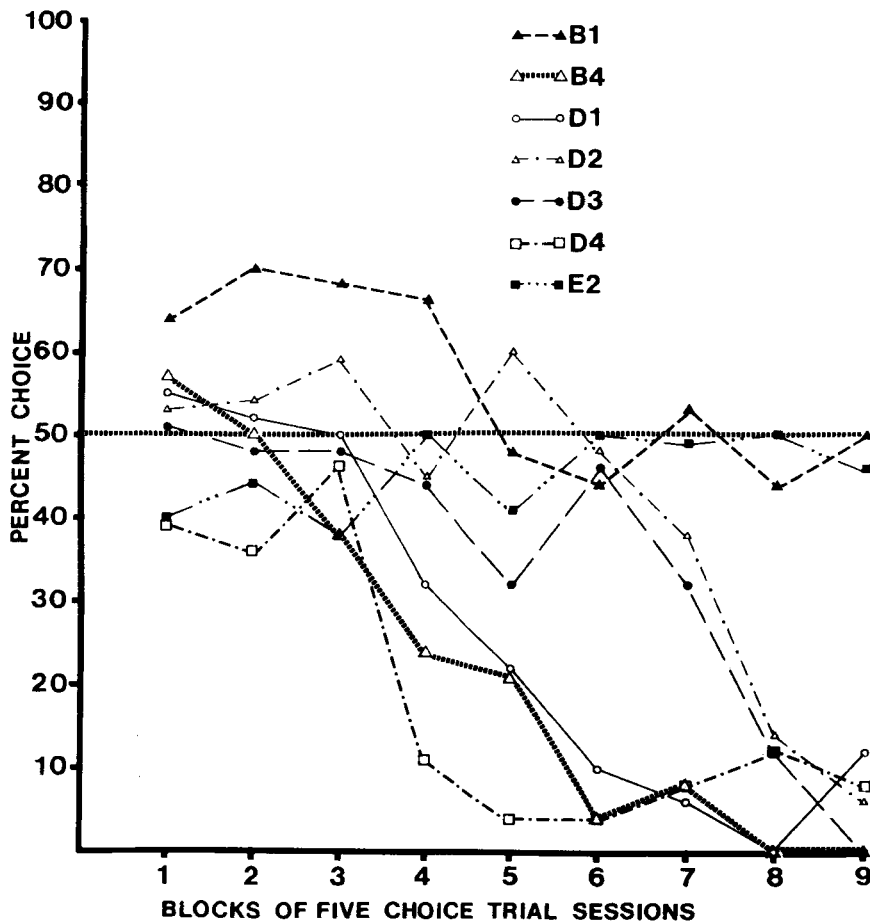this period, the percent choice of Sr+ for Pigeon B2 re-

