2005

Biased forgetting effects in the assessment of memory for filled and empty intervals: Evidence for the instructional failure/confusion hypothesis

Stephen Gagne
Wilfrid Laurier University

Biased Forgetting Effects in the Assessment of Memory for Filled and Empty Intervals: Evidence for the Instructional Failure/Confusion Hypothesis

By

Stephen Gagné

Bachelor of Arts (Honours), Psychology

University of Windsor (2003)

THESIS

Submitted to the Department of Psychology

In partial fulfillment of the requirements for

Master of Science, Psychology

Wilfrid Laurier University (2005)

Your file Votre référence
ISBN: 0-494-04872-7
Our file Notre référence
ISBN: 0-494-04872-7

Abstract

According to the instructional failure hypothesis, the contextual similarity of the intertrial interval (ITI) and the delay interval (DI) is responsible for the choose-short bias that occurs when memory for filled intervals is tested. This hypothesis may also explain the choose-long bias for empty intervals, if birds confuse an extended DI with a long empty interval. In the present study, pigeons were trained in a within-subjects design to discriminate durations of a filled interval (2-s and 8-s of light), and durations of an empty interval (2-s and 8-s bound by two 1-s light markers). In order to disambiguate the ITI, sample presentation phase, and the DI, the colour displayed on the video monitor was different during the various phases of a trial. In Experiment 1, the ITI was magenta, the sample presentation phase was dark, and the DI was grey. Although accuracy was greater for short than for long samples, the retention functions were parallel. Thus, there was no evidence of a biased-forgetting effect. In Experiment 2, additional tests indicated that the interval from the offset of the ITI to the onset of the DI, as well as the filled and empty intervals themselves affected choice responding in Experiment 1. In Experiment 3, the colour displayed on the video monitor was manipulated such that the illumination was magenta during the ITI, and dark during the sample presentation and comparison phases in order to prevent pigeons timing from the ITI offset to the DI onset, and to create confusion for one interval-type, but not the other. A choose-long bias was obtained for empty intervals, and no-bias was obtained for filled intervals. In Experiment 4, the colour displayed on the video monitor was manipulated such that the illumination was dark during the ITI and sample presentation phase, and grey during the comparison phase in order to prevent pigeons from timing from ITI offset to DI onset, and to eliminate the possibility of confusion for both filled and empty intervals. The results of this experiment revealed no biases for either filled or empty intervals. In Experiment 5, the colour displayed on the video monitor was dark for all of the phases of the trial in order to create confusion, and thus produce the choose-short and long biases. The results of this experiment revealed no bias for filled intervals, and a choose-long bias for empty intervals. The results are discussed in the context of the instructional confusion hypothesis.

Table of Contents

List of Tables

List of Figures

Biased Forgetting Effects in the Assessment of Memory for Filled and Empty Intervals: Evidence for the Instructional Failure/Confusion Hypothesis

Symbolic Matching-to-Sample

A method frequently used in comparative cognition studies is symbolic matching-to-sample. This procedure consists of presenting a choice between stimuli that symbolically represent a condition. An example of this in timing studies is after the presentation of a stimulus (i.e., light or tone), two coloured keys are illuminated as comparison stimuli: green is correct for a short sample, and red is correct for a long sample. A correct response is immediately followed by reinforcement (i.e., food), and an incorrect response results in no reinforcement.

Filled Intervals

A filled interval is a time sample signaled by a continuous stimulus (i.e., light or tone). Studies typically use a light stimulus with pigeons, although it has been done with tones (Santi, Ross, Coppa, & Coyle, 1999). Pigeons can learn this task relatively easily, and with high accuracy for both short and long samples. However, a common finding with filled intervals is that when a delay interval (DI) is introduced, pigeons are biased to choose the comparison stimulus associated with the short sample. As the delay is increased, accuracy when long samples are presented decreases, whereas the accuracy for when the short sample is presented is significantly higher (Grant & Spetch, 1991, 1993; Kraemer, Mazmanian, & Roberts, 1985; Santi, Ducharme, & Bridson, 1992; Sherburne, Zentall, & Kaiser, 1998; Spetch & Rusak, 1989, 1992; Spetch & Wilkie, 1983). This phenomenon is called the choose-short bias and it is illustrated in the top panel of Figure 1.

Possible explanations of the choose-short bias

There are several speculations as to why the choose-short bias phenomenon occurs. The following will review possible explanations of the choose-short bias such as: the subjective shortening model; analogical coding and default responding; the detection model; and the confusion hypothesis.

Subjective shortening model. One of the most well known explanations of the choose-short bias is the subjective shortening model (Spetch & Wilkie, 1983). The subjective shortening model consists of three main assumptions. The first assumption is that an event gradually foreshortens as the temporal distance between the event and the temporal decision increases. For example, an eight second sample may seem more like a five second sample after a delay. The second assumption is that temporal information is stored analogically, and that it is assessed retrospectively rather than prospectively. Analogical coding proposes that timing information is quantitatively stored in working memory, for example, the timing of events can be accomplished via pulses: more pulses accumulated would mean that it was longer. Retrospective retrieval means that the 'decision' (short or long) is made at the time the comparison stimuli are presented. The third assumption proposes that the choose-short bias is due to the comparison of a foreshortened sample in working memory to the established duration in reference memory. For example, an 8-s sample will foreshorten in working memory as the delay increases. When the comparison stimuli are presented, the representation in working memory for the 8-s sample may be closer to the reference memory for a 2-s sample, thus the pigeon will choose the 'short' comparison.

Relative-duration hypothesis. Although the subjective shortening model may account for the choose-short effect, there are some limitations in generalizing the model to other results. Spetch and Rusak (1989) trained pigeons to discriminate between short 2-s and long 8-s intervals of light with a 45-s intertrial interval (ITI) and a 10-s DI. During testing, they manipulated the length of the ITI and DI independently by increasing or decreasing their duration. They found that a choose-short bias was evident when the ITI or the DI was increased relative to baseline training, and a choose-long bias was evident when either the duration of the ITI or DI was decreased relative to baseline training. The subjective shortening hypothesis can explain the results of the choose-short and choose-long biases when the DI is manipulated. When the DI is relatively longer than the baseline, the sample subjectively shortens resulting in a choose-short bias. When the DI is relatively shorter than baseline, the sample undergoes less subjective shortening, resulting in a choose-long bias. The subjective shortening hypothesis, however cannot explain the choose-short and long biases when the ITI is manipulated: there would be no foreshortening of the samples because the ITI occurs before the sample. Spetch and Rusak (1989, 1992) proposed a relative-duration hypothesis to account for this phenomenon. They proposed that pigeons do not only time the sample, as was assumed in the subjective shortening model, rather pigeons judge the length of the event relative to the temporal background of the trial. The temporal background of a trial consists of the time surrounding the sample, such as the ITI and the DI. A choose-short bias would be observed when either or both the ITI and DI durations were longer than during training because the sample duration would be smaller in relation to the temporal background. A choose-long bias would be observed if the durations of either or both the ITI and DI were shortened because the sample duration would be greater in relation to the temporal background. As with the subjective shortening model, the relative duration hypothesis assumes that the samples are coded analogically and retrieved retrospectively. In another experiment, Spetch and Rusak (1992) manipulated the ambient conditions of the ITI and the DI. They had trained pigeons on an event discrimination task: in one group, the ambient condition of the ITI was illuminated, whereas in the other group, the ambient condition of the ITI was dark. During delay testing, both groups received two ambient condition manipulations: an illuminated DI, and a dark DI. They found that the choose-short bias was dependant on the ITI and the DI having the same illumination condition (i.e., light-light or dark-dark), and when the ITI and DI were differentiated, no bias was found. These results can be explained by the relative duration hypothesis if it is assumed that the pigeons treat the ITI and the DI as part of a common temporal background when the stimulus conditions are the same, but not when they differ, however, these results could be explained by other hypotheses that will be discussed.

Although the subjective shortening and the relative duration models might be the cause of the choose-short bias, there are some assumptions made in these theories that have been scrutinized in other studies to determine the plausibility of these theories. The following studies will aim to examine the assumptions of analogical coding and retrospective memory search.

Kraemer, Mazmanian, and Roberts (1985) proposed that one of the underlying assumptions of the subjective shortening hypothesis, namely that an event is coded analogically may be faulty. They proposed that an event is coded categorically rather than analogically, and that this coding was either done retrospectively or prospectively, as was proposed by Honig (1981). A retrospective strategy consists of the 'decision' being made at the time the comparison stimuli are presented, for example: following the presentation of the comparison stimuli, a pigeon would search memory for what type of sample was presented. A prospective strategy consists of the 'decision' being made before the presentation of the comparison stimuli, for example: either during or immediately following the sample presentation, a pigeon would determine whether a short or long sample occurred (peck 'green' or 'red'), and would retain that information until the comparison stimuli presentation. Kraemer et al. (1985) explained the choose-short bias by attributing this phenomenon to the forgetting of a categorical code in working memory. The rationale for this is that when an animal is given a forced choice when a sample is forgotten (which thus becomes the absence of a sample) stimulus generalization will occur: the 'no-sample' will be generalized to being like a short sample because they are most similar. Kraemer et al. (1985) devised an experiment to determine whether events were coded analogically or categorically. Pigeons were trained to discriminate between 'no-sample' 0-s, short 2-s, and long 8-s samples signaled by light. There were three choice comparisons, one for each of the samples. If pigeons were using analogical coding, after a long sample, a choose-short bias would be observed at a moderate delay, and then as the delay increased, the pigeons would show a 'no-sample' bias. In contrast, if a pigeon was using a categorical coding strategy after a long sample, then as the delay increased, there would be less responding to the correct 'long' comparison, and a 'no-sample' bias would be expected because the long sample would be forgotten, and would be responded to as if it had been a no-sample trial. Kraemer et al. (1985) observed evidence for pigeons using categorical coding: they had obtained a 'no-sample' bias after a delay when either short or long samples were presented. There was no gradual decline when long samples were presented (choose-short bias at moderate delay, followed by a choose-no-sample bias at greater delays). This provides evidence for categorical coding of time samples, and thus is evidence against the subjective shortening hypothesis.

This issue has created some contradictory results. Spetch and Sinha (1989) devised a study to determine whether pigeons coded time samples categorically and prospectively rather than analogically and retrospectively, which is expected with subjective shortening. Spetch and Sinha (1989) used an intratrial proactive interferance design to test this. Pigeons were trained to discriminate short 2-s and long 10-s durations of light (houselight). Following acquisition, presample trials were included. The presample trials consisted of short 2-s or long 10-s durations of light (houselight, feeder light, or keylight) that were presented after the ITI. The presample was followed by an interstimulus interval (ISI), which preceded the sample, and the sample was followed by a delay of 0, 5, or 10-s. There were four types of trials: short-short, short-long, long-short, long-long (presample-sample, respectively). It was proposed that if pigeons would in fact be categorically processing these time samples, then accuracy would be best when the presample and sample were consistent (e.g., short-short), and worse when they were inconsistent (e.g., short-long) after a delay. This is proposed because the presample would create more interference with the memory of the target event. The results of this experiment displayed a choose-long bias for all trial-types, including the short-short sample. These results suggest that pigeons were not categorically processing the time samples, rather they were likely timing through the whole trial, which suggests analogical coding of the time sample, which supports the subjective shortening hypothesis.

Given the results of the Kraemer et al. (1985) and the Spetch and Sinha (1989) experiments, it is unclear whether pigeons use analogical or categorical coding. These results also do not prove nor disprove the subjective shortening hypothesis. The following will examine the asymmetrical coding and default responding which will test another assumption made by the subjective shortening hypothesis.

Asymmetrical coding and default responding. Another hypothesis in support of categorical coding is the default response rule. It proposes that time samples are coded asymmetrically: only one sample is coded in memory (i.e., the most salient). The pigeon thus searches memory for the coded sample, and if the sample is not found, the pigeon will respond by default to the other choice. The model suggests that the choose-short bias may be explained by pigeons coding only the long samples: as the delay is increased, the memory for the long sample is forgotten, and the pigeon will thus respond short by default.

Grant (1991) devised a study that investigated asymmetrical coding and default responding in the context of a non-temporal task. Three groups of pigeons were trained to discriminate between either: food and no-food; coloured field and non-coloured field; triangle and no-triangle trials. In the food no-food group, pigeons were trained to discriminate between whether they were presented with 3-s of grain and whether they received 3-s of darkness without food. Subsequently, pigeons' memory for the event was tested after a 4 or 10-s delay. The other groups of pigeons were trained and tested in a similar fashion with the exception of the stimuli presented. It was proposed that if pigeons were coding both the event and the non-event, accuracy would decrease to chance for both sample types. If the samples were coded asymmetrically, it was expected that accuracy for the coded sample would decrease below chance as the delay was increased, whereas the accuracy for the non-coded sample would remain high. The results showed that the pigeons responded with high accuracy to the non-event sample (no-food, no-colour, no-triangle), whereas the accuracy for when the stimuli were present (food, colour, triangle) was below chance as the delay increased. These results show that pigeons may be responding by default when the coded sample was forgotten.

Grant and Spetch (1994) investigated default responding in another study that was more relevant to timing. They trained pigeons to discriminate 2 and 10-s, and 4.5 and 22.5-s samples within the same session. The 2 and 10-s samples were mapped onto different comparison stimuli than the 4.5 and 22.5-s samples (red vs. green and horizontal bar vs. vertical bar keys). The pigeons were subsequently tested with transfer trials. An example of a transfer trial is that the 2-s sample was presented, and was followed by a delay, then by the comparison stimuli associated with the 2-s and 4.5 second samples (e.g., red and vertical bar keys). The rationale of this design is that if pigeons were using a default response strategy, accuracy would be at chance because the 2-s sample was not coded, and both of the default response comparisons were presented. The results showed above chance performance during these trials, suggesting that the pigeons do not respond by default, and that all samples were coded.

Detection Model. Another possible explanation of the choose-short bias is the detection model. Gaitan and Wixted (2000) proposed that pigeons search their memory for whether the most salient sample was presented and they respond by default to the other choice when it was not shown or remembered. The main distinction between the detection model and default responding is that it is assumed that pigeons code both the more-salient and less-salient sample, although the pigeons use a strategy in which they search memory for only one sample-type. The detection model can account for the choose-short bias because if pigeons are only searching memory for long samples and responding short in its absence, then the memory for the long sample will be forgotten as the delay is increased. Gaitan and Wixted (2000) investigated this model by training pigeons to discriminate between 2 and 10-s time samples, and also presenting no-sample trials. In one experiment, the correct choice was the same for the 2-s and the no-sample trials, whereas in the other experiment the correct choice was the same for the 10-s and no-sample trials. In the first experiment, it was proposed that pigeons would search their memory for the long sample, whereas in the second experiment their strategy would be to search memory for the short sample. It was hypothesized that if the pigeons used a detection strategy, accuracy would remain high for the 2-s and no-sample trials in the first experiment, whereas in the second experiment, accuracy would remain high for the 10-s and no-sample trials. Under these conditions, decay would not affect pigeons' memory for the 'non-searched' samples, maintaining high accuracy. The results of the first experiment supported their hypotheses: the pigeons' accuracy when the 2-s and no-sample trials were presented was significantly greater than when the 10-s sample was presented. In the second experiment their hypotheses were also supported: accuracy for the 10-s and no-sample trials was significantly higher than when the 2-s sample was presented. The results of both of these experiments support the detection hypothesis because accuracy for the 'searched' sample decreased, and accuracy for the 'non-searched' sample remained high. This suggests that pigeons may be using other strategies, which may be the cause of the biases found in other experiments.

In another study, Gaitan and Wixted (2004) devised an experiment that manipulated light intensity to test the detection model in a non-temporal task. In the first experiment, pigeons were trained to discriminate between bright vs. dim house-light samples, and following the acquisition of the task; a no-light probe trial was introduced as well as delay testing trials. One of their hypotheses was that the delay would decrease the accuracy of the most salient sample (the searched sample), whereas the accuracy of the least salient sample would remain relatively unaffected. Another hypothesis was that on the no-light samples, pigeons would choose the 'dim' comparison by default. The results supported the hypotheses: there was higher accuracy for the dim samples, and accuracy well below chance for the bright samples when a delay was introduced. When the no-light samples were presented, pigeons chose the 'dim' comparison as was expected. In the second experiment, a similar procedure was used with the exception that the samples presented were the dim and no-light samples. This was to show that less salient samples could be the most salient, and that with this manipulation; the dim sample would be searched. The result of this experiment was that accuracy was maintained for the no-light samples, and decreased well below chance for the dim samples when a delay was introduced. These results provide support for the detection hypothesis in a non-temporal task, and illustrate the importance of sample salience in regards to biases.

