

## Choice biases in delayed matching-to-sample duration with pigeons: Manipulations of ITI and delay illumination

Ronald Kelly and Marcia L. Spetch

*University of Alberta, Canada*

Two experiments investigated the effects of similarity between intertrial interval (ITI) and delay illumination on the choose-short effect. Different groups of pigeons learned to match “short” (2 s) and “long” (6 or 8 s) food samples to green and red test stimuli in a matching-to-sample procedure with a 5-s training delay. Subsequent 10- and 20-s delay tests revealed choose-short effects if the ITI and delay were both illuminated (i.e., group ON–ON), if the ITI and delay were both dark (i.e., group OFF–OFF), and if the ITI was illuminated and the delay was dark (i.e., group ON–OFF). In addition, either a choose-short effect or a choose-long effect was observed if the ITI was dark and the delay was illuminated (i.e., group OFF–ON). Results are incompatible with the *confusion/instructional failure* view of the choose-short effect.

The delayed matching-to-sample (DMTS) task with duration samples has been a particularly useful procedure for examination of coding processes in pigeons' working memory (for a review, see Grant, Spetch, & Kelly, 1997). Each trial begins with a temporal sample, consisting of either a “short” (e.g., 2-s) or “long” (e.g., 8-s) duration of a stimulus event (e.g., presentation of a food hopper), followed by an opportunity to choose between two test stimuli, such as green and red pecking keys. On trials in which a short sample had occurred, one choice response—say, pecking green—is considered correct and is therefore reinforced with food, whereas the alternate choice response, pecking red, is considered incorrect and is not reinforced. Conversely, on trials in which the long sample had occurred, choice of green is incorrect and choice of red is correct. Finally, a darkened intertrial interval (ITI, e.g., 45 s) separates trials within an experimental session. Numerous studies have shown that birds quickly learn to choose accurately in DMTS duration tasks (e.g., Cohen, Calisto, & Lentz, 1981; Fetterman, 1995; Grant & Kelly, 1996; Kraemer, Mazmanian, & Roberts, 1985; Spetch & Wilkie, 1981; Wilkie & Willson, 1990).

---

Requests for reprints should be sent to Marcia L. Spetch, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada, T6G 2E9. Email: mspetch@ualberta.ca

This research was supported by a Natural Sciences and Engineering Research Council of Canada research grant held by the second author. We thank Karen Freeman and Tiana Rust for assistance with the experiments.

During an extended-delay test, when the opportunity to choose between test stimuli is postponed by a brief interval of darkness (e.g., 10- or 20-s delay), birds show an increased tendency to choose the test stimulus that is correct after the short sample, regardless of which sample had actually occurred on the trial. The result of this tendency is stable and accurate choice responding on extended-delay trials in which the short sample had been presented, but there is a precipitous decline in choice accuracy on extended-delay trials in which the long sample had been presented. This asymmetry in short- and long-sample retention functions with increasing delay is known as the *choose-short effect* (Spetch, 1987; Spetch & Wilkie, 1982, 1983).

One popular theoretical account of the choose-short effect invokes working memory processes involving analogical coding and subjective shortening (Spetch, 1987; Spetch & Wilkie, 1982, 1983). Accordingly, a duration sample is believed to establish a unique analogical code in working memory, which represents the cumulative nature of the associated temporal stimulus. For example, the short sample may be coded as the number of periodic elements that accumulate during the 2-s event, and the long sample as the number of elements that accumulate over the 8-s event. During an extended-delay test, the working-memory code established by the sample on that trial undergoes gradual weakening, conceptualized as a systematic loss of elements that had accumulated during the sample event, i.e., a “subjective shortening” of the coded duration.

As a consequence of these processes, on an extended-delay trial in which the long sample had been presented, the subjectively shortening long-sample code comes to increasingly correspond to the few elements of a short-sample code in reference memory. As a result, choice of the test stimulus that is correct on short-sample trials becomes increasingly likely. By contrast, on an extended-delay trial in which the short sample had occurred, the subjectively shortening short-sample code continues to correspond more closely to the few elements of a short-sample code in reference memory than to the many elements of a long-sample code. As a result, choice of the test stimulus that is correct on short-sample trials remains likely (Spetch, 1987; Spetch & Wilkie, 1982, 1983).

The instructional failure/confusion hypothesis provides an alternative account of the choose-short effect (Sherburne, Zentall, & Kaiser, 1998). In this model, similarity between illumination conditions of the ITI and the delay are held accountable for the phenomenon. Specifically, when darkness prevails during both the ITI and the delay, subjects have difficulty in discriminating between the two intervals. Then, when test stimuli are presented after the delay, “instructional failure” is likely to occur. Loosely speaking, the animal is uncertain about what to do on such trials because, during training, choice responding was never permitted immediately after an interval of darkness (i.e., after the ITI). Consequently, choice of the short-associated test stimulus becomes more likely, presumably because the content of working memory during instruction failure (viz., *no sample*) is judged to be more similar to the short-sample code than to the long-sample code in reference memory.

Several findings have given support to the instructional failure/confusion view. In some experiments using a DMTS duration task with a 0-s training delay, birds were given test trials in which the chamber was illuminated during the extended delay. Under these conditions, the choose-short effect did not occur and, instead, accuracy on short- and long-sample trials declined at equivalent rates as a function of increasing delay.

Presumably, illumination distinguished the delay from the dark ITI, thereby eliminating instructional failure and ensuring that choice responding was controlled by sample-activated codes in working memory (Fetterman & MacEwen, 1989; also see Spetch & Rusak, 1992b; Sherburne et al., 1998).