Figure 3. Percent choice of Sr+ for individual pigeons in Group 3 over the nine blocks of
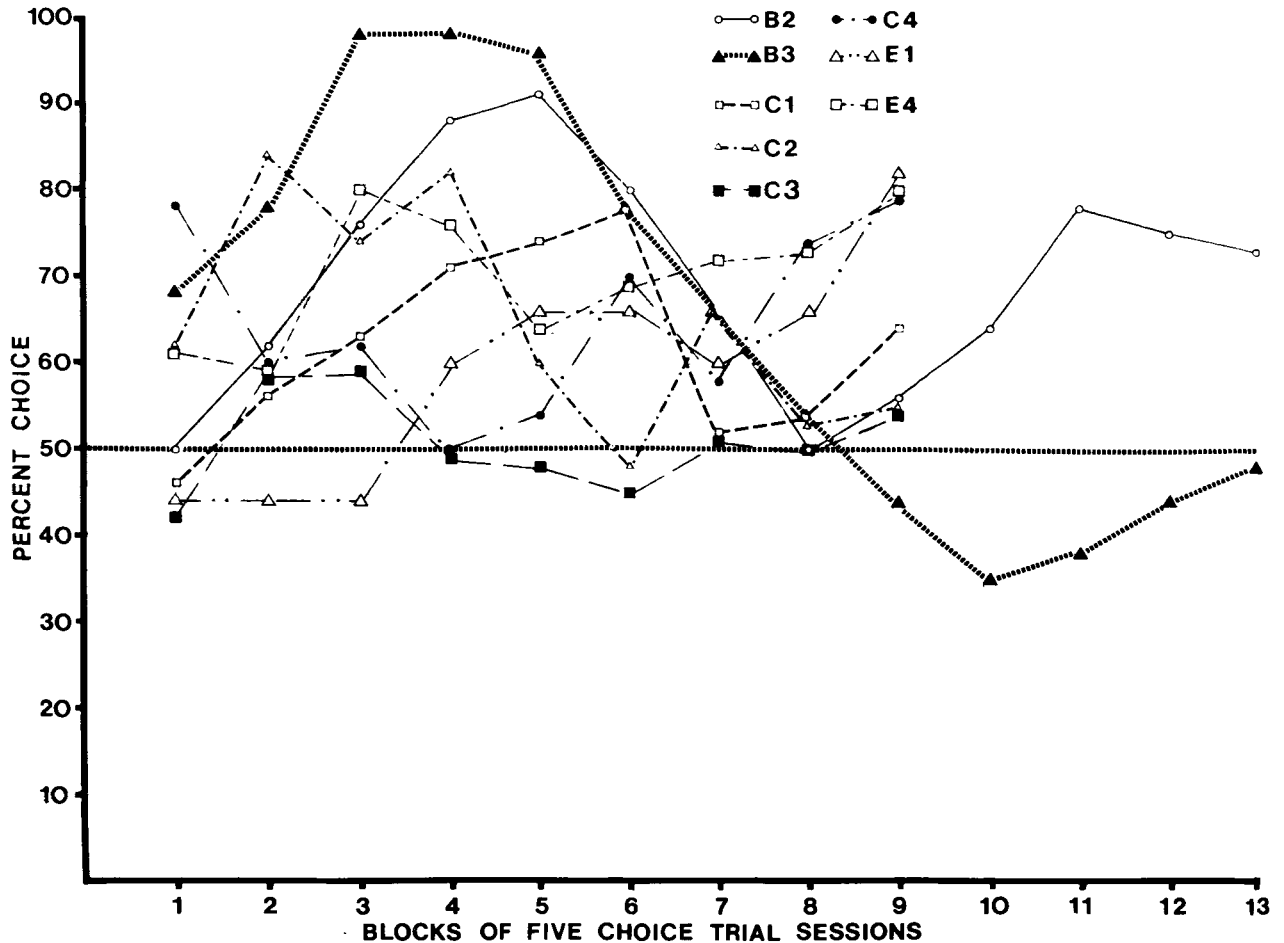Phase 2.

Figure 4. Percent choice of Sr+ for individual pigeons in Group 30 over the nine blocks of Phase 2.

mained below 50%, although there is an upward trend from Block 10 through Block 13. The percent choice of Sr+ for Pigeon B3 rose again to 88% on Block 11, and then declined slightly.

DISCUSSION

The present experiment indicates that colored keylights will demonstrate substantial differential conditioned reinforcing value within houselight contexts when they are of 30 sec duration. In Phase 2 of training, the percent choice of Sr+ was significantly greater than 50% for Group 30. When the colored keylight stimuli were of 3 sec duration, there was no evidence of conditioned reinforcing value of Sr+ on test trials. The percent choice of Sr+ was significantly below 50% for Group 3.

Stated differently, the Group 30 pigeons responded differentially to the two keylight colors dependent upon the houselight context, that is, they chose one keylight color in houselight A and the other keylight color in houselight B.

The pigeons in Group 3 also responded differentially on the basis of the houselight condition and the keylight color. In addition, however, they responded "appropriately" to the trial type. Rather than choosing the keylight color that was most often reinforced overall in a particular houselight context, they chose the keylight color that was most often reinforced on test trials within that particular houselight context.

The training trial presentations of the red and green keylights differed from the test trial presentations in their antecedents; that is, the dark keys of the ITI immediately preceded the colored keylights on training trials and the striped choice stimuli and at least one peck preceded the colored keylights on test trials. Thus, if the STM of these trial antecedents persisted throughout the duration of the colored keylight stimuli, it could have provided the basis for discriminating training from test trials. The pigeon demonstrates an STM for colored keylights of at least 3 sec in a delayed-matching-to-sample procedure (see Roberts & Grant, 1976, for a review). Thus, it is not surprising that the pigeons in Group 3 discriminated training from test trials.

The maximum of the STM function for pigeons seems to depend heavily on the events to be remembered and on the testing procedure. Under maximal conditions, well-

practiced pigeon subjects have demonstrated effectiveness of STM cues over durations as long as 30–60 sec (e.g., Grant, 1976; Honig, 1978). Two factors that seem to be important in determining the duration of effective STM cues are the salience of the stimuli to be remembered and the duration of their availability. In the present study, the dark keylight ITI was present for 60 sec before each training trial, but it may be presumed to be of low salience. Preceding the red and green keylights on each choice/forced trial were approximately 3 sec of the choice stimuli and a few pecks. Although presumably more salient than the dark-key ITI, such stimuli are brief and are not likely to produce STM cues that are effective over a long duration. It is reasonable to expect, then, that the STM of trial antecedents may no longer be an effective cue at the end of the 30-sec-long red and green keylights of this experiment, at least early in training. If this analysis is correct, we should expect the red stimulus on test trials during context A to be indistinguishable from the red stimulus on training trials during context A, but only during the final few seconds before the end of the trial. We might therefore expect that the conditioned reinforcing value of the red keylight would generalize between training and test trials in context A during those final few seconds. This means that differential conditioned reinforcing value should accrue to the red and green keylights dependent upon the context but not upon the type of trial in which the colored keylight is embedded when the colored keylights are of 30 sec duration throughout. The data for Group 30 are consistent with this hypothesis.

