



Published in final edited form as:

*Learn Behav.* 2013 September ; 41(3): 319–332. doi:10.3758/s13420-013-0107-9.

## Auditory proactive interference in monkeys: The role of stimulus set size and intertrial interval

**James Bigelow** and **Amy Poremba**

University of Iowa, Iowa City, IA

### Abstract

We conducted two experiments to examine the influence of stimulus set size (the number of stimuli that are used throughout the session) and intertrial interval (ITI, the elapsed time between trials) in auditory short-term memory in monkeys. We used an auditory delayed matching-to-sample task wherein the animals had to indicate whether two sounds separated by a 5-s retention interval were the same (match trials) or different (non-match trials). In Experiment 1, we randomly assigned a stimulus set size of 2, 4, 8, 16, 32, 64, or 192 (trial unique) for each session of 128 trials. Consistent with previous visual studies, overall accuracy was consistently lower when smaller stimulus set sizes were used. Further analyses revealed that these effects were primarily caused by an increase in incorrect “same” responses on non-match trials. In Experiment 2, we held the stimulus set size constant at four for each session and alternately set the ITI at 5, 10, or 20 s. Overall accuracy improved by increasing the ITI from 5 to 10 s, but the 10 and 20 s conditions were the same. As in Experiment 1, the overall decrease in accuracy during the 5-s condition was caused by a greater number of false “match” responses on non-match trials. Taken together, Experiments 1 and 2 show that auditory short-term memory in monkeys is highly susceptible to PI caused by stimulus repetition. Additional analyses from Experiment 1 suggest that monkeys may make same/different judgments based on a familiarity criterion that is adjusted by error-related feedback.

### Introduction

Proactive interference (PI) occurs when memory processing at a given point in time disrupts memory processing at a future time (Baddeley, 1990; Wright, Urcioli, & Sands, 1986). For example, processing a particular stimulus on a given trial of a memory task can interfere with processing on a subsequent trial that uses the same stimulus. Early studies in human verbal memory led investigators to conclude that PI is a predominant cause of mnemonic failure in laboratory experiments as well as everyday memory usage (Keppel & Underwood, 1962; Underwood, 1957).

PI appears to be pervasive in human memory, having been observed in a variety of visual (Badre & Wagner, 2005; Hartshorne, 2008; Makovski & Jiang, 2008; Mecklinger et al., 2003), motor (Burwitz, 1974; Cothros et al., 2006; Herman & Bailey, 1970), verbal (Feredoes, Tononi, & Postle, 2006; Kane & Engle, 2000; Keppel & Underwood, 1962; Postle et al., 2001), and auditory memory tasks (Ruusuvirta, 2000; Ruusuvirta, Astikainen, & Wikgren, 2002; Ruusuvirta, Wikgren, & Astikainen, 2008; Visscher, Kahana, & Sekuler, 2009). Similarly, the influence of PI has been widely reported in studies of memory processing in animals, including pigeons (Grant, 1975; Hogan, Edwards, & Zentall, 1981; Wright, Katz, & Ma, 2012), rats (De Rosa & Hasselmo, 2000; Dunnett & Martel, 1990;

Gleitman & Jung, 1963; Grant, 1981), monkeys (Mishkin & Delacour, 1975; Overman & Doty, 1980; Wright, 2006, 2007), chimpanzees (Hayes & Thompson, 1953), chickadees (Hampton, Shettleworth, & Westwood, 1998), and dolphins (Herman, 1975; Thompson & Herman, 1981). With the exception of studies in dolphins, nearly all studies of PI in animals have focused on visual short-term memory (see further discussion below).

One of the traditional approaches to assessing PI in nonhuman primates has been varying the degree to which stimuli are reused from trial to trial in a memory task. The typical finding is that memory capabilities improve when new stimuli are used for each trial because confusion arising from stimulus repetitions between trials is reduced. For instance, Hayes and Thompson (1953) found that chimpanzees committed fewer errors on a delayed-response task if new stimuli were used for each trial than if a single pair of stimuli were alternately used as the sample and comparison throughout the experiment. Similarly, Mishkin and Delacour (1975) observed that monkeys require relatively few sessions to learn visual memory tasks if trial-unique stimuli are used, whereas they require significantly more sessions or fail to learn if only two memoranda are repeatedly presented throughout a session. Using a similar visual task, Overman and Doty (1980) found that the maximum retention interval in monkeys increased from under 30 s when two stimuli were repeatedly reused to over 24 h when trial-unique stimuli were used. The positive relationship between performance and the number of stimuli used in a session (i.e., stimulus set size) has been consistently reported in several additional studies of visual memory in monkeys (Mason & Wilson, 1974; Medin, 1980; Sands & Wright, 1980; Worsham, 1975).

An additional, more direct method for examining the influence of PI is to evaluate performance on Trial  $N$  as a function of the stimuli presented on previous trials (intertrial PI). A number of studies have demonstrated that performance on Trial  $N$  can be significantly altered by memory processing on the immediately preceding trial, Trial  $N-1$  (Edhouse & White, 1988; Grant, 1975; Hogan et al., 1981; Makovski & Jiang, 2008; Moise, 1976; Reynolds & Medin, 1981; Thompson & Herman, 1981; Worsham, 1975). These studies are consistent with decreased overall performance associated with smaller stimulus set sizes because the frequency with which the same stimuli are used for both Trial  $N$  and Trial  $N-1$  increases when smaller stimulus sets are used.

Relatively few studies have examined the effect of stimulus repetition beyond non-adjacent trials, i.e., when the same stimuli are used for Trials  $N$  and  $N-2$  or Trials  $N$  and  $N-K$ . One such experiment reported that visuospatial memory in rats was influenced by PI produced by Trial  $N-1$ , but not by more distant, non-adjacent trials (Dunnett & Martel, 1990). By contrast, Hartshorne (2008) reported that human visual memory was susceptible to PI caused by stimulus repetitions across at least four trials (see also Monsell, 1978). Similarly, Wright, Katz, and Ma (2012) recently reported that significant PI in pigeon visual memory was produced by stimulus repetitions separated by as many as 16 trials, which was the longest distance tested. Finally, recent data from our laboratory indicate that stimulus repetitions separated by up to 10 trials can produce significant PI in auditory memory in monkeys (Bigelow & Poremba, in press). These studies highlight the potentially perseverative nature of PI by showing that repeating a small number of stimuli throughout a session can negatively impact performance beyond the immediately subsequent trial.

