



## Short-term memory for temporal intervals: Contrasting explanations of the choose-short effect in pigeons

Carlos Pinto\*, Armando Machado

Minho University, Portugal

### ARTICLE INFO

#### Article history:

Received 4 March 2010

Received in revised form 7 May 2010

Available online 9 June 2010

#### Keywords:

Choose-short effect

Short-term memory

Timing

Subjective shortening model

Coding model

Confusion hypothesis

Delayed matching-to-sample

Pigeon

### ABSTRACT

To better understand short-term memory for temporal intervals, we re-examined the choose-short effect. In Experiment 1, to contrast the predictions of two models of this effect, the subjective shortening and the coding models, pigeons were exposed to a delayed matching-to-sample task with three sample durations (2, 6 and 18 s) and retention intervals ranging from 0 to 20 s. Consistent with the coding model, the results suggested a sudden forgetting of memories for duration. In Experiment 2, to test the confusion hypothesis, the characteristics of the ITI and the retention interval differed. Contrary to the confusion hypothesis, a choose-short effect was obtained. In both experiments, a test with only two of the three comparison keys was performed. The results suggest three effects that may be controlling the birds' responses: stimulus generalization when no retention interval is present; an increase in random responding at longer retention intervals; and, similarly, an increase in preference for the "short-sample" key at longer retention intervals.

© 2010 Elsevier Inc. All rights reserved.

In a study aimed at understanding short-term memory for temporal intervals, [Spetch and Wilkie \(1982\)](#) exposed pigeons to a matching-to-sample procedure with time intervals as samples. When the houselight was turned on for 2 s (short sample), the pigeons were rewarded for choosing one comparison key (say, Green), and when the houselight was on for 10 s (long sample), they were rewarded for choosing another comparison key (say, Red). After the animals learned this conditional discrimination, the authors introduced a retention interval, ranging from 0 to 20 s, between the offset of the sample and the presentation of the comparison stimuli. They found that accuracy on the long-sample trials decreased with the retention interval (i.e., the pigeons were more likely to choose the short-sample comparison), whereas accuracy on the short-sample trials remained high and did not vary with the retention interval. They labelled the phenomenon the "choose-short effect".

To explain the effect, [Spetch and Wilkie \(1983\)](#) proposed the subjective shortening hypothesis. According to it, the memory for a sample is represented quantitatively on a subjective scale of duration such that the longer the sample, the higher the memory value. During training, the animal learns to choose the long-sample comparison when the memory has a high value and the short-sample comparison when the memory has a low value. To explain the choose-short effect, the authors also assumed that, when the sample stimulus is turned off, the memory decays or "shortens". Hence, on long-sample trials, as the memory for the long sample decays, its value becomes increasingly closer to the value of the short sample and therefore more responses to the short-sample comparison occur. In contrast, on short-sample trials, the decay of the memory does not result in incorrect choices because the shortened memories remain more similar to the memories for the short sample than to the memories for the long sample.

An alternative account of the choose-short effect is [Kraemer, Mazmanian, and Roberts \(1985\)](#)'s coding hypothesis. This account considers that, based on the samples' relative durations, each sample is coded categorically in a non-temporal

\* Corresponding author at: Escola de Psicologia, Universidade do Minho, 4710 Braga, Portugal.

E-mail address: [carlos.arop@gmail.com](mailto:carlos.arop@gmail.com) (C. Pinto).

dimension. During the retention interval, the animals are increasingly likely to forget the sample code. Furthermore, the authors also assumed that when the code is forgotten the animals do not choose the comparison stimuli randomly, but are biased towards the comparison associated with the short sample. They argued that the absence of a code is equivalent to the memory of a 0-s sample, which in turn is more similar to the memory for the short sample (2 s) than to the memory for the long sample (10 s) – hence the choose-short effect. The assumption is supported empirically by the observation that when no sample is presented, pigeons tend to select more frequently the response associated with the short sample (Spetch & Wilkie, 1983).

To test the two models, Kraemer et al. (1985) used a matching-to-sample procedure with three sample durations (short, intermediate and long) and three comparisons, henceforth referred to as Short, Intermediate and Long keys (i.e., the Short key is correct after the short sample, etc.). The models make significantly different predictions concerning the effect of the retention interval on the long-sample trials. Specifically, the subjective shortening model predicts that, as the retention interval increases, preference should change from the Long key to the Intermediate key and finally to the Short key. This ordinal pattern stems from the decay of the memory trace for the long sample. To better understand the prediction, let  $l(t)$ ,  $i(t)$  and  $s(t)$  be the traces of the long, intermediate and short samples, respectively, at  $t$  seconds into the retention interval. Because the samples differ in duration, we assume that  $l(0) > i(0) > s(0)$ , that is, at the sample offset ( $t = 0$  s), the trace for the long sample is greater than the trace for the intermediate sample which in turn is greater than the trace for the short sample. On the long-sample trials, the trace of the long sample decays during the retention interval, that is,  $l(t)$  decreases. Hence, for some value of  $t$ ,  $l(t)$  should be equal to  $i(0)$  and consequently the animal should prefer the Intermediate key. On the long-sample trials, then, the most preferred key should change from the Long key (for all  $t$  such that  $l(t)$  is closer to  $l(0)$  than to  $i(0)$ ), to the Intermediate key (for all  $t$  such that  $l(t)$  is now closer to  $i(0)$  than to  $l(0)$ ), to the Short key (for all  $t$  such that  $l(t)$  is closer to  $s(0)$  than to  $i(0)$ ).<sup>1</sup>

In contrast, the coding model predicts no preference for the Intermediate key because either the animal has retained the code for the long sample and it prefers the Long key – for that is what it learned during training, or it has forgotten the code and it prefers the Short key – for the absence of a code is more similar to the short-sample code than to the intermediate- or long-sample codes. In either case, the subjects should never prefer the Intermediate key.

Kraemer et al. (1985) tested the foregoing predictions with three samples of 0 s (no-sample), 2 and 10 s and retention intervals ranging from 3 to 21 s. A choose-short effect was obtained – on the long-sample trials, preference for the Long key decreased and preference for the Short key increased. More critically, they also observed that the Intermediate key was never the most preferred key. These results are consistent with the coding model but not with the subjective shortening model.

However, Kraemer et al.'s (1985) findings are difficult to interpret for the following reason. On the no-sample trials, the comparison stimuli occurred immediately after the dark ITI and therefore, during training, the pigeons learned to choose the “no-sample” key after a period of darkness. Because similar blackout periods were then used as the retention interval during test trials, it is possible that the preference for the “no-sample” key was due, not to any process of shortening or forgetting, but to the choice reinforced during training (i.e., pecking the “no-sample” key after a blackout). Hence, the first experiment reported in this paper replicated Kraemer et al. (1985) three sample, three comparison procedure but using durations greater than 0 s. Its specific purpose was to see if the pattern of findings would remain consistent with the coding hypothesis or if it would become more consistent with the subjective shortening model; its general purpose was to determine whether the forgetting of memories for duration is a sudden, all-or-none process, consistent with the coding model, or a continuous, gradual process, consistent with the subjective shortening model.

Sherburne, Zentall, and Kaiser (1998) provided a third account for the choose-short effect. Known as the confusion hypothesis, the account suggests that the choose-short effect stems from the instructional ambiguity present in the procedure. When the retention interval is introduced, it is experienced as a new event because the pigeons have never been exposed to it. Moreover, the retention interval is similar to the inter-trial interval (ITI) in the sense that both are spent in darkness. According to the authors, the similarity between the ITI and the retention interval leads the animal to interpret the retention interval as the beginning of a new trial. When the comparison stimuli are turned on, the animal chooses as if the trial had been a no-sample or 0-s sample trial. Because the shortest sample is the most similar to the 0-s sample, the choose-short effect follows. In addition, the longer the retention interval, the more similar to the ITI it appears to be and, consequently, the higher the probability of confusion and of choosing the short-sample comparison key.