As indicated earlier, however, STM cues have been known to provide an effective basis for differential behavior of pigeons for as long as 30 sec. This factor may account for the marked individual variability of the behavior of pigeons in Group 30. An increase in effectiveness of STM cues would have resulted in an increase in discrimination between training and test trials, and thus a decrease in percent choice of Sr+ as training progressed. Winter and Perkins's (1982) Experiment 2 used a delayed reinforcement procedure and stimuli similar to those in the present experiment. In that study, 2 of 4 pigeons gave evidence of an ability to bridge a 30-sec delay at the end of 60 days of training. If an increase in the effectiveness of STM cues accounts for the decrease in percent choice of Sr+ late in the training of some subjects in Group 30, then the performance of the 2 subjects given extended Phase 2 training seems puzzling. Once the discrimination between training and test trials was complete, the pigeons in Group 30 should have chosen Sr− exclusively. Instead, the percent choice of Pigeon B3 tended toward 50% on the additional 50 sessions of Phase 2, and the percent choice of Pigeon B2 returned to approximately 70% after dipping to an average of 50% on Sessions 70–80.

Studies by Williams (1982, 1984) may shed a light on the performance of Pigeons B2 and B3. Williams found that discriminative performance of pigeons was degraded when one of the elements of a conditional discrimination was intermittently presented alone. If the STM of trial antecedents was sometimes available at the end of the 30-sec colored keylights in this experiment and sometimes not, then the conditions for the pigeons in Group 30 may have been similar to that for Williams's subjects. That is, the trials in which the STM cues persisted to the end may be considered to have been like Williams's compound training trials and the trials in which the STM did not persist to the end might be considered to be like Williams's element trials which interfered with acquisition of the discrimination.

A model of the acquisition of conditioned reinforcing value is implicit in the foregoing discussion, and should be made explicit at this point. According to this model, conditioned reinforcing value is transmitted to the constellation of cues that immediately and reliably precede reward. The term *cues*, rather than *stimuli*, is intentionally used here to include such factors as proximal representations of current external stimuli (e.g., the houselight contexts), memories of past events, and other internal states (e.g., hunger, outcome expectancies). This momentary set of cues acquires conditioned reinforcing value from the reward and may transmit conditioned reinforcing value by reliably following earlier cue patterns or through generalization to similar constellations of cues. It seems likely, given this model, that the cues at the beginning of the 30-sec red stimuli on training and test trials, which *are* distinguished from one another by STM cues, acquire equal reinforcing value because they are reliably followed by the same set of cues.

Two possible alternative explanations for these results must be considered. The first possibility is that the 30-sec stimuli are more effective conditioned reinforcers than the 3-sec stimuli, because each second the conditioned reinforcer is present adds to the total reinforcing value. According to this argument, even though the 3- and 30-sec stimuli are associated with the same magnitude of primary reward, the 30-sec Sr+ will be a more effective conditioned reinforcer than the 3-sec Sr+ simply because it is of a longer duration.

Another possible alternative explanation involves the notion that primary rewards can have a direct reinforcing effect on responses that precede them by a short interval of time but not by a longer interval. According to this rationale, the pigeons in Group 3 chose Sr− on test trials because the reward 3 sec later directly reinforced the response producing Sr−. This primary reward overrode any possible generalization of conditioned reinforcing value from training trials that might have favored choice of Sr+. Thus, due to the direct reinforcing effects of food 3 sec later, the pigeons in Group 3 chose Sr−. Furthermore, this rationale specifies that the food reward delayed by 30 sec cannot directly reinforce choice of Sr− because primary reinforcement cannot bridge a gap of 30 sec. Thus, the generalized conditioned reinforcing value of Sr+ acquired on training trials produced choice of Sr+ on test trials for pigeons in Group 30.