More generally, the instructional failure/confusion hypothesis predicts that any operation that distinguishes the ITI and the delay should reduce instructional failure and, thereby, tend to eliminate the choose-short effect. In other words, the choose-short effect should be present whenever the ITI and delay are similar in illumination (and are difficult to discriminate), and absent whenever these two intervals differ in illumination (and are easy to discriminate). This prediction holds regardless of the specific illumination conditions that prevail in each of the two intervals. And in fact, Sherburne et al. (1998) found that the choose-short effect occurred if the illumination condition of the delay was the same as that of the ITI used during training (i.e., both dark or both illuminated), whereas the choose-short effect was absent if the illumination condition of the delay differed from that of the training ITI (i.e., ITI dark and delay illuminated, or ITI illuminated and delay dark).

However, other experimental results seem to be inconsistent with the instructional failure/confusion hypothesis. In particular, this account predicts greater overall accuracy on extended-delay trials under conditions in which ITI and delay illumination differ. After all, if such conditions eliminate instructional failure and enhance control by codes in working memory, then it stands to reason that accuracy should be greater than in conditions in which ITI and delay illumination are the same and, consequently, instructional failure exists and impedes control by working-memory codes. Although Sherburne et al. (1998) noted such a difference, it was observed in only one of two extended-delay phases. Moreover, neither Fetterman and MacEwen (1989) nor Spetch and Rusak (1992b) found higher overall accuracy under conditions in which ITI and delay illumination differed.

And, quite contrary to the instructional failure/confusion account, one might expect that novelty of illumination during the delay should increase, rather than decrease, instructional failure. Consider that in the DMTS duration task, aside from the temporal novelty of the delay, groups that receive different illumination during the delay and the ITI will encounter additional novelty, namely, the illumination condition of the delay. In fact, a change in delay illumination from training to testing has been shown to produce retroactive interference (i.e., reduced accuracy) in DMTS tasks using nontemporal samples, such as line orientation or colours (e.g., Cook, 1980; Maki, Moe, & Bierly, 1977).

If novelty of delay illumination retroactively interferes with sample memory in DMTS duration tasks, then overall accuracy should be lower, not higher, when delay illumination is novel. Fetterman and McEwan (1989) observed such a reduction in the novel illumination condition (decreased discriminability,  $A'$ ), but only during the first replication and only at the longest delay. Moreover, the opposite effect was seen in one of two replications conducted by Sherburne et al. (1998). In Spetch and Rusak's (1992b) study, overall accuracy was similar across conditions, but was close to chance on all extended-delay trials.

In evaluating the relative merits of the analogical coding/subjective shortening account and the instructional failure/confusion view, and the role, if any, of retroactive interference in pigeons' DMTS duration tasks, it may be important to note that previous

investigations of the relationship between ITI/delay similarity and the choose-short effect have always involved training with a 0-s delay. This procedural feature may be of some consequence. Consider that during training, the subject can gain experience with the illumination condition that prevails during the ITI, but it cannot gain experience with the illumination condition that will prevail during extended delays, because of course the training delay is 0 s in duration. In other words, the ITI/delay similarity variable is confounded with novelty of the delay illumination. Specifically, if the illumination conditions are the same during the ITI and delay, then the subject is already familiar with the illumination condition of the delay. By contrast, if the illumination conditions of the delay and ITI differ, then the delay presents a novel context to the subject.

Regardless of the extent to which previous results may reflect retroactive interference, instructional failure, both processes, or neither process, it is prudent to remove delay illumination novelty as a potentially confounding variable when investigating the effects of ITI/delay similarity. This can be done by using a non-zero delay during training, thereby giving subjects experience with the illumination condition that will prevail during extended delays. As a consequence of non-zero delay training, the illumination conditions of the extended delays should be familiar to subjects in all groups. Such a design would allow a purer test of the effect of the similarity of the ITI and the delay, because potentially confounding effects of illumination novelty would be eliminated.

In several studies, pigeons have demonstrated the choose-short effect during extended delay testing when trained with a non-zero delay (e.g., Grant & Kelly, 1998; Spetch, 1987; Spetch & Rusak, 1989, 1992a). However, in these studies, illumination conditions during the delay and ITI were the same, namely, both dark. In the present study, birds were trained in a DMTS duration task using a 5-s delay, and short and long food samples. In Experiment 1A, the illumination condition (dark vs. illuminated) was manipulated in both temporal intervals (delay and ITI) to form four groups: group OFF-OFF for which both intervals were dark, group ON-ON for which both intervals were illuminated, group ON-OFF for which the ITI was illuminated and the delay was dark, and group OFF-ON for which the ITI was dark and the delay was illuminated. Experiment 1B consisted of a replication involving only the ON-OFF and OFF-ON conditions. In all cases, acquisition at the 5-s delay was followed by test trials in which extended-delays of 10 and 20 s occurred.

The confusion view predicts instructional failure, and hence the choose-short effect, only in groups for which the ITI and delay are highly confusable (i.e., groups ON-ON and OFF-OFF). In groups for which the ITI and delay intervals can be easily distinguished (i.e., groups ON-OFF and OFF-ON), instructional failure should not occur, and hence the choose-short effect should be absent. In these latter two groups, confusion between the ITI and delay should be considerably reduced by use of distinct illumination conditions during the ITI and delay, as in the Sherburne et al. (1998) study. Moreover, any remaining tendency to confuse intervals in groups ON-OFF and OFF-ON should be prevented even further, because birds would have become familiar with the 5-s delay and its illumination condition during training. In essence, the distinct illumination conditions of the delay should acquire good stimulus control of choice behaviour during training, and this stimulus control should remain high during testing because the illumination conditions remain intact. Thus, although the confusion view allows that a choose-short

effect may occur in groups ON–ON and OFF–OFF, it predicts that this effect will not occur in groups ON–OFF and OFF–ON.