The choose-short bias and confusion. Another interpretation of the choose-short bias for filled intervals is that it can be an artifact due to confusion (Sherburne, Zentall, & Kaiser, 1998). Sherburne and colleagues argue that the choose-short bias is due to confusion between the ITI and the DI. The confusion can arise from the similarities between the ITI and a novel DI, which can produce a 'resetting' of the internal clock because in previous training, when a pigeon encountered a dark period, it meant the beginning of a trial. Once the comparison stimuli appear, the pigeon will respond to the 'short' comparison because it has 'reset' its clock in anticipation of a new trial, and does not have a sample in working memory. Past studies have shown that a 'no-sample' trial is judged to be more similar to a short sample (see Fetterman, & MacEwen, 1989; Kraemer et al., 1985), and thus the reason for the choose-short bias. Sherburne et al. (1998) devised a study to illustrate this: they had four groups of pigeons with differing conditions; two of the conditions consisted of matching illumination during the ITI and DI (both illuminated, or both dark), and the two conditions consisted of differentiated illumination of the ITI and DI (ITI dark, DI illuminated; ITI illuminated, DI dark). The results of this experiment were similar to those of Spetch and Rusak (1992), namely that the conditions with differentiated ITI and DI illumination conditions did not yield a choose-short bias whereas in the conditions where the ITI and DI were the same, a choose-short bias was found. This was interpreted to be in support of the confusion hypothesis because the pigeons were said to confuse the DI with the ITI.

In another study, Zentall, Klein, and Singer (2004) examined the effect of confusion of the DI with the ITI in the context of the detection model proposed by Gaitan and Wixted (2000, 2004). Zentall et al. (2004) replicated the Gaitan and Wixted (2000) study discussed previously and obtained similar results. They trained pigeons to discriminate between 0, 2, 10-s samples following a 1-s delay. The 2-s sample was mapped onto one choice comparison and the 0 and 10-s samples were mapped on the other choice comparison. Following acquisition, delay-testing was introduced: the delays were 1, 2, 4, and 12-s. The results were similar to what Gaitan and Wixted (2000) found:

they found a decrease in accuracy for the 2-s sample, but maintained high accuracy for the 0 and 10-s samples. This was interpreted by Gaitan and Wixted (2000) as being due to pigeons using a detection strategy in which pigeons search memory for whether one of the samples was presented. Because this sample was subject to forgetting, pigeons would subsequently choose the other alternative by default. Zentall et al. (2004) proposed that pigeons were not using a detection strategy, but rather that the results were due to the confusion of the DI with the ITI. Zentall et al. (2004) proposed that the method used by Gaitan and Wixted (2000) may be problematic because the pigeons were trained to discriminate 0-s samples, which consisted of a 15-s ITI, followed by 1-s delay, and then the comparison stimuli were presented. This method was said to be problematic because pigeons may have learned a rule that when a period of darkness immediately precedes the comparison stimuli, it was a 0-s sample. When this is generalized to delay testing, all of the samples may become '0-s' samples with respect to this. Zentall et al. (2004) tested this hypothesis by training pigeons to discriminate 0, 2, 10-s samples, and mapped the 10-s sample on one choice comparison, and the 0, and 2-s sample on the other. They included two groups: one in which the ITI was illuminated and the DI was dark (group ITI-light), and the other in which the ITI and DI were dark (group ITI-dark). If the pigeons were in fact using a detection strategy, there would be relatively unchanged accuracy following a delay for the 0 and 2-s sample, whereas the accuracy would decrease significantly when the 10-s sample was presented for both group ITI-light and ITI-dark. However, if the confusion hypothesis were correct, then a decrease in accuracy for all trial-types would be expected as the delay increased for group ITI-light, and a decrease in accuracy only for the 10-s sample for group ITI-dark. The results found supported the confusion hypothesis: accuracy decreased for all trial-types for group ITI-light, but not for group ITI-dark. This provides evidence that perhaps for temporal tasks, pigeons do not use a detection strategy.

Dorrance, Kaiser, and Zentall (2000) proposed that another source of confusion that may cause the choose-short bias is the novelty of the delay interval. Dorrance et al. (2000) devised a study that reduced the novelty effects. They trained pigeons to discriminate between short 2-s and long 8-s durations. Each training trial included a variable delay interval (0, 1, 2, 4-s). There were two groups: one group consisted of matching illuminations during the ITI and the DI (Dark-Dark), and in the other, the illumination conditions of the ITI and DI were differentiated (Light-Dark). The acquisition results were noteworthy because the pigeons in group Light-Dark acquired the task significantly faster and performed better overall than in group Dark-Dark. This is contrary to what would be expected when considering that the relative-duration hypothesis proposes that group Light-Dark would be a more difficult task to learn because the 'background' in which the sample was presented differed. After subsequent delay testing, the pigeons in group Dark-Dark did not display a significant choose-short bias, especially when the delay intervals were the same as during training. The pigeons in group Light-Dark however, displayed a choose-long bias which increased as a function of the delay. Dorrance et al. (2000) proposed that this might have been due to the pigeons timing to the comparison stimuli. These results provide support for the confusion hypothesis.

Although the confusion hypothesis explains the choose-short bias in the previous study, there have been studies that have generated contradictory results. Kelly and Spetch

(2000) trained pigeons to discriminate short 2-s and long 6 or 8-s intervals of light with a baseline delay of five seconds. There were four conditions: ON-ON, OFF-OFF, ON-OFF, and OFF-ON (ITI-DI, respectively), which were similar to the Sherburne et al. (1998) study. Subsequent delay tests (10 and 20-s delays) revealed a choose-short bias in three of the four groups: ON-ON, OFF-OFF, and ON-OFF. The OFF-ON condition did not produce a bias. Kelly and Spetch (2000) found that the majority of the birds in the OFF-ON condition showed a choose-short bias, however some birds showed a choose-long bias which cancelled out the choose-short bias. They proposed that the choose-long bias was due to the pigeons timing the total duration of the illumination. These results are contradictory to the confusion hypothesis because a choose-short bias should not occur when the ITI and DI are differentiated. These results also contradict the notion of the confusion due to the novelty of the DI because the bias was observed even when the pigeons were trained with a baseline delay.

Leblanc and Soffié (2001) also devised a study that calls into question the validity of the confusion hypothesis. They used a stepwise delay procedure with rats, and subsequently obtained a choose-short effect when delay was greater than that of the baseline delay, and a choose-long effect when it was less than the baseline delay. Leblanc and Soffié (2001) had observed the rats' behaviour during each part of the trial. They found that the rats' behaviour differed during the ITI and the DI. During the ITI, the rats walked around the operant chamber whereas during the DI, the rats faced the levers of the operant chamber. The argument considers that if the rats were in fact confusing the DI for the ITI, their behaviour would be the same in the two conditions.

The literature regarding the choose-short bias and memory for time in pigeons suggests that there is more than one possible reason that this effect is found. Although the instructional confusion hypothesis might account for some of the results, it cannot explain other phenomenon such as some of the effects explained by the relative duration hypothesis (testing memory at a delay that is less than baseline will produce a choose-long bias). It is important to consider that there may be multiple factors controlling the results of these studies.

Empty Intervals

A less extensively studied time sample in memory for time is the empty interval. An empty interval is a time sample signaled by two markers. An example of an empty interval is as follows: the beginning of the sample is signaled by the presentation of a one second marker (either tone or light), followed by the time sample, then the end of the time sample is signaled by another marker. Pigeons can learn to discriminate short and long empty intervals relatively quickly, and with great accuracy. However, after the introduction of a delay interval, the opposite of what is observed with filled intervals is found, namely a choose-long bias (Grant, 2001; Santi et al., 1999). This effect is illustrated in the bottom panel of Figure 1. Santi, Hornyak, and Miki (2003) devised an experiment that analyzed the effect of having both filled and empty intervals mixed within a session. The underlying purpose of this experiment was to determine whether the processing of events were different for filled and empty intervals. The premise of this was that since all other conditions were the same, then only one bias (if any) would be observed if event processing was the same. If a choose-short and choose-long effect were obtained for filled and empty intervals, respectively, then the results would support differing event processing. Pigeons were trained to discriminate short and long samples of both filled and empty intervals; the comparison stimuli for filled and empty intervals were differentiated by different colour comparisons. Once the pigeons acquired the task with high accuracy, delay testing was introduced (0, 1, 3, and 9-s delays). A choose-short and choose-long bias for filled and empty intervals, respectively, was attained which was consistent with the previous studies.

Possible explanations of the choose-long bias.

There are several speculations as to why the choose-long bias phenomenon occurs. The following will review possible explanations of the choose-long bias such as: confusion; summation of marker illumination; analogical coding and default responding; and the detection model.

The choose-long bias and confusion. Santi and colleagues (1999) reported one of the first investigations of empty intervals. In this experiment, either tone or light markers of 500-ms signaled the time intervals. After the acquisition of the task, the delay intervals were introduced (0, 1, 3, 9-s). The results of the experiment were the same for both the tone and light markers: they observed a moderate choose-short bias at a 1-s delay, whereas a choose-long bias was found following the 3 and 9-s delay. The bias increased as the delay increased. These results could have been interpreted as being due to the different processing of filled and empty intervals, however a different interpretation can be attributed, namely confusion. Santi et al. (1999) interpreted these results as the pigeon may have been confused when a novel DI was introduced, and was timing the duration from the second marker to the comparison stimuli, thus obtaining the choose-short bias at the 1-s delay, and a choose-long bias at longer delays.

Grant (2001) investigated the choose-long effect further by testing whether pigeons were indeed timing from the second marker or whether the effect was due to processing filled and empty intervals differently. Grant (2001) trained pigeons to discriminate short 2-s and long 8-s empty intervals in an operant chamber. For one group of pigeons, a red key light marked the onset of the empty interval, and a green key light signaled the end of the samples consistently. For the other group of pigeons, the red and green key lights were used interchangeably to signal the onset and the offset of the samples. Once this task was acquired, delay intervals were introduced. Grant (2001) found that the choose-long bias was observed for both the consistent and inconsistent groups. To determine whether the pigeons were actually attending to the markers, another experiment was devised to test for this. The consistent and inconsistent groups were given the opposite condition, for example, in group-consistent, the red usually signals the onset of the sample, but during the test trials, the green key light signaled the onset of the sample, and the red signaled the offset. The inconsistent group essentially did not have different conditions. The results showed that there was a significant drop in accuracy in the opposite condition for the consistent group. The results of these experiments suggest that there may in fact be different processes in processing filled and empty intervals, and that pigeons do attend to the markers, which suggests that the choose-long bias is not due to confusion of the markers.

Summation of marker illumination. Another possible explanation of the choose-long bias is that the pigeons were summing the duration of the two markers of the empty interval. The subjective shortening model may account for this if the two markers were summed. This is because on a long sample trial, the duration of the first marker will subjectively shorten more than on a short sample trial: thus, the long sample would be represented as a short marker summation, whereas the short sample would be represented by a long marker summation. The choose-long bias would be expected to occur following a delay because the marker summation would undergo further subjective shortening, and the marker summation would be more similar to a long sample than a short sample. Santi et al. (2003) investigated this in one experiment. Pigeons were trained to discriminate short and long empty intervals signaled by 500-ms of light. During testing, on some of the trials, the duration of the second marker was lengthened to 750-ms instead of 500-ms for both the short and long samples. On other trials, the duration of the second marker was increased to 1000-ms instead of 500-ms. No significant difference in accuracy was found for both the short and long durations, which is evidence against the pigeons using a summation strategy.

Asymmetrical coding and default responding. The asymmetrical coding and default responding rule previously discussed in the context of filled intervals might account for the choose-long bias found with empty intervals: that is the pigeons code the samples asymmetrically. The pigeons may only be coding the short samples, and as the delay increases, the memory for the sample becomes forgotten, and the pigeon will respond 'long' by default because there is no memory for a sample (the short sample). The most salient sample would be assumed to be short for empty intervals as opposed to the long sample for filled intervals. Santi et al. (2003) investigated this possibility with both filled and empty intervals. This experiment consisted of probe trials that included either the presentation of an empty-long or a filled-short sample, which would be considered non-coded samples according to asymmetrical coding. Following the presentation of the sample, either of the following comparison stimuli sets was presented: empty-long vs. filled-short which were the default responses, or empty-short vs. filled-long which were the coded responses. It was expected that if pigeons were in fact coding these samples asymmetrically, then accuracy would be at chance when either the default response, or the coded response comparison stimuli sets were presented, whereas if pigeons coded these samples, accuracy would be above chance when the default response comparison stimuli set were presented. The results showed that accuracy was significantly above chance when the default response comparison set was presented. This provides strong evidence against default responding.

Detection model. The detection model previously discussed to explain the choose-short bias with filled intervals might account for the choose-long bias with empty intervals. The detection model proposes that pigeons use a strategy that involves retrospectively searching memory for one of the sample-types (usually the most salient) (Gaitan, & Wixted, 2000). Following a delay, memory for the 'searched' sample would decay, thus creating a bias. For empty intervals, the short sample would be the sample that is searched in memory, and pigeons would respond to long when the short sample is absent in memory is absent because of decay as the delay is increased.

The review of the literature suggests many noteworthy, yet conflicting explanations of the choose-short and long biases. There are many explanations of this phenomenon; however these hypotheses do not explain all of the outcomes. An important consideration to make is that there may be more than one factor controlling this behaviour, such as variations in coding and retrieving strategies which may ultimately be the result of differing experimental procedures. It is thus important to consider the effect of procedural manipulations in order to isolate how pigeons process and remember filled and empty intervals.

Experiment 1

The purpose of this experiment was to determine whether the choose-short and long biases are due to the confusion of the novel DI with the ITI. This was accomplished by differentiating the illumination conditions of the ITI, sample presentation phase, and the comparison phase (DI and comparison stimuli presentation), which were magenta, dark, and grey respectively. The pigeons were trained using a symbolic matching to sample procedure to discriminate between short 2-s and long 8-s durations of filled and empty intervals. Pigeons' memory for time was tested following acquisition of this task by introducing a delay following the sample (1, 3, 9-s delays). A diagram of the procedure is illustrated in Figure 2. It was expected that if the confusion hypothesis is correct for explaining the choose-short and choose-long biases for filled and empty intervals, then no biases should occur.

Method

Subjects and Apparatus

The subjects used in this experiment were six Silver King pigeons, maintained at approximately 80% of their free-feeding weight. The pigeons were familiar with the apparatus, but naïve to the task. They had previous training in a delay to reward symbol discrimination study, but had never been trained to discriminate short and long durations of filled and empty intervals. The pigeons had access to constant grit and water, and were fed Purina Pigeon Chow when they were not run, and when it was necessary to maintain their weight. The pigeons were housed individually in a common room that was maintained at 22 degrees Celsius and illuminated with a 12:12 light/dark cycle. The experiment was conducted five days a week between 8:30 am and 1:00 pm.

Three touchscreen stations located in individual rooms were used for this experiment. Each station consisted of a clear Plexiglas cage (30 cm wide X 40 cm deep X 36 cm high) with an open end facing the touchscreen. On both the left and the right sides of the apparatus, was an opening of 5.7 X 5 cm that provided access to a hopper filled with grain (Coulbourn Model E14-10). A colour super VGA monitor (Mitsubishi SD4311C) with an attached touch frame (Carrol Touch, Frame 8100-9583-01, Card 8200-3224-01) was at the open end of the Plexiglas cage. An IBM compatible computer that controlled the stimulus displays, recorded peck location, and operated the hoppers was located in each room.

Procedure

Pre-training. The pigeons were trained to peck a coloured square, presented on the upper left or right side of the screen. The side that the square appeared was random; either colour was presented on either side of the monitor. The colour of the square was either blue, yellow, green, or red; the dimensions were approximately 4 cm X 4 cm. Reinforcement consisted of a 6-s access of mixed seed, presented randomly on either the right or left side of the apparatus. Once the pigeons always pecked the squares, they proceeded to the baseline-training phase of the experiment.