In addition to reducing the number of intertrial stimulus repetitions, an additional means whereby PI can be reduced is increasing the time that elapses between each trial, or intertrial interval (ITI). Increases in overall accuracy resulting from increasing the ITI have been demonstrated in a variety of memory tasks in humans and nonhuman animals (Cermak, 1970; Cohen, Reid, & Chew, 1994; Herman, 1975; Maki, Moe, & Bierley, 1977; Mason & Wilson, 1974; Roberts, 1980; Roberts & Kraemer, 1982). In some cases, the benefits of

increasing the ITI have been relatively modest. For instance, Jarrard and Moise (1971) reported that monkeys' visual short-term memory accuracy improved by approximately 5 to 10% (depending on the retention interval) after increasing the ITI from 5 to 15 s, but no significant advantage was gained by further extending the ITI to 30 or 60 s. In other cases, increasing the ITI has led to more substantial benefits, to the extent that the influence of intertrial PI has been reduced to zero. In Dunnett and Martel's (1990) study of rat visuospatial memory, PI effects from Trial  $N-1$  were eliminated by increasing the ITI from 5 s to 15 s. Similarly, two studies of pigeon visual memory reported significant intertrial PI effects when the ITI was 2 s, but not when it was 10 s or greater (Grant, 1975; Hogan et al., 1981). These results suggest that increasing the time between trials allows greater decay of irrelevant memory traces, which might otherwise compete with memory demands of the current trial.

With few exceptions, studies of PI in animals including monkeys have been concerned with the visual sensory modality. Because auditory perception and memory are crucial for key aspects of nonhuman primate ethology such as predator evasion and conspecific communication (e.g., Ghazanfar & Hauser, 2001), the relatively sparse auditory memory literature, including PI, constitutes a significant deficit in scientific understanding. One likely reason for the lack of experimental data in the auditory modality is that, unlike visual memory tasks, monkeys require extensive training to learn auditory memory tasks (Cohen, Russ, & Gifford, 2005; Colombo & D'Amato, 1986; D'Amato & Colombo, 1985; Fritz, Mishkin, & Saunders, 2005; Kojima, 1985; Scott, Mishkin, & Yin, 2012). For instance, Fritz et al. (2005) reported that monkeys required ~15,000 trials to learn an auditory memory task, whereas Mishkin and Delacour (1975) reported that only ~500 trials were needed to learn a comparable visual memory task. A related finding is that the maximum reported retention interval for auditory memory in monkeys ranges from 16 to 35 s (Colombo & D'Amato, 1986; Fritz et al., 2005; Kojima, 1985), whereas for visual memory it ranges from minutes to hours or more (Murray & Mishkin, 1998; Overman & Doty, 1980). Monkeys are also capable of retaining tactile information for at least several minutes (Bauer & Steele, 1985; Buffalo et al., 1999; Suzuki et al., 1993), suggesting that auditory memory tasks may be uniquely challenging.

Because an investigation of the basic parameters of auditory PI in monkeys is lacking, we conducted two experiments to address the effects of stimulus set size (Experiment 1) and the duration of the ITI (Experiment 2) in an auditory short-term memory task in monkeys. The experiments were designed in such a way that the results would be roughly comparable to previous studies of memory in monkeys using the visual modality. As with visual PI, we hypothesized that the influence of auditory PI would diminish as a function of the stimulus set size and the duration of the ITI.

## Experiment 1

### Method

**Subjects**—Two adult male macaque monkeys (*Macaca mulatta*) served as subjects for this experiment (Monkeys FR and SE). The monkeys had been trained to perform the auditory short-term memory task prior to the experiment. The animals were housed in individual cages with ad libitum access to water and controlled feeding schedules, under a 12:12-h light:dark cycle. Experimental sessions were conducted 3-5 days per week. The majority of food was given after the experimental session each day (Harlan monkey diet plus fresh fruit, vegetables, and treats) and each animal was maintained above 85% of their weight during the use of a controlled feeding schedule. All procedures were carried out with approval from The Institutional Animal Care and Use Committee at the University of Iowa.

**Apparatus**—Subjects were placed in a sound-attenuated chamber for the duration of each experimental session. The animal was held in a custom-made primate chair that allowed free arm movement for behavioral responses. Acoustic stimuli were delivered through a single speaker located 15 cm in front of the primate chair at eye level. Behavioral responses were made via a single acrylic button positioned 3 cm below the speaker. Small food rewards were dispensed from a pellet dispenser (Med Associates, Georgia, VT) through a copper tube into a dish located 3 cm below the response button. A dimmed 40-watt house light provided illumination throughout the duration of the experiment, and a second light provided additional illumination during the ITI. Custom-designed software (LabView, National Instruments, Dallas, TX) controlled and recorded the stimulus presentations and other task events.

### Procedure

**Task:** The short-term memory task used for this experiment was a variation of the delayed matching-to-sample (DMS) task, which is suitable for use with auditory stimuli. The typical DMS task begins with the presentation of a sample stimulus, which is followed by a retention interval, after which subjects are rewarded for identifying the sample from two test stimuli. In the same/different variation of the DMS task (Wright, 2006), a single test stimulus is presented following the retention interval and the subject must indicate whether it is identical to (match trials) or different from the sample (non-match trials). The traditional two-choice and same/different versions of the DMS task produce very similar outcomes in visual short-term memory performance in monkeys (D'Amato & Worsham, 1974). Thus visual memory experiments using the traditional DMS task can be reasonably compared to the current experiment as well as previous studies, which use the auditory same/different DMS task (Colombo & D'Amato, 1986; Kojima, 1985; Wright, Shyan, & Jitsumori, 1990).