By comparison, the analogical coding/subjective shortening account predicts choice biases in all groups of the present study. During training, birds would be given considerable experience with the delay illumination condition that would prevail during testing. Thus, in contrast to previous studies that varied ITI and delay illumination, the illumination condition of the extended-delay test would not be novel for any of the groups. Consequently, any retroactive interference due to novelty of delay illumination should be eliminated, and choose-short effects in groups ON–OFF and OFF–ON would not be obscured. Thus, the most straightforward prediction of the analogical coding/subjective shortening model is a choose-short effect in all groups.

Nevertheless, closer consideration of the trial conditions experienced during training with a constant delay suggests another possibility—namely, that some birds might incorporate the delay into the functional sample duration. Specifically, although the short and long samples provide the operational cues for correct test stimulus choice, *trial duration* (i.e., the duration of the sample plus the duration of the 5-s delay) provides a redundant cue. Learning to base choices on trial duration instead of, or in addition to, the sample durations may be particularly likely if the illumination conditions of the delay are similar to those of the sample event (i.e., illuminated), yet distinct from the ITI (i.e., group OFF–ON). In this case, a bird could learn to base choices on the duration of *illumination*. As a consequence, extending the delay would be functionally equivalent to extending the sample, and the bird should respond as if the long sample had occurred (i.e., a *choose-long effect*) just as birds do on tests in which the actual sample is lengthened beyond the training value (Spetch & Grant, 1993). It should be noted that there is evidence suggesting that under some training conditions, pigeons treat the ITI or delay interval as part of the functional sample to be timed (Santi, Coyle, Coppa, & Ross, 1998; Santi, Ross, Coppa, & Coyle, 1999).

## EXPERIMENT 1A

### Method

#### Subjects

Sixteen experimentally naive Silver King pigeons (*Columba livia*), each under 1 year of age, were individually housed in a common colony room. Each animal was reduced to and maintained at approximately 85% of its free-feeding weight and had continuous access to water and grit in its home cage. A 12-hr/12-hr diurnal schedule (light onset at 7:00 a.m.) cycled in the colony room. Birds were randomly assigned to one of four groups with the restriction that there were four subjects in each group. Sessions began at roughly the same time every day for each subject, and sessions never exceeded 1 hr in duration. Five to six sessions were given per week.

#### Apparatus

Four operant chambers with inner dimensions of approximately 34.0 cm × 30.0 cm × 34.0 cm (height × length × width) were located in a darkened, sound-attenuated experimental room. An

observation panel on one end wall held two pecking keys (each 2.5 cm in diameter), one to the left of centre and the other to the right (inner edges of keys spaced 13.3 cm). The bottom edges of the keys were raised 23.0 cm above a barred floor. A force of at least 0.25 N applied to a key could be detected. Projectors (Industrial Electronics, Inc., Van Nuys, CA) behind each key were equipped to present homogeneous fields of green and red onto the keys. A 28-V shielded houselamp was attached to the observation panel 31.0 cm above the floor, and light from the lamp was directed toward the ceiling of the chamber. A rectangular opening in the panel (5.5 cm wide  $\times$  5.0 cm high), the bottom edge of which was 9.0 cm above the floor, provided access to a retractable food magazine filled with mixed grain. When the magazine was presented, the opening was lit by a recessed and shielded 28-V bulb in the opening. Ventilation fans in each chamber produced white noise that masked extraneous sound. Stimulus events were directed from, and responses were recorded by, a micromputer located in an adjacent room.

## Procedure

*5-s DMTS duration training.* Following magazine training and autoshaping to red and green keylights, subjects received trials that began with a duration sample, consisting of either a "short" (2-s) or a "long" (8-s) presentation of the food magazine from which the animal could eat. Within each session, each sample occurred equally often in a randomly determined order. A 5-s delay followed termination of the sample, and then green and red test stimuli were presented, one on the left key and the other on the right key. The left/right position of test stimuli varied randomly over trials, with the restriction that both possible arrays occurred equally often after each sample within each session. Each session consisted of 48 trials.

For two subjects in each group, choice of the green test stimulus was designated correct following the short sample and incorrect following the long sample, whereas choice of red was correct after long and incorrect after short. For the remaining two subjects in each group, the relation between sample duration and correct/incorrect choice was reversed. Correct choices always resulted in reinforcement, consisting of presentation of the illuminated food magazine. Duration of reinforcement was constant for each subject, but varied across subjects (from 2 to 3 s), and was followed by a 45-s ITI. Incorrect choices were never reinforced, but instead immediately produced the ITI. Accuracy was computed as the proportion of correct responses on all trials within a session, multiplied by 100 (i.e., overall percentage correct), and was further separated with respect to sample type (i.e., short- and long-sample percentage correct).

Groups differed in the illumination conditions that prevailed during the ITI and during the delay. Subjects in group ON-ON received trials with the houselight illuminated throughout the session, whereas subjects in group OFF-OFF received trials with the houselight dark throughout the session. Subjects in group ON-OFF received trials in which the houselight was illuminated during the ITI, was darkened with onset of the duration sample, and then was illuminated upon completion of the trial (i.e., after reinforcement or an incorrect choice). Finally, subjects in group OFF-ON received trials in which the houselight was dark during the ITI, was illuminated with onset of the duration sample, and then was darkened upon completion of the trial (i.e., after reinforcement or an incorrect choice). Training was given until individual birds met or exceeded a criterion of 80% correct trials over five consecutive sessions.