Baseline-training. The task consisted of discriminating between short 2 s and long 8 s durations of both filled and empty intervals. The time sample cues consisted of an illumination on the touchscreen. The filled intervals consisted of presenting a continuous cue for either 2 or 8-s. An empty interval consisted of a time sample (2 or 8-s of darkness) signaled by two markers (1-s duration each): one to indicate the beginning of the sample, and another to indicate the end. The signals and markers for filled and empty intervals were differentiated: one was a horizontal brown rectangle (approx. 1.75 cm X 3.25 cm), and the other a vertical white rectangle (approx. 3.25 cm X 1.75 cm). These parameters were counterbalanced: for half of the birds the brown horizontal rectangle was the signal for filled intervals, and for the other half it was the marker for the empty intervals. Both filled and empty interval trials were presented randomly within the same session: half of the trials consisted of filled intervals, and the other half consisted of empty intervals.

Every trial began with a randomly selected ITI (4, 8, 16, 32-s), which consisted of magenta illumination on the entire video monitor. Following the ITI, there was a period of 500 ms of darkness before the onset of the signal or marker. After the signal or marker, there was another 500ms of darkness. This was followed by the comparison phase, which consisted of a grey video monitor illumination. The comparison stimuli for the empty and filled intervals were differentiated: they were either blue vs. yellow, or red vs. green. The comparisons were partially counterbalanced between the birds: for half of the birds, the comparison for the filled signals was blue vs. yellow, and for the other half it was red vs. green: the details of this are given Table 1. After a correct response, the bird was reinforced with access to mixed grain from one of the food hoppers. As in the previous phase of the experiment, the side of the hopper presentation was determined randomly. The pigeons were given 6-s access to the food hopper after a correct response for sessions 1-32. Subsequently, the hopper time was reduced to 4-s for the remainder of training due to problems maintaining the birds at 80% of their free-feeding weight. The experiment was a corrected procedure: after an incorrect response, the pigeon was not reinforced; however the grey illumination remained for either 4 or 6 seconds (depending on the hopper time setting). Following this, the same trial was repeated beginning with the initial 500 ms of darkness. The trial was repeated until the correct response was given: the pigeon was reinforced at this point. There were an equal number of filled and empty trials, and the order in which they were presented was random. Baseline training consisted of 64 trials in a session for sessions 1 and 2, and subsequently increased to 128 trials per session for the remainder of the training. The criterion for advancing to the delay-testing phase was 70% or better on all of the conditions (filled-short, filled-long, empty-short, empty-long) for at least two consecutive sessions.

Delay-testing. Delay testing was conducted for 15 sessions of 144 trials each. Ninety-six of the trials in a session were the identical to the baseline-training phase, and the other 48 trials consisted of delay testing trials (16 of each delay type). There were three delay intervals: 1, 3, and 9-s. The delay testing trials were essentially the same as the baseline training trials, with the exception of the delay intervals, which consisted of extending the grey video monitor illumination of the comparison phase prior to the presentation of the comparison stimuli). The correction procedure remained only for the 0-s baseline trials to maintain performance.

Results and Discussion

Acquisition.

All of the pigeons reached criterion within 52 sessions, and the data were blocked into sessions of 4. The mean percent correct sample discrimination with filled and empty intervals for acquisition is shown in Figure 3. Mean percent correct for the last two blocks was as follows: 86.96% for empty-short, 84.46% for empty-long, 91.67% for filled-short, and 89.69% for filled-long. In all statistical analyses reported in this study, the rejection region was $p < .05$. An interval-type X sample duration X block (2 X 2 X 13) analysis of variance (ANOVA) was conducted to analyze the data. A significant effect of sample duration was found [F (1, 5) = 20.42]: there was higher accuracy for short rather than long samples. A significant effect of block [F (12, 60) = 43.74] was also found: there was overall higher accuracy as the blocks increased. There was also a significant interval-type X sample duration interaction [F (1, 5) = 6.49]. The simple main effect of the interval-type X sample duration interaction was a sample duration effect with filled intervals [F (1, 5) = 59.41], but not with empty intervals: there was higher accuracy for short rather than long samples with filled intervals, and no difference with empty intervals. A significant interval-type X sample duration X block interaction was also found [F (12, 60) = 2.43]. The simple main effects for this interaction included a significant effect of interval-type for blocks 1, 12, 13 [F (1, 5) = 9.92, 7.79, 9.74]: there was higher accuracy for filled intervals during these blocks. A significant simple main effect of sample duration was also found at block 3, 4, 5, 7, 9, 10, 11, 13 [F (1, 5) = 11.58, 6.73, 20.50, 7.09, 11.02, 14.43, 17.28, 9.83]: during these blocks, there was higher accuracy for short samples. At block 3 a significant interval-type X sample duration interaction was found [F (1, 5) = 10.55]. The interaction included a significant effect of sample duration for filled intervals [F (1, 5) = 38.96], but not for empty intervals: there was higher accuracy for short samples with filled intervals.

These results indicate higher accuracy for the short samples, and also a non-significant trend of higher accuracy for filled intervals. This significant difference between short and long samples was also found in other studies with filled intervals (e.g., Santi, Sanford, & Coyle, 1998), but statistically significant differences have not been found with empty intervals. This difference in accuracy of short and long samples is important to consider for interpreting the results of the delay testing phase and the other phases of this study because the trend will likely be present throughout this study. The non-significant trend of higher accuracy for filled intervals was also present in the Santi et al. (2003) study.

Delay testing.

The mean percent correct sample discrimination with filled and empty intervals for delay testing is shown in Figure 4. An interval-type X sample duration X delay (2 X 2 X 4) ANOVA was conducted to analyze the data. The analysis revealed a significant main effect of delay [F (3, 15) = 160.39]: as the delay increased, accuracy decreased. There was also a significant effect of sample duration [F (1, 5) = 14.53]. There was higher accuracy for short samples compared to long samples at all delays. There were no significant interactions; the retention functions were parallel

These results suggest that both the choose-short and choose-long biases may be due to instructional confusion. These results are consistent with Sherburne et al. (1998): by differentiating the ambient illumination of the ITI, sample presentation phase, and the DI, the biases usually found with filled (choose-short) and empty (choose-long) intervals was not found. Accuracy for short samples was significantly higher than for long samples for both filled and empty intervals; however this difference was also found in the baseline and acquisition data. The retention functions were parallel, thus the difference in accuracy found was not a working memory bias.

It was possible, however that the pigeons were not timing filled and empty intervals, rather they could have discriminated short and long samples by timing from the offset of magenta to onset of grey, and thus not processing whether the sample was filled or empty. The next study will examine how the pigeons were attending to the sample stimulus, and what strategies were being used to discriminate short and long samples because it was possible to solve the task without attending to the sample signals and markers. This is an important factor to investigate because the underlying issue that this manipulation intended to discover was confusion in filled and empty intervals.

Experiment 2

The purpose of this experiment was to determine the features of the trial that the pigeons utilized to solve the task in Experiment 1. As was previously discussed, the pigeons could have been attending to other features of the trial to solve the task, such as timing from the offset of the ITI to the onset of the comparison phase. It is important to make the distinction of whether the pigeons were timing filled and empty intervals in the previous experiment. This distinction is important because there are different response biases for filled and empty intervals, which suggests that these two interval-types may be processed differently. There were three probe-test phases in this experiment: the pre- and post-sample manipulation phase; signal and marker omission phase; and the no-sample phase. The procedure for all of the phases in this experiment is illustrated in Figure 5.

The purpose of the pre- and post-sample manipulation phase was to determine whether the pigeons were timing filled and empty intervals or whether they were simply timing between the offset of the ITI and the comparison phase to solve the task. This was accomplished by keeping the duration between the ITI and the comparison phase constant for all trial types, and inserting the signal or markers within an 11-s allotted time. With these manipulations, it would be expected that if the pigeons were simply timing from the offset of the ITI to the comparison phase, then pigeons would be expected to choose-long when the sample presented was either short or long, regardless of whether they were filled or empty because the 11-s interval was more like a long sample.

The purpose of the signal and marker omission phase was to determine the amount of control the signals and markers have on solving the discrimination task. This was accomplished by omitting the signals and markers of a trial, but retaining the duration of the trial-type (a period of darkness). With this manipulation, it was expected that if the pigeons did not attend to the signals and markers to solve the task, then accuracy during these trials would not differ from the regular baseline trials. This phase of the experiment also included a delay manipulation to determine whether the same response patterns found in Experiment 1 would be replicated.

The purpose of the no-sample phase was to determine the effect of eliminating the signals and markers, as well as the time sample. This manipulation considers important theoretical implications: the confusion hypothesis proposes that pigeons will respond short when they are confused because of stimulus generalization (Sherburne et al., 1998). The no-sample trials will be generalized to the short sample response rather than to the long sample response. Previous studies have shown that when the comparison stimuli associated with filled intervals are presented on no-sample test trials, a choose-short bias was observed (Grant et al., 1994; Kraemer et al., 1985; Santi et al., 2003). Whereas, when the comparison stimuli associated with empty intervals were presented on no-sample test trials, the results have been inconsistent. These studies have reported a choose-short bias

(Santi et al., 1999), a choose-long bias (Santi et al., 2003), and no bias (Grant 2001). Consequently, while a choose-short bias was expected when the comparison stimuli associated with filled intervals were presented, predictions cannot be made when the comparison stimuli associated with empty intervals were presented.

Method

Subjects and Apparatus

The subjects and the apparatus used in this experiment were the same as in Experiment 1.

Procedure

Pre- and post-sample manipulation. The sessions of this phase of the experiment consisted of 150 trials, of which 120 were regular baseline training trials and the remaining 30 were probe trials: for every five trials, one was a probe trial. The probe trials began with a variable ITI, as in baseline training, followed by an 11-s period during which the signal was presented. The sample-types that were presented within this period of time were empty-short, filled-short, and filled-long samples. Empty-long samples were not included in this because the duration of an empty-long trial is 11-s, thus, if the pigeons were timing from the offset of the ITI to the onset of the DI, a long response would be made on all of the trials. There were two types of probe trial: pre-sample duration manipulation and post-sample duration manipulation trials (see Figure 5). In the pre-sample manipulation trials, the sample was presented at the end of the 11-s period, whereas in the post-sample manipulation trials, the sample was presented at the beginning of the 11-s period. Pigeons were reinforced for a correct response to the sample-type presented during the probe trials, however an incorrect response resulted in no reinforcement and the trial was not corrected. The type of probe trial (pre- or post-sample) was alternated between sessions: for half of the birds, the pre-sample test was the first session, whereas the other half received the post-sample test. This phase was conducted for 4 sessions both the pre- and post-sample test (2 sessions each).

Signal and marker omission. This phase of the experiment consisted of 160 trials for each session, of which 32 were probe trials: for every five trials, one was a probe trial. Half of the probe trials included a 0-s delay, and the other half included a 9-s delay. The other 128 trials were the same as the baseline trials. The probe trials were essentially the same as regular trials, with the exception of the sample presentation stage: the signals and markers were omitted (only a period of darkness). The duration of the period of darkness included the two 500-ms periods (before and after the sample in baseline trials), the markers, and the sample type. The darkness durations were as follows: 5, 11, 3, and 9-s (representing empty-short, empty-long, filled-short, filled-long respectively). The comparison stimuli that were presented on these trials matched the sample it was representing (e.g., 5-s darkness had the same comparison stimuli and correct response as an empty-short sample). Pigeons were reinforced for choosing the comparison associated with the duration presented. When errors occurred, pigeons were not reinforced, and the trial was not repeated. The signal and marker omission probe test was conducted for 10 sessions.

No-sample. There were 160 trials per session, and 32 of these were the no-sample probe trials. The no-sample probe trials began with a variable ITI, followed by 1-s of darkness (to account for the two 500-ms intervals at the beginning and the end of the sample presentation phase), then the comparison stimuli was presented. The comparison stimuli shown were either blue and yellow, or red and green. Pigeons were reinforced 50% of the time for the probe trials, regardless of response. The no-sample phase was conducted for 5 sessions.

Results and Discussion

Pre- and post-sample manipulation.

The mean percent correct response for the pre- and post-sample phase is shown in Figure 6. A condition X interval-type/duration (3 X 3) ANOVA was conducted to analyze this data. Only the empty-short, filled-short, and filled-long samples were analyzed for the three conditions: baseline, pre-sample, and post-sample. A significant effect of condition [F (2, 10) = 94.58], interval-type/duration [F (2, 10) = 221.99], and a condition X interval-type/duration interaction [F (4, 20) = 34.43] was found. The significant simple main effects of the interaction are the effect of condition for empty-short [F (2, 10) = 81.94] and filled-short samples [F (2, 10) = 67.24], but not for filled-long intervals. A multiple comparison analysis was also conducted in order to analyze whether there was a difference between the pre- and post-sample means within each interval-type/duration, and also whether there was a significant difference between the pre- and post-sample means and the baseline mean within each interval-type/duration. There were no significant effects between the pre- and post-sample manipulations for all interval-type/durations. The test revealed significantly higher accuracy for baseline trials than for the pre- and post-sample manipulation means and the baseline mean for the empty-short [F (1, 15) = 196.03] and filled-short trials [F (1, 15) = 178.79].

These results suggest that the pigeons were timing from the offset of the ITI to the comparison phase in this task. The mean percent accuracy during manipulations of the empty-short and filled-short samples was 32.08% and 27.5% compared to 91.95% and 94.86% at baseline, respectively. Although this manipulation greatly affected performance, the mean percent correct was above what would be expected if the pigeons were simply timing from the offset of the ITI to the comparison phase. The mean percent correct for the short sample probe trials would have been expected to be roughly 10% if the pigeons were only attending to the duration between the ITI and comparison phase because baseline accuracy for long samples was roughly 90%. This suggests that the pigeons' performance may have been controlled by other factors such as the signals and markers. The following phase in the experiment will investigate whether the signals and markers played a role in comparison choice.

Signal and marker omission.

The mean percent correct for the signal and marker omission phase is found on Figure 7. An interval-type X sample duration X condition (2 X 2 X 3 - baseline, 0-s delay probe, 9-s delay probe) ANOVA was conducted to analyze the data. A significant effect due to condition [F (2, 10) = 73.41] and a significant interval-type X sample duration X condition interaction was found [F (2, 10) = 7.92]. An analysis of this interaction revealed significant effects of interval-type [F (1, 5) = 7.77] and sample duration [F (1, 5) = 41.82] for the baseline trials: there was significantly higher accuracy for short samples, and higher accuracy for filled intervals at baseline. An interval-type X sample duration interaction was found for the 0-s probe trials [F (1, 5) = 7.25], which reveals a significant effect of interval-type when the sample was short [F (1, 5) = 6.49], but not when it was long: there was higher accuracy for short samples when the interval-type was filled. A multiple comparison analysis was also conducted in order to determine whether the mean of the 0-s probe trials differed from the baseline trials. There was significantly higher accuracy for the baseline trials than for the 0-s probe test for empty-short, empty-long, filled-short, and filled long samples [$F(1, 15) = 1310.84, 660.34, 700.51, 879.01$, respectively].

These results suggest that the signals and markers exerted partial control in the sample discrimination task because there was higher accuracy when the signals and markers were present. The pigeons were thus utilizing both the duration between the ITI and the comparison phase, and the signals and markers as cues. Pigeons may have learned to use both cues because of the correction procedure. The correction procedure repeats a trial that an error occurred in, forgoing the ITI. This procedure prevents the pigeons from using the offset of magenta to begin timing, which may encourage the use of the signals and markers. The pigeons could have also used the offset of the grey at the end of the comparison phase (following an incorrect response, the video monitor is illuminated for the duration that the hopper presentation would have been), and timed up to the onset of the grey illumination, that is from grey to grey. This is however a less likely scenario because the result shows control of both the signals and markers, and the offset of magenta to the onset of grey.

No-sample.

The mean percent short response for the no-sample phase is found on Figure 8. The mean percent short response for no-sample trials when the comparison stimuli associated with filled and empty intervals was presented were 76.04% and 69.38%, respectively. An interval-type X sample-type (2 X 3) ANOVA was conducted to analyze the data. A significant effect of sample-type was found [$F(2, 10) = 120.09$]. A single sample *t* test revealed that the mean percent short response for the no-sample trials was significantly above chance when the comparison stimuli associated with either filled [$t(5) = 3.27$] or empty [$t(5) = 3.14$] were presented.