The confusion hypothesis predicts that if the retention interval presented on test trials differs from the ITI used during training, the animals will not be confused and therefore similar levels of errors will be made on long and short-sample trials. The two retention functions should be similar. To test the prediction, Sherburne et al. (1998) exposed two groups of pigeons to a conditional discrimination with two durations of a keylight as samples. For one group, the houselight was on during the ITI, but for the other group it was off (i.e., the ITI was spent in darkness). During the training phase there was no retention

<sup>1</sup> We have assumed a subjective shortening model with a decision rule based on the difference between the sample's trace, say  $f(t)$ , and the values  $l(0)$ ,  $i(0)$ , and  $s(0)$ , the values the animal learned to associate with the Long, Intermediate and Short keys, respectively. A different model would assume two thresholds,  $a$  and  $b$ , with  $l(0) > a > i(0)$ , and  $i(0) > b > s(0)$ . Then, the decision rule could be: Choose Long if  $f(t) > a$ , Intermediate if  $a > f(t) > b$ , and Short if  $b > f(t)$ . However, the ordinal pattern of changes in preference with the retention interval given a long sample would not change from the pattern described in the text.

interval. Next, the pigeons were tested with both lit and dark retention intervals. Sherburne et al. (1998) showed that, if the retention interval differs from the ITI (i.e., the houselight is on during the retention interval but off during the ITI or the other way around), the choose-short effect is not obtained. Although this result is consistent with the confusion hypothesis, Kelly and Spetch (2000) failed to replicate it. In addition, the choose-short effect does not depend on the novelty of the retention interval because the effect is observed even when the animals are trained with a retention interval from the beginning of the experiment (Spetch, 1987).

The second experiment presented below further tested the confusion hypothesis. We used the same three samples and three comparison keys used in Experiment 1 but reduced the similarity between the ITI and the retention interval. The goal was to determine whether the choose-short effect would be eliminated, as the confusion hypothesis predicts, or retained, as the two other models predict.

One additional goal of the two experiments was to assess the preference hierarchy for the three comparison keys following samples of different duration. That is, we wanted to know not only which key is the most preferred, but also which is the second most preferred and which is the least preferred. To illustrate, consider the short-sample trials. Without a retention interval, most of the responses are expected to be directed to the Short (correct) key. But what is the relative preference for the Intermediate and Long keys, the incorrect choices? And how does the relative preference for these other keys change with the retention interval? One way to answer the questions is to analyze how errors are distributed to the Intermediate and Long keys. For example, are they randomly distributed or are they more concentrated on the Intermediate key due to stimulus generalization?

But a potential problem with the error analysis is that errors may be so rare as to render their analysis unreliable. One way to solve the problem is to present only two comparison keys during choice. In the foregoing example, after the short sample, the animal would have to choose only between the Intermediate and Long keys. Therefore, in addition to the standard three-key test trials, both experiments included also a second phase in which two-key test trials were run. The results from these two-key test trials revealed the full preference hierarchy for the comparison keys, which may help us to develop better models of the choose-short effect in particular and of how temporal memories are forgotten in general.

## Experiment 1

Experiment 1 used a delayed matching-to-sample procedure with three samples of short (2 s), intermediate (6 s), and long (18 s) duration, and three comparison stimuli, Short, Intermediate, and Long keys. After the pigeons learned the conditional discrimination, we introduced test trials, first with all three keys available, and, second, with only two keys available.

Consider the tests with all three keys available. The subjective shortening and the coding models predict that, on the long-sample trials, the preference for the Long key should decrease with the retention interval, either because of memory trace decay or memory code forgetting. They also predict that preference for the Short key should increase with the retention interval. The difference between the two models lies in the predicted pattern of changes in preference from the Long to the Short key. Whereas the subjective shortening model predicts that, somewhere along the retention interval, there should be a clear preference for the Intermediate key, the coding model predicts no such preference.

The two models predict similar outcomes on the short- and intermediate-sample trials. On short-sample trials, they predict a preference for the Short key regardless of the retention interval. On intermediate-sample trials, they predict a decrease in preference for the Intermediate key and an increase in preference for the Short key as the retention interval increases; choices of the Long key should remain low at all retention intervals. Therefore, the long-sample trials are the most critical to contrast the two models.

Consider now the tests with two keys available. The subjective shortening and coding models do not make very explicit predictions for these test trials. The subjective shortening model states that memories for different samples are represented on the same subjective scale according to their durations. Hence, without further assumptions, it seems that, on the two-key test trials, the model would predict results consistent with stimulus generalization (e.g., on short-sample trials, if the Short key is not available, the pigeon will prefer the Intermediate key over the Long key).

Concerning the coding model, it has not been made clear whether the different codes have an ordinal relation among them. The assumption that when an animal forgets a code it has a tendency to select the shorter comparison available because it is the one closer to a 0-s sample suggests an ordinal relation among the codes, but this assumption is not explicit in the model. If an ordinal relation exists such that the codes for the short, intermediate and long samples are placed in this order, then the model predicts the same outcome as straightforward stimulus generalization. If no ordinal relation exists between the codes for different samples, then the animal will not be able to tell which of the comparisons available is closer to the correct comparison and therefore random responding is expected to occur.

## Method

### Subjects

Four adult pigeons (*Columba livia*) participated in the experiment. The animals had previous experience with timing tasks, one (pigeon P190) with Fixed Interval schedules (Monteiro & Machado, 2009) and the three others with double bisection procedures (Oliveira & Machado, 2008). The birds were maintained at 80% of their free-feeding body weight, with mixed grain provided mostly during experimental sessions. Water and grit were freely available in their home cages. The pigeon

room was maintained in a 13:11 light/dark cycle, with the lights on at 08:00, and its temperature was kept between 20 and 22 °C. The experiment was conducted 6 days per week at approximately the same time of day for each pigeon.

### *Apparatus*

Two identical operant chambers were used. Each chamber was 34 cm high, 35 cm long and 31 cm wide. The response panel was equipped with three circular response keys, 2.5 cm in diameter, arranged in a horizontal row. The bottom edge of each key was 22.5 cm above the wire mesh floor, and the keys were 9 cm apart, center to center. Each key could be illuminated with red, green and yellow hues, or with a white vertical bar against a dark background. The food hopper was accessible through a 6 × 5 cm opening that was centered horizontally on the response panel, 8.5 cm above the floor. When the hopper was raised, a 7.5 W light illuminated its opening and grain became accessible to the pigeon. On the wall opposite the response panel, 30 cm above the floor, a 7.5 W houselight provided general illumination.

The operant chamber was encased in an outer box, which was equipped with an exhaust fan. The fan circulated air through the chamber and masked outside noises. Control of experimental conditions and data recording were performed by personal computers programmed in the C++ language.

### *Procedure*

A symbolic matching-to-sample procedure with three samples, varying in duration, and three comparison stimuli was used. The sample durations were 2 s (short), 6 s (intermediate) and 18 s (long). Because the sample durations followed a geometric sequence, the discriminability between the short and intermediate sample should be equal to the discriminability between the intermediate and long samples (Weber's law; see, e.g., Gallistel, 1990; Gibbon, 1977).