*Extended-delay testing.* Sessions were identical to those of training, except that the 5-s baseline delay occurred on only 50% of trials, whereas a 10- and a 20-s delay each occurred on 25% of trials. The order of delay trials was determined randomly within a session. For all groups, illumination condition during a 10- or a 20-s delay was identical to that of the baseline 5-s delay, which, in turn,

was identical to the illumination condition during the 5-s delay of training. Trials at each delay were counterbalanced with respect to sample type, and each of the two possible test stimulus arrays occurred equally often after each sample at each delay.

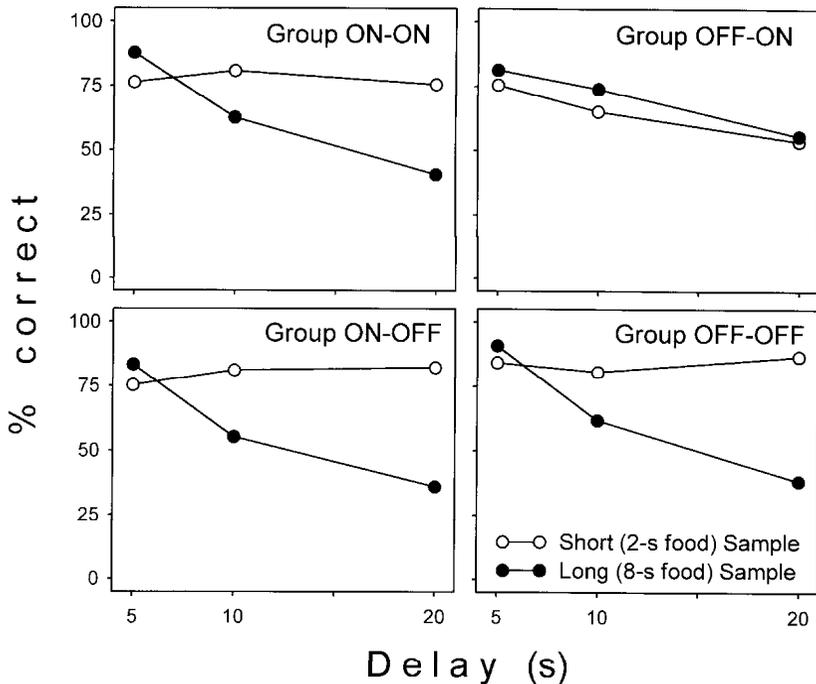
Extended-delay testing consisted of eight consecutive test sessions. Accuracy was separated with respect to sample type at each of the three delays.

## Results and discussion

The critical region for rejection of the null hypothesis was defined as  $\alpha = .05$ .

*5-s DMTS duration training.* Acquisition was fastest in group OFF-OFF ( $M = 18.5$ , range = 12 to 34 sessions to criterion), intermediate in groups ON-ON and OFF-ON (respective  $M$ s = 24.8 and 26.3, respective ranges = 17 to 35 and 21 to 37 sessions), and slowest in group ON-OFF ( $M = 42.8$ , range = 17 to 70 sessions). However, these differences were not significant in a one-way analysis of variance (ANOVA),  $F = 2.14$ .

*Extended-delay testing.* Graphs in Figure 1 display accuracy on short- and long-sample trials as a function of delay in groups ON-ON, OFF-ON, ON-OFF, and OFF-OFF. Functions indicate the presence of the choose-short effect in all groups except group OFF-ON.



**Figure 1.** Accuracy (% correct) on short- and long-sample trials as a function of delay in groups ON-ON (upper left), OFF-ON (upper right), ON-OFF (lower left), and OFF-OFF (lower right) during extended-delay testing in Experiment 1A.

A split-plot ANOVA, with group as a between-subjects factor and sample and delay as within-subjects factors, returned significant main effects for sample,  $F(1, 12) = 16.28$ , and delay,  $F(2, 24) = 190.36$ , as well as significant two-way interactions for Sample  $\times$  Group,  $F(3, 12) = 4.06$ , and Sample  $\times$  Delay,  $F(2, 24) = 29.85$ . Most important, the three-way Sample  $\times$  Delay  $\times$  Group interaction was significant,  $F(6, 24) = 3.07$ .

This three-way interaction was explored by conducting separate repeated measures ANOVAs for each group, with sample and delay as factors. These analyses revealed significant Sample  $\times$  Delay terms in groups ON-OFF, OFF-OFF, and ON-ON, respective  $F_s(2, 6) = 9.84, 22.15, \text{ and } 21.56$ , and absence of a significant Sample  $\times$  Delay term in group OFF-ON,  $F < 1$ . Thus, analyses confirmed a statistically reliable choose-short effect in all groups except group OFF-ON.

To determine whether the groups differed in overall retention accuracy, a one-way group ANOVA was conducted on overall accuracy (i.e., collapsed over short- and long-sample trials) at the 20-s delay. This analysis revealed no significant difference between groups,  $F = 1.51$ , indicating that groups did not differ in overall accuracy at the 20-s delay. Thus, the absence of a choose-short effect in group OFF-ON was not accompanied by an increase in overall accuracy at the longest delay.

## EXPERIMENT 1B

Experiment 1B was designed to replicate results for the critical two conditions, groups ON-OFF and OFF-ON, with several procedure variations. We hoped to establish the reliability and generality of the results obtained in Experiment 1A, and also to expand the database for these two conditions.

### Method

#### Subjects

Eight experimentally naive racing pigeons (*Columa livia*), each under 1 year of age, were maintained exactly as were previous animals. Birds were randomly assigned to one of groups ON-OFF or OFF-ON, with the restriction that there were four subjects in each group.