The task for Experiment 1 used a fixed retention interval, or interstimulus interval (ISI), of 5 s, and a variable ITI averaging 10 s (range: 8-12 s). Each session consisted of a total of 128 trials with an equal number of match and non-match trials presented in pseudorandom order. Following the presentation of the test stimulus, the response button was illuminated for 1 s to indicate that a response could be made. If a button press was made outside of the 1-s response window, the current trial was aborted and replaced with a new trial. For match trials, correct responses were defined by the presence of a button-press (“go” response) following the test stimulus, whereas for non-match trials, correct responses were defined by the absence of a button press (“no-go” response). The task used an asymmetric reinforcement contingency in which correct “go” responses on match trials were rewarded with a small food pellet and incorrect button presses (false alarms) on non-match trials were occasionally punished by a brief, mild air puff presented indirectly from a distance of approximately 15 cm from the animal (approximately 1/10 of incorrect non-match trials were punished on a variable schedule). Similar asymmetric reinforcement contingencies have been used in previous studies of auditory short-term memory in monkeys because they facilitate learning the match vs. non-match rule (Colombo & D'Amato, 1986; Kojima, 1985; Stepien & Cordeau, 1960).

**Stimulus sets:** Experiment 1 consisted of a systematic manipulation of the number of sounds that were recycled as the sample and test stimuli throughout each 128-trial session. For a given experimental session, one of the following stimulus set sizes was randomly selected: 2, 4, 8, 16, 32, 64, or trial unique. For the trial-unique condition, 192 stimuli were used (64 sounds used for 64 match trials, 128 sounds used for 64 non-match trials). After a stimulus set size was used for a session, it was not used again until the remaining six set sizes had been used. Each animal completed a total of 20 sessions with each stimulus set size.

The stimuli used for each session were randomly selected from a collection of 192 sounds consisting of 32 exemplars from each of the following six sound classes: conspecific monkey vocalizations, human vocalizations, animal vocalizations, natural and environmental sounds, music clips, and synthetic/abstract sounds. The monkey vocalizations were recorded at a natural monkey reserve in South Carolina, USA (by author A. P.), and included a variety of coos, grunts, screams, shrill barks, and harmonic arches. The human vocalizations consisted of a variety speech and non-speech vocalizations from a variety of speakers, including members of each gender. Animal vocalizations came from a wide variety of birds and mammals other than monkeys and humans. Natural and environmental sounds included natural phenomena, such as thunder and breaking tree branches, as well as sounds that the animals might have been exposed to in the laboratory, such as a door closing or a broom falling on the floor. Music clips comprised multi-note sequences extracted from variety of sources such as solo instrument performances, popular music recordings, and TV commercials. Synthetic and abstract sounds were artificial sounds generated by electronic synthesizers or downloaded from abstract sound categories (e.g., “science fiction”) of a commercially-available sound effects database ([www.soundsnap.com](http://www.soundsnap.com)). All sounds were trimmed to 500 ms, volume normalized, and presented at  $75 \pm 5$  decibels. Within a session, each sound had an equal chance to be presented on a given trial as the sample and/or test stimulus, depending on whether it was a match or non-match trial, with the constraint that each sound was presented an equal number of times throughout the session.

We found no evidence that the effects of PI differed among the sound types used in our study, thus, the results presented below are collapsed across sounds. It should be noted, however, that our experiment was not specifically designed to test for differences in PI effects among sound types. For instance, by randomly selecting the sounds to be used for each session, there were an unequal number of sessions using a given sound type. Thus, the question of whether PI interacts with sound type remains to be addressed by future studies.

**Analysis:** Although PI literature has traditionally focused on accuracy (percent correct) as the dependent measure, a relatively small number of publications have reported modulation of response latency by PI (Hendriks, 1986; Monsell, 1978; Wixted & Rohrer, 1993). For this reason, both accuracy and response latency were evaluated as dependent variables in our study.

The animals occasionally stopped responding before the final trial of the experiment session. For sessions in which the subject did not make a single response during the last 20 trials, we considered the final response as the end of the session. The remaining trials were rejected from accuracy and response latency analyses to ensure that any observed effects could be attributed to mnemonic rather than attentional or motivational factors. The statistical test used to evaluate all effects of PI was repeated-measures ANOVA with an alpha level of .05, using session means as individual data points. There were small but significant differences in mean overall accuracy between the monkeys in both experiments. In Experiment 1, Monkey FR averaged 79.7% correct, and Monkey SE averaged 83.5% ( $F[1,19] = 7.57, p < .05, \eta^2_p = .29$ ). In Experiment 2, Monkey FR averaged 68.0% correct, and Monkey SE averaged 74.3% ( $F[1,19] = 27.13, p < .05, \eta^2_p = .59$ ). However, accuracy for both animals was affected by PI in similar ways: we replicated each of the analyses below with *subject* (Monkey FR, Monkey SE) as an additional factor and found no significant interactions. Thus, the results below are given as the combined average of both animals. Any data points that were missing for a given analysis were substituted with the series mean (Bigelow & Poremba, in press; Roth, 1994). For example, in the unusual case that no incorrect button presses were made on non-match trials for a given session, and therefore no incorrect response latency data were available, the missing data point was estimated for the repeated

ANOVA as the mean incorrect non-match response latency of the remaining sessions within the same stimulus set size condition in which such responses occurred.

## Results and discussion

An unanticipated but interesting initial observation in our data set was that the animals were more likely to quit responding before the final trial of experimental sessions using the smallest stimulus sets. Thus, the mean number of incomplete trials per session for the two-stimulus condition was 25.7, whereas for trial-unique sessions it was 12.7. (Fig. 1). A repeated-measures ANOVA confirmed that the effect of stimulus set size on early quitting was significant:  $F(6, 234) = 3.82, p < .05, \eta^2_p = .09$ . Post-hoc tests (Fisher's LSD, alpha level:  $p < .05$ ) indicated that the number of incomplete trials was significantly greater for the two-stimulus condition than the 8, 64, or trial-unique conditions. This finding perhaps reflects the tendency of the animals to quit when the task becomes particularly difficult (see below).