The birds first learned to discriminate the short and long samples. Each trial began with the presentation of the sample stimulus, the houselight, for either 2 or 18 s. Sample offset was followed immediately (i.e., with a 0-s retention interval) by the illumination of two comparison keys (which, for different pigeons, could be red and yellow, red and green, or yellow and green). After a response, the comparison keys were turned off. If the response was correct, reinforcement was delivered and then a 30-s ITI, with all lights off, followed; if the response was incorrect, the ITI followed immediately and the trial was repeated (correction method). After three consecutive incorrect responses, only the correct comparison key was illuminated. The reinforcement duration varied across animals from 2.5 to 3 s.

Each session comprised 60 new trials (excluding correction trials), 30 for each sample. Across trials, both the sample stimulus and the location of the comparison stimuli varied pseudo-randomly, with the constraint that in each session each sample was presented the same number of times and each of the 6 spatial arrangements of the comparison stimuli (2 colors distributed over 3 keys = 6 possibilities) also appeared the same number of times.

Training continued until the pigeon made at least 90% correct responses to each sample (excluding correction trials) for 3 consecutive sessions. When this criterion was met, the intermediate-duration sample was introduced. During choice, the three comparison stimuli (e.g., red, green, and yellow) were presented simultaneously. Each session comprised 54 trials, 18 for each of the 3 samples. In addition, for each sample, each of the 6 permutations of the comparison keys occurred 3 times. Training with 3 samples continued for a maximum of 70 sessions or until an average of 80% correct responses to each sample in a block of 5 sessions was obtained. Then, testing with retention intervals began.

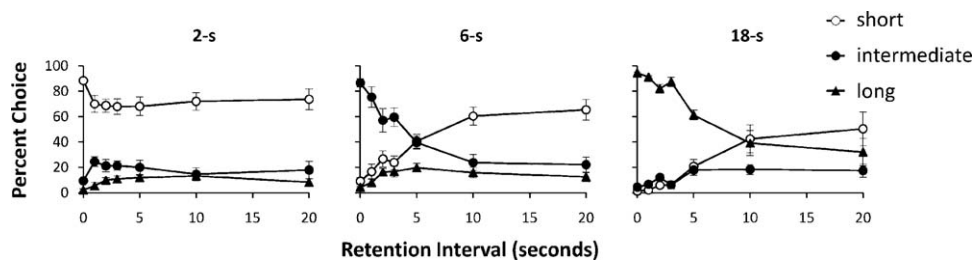
The test phase was divided into two parts. In the first, all 3 comparison keys were available for choice after the retention interval (three-key test). In the second, only 2 of the 3 comparison keys were available for choice after the retention interval (two-key test).

*Three-key test trials.* Each test trial began with the illumination of the houselight for 2, 6 or 18 s, and then a retention interval of 0, 5, 10 or 20 s followed. After the retention interval, the 3 comparison keys were illuminated, correct choices were reinforced and followed by the ITI, and incorrect choices initiated the ITI immediately. The correction method applied only on trials with 0-s retention intervals (Trials with 0-s retention intervals were the regular training trials that maintained the discrimination throughout testing.). Both the ITI and the retention interval were spent in darkness.

Each session comprised 54 trials, 36 with a 0-s retention interval (12 for each sample) and 18 with one of the other retention intervals (6 for each sample). Notice that in each session only one of the positive retention intervals was used and it was selected pseudo-randomly across sessions. Testing continued for a total of 30 sessions, 10 for each of the three positive retention intervals.

The functions obtained with the preceding series of retention intervals (from 0 to 20 s) changed the most between 0 and 5 s. Hence, it seemed appropriate to study the retention functions more closely within that range. To that end, we returned the pigeons to 5 baseline training sessions (i.e., 54 trials, all with the 0-s retention interval) and then exposed them to a new series of test sessions with retention intervals of 0, 1, 2, 3 or 5 s. Testing lasted 40 sessions, 10 for each of the four positive retention intervals.

*Two-key test trials.* The procedure remained the same, with one exception: On each trial, only two of the three comparison keys were presented. That is, choice involved one of the three pairs of keys, Short vs. Intermediate, Short vs. Long, and Intermediate vs. Long. The two comparisons presented on each trial were distributed randomly over the three response keys (6 different spatial arrangements for each pair). The retention interval was either 0 s or one of three positive values, 5, 10 or 20 s. Each session comprised 108 trials, 36 for each of the three sample durations. Each set of 36 trials comprised 3 pairs



**Fig. 1.** Mean choice percentages of the three comparison keys as a function of retention interval. Each panel corresponds to a different sample. The data come from the three-key test trials of Experiment 1.

of comparisons  $\times$  6 combinations of spatial arrangements for each pair  $\times$  2 values of the retention interval. Testing lasted 30 sessions, 10 for each of the three positive retention intervals. The correction procedure was applied only to trials with 0-s retention intervals and with the correct comparison key as one of the two choice keys. No reinforcement was given on trials with two incorrect comparison keys. Both the retention interval and the ITI were spent in darkness.

### Results and discussion

The pigeons took between 5 and 24 sessions (average = 14) to reach the learning criterion in the two-sample task and between 21 and 54 additional sessions (average = 34) to reach the criterion in the three sample task. When errors were made on the short- and long-sample trials, they consisted mainly of the choice of the Intermediate key, whereas errors on intermediate-sample trials were distributed roughly equally between the Short and Long keys. Hence, the birds' responses were under control of the sample, and the intermediate sample was perceived as between the short and long samples.

#### Three-key test trials

Pigeons were exposed to two series of tests, one with retention intervals of 0, 5, 10 and 20 s and the other with intervals of 0, 1, 2, 3 and 5 s. To determine whether performance differed between the two series, we compared proportion correct on the common intervals of 0 and 5 s. Two-way repeated-measures ANOVAs with series (2 levels) and sample (3 levels) as factors showed either no significant difference between series (i.e.,  $p > .05$ ) or a significant but small difference (under 5%). Therefore, the results from the two series were combined and the data at the 0 and 5-s intervals were averaged between the two series.

Fig. 1 shows the average percentage of trials on which the pigeons chose the Short, Intermediate and Long keys as a function of the retention interval and sample duration. The left, middle, and right panels refer to the short, intermediate, and long samples, respectively. In all conditions, the overall matching accuracy decreased with the retention interval. On the short trials (left panel, open circles), after an initial decrease, matching accuracy stabilized at a relatively high value (70%), but on intermediate (middle panel, filled circles) and long trials (right panel, filled triangles) matching accuracy decreased steadily with the retention interval.