Two of the original four subjects in group OFF-ON were dropped from the experiment for failure to acquire the baseline 5-s DMTS duration task, and were replaced by two more naive racing pigeons. One of these subjects was also subsequently dropped for failure to acquire, and was replaced by another naive racing pigeon. Thus, in total, 11 pigeons served in this experiment, but only 8 pigeons participated in delay testing.

#### Apparatus

The apparatus was the same as that in Experiment 1A.

#### Procedure

*5-s DMTS duration training.* Training for the two groups was identical to that for equivalent groups in Experiment 1A in all respects except for the following. First, short and long samples

consisted of 2- and 6-s food durations. Second, the ITI was 30 s. Third, a correction procedure was in effect, such that following an incorrect response and the ensuing ITI, all trial events of the preceding trial were repeated. Choices on correction trials were reinforced if correct, but were never included in computations of accuracy. Incorrect choices on correction trials merely reinitiated the correction trial.

After the first 20 sessions of training, accuracy was still at chance levels in both groups. In an attempt to facilitate acquisition, the delay between offset of the sample and onset of the test stimuli was reduced to 2 s. Then, over sessions, each time an accuracy criterion was met by a subject, the delay for that subject was incremented by 1 s until the original value of 5 s was restored. The accuracy criterion used was 80% correct or greater on both trial types in a session in which all 48 trials had been completed. However, some subjects routinely failed to complete all trials within the 1 hour allotted for each session. This failure to complete trials was likely due to satiation effects brought about by use of food samples combined with the relatively small size of racing pigeons. Thus, the accuracy criterion was relaxed to 80% correct or greater on both trial types in each of consecutive sessions which, when aggregated, met or surpassed the 48-trial criterion.

After the original 5-s delay had been restored, additional training was given until the number of correct trials was 80% or greater on both trial types over two consecutive blocks of four complete sessions (i.e., 348 trials in total). Again, however, because some subjects failed to complete all trials within a session, the accuracy criterion was relaxed to 80% correct or greater on both trial types in consecutive sessions that when aggregated, met or surpassed the 348-trial criterion.

*Extended-delay testing.* Testing for the two groups was identical to that for equivalent groups in Experiment 1A in all respects except for the following. Extended-delay testing consisted of two replications, each comprising four consecutive test sessions in which each of the 48 trials (i.e., 192 trials in total) had been completed. As during training, however, if subjects failed to complete all trials in a test session, additional sessions were given until the 192-trial criterion was met or surpassed. The correction procedure was removed during extended-delay test sessions. The two replications were separated by several "recovery" sessions, which were identical to 5-s DMTS duration training. The criterion for recovery was 80% correct or greater on both trial types in each of two consecutive recovery sessions in which all 48 trials (i.e., 96 trials in total) were completed or, alternatively, in each of consecutive recovery sessions that, when aggregated, met or surpassed the 96-trial criterion.

## Results and discussion

*5-s DMTS duration training.* Acquisition was extremely slow, particularly in group OFF-ON. Because many training sessions were incomplete, the following comparison of group acquisition rates used total trials to acquisition as the dependent measure. For subjects that met acquisition criteria, mean trials to acquisition were 2,710.3 (range = 2,198 to 3,264) in group ON-OFF and 5,117.3 (range = 2,026 to 8,236) in group OFF-ON. Including data from the three subjects in group OFF-ON that were dropped from the experiment for failure to acquire (using final trial number as trials to acquisition), the group mean was 7,619.0 (range = 2,026 to 12,774). A one-way ANOVA on data for all 11 subjects indicated significantly faster acquisition in group ON-OFF,  $F(1, 9) = 5.32$ .

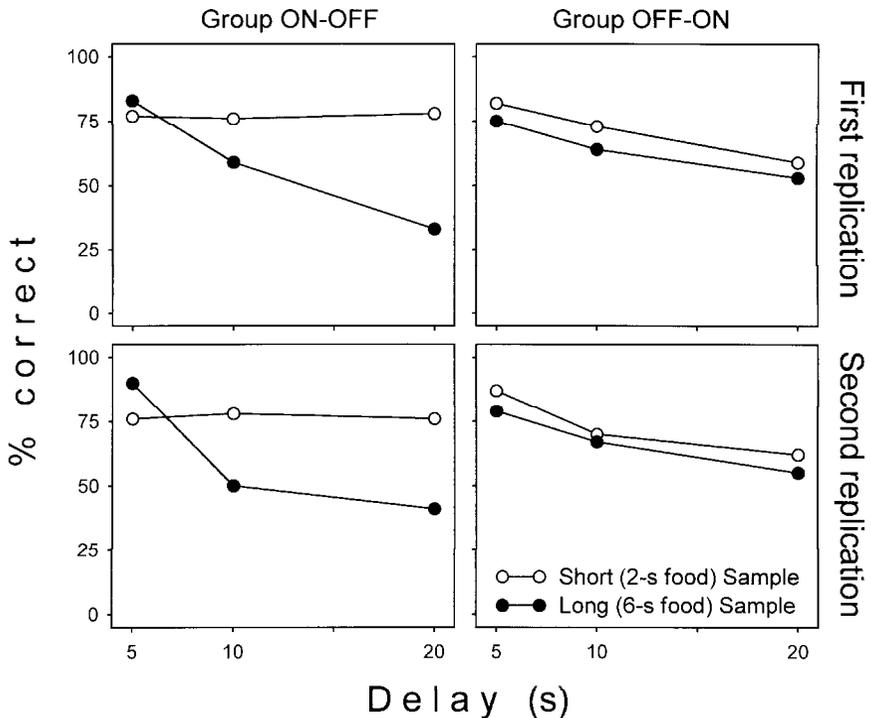
The generally slower acquisition in this experiment than in Experiment 1A could reflect subject factors: in particular, rapid satiation encountered with use of smaller racing

pigeons. In addition, slower acquisition could reflect procedural factors: in particular, use of a more difficult duration discrimination (2 vs. 6 s rather than 2 vs. 8 s).