The results show a bias to choose-short when the comparison stimuli associated with either filled or empty intervals were presented. The choose-short bias observed when the comparison stimuli associated with filled intervals were presented was consistent with previous no-sample testing (Grant et al., 1994; Kraemer et al., 1985; Santi et al., 2003), but the choose-short bias found when the comparison stimuli associated with empty intervals was presented is inconsistent with past research (Grant 2001; Santi et al., 2003). This however may be due to the results found in the delay testing of Experiment 1 and that the pigeons were timing from the ITI to the comparison phase, which makes whether they were timing empty intervals ambiguous. Although the proportion of short responses is not as high when no-sample is presented in comparison to when a short sample is presented, this is expected because of generalization decrement, and consistent with previous research (Kraemer et al., 1985; Santi et al., 2003). These results are consistent with the confusion hypothesis because the pigeons associated the no-sample trials with the short sample, as was proposed by Kraemer et al. (1985).

Experiment 3

The purpose of this experiment was to address the issue of the pigeons timing from the offset of the ITI to the onset of the comparison phase, and to determine whether the choose-long bias can be created for empty intervals, and no bias for the filled intervals. This was achieved by differentiating the video monitor illumination of the different phases of the trials: magenta ITI, a dark sample presentation phase, and a dark comparison phase. The procedure of this experiment is illustrated in Figure 9. If the choose-long bias found in previous studies was due to the confusion of the delay interval with a long empty interval, with these conditions, a choose-long bias would be expected. A choose-short bias for filled intervals is not expected because the ITI is differentiated from the DI, and thus cannot be confused.

Subjects and Apparatus

The subjects and the apparatus used in this experiment were the same as in Experiments 1 and 2.

Procedure

Baseline-training. The procedure was essentially the same as in Experiment 1 with the exception of the video monitor illumination during the various phases: there was a magenta ITI, dark sample presentation, and comparison phases. There were 128 trials per session, and the criterion for advancing to the delay-testing phase was 70% or better on all of the conditions (filled-short, filled-long, empty-short, empty-long) for at least two consecutive sessions.

Delay-testing. Delay testing was conducted for 20 sessions of 144 trials each. Ninety-six of the trials in a session were identical to the baseline-training phase, and the other 48 trials consisted of delay testing trials. The delay durations were 1, 3, and 9-s. Other than the change in the illumination of the comparison phase, the procedure was identical to the delay testing in Experiment 1.

Results and Discussion

Acquisition.

All of the pigeons reached criterion within 20 sessions. The mean percent correct sample discrimination with filled and empty intervals for acquisition is shown in Figure 10. Mean percent accuracy for the last two sessions was as follows: 89.06% for empty-short, 85.16% for empty-long, 93.49% for filled-short, and 91.67% for filled-long. An interval-type X sample duration X session (2 X 2 X 20) ANOVA was conducted to analyze the data. A significant effect of interval-type [F (1, 5) = 7.33], sample duration [F (1, 5) = 31.26], and session [F (19, 95) = 11.30] was found. There was higher accuracy for filled intervals, for short samples, and accuracy improved over sessions. An interval-type X sample duration [F (1, 5) = 8.05] and an interval-type X sample duration X session interaction [F (19, 95) = 1.87] were found. The effects of the interval-type X sample duration interaction indicate that there was a significant difference between filled and empty intervals when the sample was long, but not when it was short [F (1, 5) = 8.82]. There was higher accuracy for filled intervals when the sample was long. The simple main effects of the interval-type X sample duration X session interaction reveal a significant effect of interval-type for sessions 4, 15, 16, 19 [F (1, 5) = 8.26, 7.83, 13.77, 35.27, respectively]: there was higher accuracy for filled intervals during these sessions. An effect of sample duration was also found for sessions 1, 7, 8, 9, 10, 13, 14, 16, 17, 18 [F (1, 5) = 7.75, 7.10, 8.05, 10.23, 34.03, 105.57, 9.24, 7.47, 8.06, 31.25, respectively]: there was higher accuracy for short samples for these sessions. An interval-type X sample duration interaction was also found for sessions 1, 4, 5 [F (1, 5) = 16.55, 25.01, 6.36, respectively], and the simple main effects of this interaction reveals a significant effect of sample duration when the interval-type was empty [*F* (1, 5) = 20.59, 10.00, 63.88, respectively], but not when it was filled: there was higher accuracy for short samples when the interval-type was empty.

The results of the acquisition of this task show similar, yet more pronounced effects of the acquisition in Experiment 1. The effect of sample duration was significant, as in Experiment 1, however the effect of interval-type in this manipulation was found significant in comparison to the non-significant trend in Experiment 1.

Delay testing.

The mean percent correct sample discrimination with filled and empty intervals for delay testing is shown in Figure 11. An interval-type X sample duration X delay (2 X 2 X 4) ANOVA conducted on these data shows a significant main effect of delay [*F* (3, 15) = 96.70]: as the delay increased, accuracy decreased. A significant effect of interval-type was also found [*F* (1, 5) = 22.75]: there was higher accuracy for filled intervals in comparison to empty intervals. A significant sample duration X delay interaction [*F* (3, 15) = 5.5] was also present. The effects found in the sample duration X delay interaction reveal that accuracy for short samples was significantly higher than the long samples at the 0-s delay [*F* (1, 5) = 24.18]. A marginally significant interval-type X sample duration X delay interaction was found [*F* (3, 15) = 3.02] which reveals a significant sample duration X delay interaction for empty intervals [*F* (3, 15) = 8.52], but not for filled intervals. The simple main effects for the sample duration X delay interaction for empty intervals reveals significantly higher accuracy for short samples at the 0-s delay [*F* (1, 5) = 13.19] and choose-long bias at the 9-s delay was marginally significant in this analysis, however, subsequent analyses revealed significant results. A one-tailed single sample t-test revealed that accuracy for the data for empty short were significantly below chance (50%) $t(5) = -2.107$. A subsequent single sample t-test revealed that accuracy for the empty long sample was significantly above chance $t(5) = 2.210$. A one-tailed paired sample t-test revealed that accuracy for empty short samples was significantly lower than the empty long samples following a 9-s delay $t(5) = -2.481$.

The results of this experiment show a significant choose-long bias for empty intervals at the 9-s delay, and no bias for filled intervals. There was, however a non-significant tendency to choose-long at longer delays for filled intervals. Four of the six birds showed a choose-long tendency, whereas the other two birds showed a choose-short tendency. The criteria for a bird tending to choose-long or short was a 10% difference at the 9-s delay, otherwise it was not considered a tendency. The tendency to choose-long may have been due to pigeons timing up to when the comparisons were presented. This effect was also found in previous studies (see Dorrance et al., 2000; Kelly & Spetch, 2000). Although the confusion hypothesis could explain the results of the choose-long tendency for filled intervals, it cannot account for the choose-short tendency found for two of the birds. Kelly and Spetch (2000) attributed their results to pigeons timing up to the comparisons, and also found that some of the birds showed a choose-short bias. They proposed that all of the birds coded the samples analogically, and that for some birds, they timed up to the comparisons, and for the others, the samples were subjectively shortened as the delay increased. This is a possible explanation for the filled data because some of the birds showed a choose-long tendency, and the others, a choose-short tendency despite the manipulations in this experiment that made the ITI and DI distinct.

The results of this experiment show that with these conditions, it was possible to create confusion for empty intervals, however the results for the filled data remain ambiguous. The empty interval data support the instructional confusion hypothesis, whereas the data for filled intervals is inconclusive.

Experiment 4

The purpose of this experiment was to ensure the pigeons were not timing from the offset of the ITI to the onset of the comparison phase. This was achieved with a similar method as in Experiment 1 with the exception that the background illumination of the video monitor differed in colour: there was a dark ITI, dark sample presentation phase, and a grey comparison phase illumination. An illustration of this procedure can be found in Figure 12. This manipulation prevented the pigeons from timing from the offset of the ITI to the comparison phase because the video monitor illumination was the same during the ITI as in the sample presentation phase. Pigeons were not able to use other cues such as timing from the offset of the grey after the hopper presentation to the onset of the grey (from grey to grey) because the ITI duration was varied such as in the previous experiments. This manipulation also prevented the pigeons from confusing the ITI with the delay interval, and the delay interval for a long empty sample. A decrease in accuracy was expected for the acquisition of this task if the offset of the ITI was an important factor in signalling the beginning of the time sample. The biases for both filled and empty intervals were not expected with these manipulations if the biases are due to instructional confusion.

Subjects and Apparatus

The subjects and the apparatus to be used in this experiment were the same as in the previous experiments.

Procedure

Baseline-training. The procedure was essentially the same as in Experiment 1 with the exception of the illumination condition: the ITI and the sample presentation phase were dark, and the comparison phase was grey. There were 128 trials per session, and the criterion for advancing to the delay-testing phase was 70% or better on all of the conditions (filled-short, filled-long, empty-short, empty-long) for at least two consecutive sessions.

Delay-testing. Delay testing was conducted for 20 sessions of 144 trials each. Ninety-six of the trials in a session were the identical to the baseline-training phase, and the other 48 trials consisted of delay testing trials. The delay durations were 1, 3, and 9-s. Other than the change in the illumination, the procedure was identical to the delay testing in Experiment 1.

Results and Discussion.

Acquisition.

The change in video monitor illumination greatly affected the accuracy of all sample-types. The mean percent correct for the last session of Experiment 3 and the first session of the current experiment is shown in Figure 13. An interval-type X sample duration X condition (whether it was magenta, dark, dark (MDD) or dark, dark, grey (DDG)) (2 X 2 X 2) ANOVA was conducted to analyze the data. A significant effect of interval [$F(1, 5) = 38.95$] and condition [$F(1, 5) = 391.95$] was found: there was generally higher accuracy for filled intervals relative to the empty intervals, and higher accuracy on the last day of the MDD condition than the first day of the DDG condition. A significant interaction of sample duration X condition was found [F (1, 5) = 8.71]. There were no significant simple main effects for this interaction, however the data suggest that accuracy for short samples were more disrupted than the long when the DDG experiment was implemented: that is, the pigeons tended to choose the comparison associated with long samples.

The majority of the birds acquired the task by the 25th session. The mean percent correct sample discrimination with filled and empty intervals for acquisition for the first 25 sessions is shown in Figure 14. The mean percent correct for the 25th session was as follows: 79.95% for empty-short, 80.99% for empty-long, 89.37% for filled-short, and 92.50% for filled-long. An interval-type X sample duration X block (2 X 2 X 5) ANOVA was conducted to analyze the data of the first 25 sessions. There were 5 blocks of 5 sessions, however for one bird; there was only 2 sessions included in the last block due to experimenter error. A significant overall effect of sample duration [F (1, 5) = 50.52] and block [F (4, 20) = 33.36] was found. There was higher accuracy for long samples, and also gradual improvements in accuracy over blocks. A significant interaction of sample duration X block was found [F (4, 20) = 6.33]. This interaction reveals significant simple main effects of sample duration at block 1, 2, and 4 [F (1, 5) = 28.61, 6.78, 19.05, respectively] in which there was significantly higher accuracy for the long samples during these blocks. Although the figure suggests that there was an effect of interval-type, this was not found to be significant because there were two birds that did not acquire the empty interval trials, which caused variability. The two birds that did not acquire the task to criterion on the empty interval trials (average of 61.72 and 60.16 for empty short and long) were given 7 sessions consisting of only empty interval trials followed by 3 regular sessions. The other birds were also given 10 more sessions during this time. All of the pigeons reached criterion by the 35th session. The mean percent correct for the last two sessions was as follows: 82.55% for empty-short, 83.59% for empty-long, 88.28% for filled-short, and 89.32% for filled-long.

The change in video monitor illumination from magenta to dark during the ITI and reinstating the grey illumination during the comparison phase clearly disrupted pigeons' effectiveness in solving the task, furthermore the trend for accuracy of short and long samples had been reversed relative to the past phases in the experiment: there was generally higher accuracy for the long samples relative to the short samples. A portion of the decrease in accuracy can be attributed to generalization decrement; however the decrease in accuracy can also be due to the pigeons learning to discriminate the duration of the signals, and the duration between the markers. Previous experiments in this study have shown that the offset of magenta and the onset of grey have exerted some control on timing samples in this task. These results reflect this control, however it is uncertain whether the change from magenta to dark, or the change from dark to grey was the greatest influence in the decrease in accuracy. The trend of higher accuracy for long samples ought to be noted for the results of the delay-testing results in order to consider possible biases.

Delay-testing.

The mean percent correct sample discrimination with filled and empty intervals for delay testing is shown in Figure 15. An interval-type X sample duration X delay (2 X

2 X 4) ANOVA conducted on these data showed a significant main effect of delay [F (3, 15) = 141.79]: as the delay was increased, accuracy decreased. No other effects were found statistically significant.

These results were found to be consistent with the confusion hypothesis because there were no biases found following a delay, only a decrease in accuracy as the delay was increased. This manipulation prevented pigeons from confusing the DI with the ITI (hypothesized cause for the choose-short bias for filled intervals), and from confusing the DI with an empty interval (hypothesized cause of the choose-long bias). This manipulation also prevented the pigeons from using other cues for timing the empty and filled samples, as was possible in the previous experiments of this study. Although these results provide strong evidence in support of the confusion hypothesis, the biases should be able to be produced with the same pigeons in order to fully support the confusion hypothesis as an explanation of these results.

Experiment 5

The purpose of this experiment was to replicate the Santi et al. (2003) experiment in order to produce the choose-short and choose-long biases for filled and empty intervals in a within-session design. This experiment is necessary in order to conclude that the manipulations of the previous experiments produced the 'no bias' results, and not individual differences in these pigeons. In this experiment, the background illumination of the video monitor was dark for the duration of the trial. A diagram of the procedure is found in Figure 16. A decrease in accuracy for all sample types was expected on the first session of this experiment, due to generalization decrement, but should not be as disrupted as in the previous experiment because the pigeons have already learned to solve the task without the magenta monitor illumination during the ITI. A choose-short bias was expected with filled intervals and a choose-long bias was expected with empty intervals following a delay if these biases were due to instructional confusion. The choose-short bias was expected because pigeons would confuse the delay interval with the ITI, and the choose-long bias would be expected because the pigeons would confuse the DI for an empty sample.

Subjects and Apparatus

The subjects and the apparatus to be used in this experiment will be the same as in the previous experiments.

Procedure

Baseline-training. The procedure was essentially the same as in Experiment 1 with the exception of the background illumination of the video monitor: the ITI, the sample presentation phase, and the comparison phase were dark. There were 128 trials per session, and the criterion for advancing to the delay-testing phase was 70% or better on all of the conditions (filled-short, filled-long, empty-short, empty-long) for at least two consecutive sessions.

Delay-testing. Delay testing was conducted for 20 sessions of 144 trials each. Ninety-six of the trials in a session were the identical to the baseline-training phase, and the other 48 trials consisted of delay testing trials. The delay durations were 1, 3, and 9-s. Other than the changes in the illumination, the procedure was identical to the delay testing in Experiment 1.

Results and Discussion.

Acquisition.

The change in procedure from a dark ITI, dark sample presentation phase, and grey comparison phase (DDG) to dark ITI, sample presentation and comparison phases (DDD) also affected pigeons' performance in the discrimination task. The mean percent correct for the last session of Experiment 4 and the first session of the current experiment is shown in Figure 17. An interval-type X sample duration X condition (2 X 2 X 2) ANOVA was conducted to analyze the data of the baseline trials of the last session of delay testing (0-s delay) for the DDG experiment with the first session of the DDD experiment. An overall significant effect of condition was found [F (1, 5) = 39.57]: there was higher accuracy on the last session of DDG compared to the first session of DDD. An interval-type X condition interaction was also found [F (1, 5) = 6.46]. The simple main effect of this interaction reveals a significant effect of interval-type effect when it was the DDG condition [F (1, 5) = 15.43]: There was higher accuracy for filled intervals when the condition was DDG, but not DDD. A significant interaction of sample duration X condition was found [F (1, 5) = 6.46]: although there were no significant simple main effects of this interaction, the data suggest that there was more disruption for short samples in comparison to long samples.

All of the pigeons reached criterion within 10 sessions. The mean percent correct sample discrimination with filled and empty intervals for acquisition is shown in Figure 18. Mean percent accuracy for the last two sessions was as follows: 86.20% for empty-short, 83.33% for empty-long, 91.67% for filled-short, and 92.45% for filled-long. An interval-type X sample duration X session (2 X 2 X 10) ANOVA was conducted to analyze the data. The analysis revealed a significant effect of session [F (9, 45) = 14.99]: Accuracy increased with the sessions.