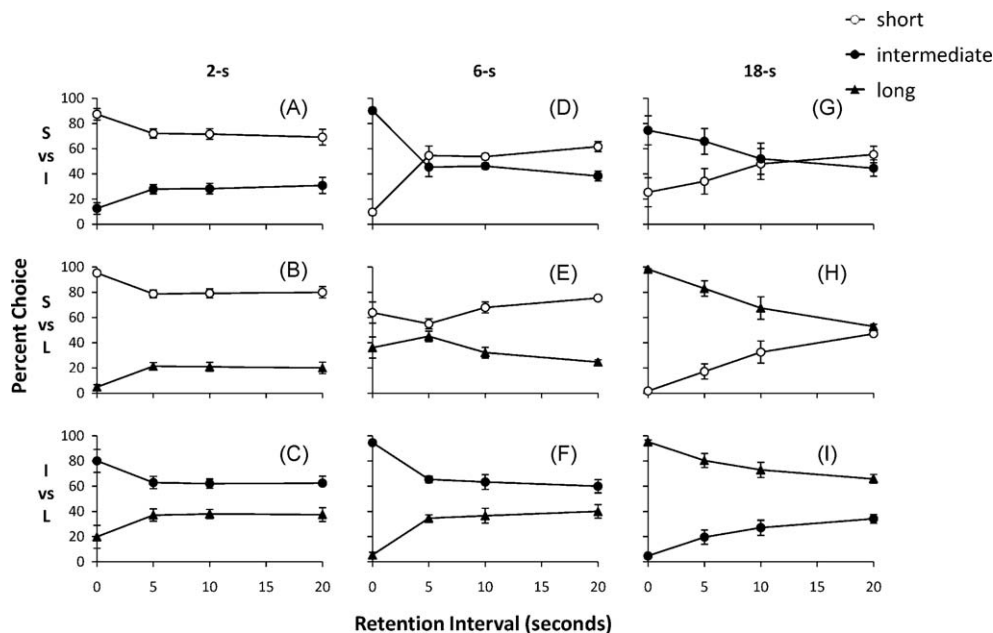
Consider now the average curves for individual types of trial. As mentioned above, when the short sample was presented, accuracy remained high. Moreover, the Short key remained the most preferred key at all retention intervals, and only a small percentage of choices were made to the Intermediate and Long keys. On the intermediate-sample trials, the pattern of results was different. The strong initial preference for the Intermediate key decreased with the retention interval while the percentage of choices of the Short key increased, with the two curves intersecting at 5 s. The percentage of choices of the Long key increased initially and then decreased slightly with the retention interval. On the long-sample trials, the pattern of results was similar to that just described: Preference for the correct (Long) key decreased and preference for the Short key increased; the percentage of choices of the Intermediate key increased slightly and then stabilized at a low value.

A repeated-measures ANOVA for matching accuracy (proportion of correct responses) with retention interval (7 levels) and sample (3 levels) as factors showed a significant interaction [ $F(12, 36) = 7.14$ ,  $p < .001$ ]. The interaction stems from the fact that, as described above, the three samples did not have similar retention functions.

To examine the errors in greater detail, we divided them into two sets, errors following the intermediate and long samples, and errors following the short sample. The percentage of errors on the 6 and 18-s trials was analyzed using a three-way repeated-measures ANOVA with retention interval (7 levels), sample (2 levels, 6 s vs. 18 s), and error (2 levels, Short key vs. other key) as factors. A significant effect of error was found [ $F(1, 3) = 10.42$ ;  $p < .05$ ] because errors to the Short key were more numerous. In addition, an interaction between sample and error also was found [ $F(1, 3) = 43.29$ ;  $p < .05$ ] because the difference between the two types of errors was significantly larger on 6-s trials than on 18-s trials. Finally, there was a significant interaction between error and retention interval [ $F(6, 18) = 6.15$ ;  $p = .001$ ] because responses to the short-sample key increased with the retention interval.

Errors on the 2-s trials were analyzed using a repeated-measures ANOVA with retention interval (7 levels) and error (responses to the Intermediate or Long keys) as factors. All effects were significant (retention interval:  $F(6, 18) = 2.92$ ;  $p < .05$ ; error:  $F(1, 3) = 10.33$ ;  $p < .05$ , and interaction:  $F(6, 18) = 3.18$ ;  $p < .05$ ). Incorrect





**Fig. 2.** Mean choice percentages of the three comparison keys as a function of retention interval. The data come from two-key test trials of Experiment 1.

choices were more concentrated on the Intermediate key and varied in a non-systematic way with the retention interval.

The average curves in Fig. 1 represent well the individual data. Among the 3 types of trials, the 18-s sample trials had the most individual variation, but, critically to test the two models, for no pigeon was the Intermediate key the most preferred key.

The foregoing results replicate those of Kraemer et al. (1985) despite the fact that the duration of the shortest sample was greater than 0 s. First, we managed to replicate the choose-short effect because the decrease in correct responses with the retention interval was significantly more pronounced on the long- than the short-sample trials. Second, as the coding hypothesis predicted, errors following the long samples consisted mainly of choices of the Short key. The pattern predicted by the subjective shortening hypothesis (as the retention interval increases, the most preferred key changes from the Long to the Intermediate to the Short key) was not observed. Hence, the present study supports the coding hypothesis, suggesting that the forgetting of memories for duration of events is an all-or-none process.

#### Two-key test trials

Fig. 2 shows the average across pigeons of the percentage of choices to each of the two keys. Each panel represents a different combination of sample (columns) and pair of comparison keys (rows). For example, the top left panel (labelled A) shows the percentage of choices of the Short and Intermediate keys when the sample was 2-s long. The following findings are noteworthy. First, although the test trials used only two keys, the observed response patterns were consistent with the choose-short effect. Thus, on 2-s trials, the percent of correct responses decreased slightly and then stabilized at a high value (cf. panels A and B, open circles). On the 6 and 18-s trials, when the two comparisons included the correct key and the Short key (see panel D, filled circles, and panel H, filled triangles), the percentage of correct responses decreased and responses to the Short key increased. A repeated-measures ANOVA for matching accuracy with retention interval (4 levels) and sample (3 levels) as factors yielded a significant effect of sample [ $F(2, 6) = 17.53$ ;  $p < .05$ ], retention interval [ $F(3, 9) = 72.12$ ;  $p < .001$ ], and interaction [ $F(6, 18) = 7.41$ ;  $p < .001$ ], confirming the difference in the retention functions of the three samples.

Second, on trials without the correct comparison key (panels C, E and G), the pattern of choices at the 0-s retention interval seems consistent with stimulus generalization. Thus, on 2-s trials (see panel C), when the choice was between the Intermediate and Long keys, the four pigeons preferred the Intermediate key, the key associated with the sample closest to the 2-s sample. A 2-tailed normal approximation to the binomial (using a significance level of .05) showed that these preferences were significantly above chance in three of the four birds. Similarly, on 18-s trials, when the choice was between the Short and Intermediate keys (see panel G), three of the pigeons also preferred strongly the Intermediate key, the key whose sample was closest to the 18-s sample (2-tailed normal approximation to the binomial). The fourth bird showed a non-significant preference for the Short key. Finally, on 6-s trials, when the choice was between the Short and Long keys (panel E), the two percentages were closer to indifference than in the previous cases. Because 6 s is the geometric mean of 2 and 18 s, stimulus generalization would predict indifference between the two keys, but a 2-tailed normal approximation to the binomial showed that, in all four birds, the proportions of the preferred key were significantly above chance. Although

indifference was not found, the preferences on these trials were not as extreme as on 2 and 18-s trials. Moreover, not all birds preferred the same key (3 preferred the Short key and 1 preferred the Long key).

Third, another subset of findings suggests a combination of stimulus generalization and the choose-short effect. Consider panel G. The sample was 18-s long but the Long (correct) key was unavailable for choice. If one assumes that the Intermediate key becomes functionally similar to the Long key due to stimulus generalization (albeit with some generalization decrement), the choose-short effect would predict that, as the retention interval increases, preference for the Intermediate key should decrease, whereas preference for the Short key should increase. This result was obtained. Consider now panel C. The sample was 2-s long but the Short (correct) key was unavailable for choice. If because of stimulus generalization the Intermediate key becomes functionally similar to the Short key, then the choose-short effect would predict that the preference for the Intermediate key would either not change or decrease slightly with the retention interval (compare with Fig. 1, left panel, open circles). This result also was observed. Finally, consider panel E. When the sample was 6-s long but the Intermediate key was unavailable, stimulus generalization would predict no preference, but the choose-short effect would predict an increasing preference for the Short key with the retention interval. That is, the two curves should start at 50% and then diverge, with the curve for the Short key increasing and that for the Long key decreasing. The obtained result is roughly consistent with this prediction.