Both of these alternative explanations are made less likely by an experiment by Astley and Perkins (1985, Ex-

periment 2). Their procedure was like that within the white houselight context of the present experiment, except that the probability of food on test trials was 50% regardless of the alternative chosen. In addition, the duration of the potential conditioned reinforcers on training and test trials was varied factorially. There were four groups of pigeons: 3-30, 30-30, 30-3, 3-3, where the first number represents the duration of the colored keylights on training trials in seconds and the second number represents the duration of the colored keylights on choice and forced trials. According to both of the alternative hypotheses described above, the relative choice of Sr+ should be determined by the relative percentage of primary reward for Sr+ on test trials when the colored keylights are 3 sec long on test and by the relative percentage of reinforcement overall when they are 30 sec long on test. The duration of the stimuli on training trials is irrelevant to these alternative hypotheses. The alternatives hypothesize that choice on test trials should approximate 50% for Groups 30-3 and 3-3 and that choice of Sr+ should be high for Groups 3-30 and 30-30. Astley and Perkins found that only the pigeons in Group 30-30 chose Sr+ more than 50% of the time. The pigeons in Groups 3-3, 30-3 and 3-30 demonstrated no preference for either Sr+ or Sr−. These findings indicate that there is no generalization between the two trial types if the duration of the colored keylights is shorter than the STM on either test or training trials, in accord with the present experimental hypothesis.

The current findings have important implications for application of the concept of conditioned reinforcement. First, conditioned reinforcing value must be considered to be a property of patterns of cues rather than of component stimuli alone. Although referring to the red keylight as the stimulus that acquires conditioned reinforcing value may serve as a convenient shorthand, we must be cautious not to take our shorthand terminology too seriously. Second, the current data are consistent with the notion that the cue pattern that immediately precedes reward acquires conditioned reinforcing value, and that this value is transmitted to earlier cue patterns by means of back-chaining and/or generalization. This is certainly not a novel idea, inasmuch as it was first stated by Kenneth Spence in 1947, but it is one that bears repeating.

Several studies demonstrate that pigeons can successfully perform on a discrimination involving three factors (e.g., Wassserman, Nelson, & Larew, 1980; Weisman, Dodd, & Larew, 1980). The present results indicate that

conditioned reinforcing value can accrue to a simulus complex involving three cues. These results also indicate that a salient external stimulus such as a houselight does not necessarily overshadow memory cues in acquisition of conditioned reinforcing value, as evidenced by the performance of the pigeons in Group 3. It may be that conditioned reinforcing value will accrue to any discriminable cue pattern. Further research is necessary to determine under what conditions various sorts of cues might enter into the conditioned reinforcing complex.

REFERENCES

ASTLEY, S. L., & PERKINS, C. C. (1985). Stimulus duration and conditioned reinforcing value measured by a learning-tests procedure. Animal Learning & Behavior, 13, 18-24.
FEDORCHAK, P. M., & BOLLES, R. C. (1986). Differential outcome effect using a biologically neutral outcome difference. Journal of Experimental Psychology: Animal Behavior Processes, 12, 125-130.
GRANT, D. S. (1976). Effect of sample presentation time on long-delay matching in the pigeon. Learning & Motivation, 1, 580-590.
HANCOCK, R. A. (1982). Tests of the conditioned reinforcing value of sequential stimuli in pigeons. Animal Learning & Behavior, 10, 46-54.
HONIG, W. K. (1978). Studies of working memory in the pigeon. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), Cognitive processes in animal behavior (pp. 211-248). Hillsdale, NJ: Erlbaum.
ROBERTS, W. A., & GRANT, D. S. (1976). Studies of short-term memory in the pigeon using the delayed matching to sample procedure. In D. L. Medin, W. A. Roberts, & R. T. Davis (Eds.), Processes of animal memory (pp. 79-112). Hillsdale, NJ: Erlbaum.
SPENCE, K. W. (1947). The role of secondary reinforcement in delayed reward learning. Psychological Review, 54, 1-8.
WASSERMAN, E. A., NELSON, K. R., & LAREW, M. B. (1980). Memory for sequences of stimuli and responses. Journal of the Experimental Analysis of Behavior, 34, 49-59.
WEISMAN, R. G., DODD, P. W., & LAREW, M. B. (1980). Representation and retention of two-event sequences in pigeons. Journal of Experimental Psychology: Animal Behavior Processes, 6, 312-325.
WILLIAMS, B. A. (1982). On the failure and facilitation of conditional discrimination. Journal of Experimental Analysis of Behavior, 38, 265-280.
WILLIAMS, B. A. (1984). Relative stimulus validity in conditional discrimination. Animal Learning & Behavior, 12, 117-121.
WINTER, J., & PERKINS, C. C. (1982). Immediate reinforcement in delayed reward learning in pigeons. Journal of the Experimental Analysis of Behavior, 38, 169-179.
ZENTALL, T. R., HOGAN, D. E., & EDWARDS, C. A. (1984). Cognitive factors in conditional learning by pigeons. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), Animal cognition (pp. 389-405) Hillsdale, NJ: Erlbaum.