There was a similar pattern of acquisition of this task in comparison to the acquisition of the previous experiment. These results can be attributed to generalization decrement, however the manipulations of the experiment may not have affected the timing strategy as much as in the previous experiment because the pigeons acquired the task much faster. There was a non-significant trend, however for pigeons to choose the comparison associated with the long samples on the first days of training. The effects of interval-type and sample duration were not significant with this data analysis.

Delay-Testing.

The mean percent correct sample discrimination with filled and empty intervals for delay testing is shown in Figure 19. An interval-type X sample duration X delay (2 X 2 X 4) ANOVA conducted to analyze these data. A significant effect of delay was found: that is as the delay increased, accuracy decreased [F (3, 15) = 90.25]. An interval-type X delay interaction was also found to be significant [F (3, 15) = 3.61]: there was higher accuracy for filled intervals at 0, 1, 9-s delays [F = (1, 5) = 9.19, 10.00, 22.50, respectively]. An interval-type X sample duration X delay interaction was also found [F (3, 15) = 5.50]. There was a significant effect of delay at all sample types: empty-short, empty-long, filled short, and filled long [F (3, 15) = 28.19, 5.99, 30.95, 15.96 respectively]: as the delay increased, accuracy decreased. A significant effect of sample duration for empty intervals was found at the 0-s delay [F (1, 5) = 30.11]: there was higher accuracy for short samples. A significant effect of interval for short samples was also found at the 9-s delay [F (1, 5) = 19.31]: there was higher accuracy for filled-short samples than empty-short. A significant effect of interval for long samples was found at the 1-s delay [F (1, 5) = 14.52]: there was higher accuracy for the filled-long samples than empty-long. The interaction revealed a significant simple interaction effect of sample duration X delay for empty intervals, but not for filled intervals [F (3, 15) = 5.09]. A choose-long bias for empty intervals was marginal at the 9-s delay with this analysis, however subsequent analyses of the data for the 9-s delay reveals significant results. A one-tailed single-sample t-test revealed that the data for the empty-short samples was significantly below chance (50%) $t(5) = -2.188$. A subsequent one-tailed t-test revealed that accuracy for the empty-long sample was significantly above chance $t(5) = 2.476$. A one-tailed paired samples t-test revealed that the accuracy for short samples was significantly lower than the long samples following a 9-s delay for empty intervals $t(5) =$ -2.373. This provides evidence for a choose-long bias for empty intervals at a 9-s delay.

There was no evidence of a choose-short bias for filled intervals; however, a significant choose-short bias was found for empty intervals at the 0-s delay, and a significant choose-long bias was found following a 9-s delay. These results raise questions regarding the reason for the choose-long bias being obtained for empty intervals and no bias found with filled intervals. The individual data for filled intervals show that two birds tended to choose-short, and two birds tended to choose-long, whereas the other two showed no tendency for either. The criteria for a bird tending to choose-long or short was a 10% difference following a 9-s delay, otherwise it was not considered a tendency. These results are similar to Experiment 3 because the choose-short tendency was found with two of the birds. It is possible that the choose-long tendency with some of the pigeons was an artifact of the pigeons timing through the DI. Another possible explanation for the lack of bias for filled intervals is that the previous training made the DI and the empty interval more confusable than the DI and ITI. A possible reason for this is that there were more trials without the magenta during the ITI compared to the grey during the comparison phase: the magenta during the ITI was omitted following Experiment 3, and the grey during the comparison phase was omitted following Experiment 4. There were more consecutive trials without the magenta monitor illumination compared to the grey illumination, which may have caused more confusion for empty intervals when tested with a delay.

General Discussion

The purpose of this study was to determine whether the choose-short and choose-long biases found with filled and empty intervals were an artifact of the procedure, namely whether the pigeons were confusing the DI for the ITI, and confusing the DI with an empty interval. This has important theoretical implications because if the biases are simply due to instructional confusion, theories such as subjective shortening proposed by Spetch and Wilkie (1983) become unnecessary. This study manipulated the video monitor illumination of the various stages of a trial in order to eliminate possible confusion of the DI with the ITI, as well as the DI with the empty sample.

In the first experiment, the video monitor illumination during the ITI, sample presentation phase, and the comparison stage were differentiated (magenta, dark, grey respectively). These conditions were differentiated in order to eliminate the possibility of confusion. When delays of 0, 1, 3, 9-s were introduced, pigeons had higher accuracy for short samples, however this was consistent throughout all of the delays: the retention functions were parallel. The results for filled intervals were similar to the results found in the Sherburne et al. (1998) study in which they differentiated the house-light illumination of the ITI and DI. The results for empty intervals were consistent with what was proposed by Santi et al. (1999), namely that pigeons may be confusing the DI with an empty interval. Although this provided evidence for the instructional confusion hypothesis, there was the possibility that with this procedure, pigeons were timing from the offset of magenta to the onset of the grey. This was possibly problematic because if they were using this strategy, the distinction between filled and empty intervals would be lost, as well as uncertainty of how the timing was being processed. The interval in question could have been processed as an empty interval (i.e. magenta and grey are markers). It was thus important to determine whether pigeons were timing from magenta to grey, and the extent to which they were being used.

The purpose of Experiment 2 was to determine whether pigeons were timing from magenta to grey, as well as to determine the amount of control the magenta and grey illumination, as well as the signals and markers have on the discrimination. The first phase of this experiment included probe trials which began with a variable magenta ITI, and subsequently a filled short, filled-long, or an empty-short sample was presented within an 11-s period (either at the beginning or the end of the 11-s period). This was followed by the presentation of the comparison stimuli. This manipulation tested whether pigeons were timing from ITI offset to DI onset. The results displayed a high tendency to choose the comparison associated with the long sample. This provided support for the speculation that the pigeons were timing from magenta to grey to discriminate between short and long samples.

The second phase of the experiment was devised to determine whether the signals and markers exerted any control in discriminating short and long samples. This experiment included probe trials in which the signals and markers were eliminated in the trial, however the duration was maintained (period of darkness for the sample-type duration). The results showed a significant decrease in accuracy during these trials. This suggested that there was dual control in this discrimination task: magenta to grey, and the signals and markers.

The purpose of the third phase in Experiment 2 was to test whether pigeons would generalize a no-sample presentation to a short sample as was proposed by Kraemer et al. (1985). This was achieved by presenting probe trials that included a variable magenta ITI, followed by 1-s of darkness, which was followed by comparison stimuli either for filled or empty. When the comparison stimuli for either filled or empty intervals were presented on these trials, pigeons chose the stimuli associated with the short sample. This provided support for the notion that pigeons generalize no-sample to a short sample, however it may be argued that the probe trials may not have been no-sample trials because pigeons could have been timing from offset of magenta to onset of grey. This would be a shorter sample than usual 3 or 5-s, but it may not be able to be considered 'no-sample'.

The results of Experiment 2 illustrated that the pigeons were discriminating short and long samples by using the offset of magenta to grey, as well as the signals and markers. This was problematic for reasons previously discussed, therefore changes in the experimental procedure was necessary. The purpose of Experiment 3 was to address these issues, and to test whether the choose-short and long biases would be found following a delay.

The purpose of Experiment 3 was to prevent the pigeons from timing from the offset of magenta to grey, and to create confusion for empty intervals, and not for filled intervals. This was accomplished by including magenta ITI, dark sample presentation, and dark comparison phases. A choose-long bias was expected because pigeons would confuse the DI with the empty interval because they both were presented within a dark video monitor illumination. There was a significant choose-long bias for empty intervals, and no bias was found for filled intervals. There was a non-significant trend, however for pigeons to choose the comparison associated with the long sample. The choose-long trend was attributed to the pigeons timing through the DI up to the presentation of the comparison stimuli because the grey used to signal the end of the sample was eliminated.

In Experiment 4, the video monitor illumination was dark during the ITI and sample presentation phase, and grey during the DI. This manipulation prevented pigeons from using the offset of magenta to discriminate short and long samples. It was thus necessary to attend to the signals and markers to solve the discrimination task. This manipulation also prevented pigeons from confusing the ITI with the DI as well as confusing the DI with an empty interval. As was expected, no biases were found following a delay. This experiment strengthened the previous evidence in favour of the instructional confusion hypothesis. These results were also consistent with the results found by Sherburne et al. (1998) when they differentiated the house-light illumination of the ITI and the DI. Because the choose-short and long biases were not found throughout this study, an experiment testing whether the biases could be produced was necessary in order to solidify the evidence for the instructional confusion hypothesis.

The purpose of Experiment 5 was to determine whether the choose-short and long biases could be replicated with these pigeons and the procedure. The procedure included a dark video monitor illumination throughout the trial in order to create confusion. This was a similar procedure to that of Santi et al. (2003), who obtained the choose-short and long biases for filled and empty intervals. The results of this experiment show a significant choose-long bias for empty intervals; however only a decrease in accuracy for short and long samples as the delay increased was found filled intervals. The choose-long bias obtained for empty intervals was similar to that of Santi et al. (1999), Grant (2001), and Santi et al. (2003). This finding supports the instructional confusion hypothesis because the bias could be replicated, however the not obtaining a bias with filled intervals was puzzling. A possible reason for this result could be that the previous training experiences prevented the birds from confusing the DI with the ITI more than the DI and empty intervals despite dark illumination during all phases of the trial. A possible reason for this is that there were more trials without the magenta during the ITI compared to the grey during the comparison phase: the magenta during the ITI was omitted following Experiment 3, and the grey during the comparison phase was omitted following Experiment 4. There were more consecutive trials without the magenta monitor illumination compared to the grey illumination, which may have caused more confusion for empty intervals when tested with a delay.

Although the results of this study fit the predictions of the instructional confusion hypothesis best, there are some findings that make these results ambiguous. The confusion hypothesis encountered difficulties explaining why the results of Experiment 5 did not show a choose-short bias for filled intervals. However, when the individual data of Experiment 5 were considered, other explanations could be made. A choose-short tendency was found in two birds, and the choose-long tendency which was found for two of the other birds could have been due to these birds timing through the DI. It is possible that some of the pigeons experienced DI-ITI confusion with filled intervals, whereas some pigeons may have been confused in a similar way as with empty intervals (timing through the DI). This may have been due to the training procedure, which caused the onset of grey to be an important cue to indicate the end of the sample. The instructional confusion hypothesis also has difficulty explaining the individual data for filled intervals in Experiment 3. Two pigeons showed a choose-short tendency at the 9-s delay, which is contrary to what would be expected with the confusion hypothesis. The result for individual birds is inconsistent with the confusion hypothesis, while the overall results are inconsistent with the subjective shortening hypothesis. The subjective shortening hypothesis, asymmetrical coding, and the detection model are not good fits for the data of the present study because they do not make predictions as to why biases do not occur when the various phases of the trial differ in illumination.

The results of this study provide strong evidence for the instructional confusion hypothesis for empty intervals, which was that the pigeons confused the DI with the empty sample, whereas more ambiguous evidence for the confusion hypothesis was found for the filled interval data. Although the confusion hypothesis offers the best fit for the data of the present study, the confusion hypothesis cannot explain all of the results of other studies.

An example of results of a study that cannot simply be explained by the confusion hypothesis is the Kelly and Spetch (2000) study. They trained pigeons to discriminate between 2 and 6 or 8-s food samples with a 5-s baseline delay. There were four groups: OFF-OFF, ON-ON, ON-OFF, and OFF-ON (ITI-DI respectively). The noteworthy aspect of this study is the difference in speed of acquisition for each group. The confusion hypothesis would predict that the acquisition of the task when the ITI and DI differ would be faster because pigeons would not confuse the ITI with the DI, thus the task would be simpler. The results of the acquisition, however revealed the fastest acquisition for group OFF-OFF, followed by ON-ON, OFF-ON, and ON-OFF (fastest to slowest). These results were puzzling in terms of the confusion hypothesis because they were contrary to what was expected.

The Leblanc and Soffié (2001) proposed an argument which argued that rats behaved differently during the ITI and the DI, therefore cannot be confusing the two. They found that during the ITI, the rats walked around the chamber, whereas during the DI the rats faced the levers. This is problematic in terms of the instructional confusion hypothesis because if the rats were confusing the DI for an ITI, it would be assumed that the rats would behave in the same way during the DI as in the ITI.

Another problematic finding for the instructional confusion hypothesis is the relative duration hypothesis. Spetch and Rusak (1989) found that when the ITI and/or the DI were extended relative to baseline, a choose-short bias occurred, whereas when the ITI and/or DI were shortened relative to baseline, a choose-long bias occurred. The instructional confusion hypothesis can explain the choose-short bias when the DI is extended (i.e., confusing the DI for the ITI). However, the confusion hypothesis cannot explain the choose-long bias when the DI is shortened because if confusion for an ITI occurred, there would be no reason why the pigeons would be biased to choose-long. The confusion hypothesis also has difficulty explaining why the choose-short bias occurs when the ITI is extended, and the choose-long bias that occurs when the ITI is shortened. Manipulations to the ITI should be irrelevant to the response biases because these manipulations occur before the sample is presented, and thus should not impact responding. This provides evidence for other processes involved in timing.

The instructional confusion hypothesis also has difficulty explaining why the choose-short bias is not observed in experimental designs that should facilitate confusion. One type of procedure in which the confusion hypothesis would predict confusion, but reliably produces no bias is the many-to-one procedure. Grant and Spetch (1993) used a many-to-one procedure to investigate the effect of mapping a temporal and a non-temporal sample onto one comparison stimulus. In one group, pigeons were trained to discriminate between 2 and 10-s, as well as discriminating keys with a horizontal and vertical bar. One of the temporal samples and one of the non-temporal samples were mapped onto one comparison stimulus, whereas the other two were mapped onto the other. Two other groups of pigeons were trained in a one-to-one procedure: one in which the pigeons discriminated only the two temporal samples, and the other in which the pigeons discriminated only non-temporal samples. Delay testing revealed that only the pigeons that were trained to discriminate only temporal samples showed a choose-short bias. The instructional confusion hypothesis would predict that regardless of comparison mapping, a choose-short bias should occur on temporal samples because the DI can be confused with the ITI.

Another area in which the instructional confusion cannot explain its results is memory for number. Hope and Santi (2004) trained pigeons to discriminate between small and large amounts of flashes of light within the same duration (2 or 8 flashes within 4-s). There were two groups: one that included an illuminated ITI and a dark DI, and the other that included a dark ITI and DI. The instructional confusion hypothesis would predict that pigeons would confuse the DI for an ITI in the group in which they were both dark, and treat it as a no-sample, and thus choose the comparison associated with the small sample. The results of this experiment revealed a choose-small bias for both groups, which is contrary to what the instructional confusion hypothesis would predict.

Previous studies have also shown that when using a successive discrimination procedure for discriminating short and long samples, the choose-short bias is not observed. Spetch and Grant (1993) trained pigeons to discriminate between 2 and 10-s samples. There were two groups: one in which was a choice task (choice between two comparisons), and the other was a successive discrimination procedure. The successive discrimination procedure involves pigeons being presented a sample (either short or long), which is followed by one of the comparison stimuli (random presentation of either the comparison associated with the short or long sample). Pigeons were reinforced after the pigeon pecked at least 5-s after the key was illuminated only when the illuminated key was correct for the sample presented. If it was the incorrect key, the pigeons were not reinforced. The confusion hypothesis would predict that pigeons in both groups would show a choose-short bias following a delay because the ITI and DI were equally confusable. The results showed a choose-short bias for the choice group, and no bias for the successive discrimination group. These results are not only inconsistent with the instructional confusion hypothesis, but also the subjective shortening hypothesis.

The literature suggests many conflicting results in relation to the confusion hypothesis; however, this study provided solid evidence for the confusion hypothesis with empty intervals, and weaker evidence for instructional confusion for filled intervals. It is possible that both confusion and subjective shortening cause the choose-short bias. It is clear that the results of this study, as well as previous studies cannot be explained by a single theory.