To summarize, the results from the three-key test trials supported the coding model, which proposes that different codes are created for each of the sample durations and that these codes may be forgotten in an all-or-none fashion. But the results from the two-key test trials cannot be explained using only categorical information. Hence, if the coding model is correct and the different samples are translated into categorical codes, these codes need to have some ordinal relation between them, a relation consistent with stimulus generalization effects.

As mentioned in the introduction, the goal of the two-key test trials was to assess the full preference hierarchy for the comparison keys following different samples and retention intervals. But this goal assumes that preference is unaffected by the number of keys available for choice. To test the assumption, we attempted to predict the two-key test results based on the three-key test results. Perhaps the simplest way to do it is as follows. Consider a three-key test trial with a 2-s sample and a 0-s retention interval. We assume that the preferences for the Short, Intermediate and Long keys are represented by variables  $pS_{2,0}$ ,  $pI_{2,0}$ , and  $pL_{2,0}$ , respectively, with the subscripts identifying the sample and retention interval values. We also assume that choice proportion equals relative preference. Hence, choice proportion for the Short key equals  $pS_{2,0}/(pS_{2,0} + pI_{2,0} + pL_{2,0})$ , for the Intermediate key it equals  $pI_{2,0}/(pS_{2,0} + pI_{2,0} + pL_{2,0})$  and for the Long key it equals  $pL_{2,0}/(pS_{2,0} + pI_{2,0} + pL_{2,0})$ . From the observed choice proportions during the three-key test trials with 2-s samples and 0-s retention intervals, we estimated  $pS_{2,0}$ ,  $pI_{2,0}$ , and  $pL_{2,0}$  and then used these values to predict all choice proportions during the two-key test trials. For example, on two-key test trials with the Short and Intermediate keys, choice proportion for Short should equal  $pS_{2,0}/(pS_{2,0} + pI_{2,0})$ . Similarly, on trials with the Short and Long keys, choice proportion for Short should equal  $pS_{2,0}/(pS_{2,0} + pL_{2,0})$ .

Fig. 3 compares the predictions (lines) with the obtained data (symbols). For the test trials that included the correct comparison key, only the data and predictions for that key are presented; for the test trials that excluded the correct comparison key (i.e., both choices are incorrect – see panels C, E and G), the data and predictions for both keys are presented.

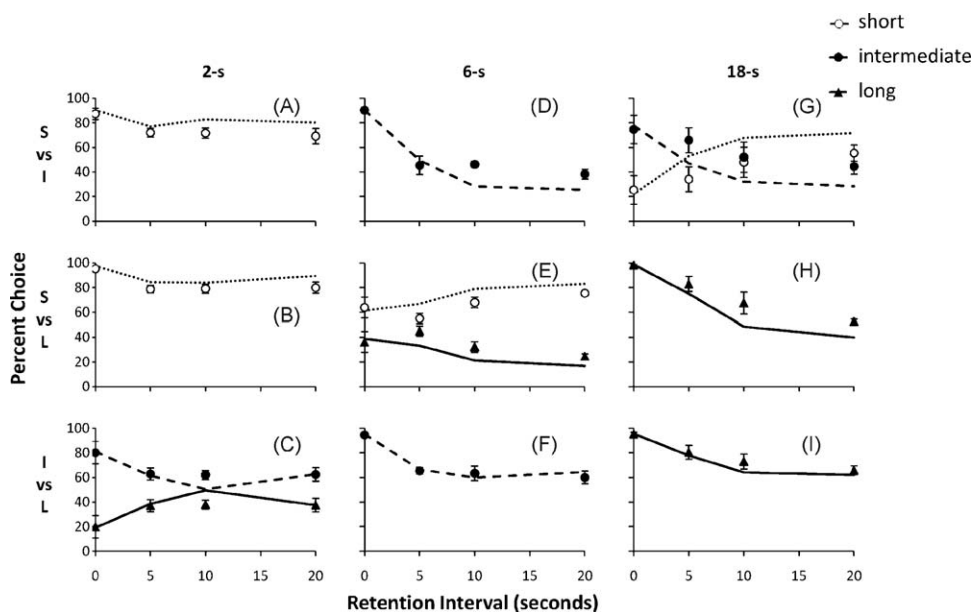


Fig. 3. Predicted (lines) and observed (symbols) choice percentages as a function of retention interval in the two-key test trials of Experiment 1.

Despite some discrepancies, predicted and obtained values followed the same general trends. Therefore, we can argue that the number of comparisons presented in testing does not seem to influence significantly the birds' performance. Two- and three-key tests reveal the same preference structure.

## Experiment 2

The conclusion that Experiment 1 supports the coding hypothesis relies on the assumption that the choose-short effect is not simply due to the similarity between the ITI and the retention interval, as the confusion hypothesis argues. Hence, in Experiment 2, we reproduced the conditions of Experiment 1 but with different stimuli signalling the ITI and the retention interval. We asked whether the choose-short effect would still hold when the ITI and the retention intervals were less likely to be confused.

### Methods

#### Subjects and apparatus

Four adult pigeons (*C. livia*) participated in the experiment. The birds had previous experience with timing procedures, one (pigeon 205) with double-bisection procedures (Oliveira & Machado, 2008) and three with Fixed Interval schedules (Monteiro & Machado, 2009). They were maintained in the same conditions as in Experiment 1. The apparatus also was the same as in Experiment 1.

#### Procedure

The procedure remained the same as in Experiment 1, except that during the ITI each of the three response keys displayed a white vertical bar against a dark background. The reinforcement duration varied across birds from 2 to 5 s.

### Results and discussion

The pigeons required an average of 18 sessions (range: 9–29) to meet the criterion with two samples, and 39 additional sessions (range: 17–52) to meet the criterion with three samples. One pigeon failed to meet the criterion with the intermediate sample. Although it gave more than 80% correct responses in a block of 5 sessions following the short and long samples, percent correct following the intermediate samples stabilized around 75%.

#### Three-key test trials

As in Experiment 1, the data from the two test series with different retention interval ranges were combined because a preliminary analysis showed that proportion correct at the 0 and 5-s intervals did not differ significantly between them.

Fig. 4 plots choice percentage for each key as a function of the retention interval. On short-sample trials (left panel), choice percentage for the Short key decreased as the retention interval increased from 0 to approximately 5 s and then stabilized around 60% for longer intervals. Incorrect choices revealed a consistent preference for the Intermediate key. On intermediate-sample trials (middle panel), the choices of the Intermediate key decreased and the choices of the Short key increased with the retention interval. On the average, for intervals greater than 3 s the Short key became the most preferred key. On long-sample trials (right panel), choices of the Long key decreased, whereas choices of both the Short and the Intermediate keys increased with the retention interval. At the longest intervals, the responses were distributed roughly equally between the three keys.