As expected, overall accuracy was significantly affected by stimulus set size:  $F(6, 234) = 37.38, p < .05, \eta^2_p = .49$ . There was also a significant linear trend indicating that accuracy increased as a function of stimulus set size:  $F(1, 39) = 141.82, p < .05, \eta^2_p = .78$ . Accuracy was poorest (72%) during sessions for which only two stimuli were used and increased for the larger stimulus set sizes, reaching its maximum (88%) during sessions with trial-unique stimuli (Fig. 2). By comparison, Mishkin and Delacour (1975) reported that visual DMS accuracy fell from 90% when trial-unique stimuli were used to 65% when only two stimuli were used. The differences in accuracy between sessions using only two stimuli versus trial-unique stimuli for the visual (25%) and auditory (16%) DMS tasks suggest that the influence of PI is roughly comparable, if perhaps somewhat less severe for auditory short-term memory. However, because this comparison is limited by differences in subjects and experimental design, future research is needed to substantiate any difference in PI between sensory modalities.

In order to examine the influence of stimulus set size on response latency, it was first necessary to separate the data by trial type, since button presses on match trials reflected correct responses whereas button presses on non-match trials reflected incorrect responses. A repeated-measures ANOVA with trial type and stimulus set size as factors resulted in significant main effects (trial type:  $F(1, 39) = 109.19, p < .05, \eta^2_p = .74$ ; stimulus set size:  $F(6, 234) = 5.38, p < .05, \eta^2_p = .12$ ) as well as a significant interaction ( $F(6, 234) = 5.73, p < .05, \eta^2_p = .13$ ). We also observed significant trends in the non-match and match response latency data. For incorrect responses on non-match trials, there was a significant linear trend indicating that the latency of incorrect responses increased as a function of stimulus set size ( $F(1, 39) = 29.02, p < .05, \eta^2_p = .43$ ). For correct responses on match trials, there was a significant quadratic trend ( $F(1, 39) = 7.65, p < .05, \eta^2_p = .16$ ). This outcome suggests that correct responses are made more slowly under high compared to moderate PI conditions. However, when PI is very low or absent, response time is greater than under moderate amount of PI is present. Correct match responses were faster on average (562 ms) than incorrect responses (701 ms). Post-hoc analyses (Fisher's LSD, alpha level:  $p < .05$ ) revealed that non-match errors were slower for the trial-unique condition than all other stimulus set sizes (Fig. 3). By contrast, correct match responses were slower for the two-stimulus set than several of the larger stimulus sets (4, 8, and 32). These results suggest that PI negatively impacts performance by increasing the speed with which non-match errors are made and by slowing correct match responses.

Evaluating accuracy for match and non-match trials separately revealed that the majority of the deficit in overall accuracy associated with the smaller set sizes (Fig. 2) is due to increased errors or false alarms on non-match trials (Fig. 4). A two-factor ANOVA testing

these differences resulted in significant main effects of trial type ( $F[1, 234] = 19.13, p < .05, \eta^2_p = .33$ ) and set size ( $F[6, 234] = 35.65, p < .05, \eta^2_p = .48$ ), as well as a significant interaction of these factors ( $F[6, 234] = 4.91, p < .05, \eta^2_p = .11$ ). Post-hoc analyses (Fisher's LSD, alpha level:  $p < .05$ ) indicated that non-match accuracy declined steadily from 89% for trial-unique sessions to 67% for sessions repeatedly reusing the same two stimuli. This outcome is consistent with the view that increasing the frequency with which stimuli are presented from trial to trial leads subjects to commit more false positive errors because they might have heard a sound on a recent trial that “matches” the test stimulus on the current trial.

In contrast to the effects of PI on non-match trials, accuracy for match trials was only significantly reduced for the smallest set sizes (set size two: 77%; set size four: 84%), with no differences among set sizes eight through trial unique (range: 86-87%). One possible interpretation for why match accuracy might decrease under high PI conditions involves the concept of feedback-related changes in the criterion of familiarity with which same/different judgments are made as discussed by Wright (2006, 2007). According to Wright, subjects will make a “match” response only if the degree of familiarity evoked by a test stimulus exceeds the animal's familiarity criterion; otherwise a non-match response will be made. When trial-unique stimuli are used, a test stimulus will only be familiar if it matches the sample stimulus from the same trial. In this context, adopting any criterion level of familiarity will suffice in making accurate same/different judgments. However, when stimuli are recycled from trial to trial, a test stimulus might evoke a certain level of familiarity by virtue of having been presented on a recent (and now irrelevant) trial, and not because it matches the sample stimulus of the current trial. Under these conditions, adopting a relatively lax criterion of familiarity will result in a high rate of false matches. Thus, one strategy for coping with a high degree of PI is to rely on a more rigid familiarity criterion such that only the most familiar test stimuli are accepted as matches. One of the predicted consequences of increasing the familiarity criterion is that, along with a decrease in false matches, the frequency with which true matches are rejected will also increase. Our data fit well with this prediction, inasmuch as more false negative errors were observed for the sessions with the greatest amount of PI.

To provide more direct evidence for shift in familiarity criterion resulting from PI, we investigated whether the rate of false matches (non-match errors) and false rejections (match errors) changed as the experimental session progressed for each stimulus set size. Accuracy data were separated by trial type and averaged for the first, second, third, and fourth quarters of the session (i.e., successive blocks of 32 trials). A three-way repeated ANOVA produced a significant interaction among trial type, stimulus set size, and trial block ( $F[18, 702] = 3.12, p < .05, \eta^2_p = .07$ ). As seen in Fig. 5, the higher PI conditions led to a steady decrease in false matches throughout the session as well as a corresponding increase in false rejections. The magnitude of these reciprocal trends diminished as a function of stimulus set size, to the extent that there was no significant *trial type*  $\times$  *trial block* interaction for the trial-unique condition. The absence of a progressive change in accuracy during trial-unique sessions is helpful in interpreting the *trial type*  $\times$  *trial block* interactions observed for the remaining conditions because it argues against attentional or motivational explanations for the changes in error rate associated with the smaller set sizes. For example, it is unlikely that the observed decrease in match responses for the smaller set sizes reflects reward satiation because these changes were not observed when new stimuli were used for each trial. Rather, it seems likely that the shift toward fewer “match” responses (for both trial types) reflects the number of negative outcomes associated with false “match” responses. This interpretation lends support to Wright's (2006, 2007) suggestion that PI will gradually produce a proportional increase in the familiarity criterion that forms the basis for the match vs. non-match decision.