*Extended-delay testing.* Graphs in Figure 2 display accuracy on short- and long-sample trials as a function of delay in group ON-OFF and OFF-ON in the first and second replications. Functions indicate the presence of the choose-short effect in both replications for group ON-OFF, and in neither replication for group OFF-ON.

Split-plot ANOVAs with group as a between-subjects factor, and sample and delay as within-subjects factors revealed significant main effects of delay in both the first and the second replications, respective  $F_s(2, 12) = 31.37$  and  $20.00$ . In addition, the main effect of sample was significant in the second replication,  $F(1, 6) = 7.53$ , and the Sample  $\times$  Delay interaction was significant in the first replication,  $F(2, 12) = 5.17$ . Most important, the Sample  $\times$  Delay  $\times$  Group interaction was significant in both replications, respective  $F_s(2, 12) = 5.45$  and  $4.74$ .

The three-way interactions were explored by conducting separate repeated measures ANOVAs for each group and replication, with sample and delay as factors. These analyses indicated presence of significant Sample  $\times$  Delay terms in both replications for group ON-OFF, respective  $F_s(2, 6) = 8.88$  and  $8.81$ , and absence of significant Sample  $\times$  Delay terms in both replications for group OFF-ON,  $F_s < 1$ . Thus, as in Experiment 1A,



**Figure 2.** Accuracy (% correct) on short- and long-sample trials as a function of delay in groups ON-OFF (left column) and OFF-ON (right column) in the first replication (upper row) and second replication (lower row) of extended-delay testing in Experiment 1B.

analyses confirmed a statistically reliable choose-short effect in group ON-OFF but not in group OFF-ON.

To determine whether the groups differed in overall retention accuracy, a one-way group ANOVA was conducted on overall accuracy (i.e., collapsed over short- and long-sample trials and over the two delay tests) at the 20-s delay. This analysis revealed no significant difference between groups,  $F < 1$ , indicating that groups did not differ in overall accuracy at the 20-s delay.

Thus, despite the greater difficulty with acquisition of the task, and despite the many procedural changes, the extended delay results for this experiment are remarkably consistent with those of Experiment 1A. Group ON-OFF showed a robust choose-short effect whereas group OFF-ON did not. The difference in choice bias following a long retention interval was not accompanied by a difference in overall accuracy.

## GENERAL DISCUSSION

The confusion view maintains that the choose-short effect is due to instructional failure brought about by similarity in, and hence difficulty in discriminating between, the ITI and the delay (Sherburne et al., 1998; Zentall, 1997). Thus, this account correctly anticipates presence of choose-short effects in groups OFF-OFF and ON-ON, for which ITI and delay illumination were identical, and absence of the choose-short effect in group OFF-ON, for which the dark ITI and illuminated delay differed. However, this view also predicts absence of the choose-short effect in group ON-OFF, for which the illuminated ITI and dark delay also differed. Clearly, as seen in both Figures 1 and 2, presence of the choose-short effect in group ON-OFF violates this prediction.

Other features of the present results are also consistent with the instructional failure/confusion account. As mentioned previously, this view predicts greater overall accuracy on extended-delay trials under conditions in which the ITI and delay illumination differed (i.e., groups ON-OFF and OFF-ON). After all, if such conditions eliminate instructional failure, then it stands to reason that accuracy should be greater on extended-delay trials than in conditions in which the ITI and delay illumination are the same (i.e., groups ON-ON and OFF-OFF) and, consequently, for which instructional failure exists. Contrary to this prediction, however, groups did not differ in overall accuracy at the 20-s delay in Experiment 1A. The two groups in Experiment 1B also did not differ in overall accuracy at the 20-s delay, even though only one group showed a choose-short effect. If for some unforeseen reason the lack of a choose-short effect in group OFF-ON in Experiment 1B was due to reduction of instructional failure, then overall accuracy on extended-delay trials should have been greater than that of group ON-OFF. The results were clearly contrary to this prediction. Although lack of improvement in overall accuracy on extended-delay trials when instructional failure was presumably eliminated conflicts with some of the data reported by Sherburne et al. (1998), it is consistent with results from other studies (Fetterman & MacEwen, 1989; Spetch & Rusak, 1992b) and provides further evidence against the instructional failure/confusion account.

The analogical coding/subjective shortening view predicts a choose-short effect during extended-delay testing, but allows that one of several factors may prevent occurrence

of the phenomenon. For example, because the choose-short effect is assumed to depend upon changes in an analogical representation of event duration, training conditions that encourage non-analogical coding of sample duration (e.g., a many-to-one mapping between samples and test stimuli) can eliminate the choose-short effect (Grant & Spetch, 1993; Santi, Bridson, & Ducharme, 1993). In addition, a choose-short effect may not occur if choices are based on trial duration rather than sample duration, because lengthening the delay alters the functional sample duration (Santi et al., 1998, 1999). As another possibility, the choose-short effect may not occur if illumination conditions during the delay are novel. In this case, retroactive interference may disrupt retention of the analogical code and/or stimulus control by sample duration, thereby resulting in random (non-biased) choice responding on some trials.

In previous investigations of delay/ITI similarity, comparisons between the same and different illumination conditions have been confounded with novelty of delay illumination. In the present study, this confound was eliminated because subjects in all groups had extensive experience with the extended-delay illumination condition, through use of a nonzero (5-s) delay during training. Thus, the analogical coding/subjective shortening model predicts that choice biases should emerge during extended-delay testing, even in conditions for which the delay and ITI were distinguished by different illumination conditions. The choose-short effect seen in the ON-OFF groups of both experiments, for which the dark delay was not novel but was nevertheless distinguished from the ITI, confirms the analogical coding/subjective shortening account and contradicts the confusion explanation.