The confusion hypothesis argues that, when the stimulus conditions differ between the ITI and the retention interval, the matching accuracy curves for the three samples should be similar. Specifically, in Fig. 4, the open circles on the left panel, the filled circles on the middle panel, and the filled triangles on the right panel should follow a similar trend. To determine whether this was the case, we submitted the percentage of correct responses to a two-way, repeated-measures ANOVA with sample (3 levels) and retention interval (7 levels) as factors. The ANOVA yielded a significant interaction between the two

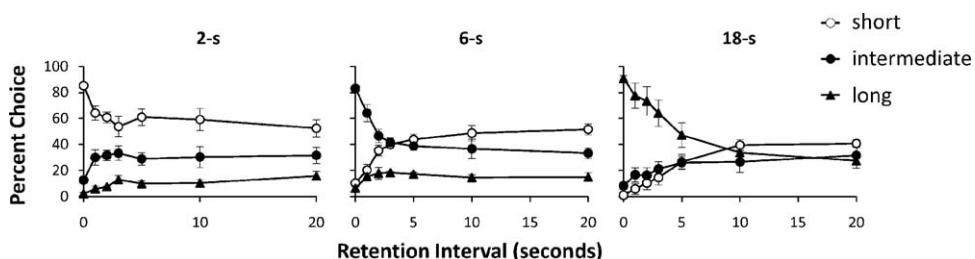


Fig. 4. Mean choice percentages of the three comparison keys as a function of retention interval. Each panel corresponds to a different sample. The data come from the three-key test trials of Experiment 2.



factors [ $F(12, 36) = 4.25, p < .001$ ]. As Fig. 4 already suggested, the retention functions for the correct response varied with the sample. This result is inconsistent with the confusion hypothesis.

The trends observed in Fig. 4 are similar to the trends observed in Fig. 1. To further examine the similarities and differences between the two experiments, we analyzed in greater detail the pattern of errors. The distribution of errors on 6 and 18-s trials was analyzed using a three-way, repeated-measures ANOVA with retention interval, sample, and error as factors. The errors were divided into two sets, responses to the Short key and responses to the other key. There was no main effect of error [ $F(1, 3) = 7.07; p > .05$ ], but there was a significant interaction between error and retention interval [ $F(6, 18) = 11.21; p < .001$ ] because the number of responses to the Short key increased with the retention interval. In addition, the difference between the two types of errors was significantly larger on 6-s trials than on 18-s trials [ $F(1, 3) = 91.04; p < .05$ ]. That is, on 6-s trials most errors were made to the Short key, whereas on 18-s trials, the errors were more evenly distributed.

The percentage of errors on 2-s trials was analyzed with a two-way, repeated-measures ANOVA with retention interval and error as factors. No significant interaction of the two factors was found [ $F(6, 18) = 1.78; p > .05$ ], but choices to the Intermediate key were significantly more frequent [ $F(1, 3) = 16.29; p < .05$ ]. Also, responses to the Intermediate and Long keys increased significantly with the retention interval [ $F(6, 18) = 19.06; p < .001$ ].

On 2 and 6-s trials the average curves represent well the individual data, but on 18-s trials, the individual results were more variable. For all pigeons, preference for the Long key decreased with the retention interval, but the pattern of choices to the other comparison keys differed across subjects. Pigeons P284 and P343 chose all three keys at chance level at the longest retention interval; pigeon P334 showed an increase in preference for both the short and intermediate-sample keys, and pigeon P205 showed an increase in preference only for the Short key.

Although the data from Experiment 2 were roughly in line with the data from Experiment 1 – in particular, the choose-short effect was obtained also in Experiment 2 – the magnitude of the choose-short effect was not as large as in Experiment 1. Specifically, in Experiment 1, there was a larger difference in the proportion of correct responses between the 2-s and the 6-s trials (cf., left and middle panels of Figs. 1 and 4), as well as between the 2 and 18-s trials (cf. left and right panels of Figs. 1 and 4). Moreover, on the 18-s trials, the difference in the error curves was greater in Experiment 1 than in Experiment 2 (cf. right panels of Figs. 1 and 4, open and filled circles).

#### Two-key test trials

Fig. 5 shows the average choice percentage to each of the two comparison keys. Although the ITI and retention interval were differentiated, evidence of the choose-short effect is still present. First, on 2-s trials (see panels A and B), as the retention interval increased, correct responses initially decreased but then stabilized at a relatively high value. Second, on the 6 and 18-s trials (panels D and H), when the choice was between the correct key and the Short key, the percentage of errors to the Short key increased with the retention interval. A repeated-measures ANOVA for matching accuracy with retention interval (4 levels) and sample (3 levels) as factors showed a significant effect of sample [ $F(2, 6) = 6.17; p < .05$ ], retention interval [ $F(3, 9) = 58.18; p < .001$ ], and interaction [ $F(6, 18) = 5.03; p < .05$ ], revealing that the retention functions following the three samples differ reliably.

The general trends observed in Fig. 5 are similar to those observed in Fig. 2 and, again, at the 0-s retention interval, they seem to be consistent with stimulus generalization. Thus, on 2 and 18-s trials, when the choice did not include the correct

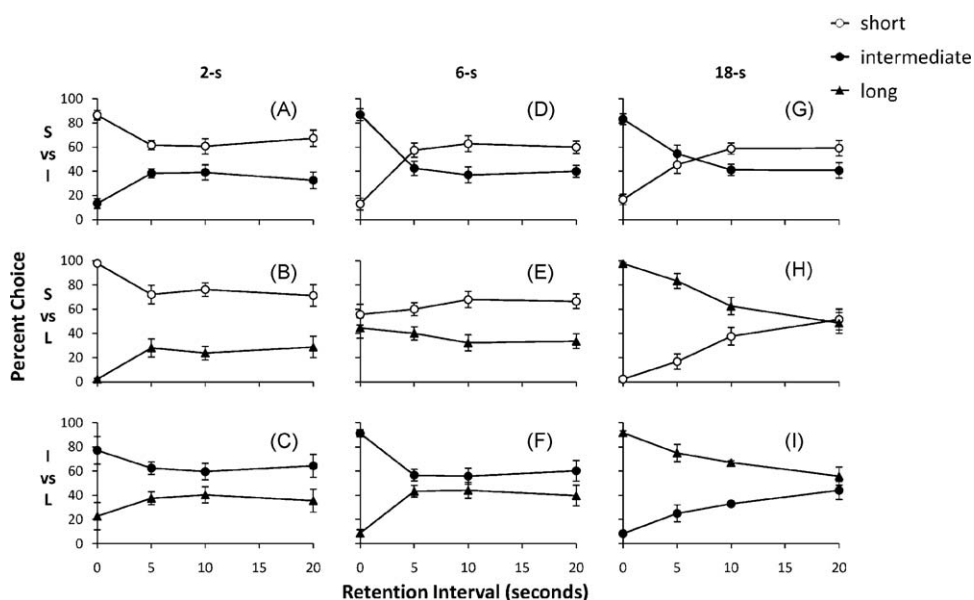
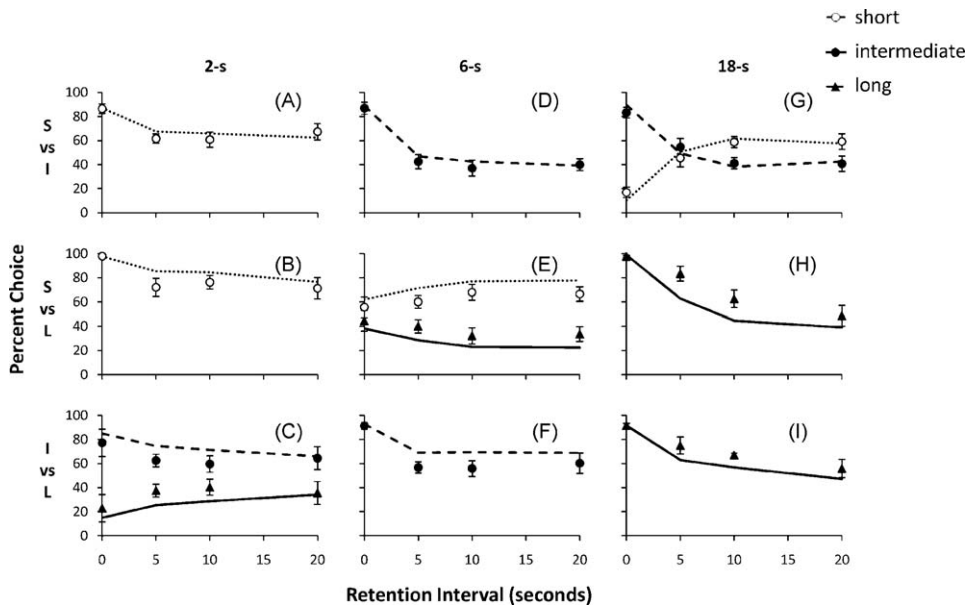


Fig. 5. Mean choice percentages of the three comparison keys as a function of retention interval. The data come from two-key test trials of Experiment 2.