Curiously, the main effect of trial block and the interaction between trial block and stimulus set size were both non-significant ( $F[3, 117] = 1.36, p > .05, \eta^2_p = .03$ , and  $F[18, 702] = 1.40, p > .05, \eta^2_p = .04$ , respectively). This indicates that, although the rate of false match responses decreases throughout the session, the rate of true match responses decreases at an approximately equal rate. Thus, the benefits of reducing non-match errors are offset completely by the costs of increasing match errors, leading to zero net improvement during the session. We also examined overall accuracy for each stimulus set over the course of four successive blocks of five experimental sessions, but found no evidence of significant improvement: a repeated-measures ANOVA returned neither a main effect of the block of five experimental sessions ( $F[3, 27] = 0.87, p > .05, \eta^2_p = .09$ ) nor an interaction between stimulus set size and experiment block ( $F[18, 162] = 1.22, p > .05, \eta^2_p = .12$ ). In several previous visual short-term memory experiments, pigeons and monkeys have shown improvements in overall accuracy despite high PI conditions that gradually emerged with consistent, extended experience with small stimulus sets (D'Amato, 1973; Grant, 1975, 1976; Wright, 2007). It is possible that, since sessions with small and large stimulus sets were interleaved in our study, the animals lacked sufficient, consistent experience with high PI conditions to result in an adaptive adjustment of their familiarity criterion for match responses. A design presenting a block of multiple, consecutive sessions using a given stimulus set size, followed by a subsequent block using a different set size could be useful in determining if this is the case. Alternatively, it is possible that the lack of improvement over the course of the experiment is related to the difficulty monkeys have in establishing enduring memories in the auditory modality.

Two additional predictions are made by Wright's (2006, 2007) suggestion that animals' same/different decisions are under the influence of a familiarity criterion, which can be modulated by error-related feedback. The first is that intertrial PI should have a graded effect on accuracy, such that stimulus repetitions separated by a large number of trials should have a smaller effect than stimulus repetitions separated by only a few trials or between adjacent trials. For example, fewer false match responses should occur if the test stimulus on Trial  $N$  was most recently presented on Trial  $N-20$  than if it was presented on Trial  $N-1$  or  $N-2$ . The second prediction is that the more rigid familiarity criterion that results from frequent exposure to PI should result in fewer non-match errors on trials for which the test stimulus has been presented on a relatively recent trial. Thus, even though training with smaller stimulus sets should yield relatively poor overall non-match accuracy, the frequency with which non-matching test stimuli are erroneously accepted as matches by virtue of having appeared on a recent trial such as  $N-1$  or  $N-2$  should decrease (see figure 9.5 in Wright, 2006 for hypothetical relationship between susceptibility to intertrial PI and familiarity criterion).

To test these predictions, we evaluated non-match accuracy on trials for which the test stimulus on Trial  $N$  had most recently been presented on Trials  $N-1$ ,  $N-2$ , or  $N-3$ . We only evaluated the effects of PI from trials  $N-1$  through  $N-3$  because, for the two-stimulus set size, the number of non-match trials for which the test stimulus most recently occurred on Trial  $N-4$  or greater was insubstantial (0.1% of total trials). Similarly, we did not include sessions using stimulus set sizes of 16 or greater in this analysis because there were too few non-match trials for which the test stimulus had been presented on Trials  $N-1$  through  $N-3$ . Repeated-measures ANOVA indicated significant main effects of stimulus set size ( $F[2, 78] = 9.42, p < .05, \eta^2_p = .20$ ) and PI location ( $F[2, 78] = 70.72, p < .05, \eta^2_p = .65$ ); the interaction was not significant ( $F[4, 156] = 0.97, p > .05, \eta^2_p = .02$ ). Consistent with similar analyses in several previous studies (Hartshorne, 2008; Wright et al., 2012), non-match accuracy increased steadily according to the distance of the most recent test-stimulus repetition (Fig. 6). Further, in confirmation of Wright's (2006, 2007) expectations, false match responses on trials with PI from trials  $N-1$ ,  $N-2$ , and  $N-3$  were less likely for sessions



that used only two stimuli compared to sessions that used four or eight stimuli. It should be noted that overall non-match accuracy was better for the larger set sizes (Fig. 4) because of the larger number of trials in those sessions that did not have PI from Trials  $N-1$ ,  $N-2$ , or  $N-3$ . Nevertheless, for the two-stimulus set size condition, accuracy on Trial  $N$  when the test stimulus hadn't been presented since Trial  $N-3$  reached a similar level (91%) to that observed for trial-unique sessions (89%).

To summarize the results of Experiment 1, PI produced by reusing a relatively small number of sounds from trial to trial decreases overall accuracy primarily by producing more false alarms on non-match trials. These non-match errors tend to be committed faster under high PI conditions than when trial-unique stimuli are used. The smallest set sizes, and therefore those that produced the most pervasive PI, also reduced the number of correct "same" decisions on match trials, and increased the amount of time before these decisions were made. These outcomes are consistent with Wright's (2006, 2007) prediction that subjects will adopt a more stringent criterion of familiarity for "same" judgments when PI becomes highly saturated. Wright's notion of a familiarity criterion is further supported by several additional results from our study. First, as the experimental session progressed, subjects committed fewer false alarms on non-match trials, but fewer correct "same" responses on match trials. This effect was roughly proportional to the degree of PI caused by stimulus repetitions. Second, as with several previous studies (Bigelow & Poremba, in press; Hartshorne, 2008; Wright et al., 2012), PI originating from progressively more distant trials produced a graded effect on non-match accuracy. Finally, a greater degree of PI throughout the session (resulting from a smaller stimulus set size) resulted in fewer non-match errors on trials with PI originating from one of the three most recent trials.

## Experiment 2

### Method

**Subjects**—The subjects were the same as those used in Experiment 1.

**Apparatus**—The apparatus was the same one used in Experiment 1.

### Procedure

**Task:** The task was similar to the one used in Experiment 1 except that the stimulus set size was held constant while the ITI was manipulated between sessions. A set of four stimuli was used to produce an intermediate amount of PI. Stimuli for each session were randomly drawn from the same stimulus population that was used in Experiment 1.