There were some limitations of this study, such as a low sample size that ought to be considered if related future studies would be conducted. There were some statistically non-significant trends found in this study that might have been significant provided there was a larger sample size. There are also some questions that remain unanswered after the series of experiments in this study that generate ideas for possible future studies. One question that remains unknown is why the choose-short bias was not found when all of the phases of the trial were dark. A possible experiment would be to test whether the differentiation of trial phases in previous experiments affect how pigeons process time samples in subsequent manipulations. This would be accomplished by training one set of birds with Dark-Dark-Grey (DDG) (ITI-sample presentation-comparison phase) followed by Dark-Dark-Dark (DDD) training, and the other set of birds on the reverse order: DDD then DDG. The DDG manipulation should not create confusion for either filled or empty intervals because the ITI and DI are different, as well as the DI and sample presentation phase, whereas in the DDD manipulation confusion should occur because the various phases of the trial are confusable. The expected outcome of this experiment would be that when training would occur with the DDD manipulation first, there would be the usual response biases (choose-short and long biases), and in the subsequent experiment (DDG), the response biases would be eliminated. In the other group of birds that would encounter the DDG training first, no-biases would be predicted in this experiment, however in the DDD experiment, if the training sequence selectively affected performance on filled intervals, a choose-long bias for empty intervals would be expected, and no-bias or a choose-long bias would be expected with filled intervals. This would be expected because in the preceding experiment (DDG), the grey may signal pigeons to stop timing when a delay interval is introduced, whereas when that cue is eliminated in the DDD experiment, the pigeons will be more likely to time through the delay interval for both filled and empty intervals. This would not be expected with pigeons trained with the DDD manipulation first because pigeons are more likely to use the filled sample offset to stop timing when a delay is presented.

References

Dorrance, B. R., Kaiser, D. H., Zentall, T. R. (2000). Event-duration discrimination by pigeons: The choose-short effect may result from retention-test novelty. *Animal Learning & Behavior. 28,* 344-353.

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.

Gaitan, S. C., & Wixted, J. T. (2000). The role of "nothing" in memory for event duration in pigeons. *Animal Learning & Behavior, 28,* 147-161.

Grant, D. S. (1991). Symmetrical and asymmetrical coding of food and no-food samples in delayed matching in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 17,* 186-193.

Grant, D. S. (2001). Memory for empty time intervals in pigeons. *Animal Learning & Behavior, 29,* 293-301.

Grant, D. S., & Spetch, M. L. (1991). Pigeons' memory for event duration: Differences between choice and successive matching tasks. *Learning and Motivation. Special Issue: Animal Timing, 22,* 180-199.

Grant, D. S., & Spetch, M. L. (1993). Analogical and nonanalogical 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. (1994). The role of asymmetrical coding of duration samples in producing the choose-short effect in pigeons. *Learning and Motivation, 25,* 413-430.

Honig, W. K. (1981). Working memory and the temporal map. In N. E. Spear & R. R. Miller (Eds.), *Information Processing in Animals: Memory Mechanisms* (pp. 167-197). Hillsdale, NJ: Erlbaum.

Hope C., & Santi A. (2004). Pigeons' memory for number of events: Effects of intertrial interval and delay interval illumination. *Learning and Motivation, 35,* 348-370.

Kelly, R. & Spetch, M. L. (2000). Choice biases in delayed matching-to-sample duration with pigeons: Manipulation of ITI and delay illumination. *Quarterly Journal of Experimental Psychology: Comparative & Physiological Psychology, 53,* 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.

Leblanc, P., & Soffie, M. (2001). The choose-short effect in rat memory for event duration: the subjective-shortening model. *Behavioural Processes, 56,* 31-40.

Santi, A., Ducharme, M. J., & Bridson, S. (1992). Differential outcome expectations and memory for temporal and nontemporal stimuli in pigeons. *Learning & Motivation, 23,* 156-169.

Santi, A., Hornyak, S. A., & Miki, A. (2003). Pigeons' memory for empty and filled time intervals signaled by light. *Learning & Motivation. 34,* 282-302

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.

Santi. A., Stanford, L., & Coyle, J. (1998). Pigeons' memory for event duration: differences between visual and auditory signals. *Animal Learning & Behavior, 26,* 163-171.

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., & Grant D. S. (1993). Pigeons' memory for event duration in choice and successive matching-to-sample tasks. *Learning & 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. (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. & Sinha, S. S. (1989). Proactive effects in pigeons' memory for event duration: Evidence for analogical retention. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 347-357.

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.

Wixted, J. T., & Gaitan, S. C. (2004). Stimulus salience and asymmetrical forgetting in the pigeon. *Learning & Behavior, 32,* 173-182.

Zentall, T. R., Klein, E. D., & Singer, R. A. (2004). Evidence for detection of one duration sample and default responding to other duration samples by pigeons may result from an artifact of retention-test ambiguity. *Journal of Experimental Psychology: Animal Behavior Processes. 30,* 129-134.

Table 1

The partial counterbalancing parameters

		Empty		Filled	
Bird	Empty Markers	Short	Long	Short	Long
1	Horizontal Brown	Green	Red	Yellow	Blue
2	Horizontal Brown	Blue	Yellow	Red	Green
3	Horizontal Brown	Red	Green	Yellow	Blue
4	Vertical White	Red	Green	Blue	Yellow
5	Vertical White	Blue	Yellow	Green	Red
6	Vertical White	Yellow	Blue	Green	Red

Figure Captions

Figure 1. An example of a choose-short bias in the top panel, and an example of a choose-long bias in the bottom panel.

Figure 2. A diagram of the procedure in Experiment 1.

Figure 3. The mean percent correct for filled and empty intervals for the baseline data in Experiment 1. The results are shown in blocks of 4 sessions. The error bars represent the standard error of the mean.

Figure 4. The mean percent correct for filled and empty intervals for the delay-testing phase of Experiment 1. The delays were 0, 1, 3, and 9-s. The error bars represent the standard error of the mean.

Figure 5. A diagram of the procedure in Experiment 2 for all phases.

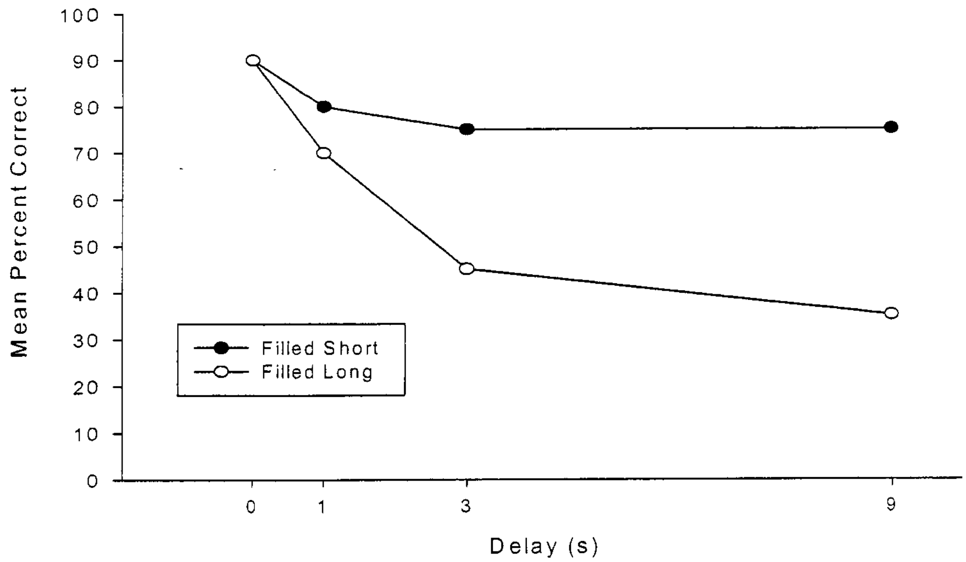
Figure 6. The mean percent correct for the pre- and post-sample manipulation phase in Experiment 2. Results are shown for empty-short, filled-short, and filled-long trials. The error bars represent the standard error of the mean.

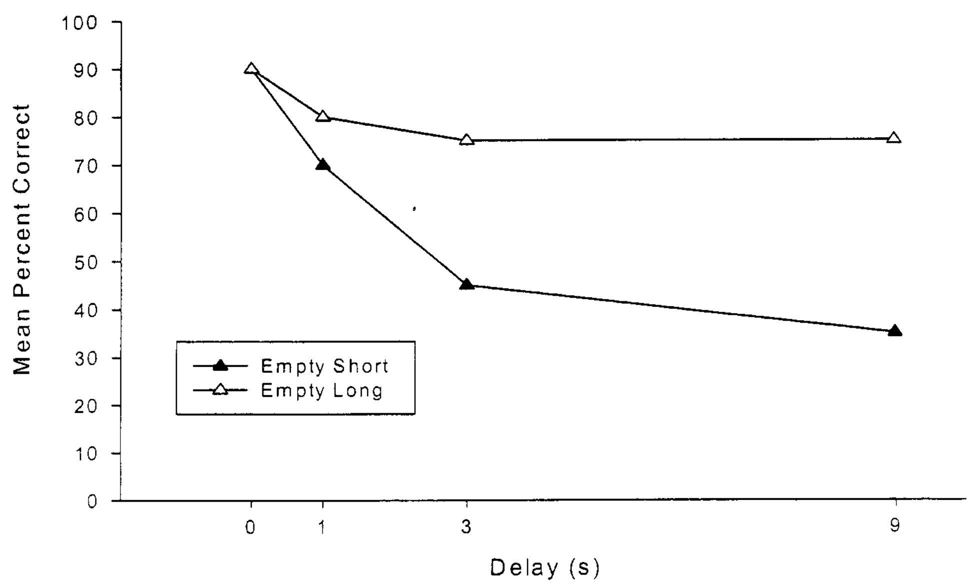
Figure 7. The mean percent correct for the filled and empty interval signal and marker omission phase in Experiment 2. The results include the 0 and 9-s delay probe trials. (ES = empty short, EL = empty long, FS = filled short, FL = filled long, and the numbers represent the length in seconds of the probe trial sample for the respective sample-type). The error bars represent the standard error of the mean.

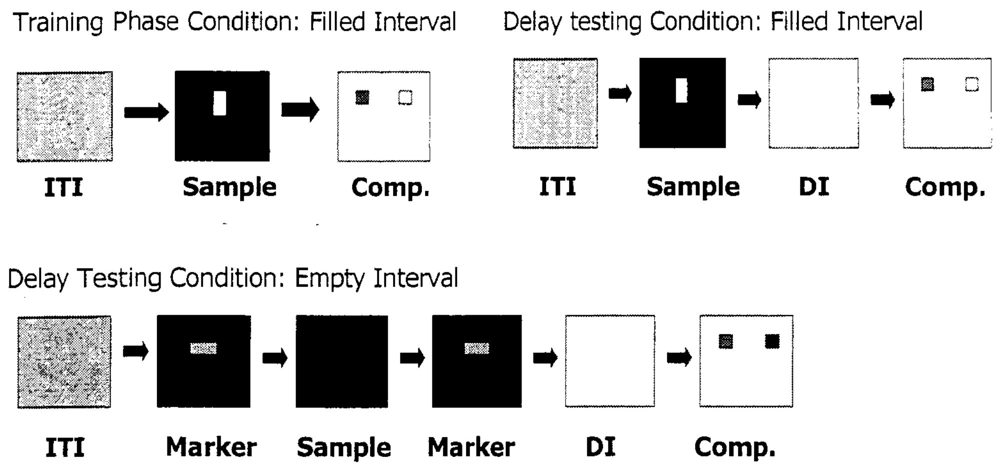
Figure 8. The mean percent short response for filled and empty intervals for the no-sample phase of Experiment 2. The error bars represent the standard error of the mean.

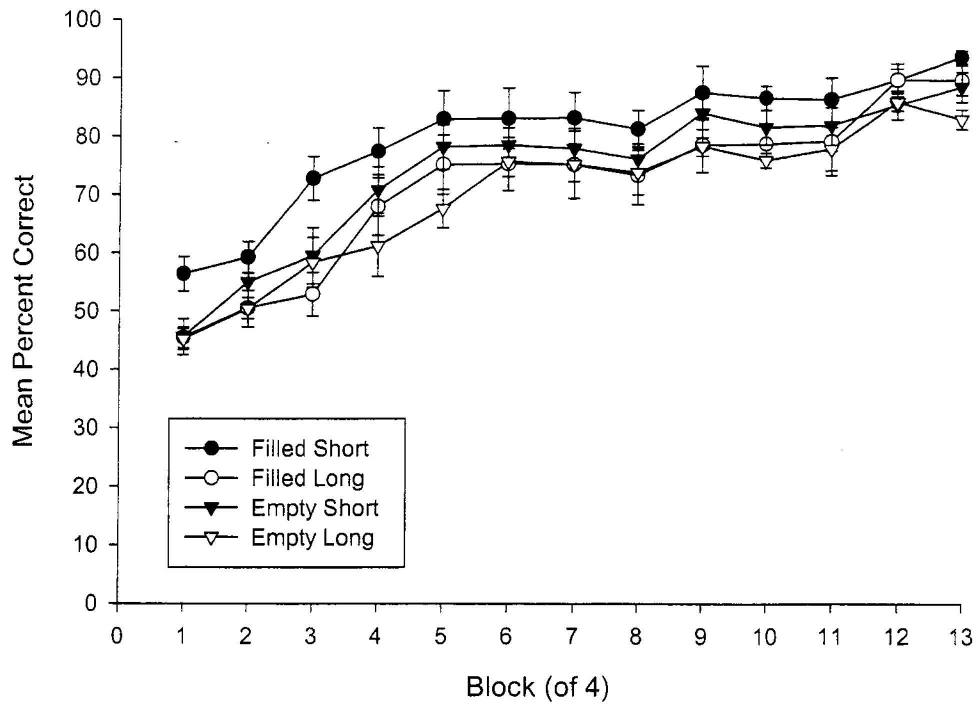
Figure 18. The mean percent correct for filled and empty intervals for the baseline acquisition data for Experiment 5. The error bars represent the standard error of the mean.

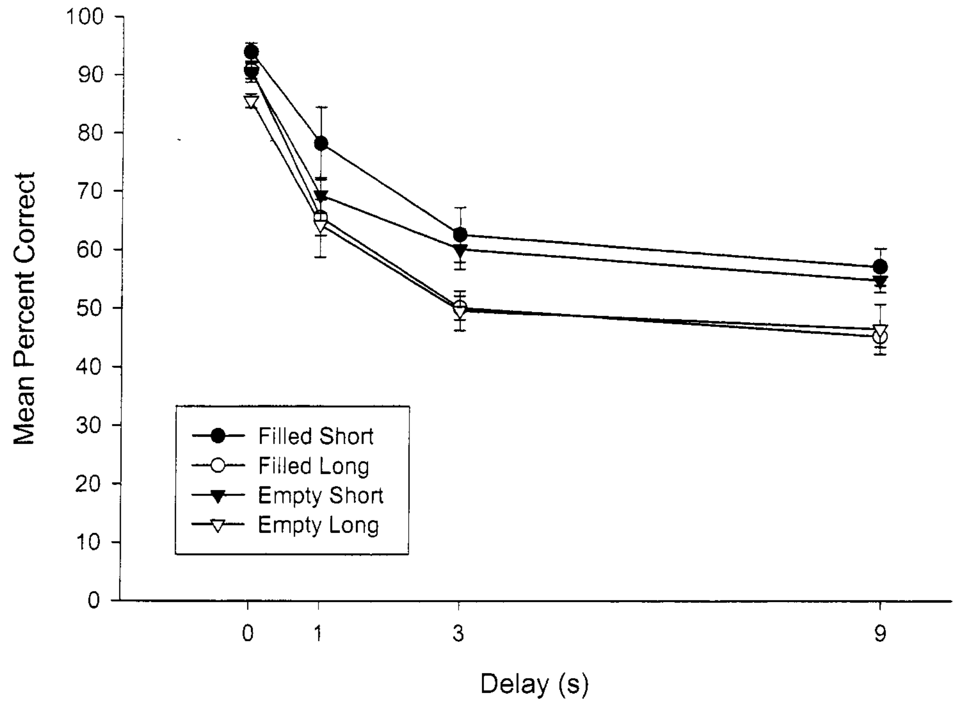
Figure 19. The mean percent correct for delay-testing phase of Experiment 5. The delays were 0, 1, 3, and 9-s. The results for filled intervals are displayed in the top panel. The results for empty intervals are displayed in the bottom panel. The error bars represent the standard error of the mean.


100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
0
1
3
9
Delay (s)
Filled Short
Filled Long

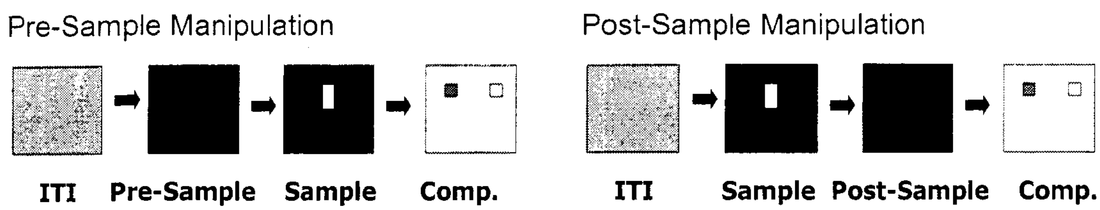

100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
0
1
3
9
Delay (s)
Empty Short
Empty Long


Training Phase Condition: Filled Interval
ITI
Sample
Comp.
Delay testing Condition: Filled Interval
ITI
Sample
DI
Comp.
Delay Testing Condition: Empty Interval
ITI
Marker
Sample
Marker
DI
Comp.