**Fig. 6.** Predicted (lines) and observed (symbols) choice percentages as a function of retention interval in the two-key test trials of Experiment 2.

key (panels C and G), the pigeons preferred the Intermediate key at the 0-s retention interval. Using a significance level of .05, a 2-tailed normal approximation to the binomial showed that these preferences were significantly above 50 percent in three birds on 2-s trials and in all four birds on 18-s trials. On 6-s trials, when the choice was between the Short and Long keys (panel E), the choice percentages were closer to indifference than in the previous cases, but they were significantly above chance in all but one bird. In addition, in the birds that showed a statistically significant preference for one of the keys, two birds preferred the Short key whereas one preferred the Long key. In sum, the preferences on 6-s trials were not as extreme or as clear as in 2-s and 18-s trials.

Fig. 5 also suggests a combination of stimulus generalization and the choose-short effect. To illustrate, consider the 18-s samples with the Long (correct) key unavailable for choice (panel G): If one assumes that the Intermediate key became functionally similar to the Long key, then the choose-short effect predicts that, as the retention interval increases, the preference for the Intermediate key should decrease and the preference for the Short key should increase. This result was obtained. Similar considerations hold for panel C. If the Intermediate key becomes functionally similar to the Short key, then the choose-short effect predicts that the preference for the Intermediate key either will not change or decrease slightly with the retention interval. This result also was observed. Finally, in panel E, the slight increase in the preference for the Short key is also consistent with a combination of stimulus generalization and the choose-short effect.

In summary, the results of Experiment 2 go against the confusion hypothesis. Although stimuli were used to differentiate the ITI from the retention interval, a choose-short effect was observed. In fact, the results obtained in the three- and two-key test trials were similar to those obtained in Experiment 1 (cf. Figs. 2 and 5), even though the choose-short effect was stronger in Experiment 1 than in Experiment 2. Thus, the similarity or dissimilarity between the ITI and retention interval did not markedly affect the birds' performance.

One could argue that the similarity of the results in the two experiments is due to the fact that the stimuli used to distinguish the ITI and the retention interval were not sufficiently salient, hence leaving some of the instructional ambiguity that Sherburne et al. (1998) mentioned. In the study in which these authors failed to observe a choose-short effect, the houselight was used to differentiate the ITI from the retention interval whereas we used white vertical bars on the three keys. Although we cannot rule out this possibility, we note that our results are consistent with those reported by Kelly and Spetch (2000) which show a choose-short effect even though a houselight was used to distinguish the two intervals.

As in Experiment 1, in Experiment 2 we also tried to predict the responses on the two-key test trials from the three-key tests data (Fig. 6). Despite some discrepancies, mostly in terms of the magnitude of some preferences, the predictions matched the data trends. Preference does not seem to change substantially with the number of choice keys available.

## General discussion

The choose-short effect is expressed in different retention functions following the short and long samples. With increasing retention intervals, preference for the Short key remains high following short samples, but preference for the Long key decreases following long samples. In this paper we used the choose-short effect to study how memories for the duration of events change with time.

The goal of Experiment 1 was to assess whether the forgetting of memories for the duration of an event was a continuous, gradual process, or a sudden, all-or-none process. Two models of the choose-short effect were contrasted, one that proposed a gradual forgetting (subjective shortening model) and another that proposed a sudden forgetting (coding model). To that end, a delayed matching-to-sample procedure with three sample stimuli (2, 6 and 18 s) and three comparison keys (Short, Intermediate, and Long) was used. After the pigeons learned the conditional discriminations, we introduced retention intervals between the offset of the sample stimuli and the onset of the comparison keys. The results argued against the subjective shortening model, suggesting that short-term memories for event duration are forgotten abruptly.

Even though the results of the critical test between the two models' predictions (shifts in preference on 18-s sample trials) supported the coding model, this model could not explain all the experimental findings. For example, the decay in choices of the Short key on 2-s trials was not predicted by this model. According to the coding model, two outcomes are expected on 2-s trials. Either the animal remembers the sample code and selects the Short key, or it forgets the code and selects the Short key (because the memory for the short sample is the closest to the memory for no sample). In both cases, and regardless of the retention interval, responses to the Short key would be expected to remain high and not decrease. This prediction is at odds with the data.

Additionally, it is important to note that some results obtained in previous experiments seem to be better explained by the subjective shortening model. For example, [Spetch \(1987\)](#) trained pigeons with a retention interval greater than zero (e.g., 10 s). In this situation, when tested with a longer retention interval (e.g., 20 s), a choose-short effect was obtained; but when the retention interval was shorter (e.g., 0 s), choose-long errors predominated. The choose-long errors are the opposite of the choose-short errors, and consist in a tendency to incorrectly choose the long-sample key on short-sample trials. According to the subjective shortening model, during testing with shorter retention intervals, the sample memory traces decay less and therefore they seem longer than the samples used in training—hence, the choose-long errors. Since the coding model assumes a bias towards the “short” response when the code is forgotten, it cannot easily explain the choose-long effect.

Another difference between the two models is that whereas the subjective shortening model assumes a retrospective encoding of the sample stimulus, the coding model assumes a prospective encoding of that stimulus ([Kraemer et al., 1985](#)). To examine this difference, [Wilkie and Willson \(1990\)](#) designed a study in which the correct choice after 2-, 8-, and 10-s samples were, respectively, red, orange and green comparison keys. Thus, easy-to-distinguish samples (2 and 8 s) were mapped onto hard-to-distinguish comparisons (red and orange keys) and, conversely, hard-to-distinguish samples (8 and 10 s) were mapped onto easy-to-distinguish comparisons (orange and green keys). Results showed that errors were more likely between easy comparison choices (orange vs. green) but hard sample discrimination (8 s vs. 10 s) than between hard comparison choices (red vs. orange) but easy sample discrimination (2 s vs. 8 s). This pattern of errors is more consistent with a retrospective encoding of the sample durations (see also [Santi, Stanford, & Symons, 1998](#)).

In sum, our data support the coding model, but this model also has limitations.

The test trials with only two comparison keys yielded two sets of results. On the one hand, there was evidence of stimulus generalization, particularly at the 0-s retention intervals. On the other hand, there was some evidence of what might be called a combination of stimulus generalization and the choose-short effect. To recapitulate, on both short- and long-sample trials without the correct key available for choice, the Intermediate key seemed to take the role of the correct key (due to stimulus generalization) such that the preference for the Intermediate key remained relatively high when the sample was short, but decreased when the sample was long (choose-short effect). Neither model accommodates this finding.

In sum, the coding model seems to be the most accurate, in that our results suggest that forgetting of memories for duration is a sudden process, but the choose-short effect does not seem to be reducible to a strictly categorical process. It seems to us that the full set of results from Experiment 1 requires that the overall mechanism proposed by the coding model be supplemented with an ordinal relation among the codes.