**Intertrial Intervals:** Experiment 2 consisted of a parametric manipulation of the ITI duration. A fixed ITI of 5, 10, or 20 s was randomly selected for each session. Note that these values correspond to ISI:ITI ratios of 1:1, 1:2, and 1:4, respectively (see D'Amato, 1973). After a session was completed using one of the ITI values, it was not used again until the monkeys had completed sessions using the remaining two ITIs. As in Experiment 1, each animal completed a total of 20 sessions using each ITI.

**Analysis:** Analyses were similar to Experiment 1 except that the independent variable was the ITI.

### Results and discussion

Repeated-measures ANOVA indicated that ITI had a significant effect on overall accuracy ( $F[2, 78] = 7.41, p < .05, \eta^2_p = .16$ ). As revealed by post-hoc tests (Fisher's LSD, alpha level:  $p < .05$ ), the 5-s ITI resulted in lower overall accuracy (67%), whereas the 10-s and

20-s ITI conditions were equal (73%; Fig. 7). Evaluating the differential effect of ITI on match and non-match accuracy again revealed that the decrease in overall accuracy was caused primarily by an increase in non-match errors at the shortest ITI (Fig. 8). Repeated-measures ANOVA revealed significant main effects of ITI ( $F[2, 78] = 6.50, p < .05, \eta^2_p = .14$ ) and trial type ( $F[1, 39] = 7.44, p < .05, \eta^2_p = .16$ ), but there was no significant interaction between the two factors ( $F[2, 78] = 0.60, p > .05, \eta^2_p = .02$ ). Post-hoc analyses indicated that there were no significant differences among the three ITI conditions for match accuracy. As with overall accuracy, non-match accuracy was significantly reduced for the 5-s ITI condition, whereas the 10-s and 20-s ITI conditions did not differ from each other.

As in Experiment 1, we examined response latency by separating the data by trial type. Repeated-measures ANOVA revealed significant effects of trial type ( $F[1, 39] = 95.04, p < .05, \eta^2_p = .71$ ) and ITI ( $F[2, 78] = 6.45, p < .05, \eta^2_p = .14$ ), as well as a significant interaction ( $F[2, 78] = 4.22, p < .05, \eta^2_p = .10$ ). As in Experiment 1, correct match responses were faster (582 ms) than erroneous responses on non-match trials (690 ms). Unlike Experiment 1, post-hoc tests revealed no differences for non-match errors as a function of ITI (Fig. 9). For correct match responses, slower response latency was observed for the 5-s ITI condition, but no differences were observed between sessions using a 10-s and 20-s ITI. The latter outcome is consistent with results from Experiment 1 in suggesting that the decision time for correct match trials increases when PI becomes saturated.

Although non-match accuracy increased significantly by extending the ITI from 5 s (62%) to 10 s (69%) or 20 s (70%), it was still well below non-match accuracy achieved in trial-unique sessions in Experiment 1 (89%). This outcome suggested that there might still be substantial PI caused by reusing stimuli even when trials have been separated by as much as 20 s. Thus, as in Experiment 1, we directly investigated the influence of intertrial PI by evaluating non-match accuracy on Trial  $N$  as a function of the trial on which the test stimulus was most recently presented. Since a four-stimulus set was used in all conditions, it was possible to evaluate the influence of PI originating from Trials  $N-1$  through  $N-5$  (the most recent PI stimulus was found on Trial  $N-6$  or greater for only 9.0% of the non-match trials). As seen in Fig. 10, accuracy improved for each ITI condition as the number of trials between stimulus repetitions increased. A repeated-measures ANOVA resulted in a main effect of ITI that was of borderline significance ( $F[2, 78] = 3.10, p = .05, \eta^2_p = .07$ ) and a main effect of PI location ( $F[4, 156] = 34.46, p < .05, \eta^2_p = .47$ ), but the interaction of these factors was not significant ( $F[8, 312] = 0.84, p > .05, \eta^2_p = .02$ ). These results indicate that, unlike several experiments in pigeons and rats, which eliminated visual PI by increasing the ITI to 15 or 20 s (Dunnett & Martel, 1990; Grant, 1975; Hogan et al., 1981, but see Wright et al., 2002), auditory PI in monkeys can influence subsequent trials even when separated by up to 20 s.

Post-hoc tests indicated that, for each condition, non-match accuracy improved when the test stimulus had most recently occurred on Trials  $N-2$  through  $N-5$  compared to when it had occurred on the previous trial,  $N-1$ . For sessions using the 5-s ITI, accuracy on trials for which the test stimulus had most recently occurred on Trial  $N-5$  was significantly higher than when it had occurred on either Trial  $N-1$  or  $N-2$ . Similarly, for the 10-s ITI condition, accuracy was greater when the test stimulus hadn't been presented since Trial  $N-4$  or  $N-5$  than when it occurred on Trial  $N-1$  or  $N-2$ . However, for the 20-s ITI condition, there was no statistically significant difference in accuracy when the test stimulus had most recently occurred on Trial  $N-2$  through  $N-5$ . This suggests that increasing the ITI from 10 to 20 s may slightly reduce the extent of intertrial PI, although this difference was insufficient to result in any significant change in overall accuracy (Fig. 7) or averaged non-match accuracy (Fig. 8).

In a prior study, we examined the extent of intertrial PI in monkeys performing an auditory DMS task with a stimulus set size of eight and a retention interval of 5 s (Bigelow & Poremba, in press). A variable ITI averaging approximately 10 s was used, making it most similar to the 10-s ITI condition in Experiment 2. In that study, we found that the monkeys were more likely to commit errors on non-match trials when the test stimulus was repeated after as many as 10 trials. By contrast, in the current study (Experiment 2, 10-s ITI), the effects of PI appear to reach asymptote by about Trial  $N-4$ . This difference could plausibly result from an increase in the familiarity criterion for “match” responses resulting from the relatively high PI conditions produced by the smaller stimulus set size (four compared to eight). Specifically, if the monkeys were using a relatively lax criterion of familiarity for “match” decisions in the eight-stimulus set condition, they would be more likely to erroneously accept a non-matching test stimulus as a “match” if it had occurred on Trial  $N-4$  or  $N-5$ , or even earlier.