On the other hand, absence of the choose-short effect in group OFF-ON, for which the illuminated delay was not novel, seems to contradict the analogical coding/subjective shortening explanation. Nevertheless, closer examination of the data suggests that the present results can be reconciled with this model. Specifically, although the mean data from group OFF-ON showed a symmetrical decline in accuracy on short- and long-sample trials, this pattern was not typified by individual subjects' data. Instead, as shown in Table 1, some subjects displayed a choose-short tendency on extended-delay trials (i.e., Bird 229 in Experiment 1A, and Birds 1552 and 394 in Experiment 1B), whereas other subjects displayed a choose-long tendency (i.e., Birds 251 and 244 in Experiment 1A, and Bird 398 in Experiment 1B). In fact, only two of eight birds in the OFF-ON groups of the present study showed a generally symmetrical decline in accuracy on short- and long-sample trials with extended delays (i.e., Bird 207 in Experiment 1A, and Bird 395 in Experiment 1B). Thus, absence of increasingly biased test stimulus responding as a function of extended delay in group OFF-ON may be, at least in part, an artifact of opposing biases within this group cancelling each other out. By comparison, all individuals in the other three conditions typified their group choose-short effects (data not shown).

Our interpretation of the individual subject data in group OFF-ON is that some birds coded the short and long food samples analogically. These subjects demonstrated a choose-short tendency, because the working-memory code subjectively shortened on extended-delay trials. Other birds, however, coded the short and long trial durations analogically. These subjects demonstrated a choose-long tendency, because an extended delay test effectively lengthened the duration being timed. Although our analysis of

TABLE 1

Short- and long-sample % correct as a function of delay in group OFF-ON during extended-delay testing in Experiments 1A and 1B

	Subject number	5-s (training) delay		10-s delay		20-s delay	
		short %	long %	short %	long %	short %	long %
Experiment 1A	207	61.5	87.5	54.2	75.0	52.1	58.3
	251	87.5	79.2	75.0	77.1	56.3	66.7
	244	81.3	71.9	60.4	79.2	41.7	62.5
	229	71.9	86.5	70.8	64.6	64.6	35.4
	Group mean	75.6	81.3	65.1	74.0	53.7	55.7
Experiment 1B	395	81.9	75.3	82.3	69.2	55.5	52.8
	398	86.5	86.5	68.8	64.6	41.7	70.8
	1552	86.7	71.9	77.6	72.9	78.2	41.7
	394	82.3	74.0	56.3	56.3	66.7	50.0
	Group mean	84.3	76.9	71.2	65.7	60.5	53.8

Note: Collapsed over replications in Experiment 1B.

individual subject data is admitted post hoc, the possibility that some birds in group OFF-ON might attend to duration of overall illumination, rather than to duration of the food sample, was anticipated on the basis of other research (Santi et al., 1998, 1999).

Although the use of a non-zero delay during training eliminated the novelty of the illumination conditions during the delay interval, the duration of the extended delays was novel, and one might argue that novelty of delay duration produced instructional failure and hence was responsible for the choose-short effect. However, this possibility would not explain the differences between groups in this or in the previous experiments on delay and ITI illumination because all groups experienced novel delay durations. Moreover, some results by Spetch and Rusak (1992a) further argue against the possibility. They trained two groups of pigeons with non-zero delays. For one group, the training delay was constant at 5 s, whereas for the other group the training delay varied across trials from 2 to 8 s, with a mean of 5 s. If instructional failure due to novelty of the delay is responsible for the choose-short effect, one might expect that the group trained with variable delays would show less of a choose-short effect during subsequent extended delay testing than the group trained with a fixed delay. This was not the case because both groups showed equivalent and significant choose-short effects during delay testing. Moreover, even during training, the group trained with variable delays showed a significant and systematic increase in their tendency to make short choices across the four training delays. Specifically, the tendency to make short choices increased from below 50% at the shortest training delay to above 50% at the longer training delays. Given that all of the training delays were equally familiar to the birds, this result clearly suggests that novelty of the delay duration per se is not responsible for the choose-short effect.

It should be noted that the analogical coding view of memory for duration does not specify the exact nature of the analogical code. Although it is often described in terms of internal processes, such as counts in an accumulator (Grant et al., 1997), the code could be any behaviour that correlates with duration. For example, there is good evidence that

differential sample responding can control choice behaviour in delayed matching-to-sample tasks (e.g., Urciuoli & DeMarse, 1994; Weaver, Dorrance, & Zentall, 1999), and this could also be the case in a duration matching task. That is, in the present study, the birds' memory of the short and long samples could have been coded in terms of the amount of food eaten during each sample or the number of food pecks made during the sample interval. To the extent that the memory generated on long sample trials changes over the delay in a way that makes it systematically more similar to the memory generated on short-sample trials, such a possibility is completely compatible with the analogical coding/subjective shortening account.

In summary, the overall pattern of results was not directly predicted by either the instructional failure/confusion account or the analogical coding/subjective shortening view. However, an examination of the individual subject data suggests that the results can be reconciled with the latter model, whereas the data seem less amenable to the former view. Specifically, the instructional failure/confusion interpretation that has been proposed to explain ITI/delay similarity effects in DMTS duration tasks with pigeons (Sherburne et al., 1998) cannot explain why instructional failure occurred (i.e., why the choose-short effect was present) under one condition in which ITI and delay illumination conditions differed (i.e., group ON-OFF) but not under a second condition in which ITI and delay illumination conditions differed (i.e., group OFF-ON). Moreover, if for some undetected reason instructional failure was reduced only in group OFF-ON, it is unclear why overall accuracy was not higher for this group than for groups that presumably experienced instructional failure. Although we agree that instructional failure/confusion may influence choice responding during extended-delay testing (Sherburne et al., 1998; Zentall, 1997), this view by itself does not appear to provide an adequate explanation for the relationship between ITI/delay similarity and the choose-short effect.