Experiment 2 focused on the confusion hypothesis of the choose-short effect, which argues that this effect is the result of the similarity between ITI and retention interval in the delayed matching-to-sample procedure. This experiment used a procedure similar to the one used in Experiment 1, the only exception being the illumination of three keys during the ITI to distinguish it from the dark retention interval. On the three-key test trials, and contrary to the confusion hypothesis' predictions, a choose-short effect was obtained. It remains to be seen if an ITI stimulus more salient than three vertical bars eliminates the choose-short effect and yields similar retention functions following the various samples. In the meantime, we tentatively conclude that the choose-short effect is not the result of an experimental artifact.

Although our results cannot be explained by the confusion hypothesis, it seems that the difference between the ITI and retention interval influenced the pigeons' responses. In particular, the choose-short effect was stronger in Experiment 1 than in Experiment 2 because, in the latter, (a) the percentage of correct responses on the short-sample trials did not remain as high (cf. open circles in the left panels of [Figs. 1 and 4](#)), and (b) the difference between the percentage of choices of the Short key and the percentage of choices of the other incorrect key also was not as high (cf. open circles and filled triangles in the middle panels of [Figs. 1 and 4](#) as well as open and filled circles in the right panels of the same figures). These results suggest that the performance on the three-key test trials depended not only on the sample being timed, but also on the context in which the sample was presented. Specifically, the stimuli filling the intervals that preceded (ITI) and followed (retention interval) the sample seemed to affect choice performance.

Another example of this contextual influence is the finding that the duration of the ITI can also affect the judgments of the duration of the sample stimulus ([Spetch & Rusak, 1992](#)). More specifically, if the ITI preceding the sample is shorter than usual, the sample is judged longer; and if the ITI is longer than usual, the sample is judged shorter. Therefore, both the

ITI and the retention interval have an influence on the judgment of the duration of the sample. Moreover, the influence of these two intervals appears to be additive. For example, it has been found that an overestimation of the sample duration as a consequence of a shortened ITI can be eliminated if the retention interval is lengthened (Spetch & Rusak, 1989).

The results of the two-key test trials in Experiment 2 provided another test of the confusion hypothesis. Once again, evidence of a choose-short effect was obtained, which is at odds with the hypothesis. Comparing Experiments 1 and 2, it is worth noting that whereas on the three-key test trials the choose-short effect obtained in Experiment 2 was not as strong as in Experiment 1, on the two-key test trials the results were almost the same in both experiments. That is, the difference in stimulus conditions between the ITI and retention interval did not seem to affect the birds' performance when only two keys were available for choice. Hence, the precise effect of the stimulus used to signal the ITI remains unclear.

With respect to previous studies (e.g., Kraemer et al., 1985), the present study introduced the novelty of two-key test trials. These trials revealed three main effects that future models and theories of timing and temporal memory must account for: (a) the stimulus generalization effect, (b) the retention interval effect, and (c) the choose-short effect. We have already discussed the stimulus generalization effect, which is most visible at the 0-s retention interval (cf. Figs. 2 and 5). We examine the other two next.

With respect to the retention interval effect, we note first that, as Figs. 2 and 5 also illustrate, matching accuracy for all three samples decreased, although to different extents, with the retention interval. Notice that errors increase not only on trials in which the incorrect alternative is the Short key (the main characteristic of the choose-short effect), but also when the incorrect key is either the Intermediate or the Long key (cf. panels F and I in Figs. 2 and 5). In addition, this effect is not exclusive to trials that have the correct comparison available. When only incorrect comparisons are available, the initially preferred key also becomes less preferred at longer retention intervals (cf. panels C and G in Figs. 2 and 5). Thus, it seems that the retention interval disrupts timing and introduces some randomness in the subjects' choices. This effect can also explain the decrease in matching accuracy with the retention interval observed in both Experiments 1 and 2 on 2-s trials, a decrease that no model predicted.

The Short key bias effect seems to be modulated by stimulus generalization. Consider the 6 and 18-s trials in which the correct key was paired with the Short key (panels D and H in Figs. 2 and 5). The percentage of Short key choices increased faster on the 6-s trials (panel D) than on the 18-s trials (panel H). Repeated-measures ANOVAs with sample (2 levels) and retention interval (4 levels) as factors yielded a significant interaction between the factors for Experiment 1 [ $F(3, 9) = 4.84$ ;  $p < .05$ ] and Experiment 2 [ $F(3, 9) = 6.3$ ;  $p < .05$ ]. Hence, longer samples seem to be more resistant to the Short key bias, with their matching accuracy decreasing slower. In their experiment with three samples, Kraemer et al. (1985) obtained a similar effect; that is, the matching accuracy for the intermediate-sample decreased faster than the accuracy for the long-sample. A similar tendency seems to exist also on the three-key test trials in Experiments 1 and 2 (see Figs. 1 and 4, respectively). In these figures, the average curves seem to indicate that matching accuracy for the intermediate sample decreases faster than the matching accuracy for the long sample, especially for retention intervals from 0 to 5 s. However, this difference was not statistically significant (Experiment 1: [ $F(4, 12) = 3.078$ ;  $p > .05$ ]; Experiment 2: [ $F(4, 12) = 1.219$ ;  $p > .05$ ]). This pattern of results is of interest because if each sample was coded categorically, equivalent rates of forgetting would be expected for all samples. Differences in matching accuracy suggest that codes for different samples may have different strengths. More specifically, the longer the sample, the stronger its categorical representation in memory and therefore the harder it is to disrupt by the retention interval. As Kraemer et al. (1985) suggested, longer samples – simply by being longer – could allow the pigeons to process and rehearse the code for longer, which could result in a stronger or more persistent representation in memory of the code. If code strength is proportional to sample duration, then the relative strengths of the codes induce an ordinal relation among them.

To summarize, the novel tests with only two comparison keys allowed us to identify some effects that may be controlling the pigeons' responses on a delayed matching-to-sample task. The effect of stimulus generalization on the pigeons' responses was most noticeable when no retention interval was present. When retention intervals were introduced, errors increased, and two different effects were then identified. On the one hand, the retention interval seems to introduce some randomness in the animal's choices. On the other hand, with retention intervals a bias towards the Short key appears. At long retention intervals, the Short key may become the most preferred key. The challenge to our models of timing and memory is to account for these effects and their interaction.

## Acknowledgements

We would like to thank Jeremie Jozefowicz for his comments on an earlier version of this paper. Research funded by a grant from the Portuguese Foundation for Science and Technology (FCT) to Armando Machado (PTDC/PSI/65678/2006).

## References

- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84, 279–325.
- Kelly, R., & Spetch, M. L. (2000). Choice biases in delayed matching-to-sample duration with pigeons: Manipulations of ITI and delay illumination. *The Quarterly Journal of Experimental Psychology*, 53B, 309–323.
- 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.

- Monteiro, T., & Machado, A. (2009). Oscillations following periodic reinforcement. *Behavioural Processes*, 81, 170–188.
- Oliveira, L., & Machado, A. (2008). The effect of sample duration and cue on a double temporal discrimination. *Learning and Motivation*, 39, 71–94.
- Santi, A., Stanford, L., & Symons, J. (1998). An analysis of confusion errors in many-to-one matching with temporal and nontemporal samples. *Animal Cognition*, 1, 37–46.
- 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., & Rusak, B. (1989). Pigeons' memory for event duration: Intertrial interval and delay effects. *Animal Learning & Behavior*, 17, 147–156.
- Spetch, M. L., & Rusak, B. (1992). 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. (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.
- 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.