To summarize the results of Experiment 2, we observed that overall accuracy increased, if only to a small extent, by extending the ITI from 5 s to 10 s, but no additional improvement was gained by increasing the ITI to 20 s. These outcomes are comparable to a study of visual short-term memory in monkeys reported by Jarrard and Moise (1971), in which a small increase in accuracy was produced by increasing the ITI from 5 to 15 s, but not by further extending the ITI to 30 or 60 s. We further found that the effects of intertrial PI are substantial even when trials are separated by 20 s. Although these results are generally consistent with previous literature, the decay of PI from previous trials produced by increasing the ITI was less than we had initially expected. Given the significant intertrial effects of PI in each condition, it is likely that future studies will require an ITI substantially longer than 20 s in order to completely eliminate auditory PI in monkeys.

## General Discussion

Like other forms of memory, our experiments show that auditory short-term memory in monkeys is susceptible to PI caused by reusing the same stimuli for multiple trials within an experimental session. In Experiment 1, PI was most severe when only a very small number of stimuli (e.g., two or four) were used throughout the session. The effects of PI diminished steadily as the stimulus set size increased, and maximum accuracy was observed when new stimuli were used for each trial (Fig. 2). In Experiment 2, PI was modestly attenuated by increasing the ITI from 5 s to 10 s, but no additional advantage was gained by increasing the ITI to 20 s (Figs. 7-10).

In both experiments, PI reduced overall accuracy primarily by increasing the number of erroneous “match” responses on non-match trials (Figs. 4 and 8). This finding is similar to what has been reported in experiments using list memory tasks in which a small number of stimuli are reused from trial to trial (Wright, 1998). In these tasks, a list of several sample stimuli is presented, each separated by a brief ISI. The last item in the list is followed by a retention interval, after which a single probe stimulus is presented, and the subject must indicate whether the probe had been presented in the list (“same”) or not (“different”). In both visual and auditory list memory tasks in monkeys, using a small stimulus set throughout the session, and thereby increasing item repetition among trials, increases the rate of erroneous “same” responses on trials in which the probe was different from the items presented in the list (Sands & Wright, 1980; Wright, 1999). This is likely because, although the probe did not match one of the list items from the current trial, it may have matched a list item from a recent trial. In other words, the error being committed by the subjects seems to be forgetting whether the probe occurred on the current trial, or some previous, now irrelevant trial (Wright, 2006). In both the current study as well as previous DMS and list memory experiments, errors on non-match trials are most likely if the probe had been

presented on the immediately previous trial, and become less likely as the number of trials since the probe had been presented increases (Bigelow & Poremba, in press; Hartshorne, 2008; Wright et al., 1986, 2012).

Several results from Experiment 1 provide support for Wright's (2006, 2007) view that a criterion threshold of familiarity is used to make the same vs. different choice, and that this familiarity criterion can be increased as a result of committing frequent false match errors. Although PI consistently increased the false alarm rate on non-match trials, the extremely saturated PI conditions also reduced the number of correct "same" judgments on match trials (Fig. 4). Additional analyses revealed a steady decrease in "match" responses as the session progressed for both match and non-match trials that was most pronounced for the highest PI conditions (Fig. 5). Further, although overall non-match accuracy was lowest for the two-stimulus set, non-match accuracy on trials with recent PI (originating from Trials  $N-1$ ,  $N-2$ , or  $N-3$ ) was higher for the two-stimulus set than the four- or eight-stimulus sets (Fig. 6). Each of these results can be seen as a consequence of increasing the familiarity criterion for "match" responses that reflects the degree of PI within an experimental session. These observations fit well with several previous animal studies showing gradual improvement over time under consistently high PI conditions (D'Amato, 1973; Grant, 1975, 1976; Wright, 2007), and with data from humans showing that their familiarity criterion can be modified by changes in stimulus presentation frequency (Yonelinas, 2002).

In general, our results show that the effects of stimulus set size and ITI in auditory short-term memory in monkeys are similar to that which has been previously reported in visual studies. In Experiment 1, subjects reached 88% overall accuracy for sessions that used trial-unique stimuli, which compares favorably with the 90% accuracy reported by Mishkin and Delacour (1975) for monkeys performing a trial-unique visual DMS task. When only two stimuli were repetitively used throughout the session, accuracy fell to 72% in our study and 65% in Mishkin and Delacour's visual study. In Experiment 2, overall accuracy increased from 67% when the ITI was 5 s to 73% when the ITI was increased to 10 or 20 s. This modest increase in accuracy is similar to that observed by Jarrard and Moise (1971), who reported that increasing the ITI from 5 to 15 s increased accuracy by 5 to 10% for various retention intervals in monkeys performing visual DMS. Moreover, like our study, in which no additional benefit was gained by extending the ITI from 10 to 20 s, Jarrard and Moise also observed no significant increase in accuracy after increasing the ITI from 15 to 30 or 60 s. In both experiments, we observed significant effects of intertrial PI that extended beyond immediately adjacent trials. PI in visual short-term memory in monkeys has similarly been observed to span multiple trials (Wright, 2007). In Wright's study as well as ours, the effects of PI diminished as a function of number of trials separating the current trial from the source of the PI. However, a direct comparison of the extent and impact of intertrial PI in these studies is complicated by differences in experimental parameters. For instance, in Wright's study, the influence of PI was far more severe when a relatively long (20 s) compared to a short (1 s) retention interval was used, whereas our study used an intermediate retention interval (5 s). These differences notwithstanding, the foregoing results imply that PI in auditory and visual short-term memory in monkeys are at least qualitatively similar. Additional studies using similar task parameters, and ideally the same subjects, are needed to determine whether any quantitative differences exist between PI in auditory and visual short-term memory.