## REFERENCES

- Cohen, S.L., Calisto, G., & Lentz, B.E. (1981). Comparisons of sample stimuli in delayed symbolic matching-to-sample: Some results and implications. *The Psychological Record*, *31*, 77-93.
- Cook, R.G. (1980). Retroactive interference in pigeon short-term memory by a reduction in ambient illumination. *Journal of Experimental Psychology: Animal Behavior Processes*, *6*, 326-338.
- Fetterman, J.G. (1995). The psychophysics of remembered duration. *Animal Learning & Behavior*, *23*, 49-62.
- Fetterman, J.G., & MacEwen, D. (1989). Short-term memory for responses: The "choose-small" effect. *Journal of the Experimental Analysis of Behavior*, *52*, 311-324.
- Grant, D.S., & Kelly, R. (1996). The role of minimum wait time and sample discriminability in the coding of event duration in pigeons. *Learning and Motivation*, *27*, 243-259.
- Grant, D.S., & Kelly, R. (1998). The effect of variable-delay training on coding of event duration in pigeons. *Learning and Motivation*, *29*, 49-67.
- Grant, D.S., & Spetch, M.L. (1993). Analogical and non-analogical coding of samples differing in duration in a choice-matching task in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *19*, 15-25.
- Grant, D.S., Spetch, M.L., & Kelly, R. (1997). Pigeons' coding of event duration in delayed matching-to-sample. In C.M. Bradshaw & É. Szabadi (Eds.), *Time and behavior: Psychological and neuro-behavioral analyses* (pp. 217-264). Elsevier: Amsterdam.
- Kraemer, P.J., Mazmanian, D.S., & Roberts, W.A. (1985). The choose-short effect in pigeon memory for

- stimulus duration: Subjective shortening versus coding models. *Animal Learning & Behavior*, *13*, 349–354.
- Maki, W.S., Jr., Moe, J.C., & Bierley, C.M. (1977). Short-term memory for stimuli, responses, and reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*, 156–177.
- Santi, A., Bridson, S., & Ducharme, M.J. (1993). Memory codes for temporal and nontemporal samples in many-to-one mapping by pigeons. *Animal Learning & Behavior*, *21*, 120–130.
- Santi, A., Coyle, J., Coppa, R., & Ross, L. (1998). The timing of auditory and visual signals by pigeons: Differential sensitivity to intertrial interval duration. *Learning and Motivation*, *29*, 345–365.
- Santi, A., Ross, L., Coppa, R., & Coyle, J. (1999). Pigeons' memory for empty time intervals marked by visual or auditory stimuli. *Animal Learning & Behavior*, *27*, 190–205.
- Sherburne, L.M., Zentall, T.R., & Kaiser, D.H. (1998). Timing in pigeons: The choose-short effect may result from pigeons "confusion" between delay and intertrial intervals. *Psychonomic Bulletin & Review*, *5*, 516–522.
- Spetch, M.L. (1987). Systematic errors in pigeons' memory for event duration: Interaction between training and test delay. *Animal Learning & Behavior*, *15*, 1–5.
- Spetch, M.L., & Grant, D.S. (1993). Pigeons' memory for event duration in choice and successive matching-to-sample tasks. *Learning and Motivation*, *24*, 156–174.
- Spetch, M.L., & Rusak, B. (1989). Pigeons' memory for event duration: Intertrial interval and delay effects. *Animal Learning & Behavior*, *17*, 147–156.
- Spetch, M.L., & Rusak, B. (1992a). Temporal context effects in pigeons' memory for event duration. *Learning and Motivation*, *23*, 117–144.
- Spetch, M.L., & Rusak, B. (1992b). "Time present and time past." In W.K. Honig & J.G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 47–67). Hillsdale, NJ: Erlbaum.
- Spetch, M.L., & Wilkie, D.M. (1981). Duration discrimination is better with food access as the signal than with lights as the signal. *Learning and Motivation*, *12*, 40–64.
- Spetch, M.L., & Wilkie, D.M. (1982). A systematic bias in pigeons' memory for food and light durations. *Behavior Analysis Letters*, *2*, 267–274.
- Spetch, M.L., & Wilkie, D.M. (1983). Subjective shortening: A model of pigeons' memory for event duration. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 14–30.
- Urciuoli, P.J., & DeMarse, T. (1994). On the relationship between differential outcomes and differential sample responding in matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*, 249–263.
- Weaver, J.E., Dorrance, B.R., & Zentall, T.R. (1999). Present/absent sample matching in pigeons: Is comparison choice controlled by the sample stimulus or by differential sample responding? *Animal Learning & Behavior*, *27*, 288–294.
- Wilkie, D.M., & Willson, R.J. (1990). Discriminal distance analysis supports the hypothesis that pigeons retrospectively encode event duration. *Animal Learning & Behavior*, *18*, 124–132.
- Zentall, T.R. (1997). Animal memory: The role of "instructions." *Learning and Motivation*, *28*, 280–308.

*Manuscript received 16 December 1999*  
*Accepted revision received 11 May 2000*

Copyright of Quarterly Journal of Experimental Psychology: Section B is the property of Psychology Press (T&F) and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.