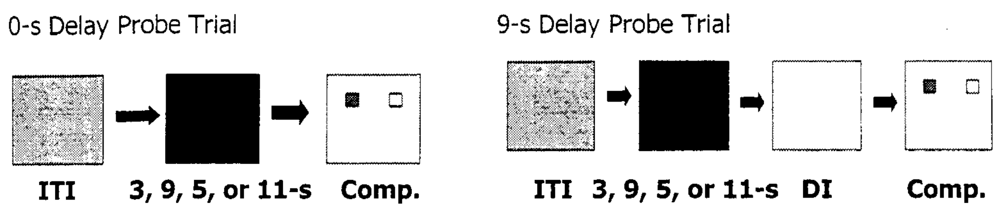

100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Filled Short
Filled Long
Empty Short
Empty Long
0
1
2
3
4
5
6
7
8
9
10
11
12
13
Block (of 4)


100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Filled Short
Filled Long
Empty Short
Empty Long
0
1
3
9
Delay (s)

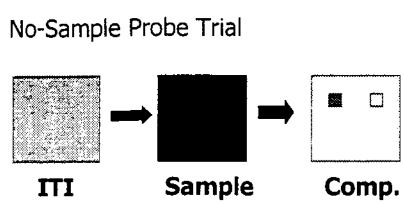
Experiment 2: Pre- and Post-Sample Manipulation Phase

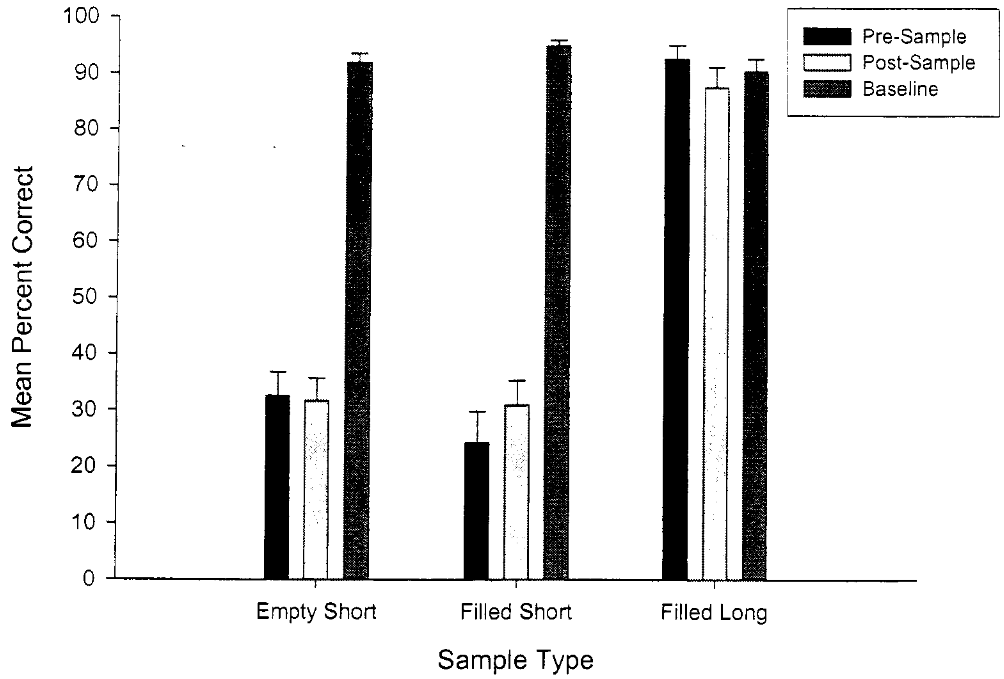


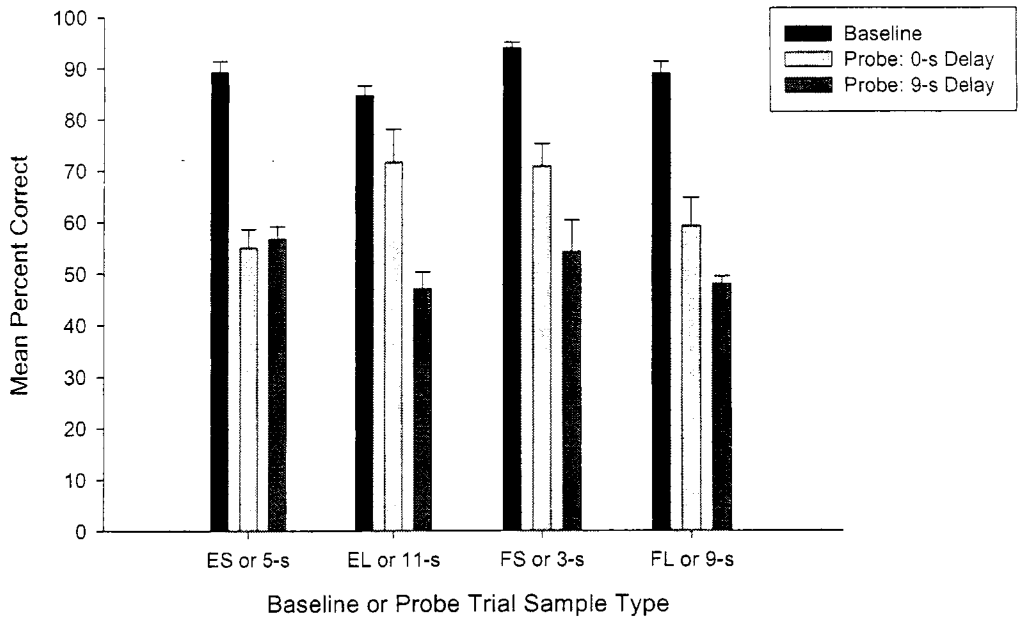
Experiment 2: Signal and Marker Omission Phase



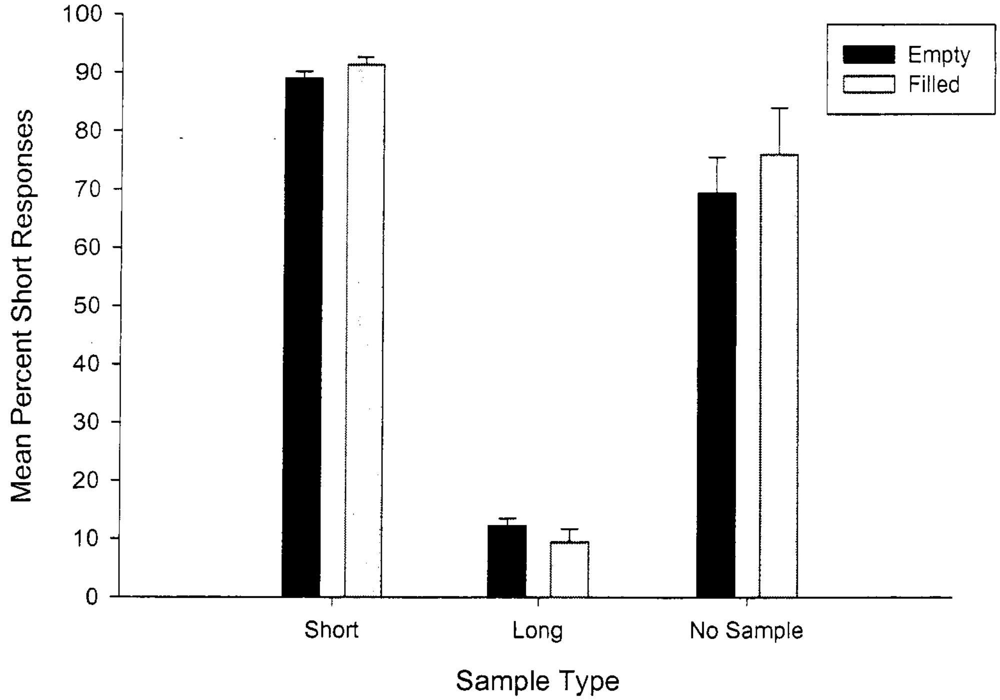
Experiment 2: No Sample Phase

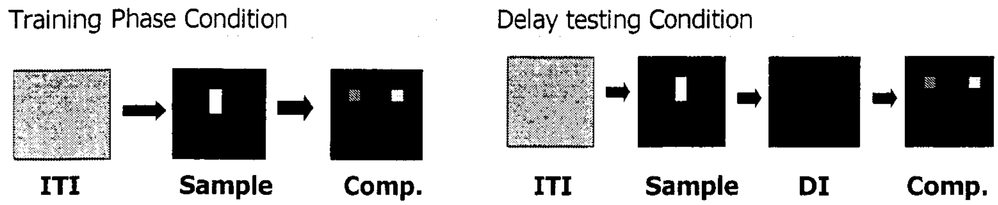


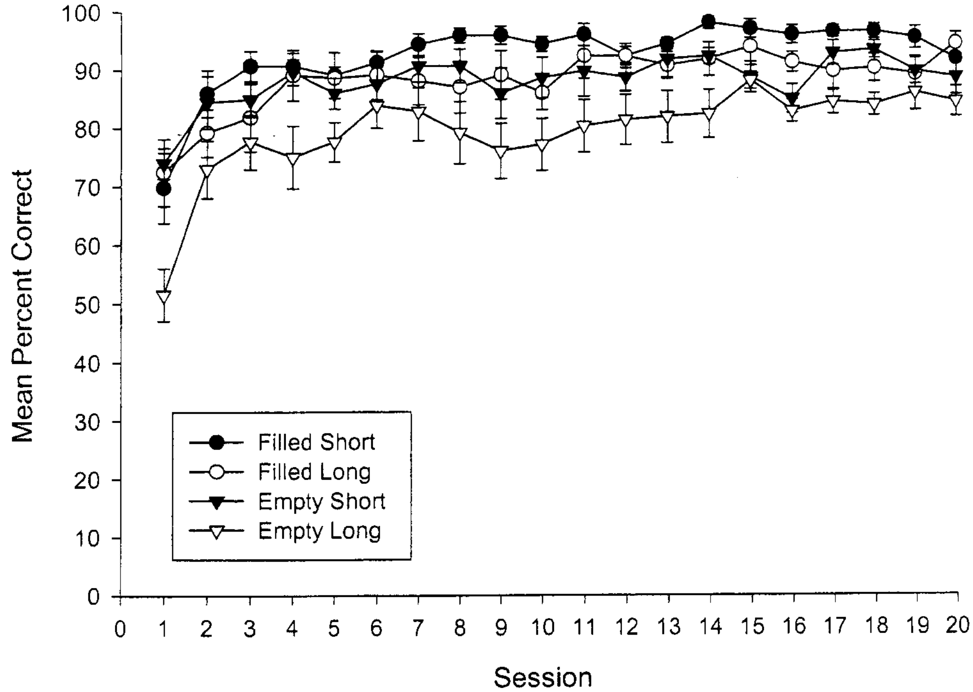

Mean Percent Correct
100
90
80
70
60
50
40
30
20
10
0
Pre-Sample
Post-Sample
Baseline
Empty Short
Filled Short
Filled Long
Sample Type

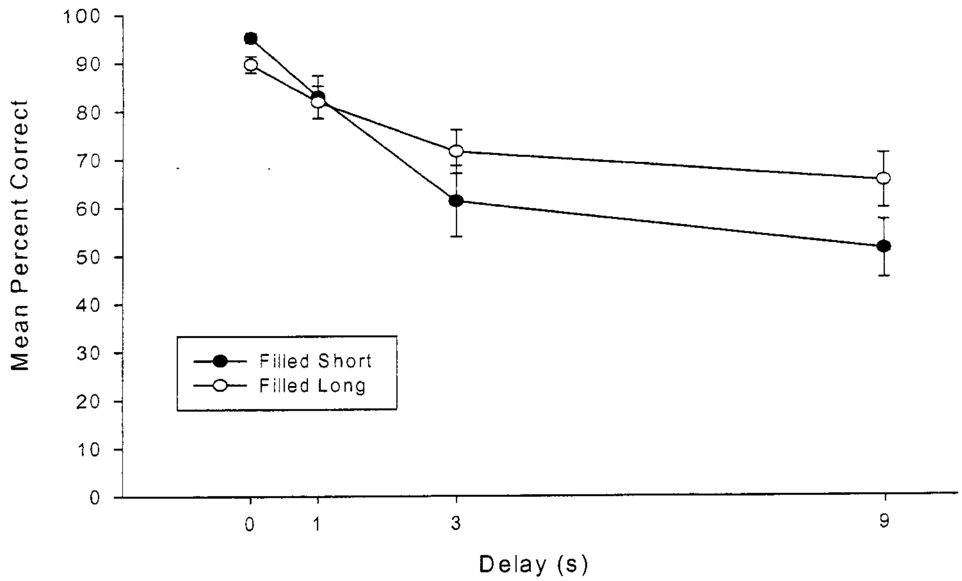


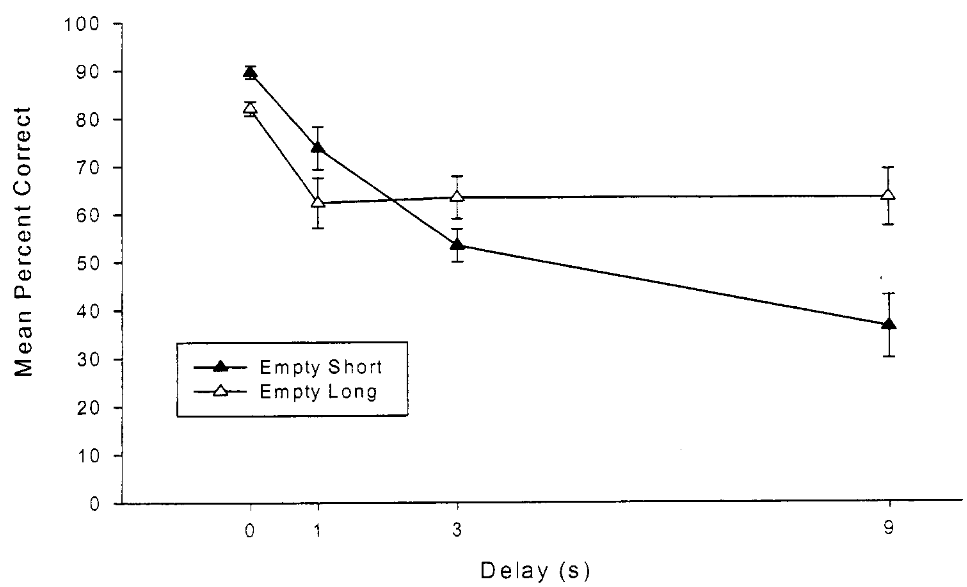
Baseline
Probe: 0-s Delay
Probe: 9-s Delay
Mean Percent Correct
100
90
80
70
60
50
40
30
20
10
0
ES or 5-s
EL or 11-s
FS or 3-s
FL or 9-s
Baseline or Probe Trial Sample Type

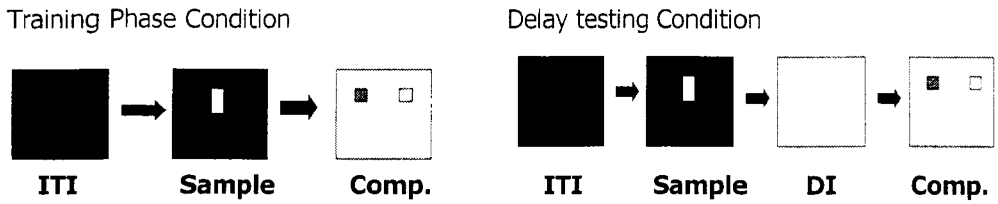

Mean Percent Short Responses
Sample Type
Short
Long
No Sample
Empty
Filled
0
10
20
30
40
50
60
70
80
90
100

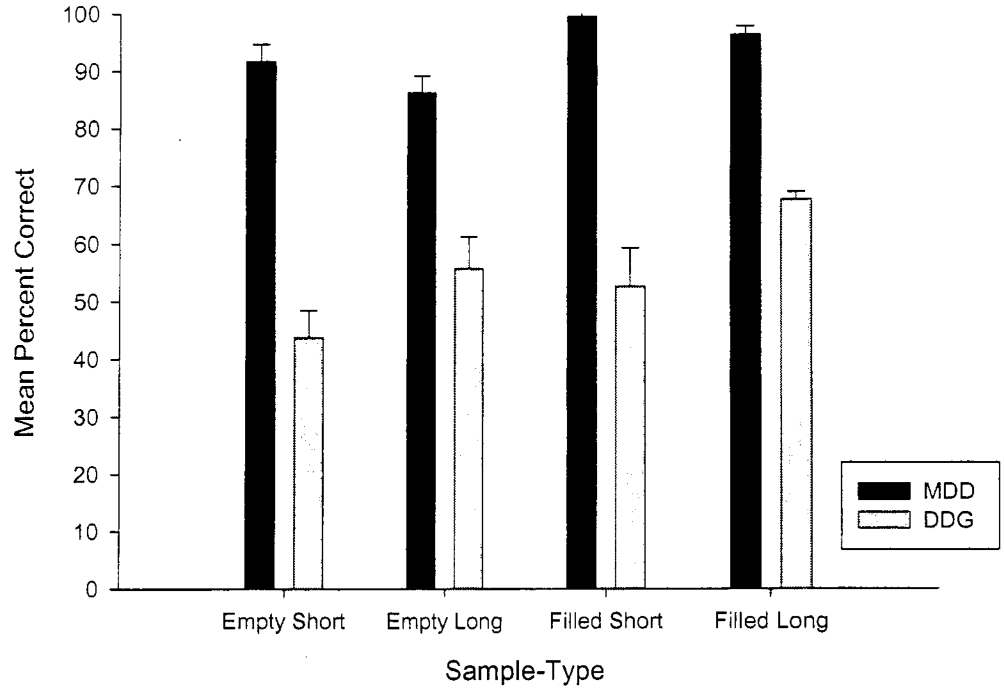

Training Phase Condition
ITI
Sample
Comp.
Delay testing Condition
ITI
Sample
DI
Comp.

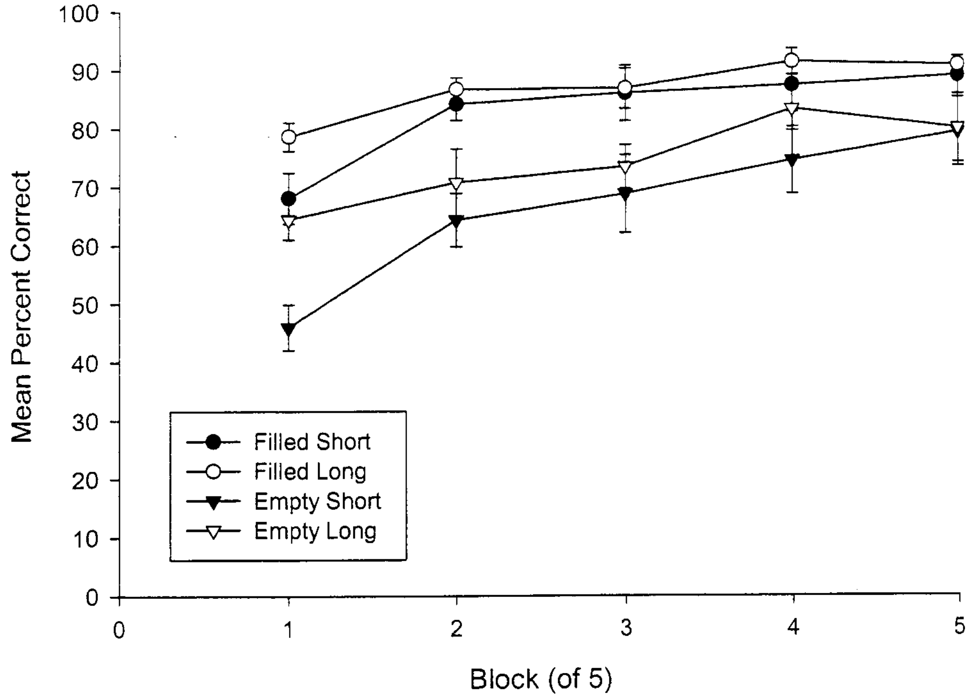

100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Filled Short
Filled Long
Empty Short
Empty Long
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Session


100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Filled Short
Filled Long
0
1
3
9
Delay (s)

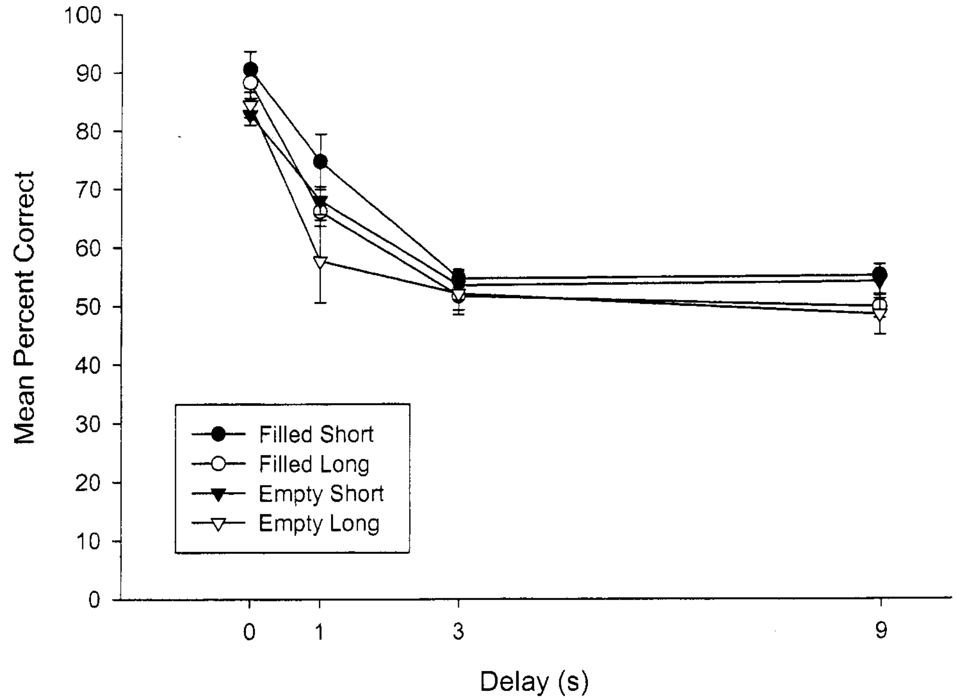

100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Empty Short
Empty Long
0
1
3
9
Delay (s)

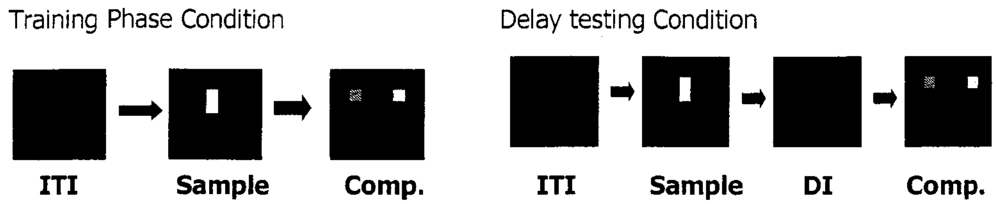

Training Phase Condition
ITI
Sample
Comp.
Delay testing Condition
ITI
Sample
DI
Comp.

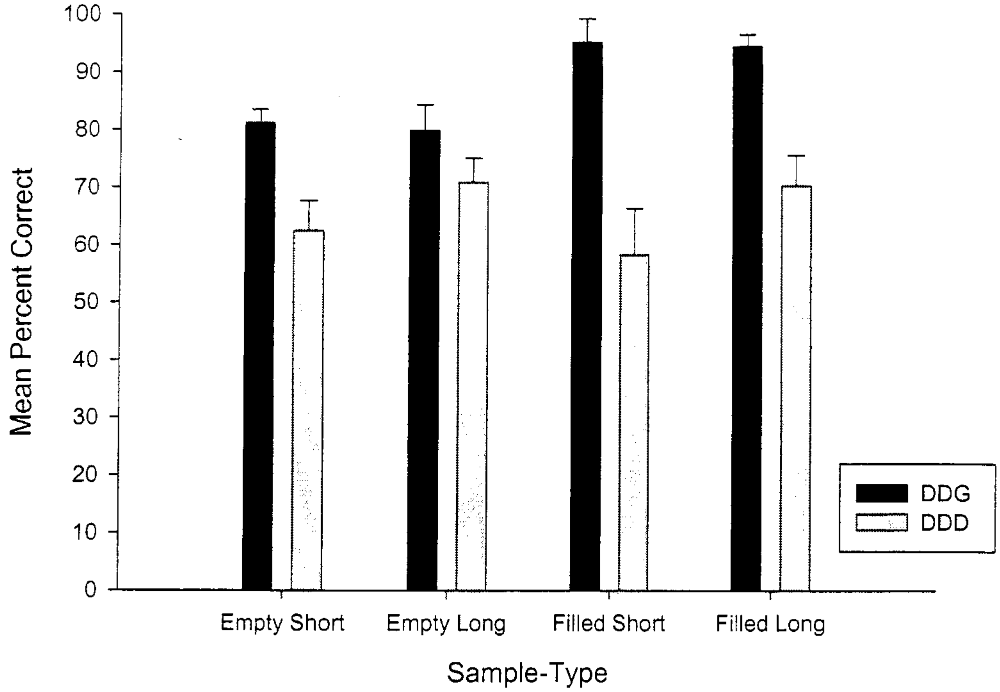

100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
MDD
DDG
Empty Short
Empty Long
Filled Short
Filled Long
Sample-Type



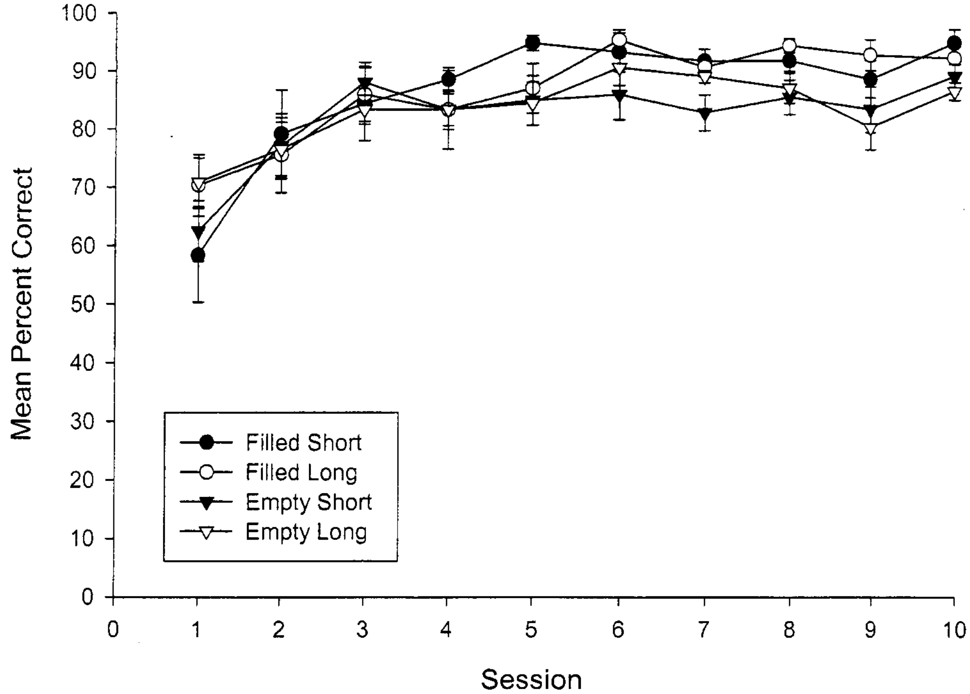
Mean Percent Correct
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
Block (of 5)
Filled Short
Filled Long
Empty Short
Empty Long


100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Filled Short
Filled Long
Empty Short
Empty Long
0
1
3
9
Delay (s)

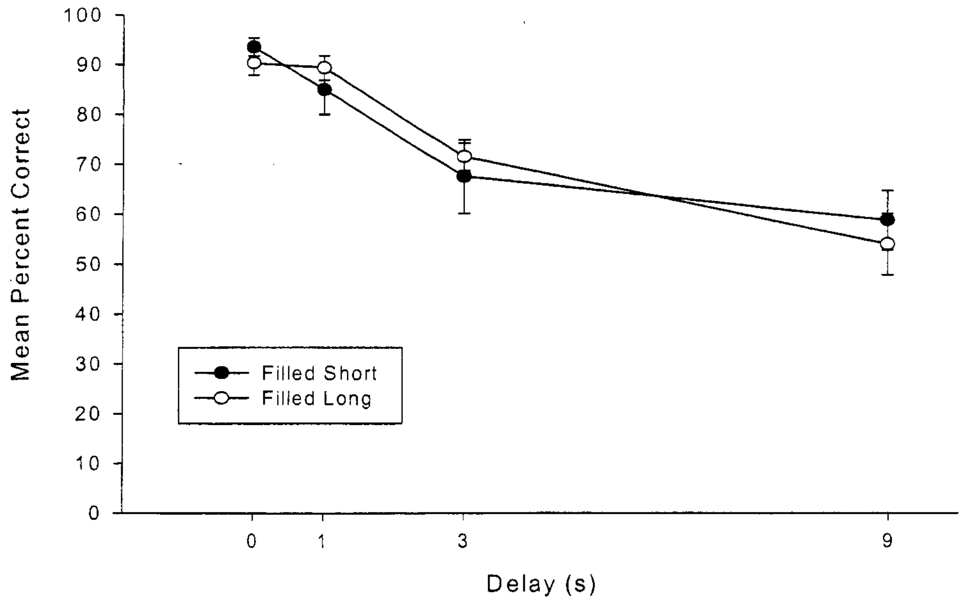

Training Phase Condition
ITI
Sample
Comp.
Delay testing Condition
ITI
Sample
DI
Comp.

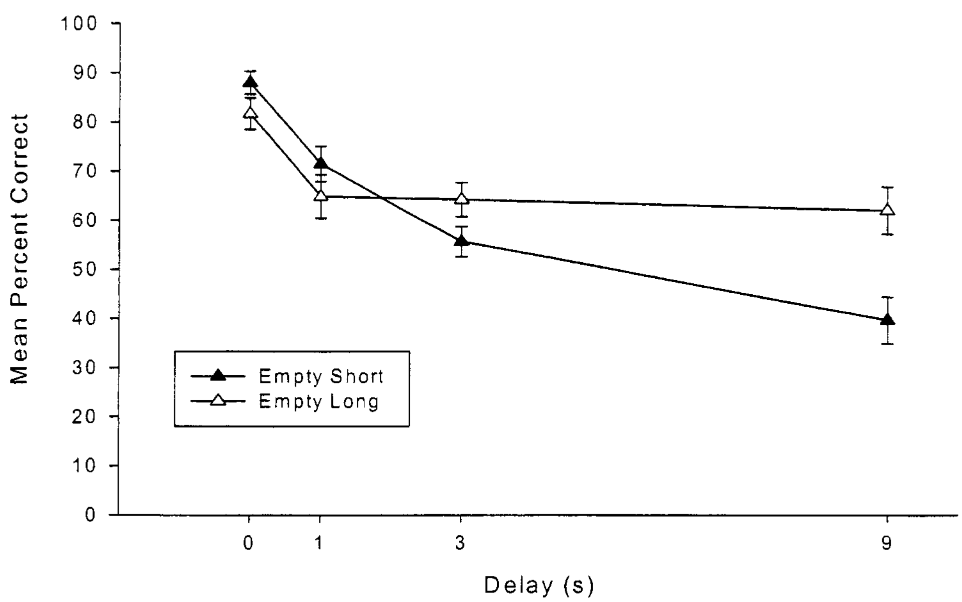


100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Empty Short
Empty Long
Filled Short
Filled Long
Sample-Type
DDG
DDD



100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Filled Short
Filled Long
Empty Short
Empty Long
0
1
2
3
4
5
6
7
8
9
10
Session


100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Filled Short
Filled Long
0
1
3
9
Delay (s)


100
90
80
70
60
50
40
30
20
10
0
Mean Percent Correct
Empty Short
Empty Long
0
1
3
9
Delay (s)