In view of our experimental outcomes as well as previous studies, future attempts to maximize accuracy by decreasing PI should avoid recycling stimuli between trials as much as possible. Ideally, new stimuli should be used for each trial. However, for some experimental paradigms, presenting multiple trials using the same stimuli is unavoidable. For instance, neurophysiological investigations of visual and auditory short-term memory

require multiple repetitions of each stimulus in order to establish reliable stimulus-evoked neuronal responses (Bigelow & Poremba, in press). Under such circumstances, a relatively long ITI may help reduce PI by increasing the decay time for irrelevant memory traces from previous trials. In addition to these factors, previous studies have indicated that PI may be reduced by increasing the stimulus exposure time (Grant, 1975) and by reducing the retention interval (Meudell, 1977; Wright, 2007; Wright et al., 2012; Zentall & Hogan, 1974). Thus, at minimum, optimal performance on DMS and similar tasks depends on the stimulus set size, ITI, retention interval, and stimulus exposure time, as well as interactions among these variables (see also van Hest & Steckler, 1996).

In summary, auditory memory in monkeys is highly susceptible to PI, which can be minimized by increasing the number of new stimuli that are presented throughout the trial. To a lesser extent, PI may be reduced by allowing an adequate interval of time between trials for sessions that reuse stimuli from trial to trial. Whether or not the monkeys will make a “same” judgment (whether correct or incorrect) may depend on whether the test stimulus exceeds a threshold level of familiarity, which may result either from having been recently presented as the sample for the current trial or from a presentation on a prior, and currently irrelevant trial. Following error-related feedback from having incorrectly chosen a non-matching test stimulus as a “match”, this threshold of familiarity may become more stringent in order to minimize future non-match errors. These findings expand our understanding of the variables governing auditory short-term memory in monkeys. Further studies directly comparing auditory, visual, and tactile short-term memory are needed in order to reveal the extent to which these findings generalize across sensory modalities.

## Acknowledgments

We thank Breein Rossi, Iva Zdilar, and Amy Opheim for assistance with data collection and Drs. Mortimer Mishkin and Richard Saunders as well as Megan Malloy at the National Institute of Mental Health for their support of the original monkey task acquisition. This research was supported by funding awarded to Amy Poremba from the University of Iowa and a grant from the National Institute on Deafness and Other Communication Disorders (DC0007156).

## References

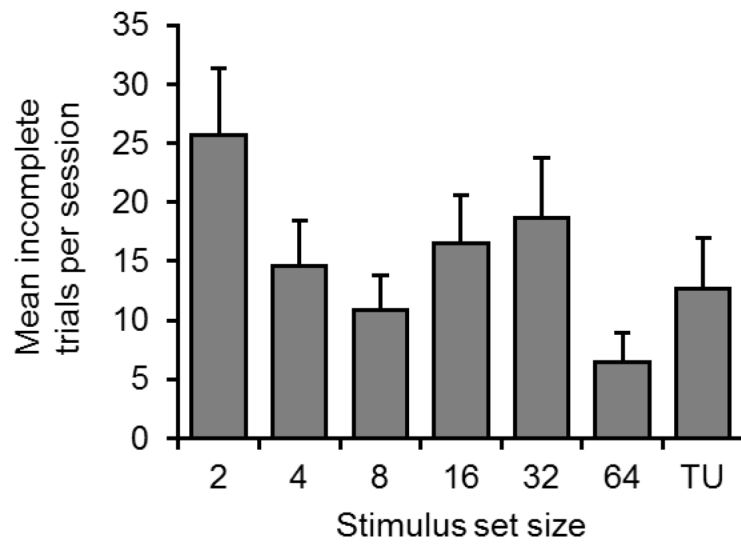
- Baddeley, AD.; Baddeley, AD. Human memory: Theory and practice. Allyn & Bacon; Needham Heights, MA: 1990. When memory fails; p. 169-189.
- Badre D, Wagner AD. Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex*. 2005; 15(12):2003–2012. [PubMed: 15788702]
- Bauer RH, Steele TL. Short-term memory for haptic cues in monkeys (*Macaca mulatta*). *Animal Learning & Behavior*. 1985; 13(3):291–302.
- Bigelow J, Poremba A. Auditory memory in monkeys: Costs and benefits of proactive interference. *American Journal of Primatology*. (in press). (in press).
- Buffalo EA, Ramus SJ, Clark RE, Teng E, Squire LR, Zola SM. Dissociation between the effects of damage to perirhinal cortex and area TE. *Learning and Memory*. 1999; 6(6):572–599. [PubMed: 10641763]
- Burwitz L. Proactive interference and directed forgetting in short-term motor memory. *Journal of Experimental Psychology*. 1974; 102(5):799–805.
- Cermak LS. Decay of interference as a function of the intertrial interval in short-term memory. *Journal of Experimental Psychology*. 1970; 84(3):499–501.
- Cohen JS, Reid S, Chew K. Effects of varying trial distribution, intra- and extramaze cues, and amount of reward on proactive interference in the radial maze. *Animal Learning & Behavior*. 1994; 22(2): 134–142.
- Cohen YE, Russ BE, Gifford GW III. Auditory processing in the posterior parietal cortex. *Behavioral and Cognitive Neuroscience Review*. 2005; 4(3):218–231.

- Colombo M, D'Amato MR. A comparison of visual and auditory short-term memory in monkeys (*Cebus apella*). *Quarterly Journal of Experimental Psychology B*. 1986; 38(4):425–448.
- Cothros N, Köhler S, Dickie EW, Mirsattari SM, Gribble PL. Proactive interference as a result of persisting neural representations of previously learned motor skills in primary motor cortex. *Journal of Cognitive Neuroscience*. 2006; 18(12):2167–76. [PubMed: 17129198]
- D'Amato, MR. Delayed matching and short-term memory in monkeys. In: Bower, GH., editor. *The psychology of learning and motivation: Advances in research and theory*. Vol. 7. Academic Press; New York: 1973. p. 227-269.
- D'Amato MR, Colombo M. Auditory matching-to-sample in monkeys (*Cebus apella*). *Animal Learning & Behavior*. 1985; 13(4):375–382.
- D'Amato MR, Worsham RW. Retrieval cues and short-term memory in capuchin monkeys. *Journal of Comparative and Physiological Psychology*. 1974; 86(2):274–282.
- De Rosa E, Hasselmo ME. Muscarinic cholinergic neuromodulation reduces proactive interference between stored odor memories during associative learning in rats. *Behavioral Neuroscience*. 2000; 114(1):32–41. [PubMed: 10718260]
- Dunnett SB, Martel FL. Proactive interference effects on short-term memory in rats: I. Basic parameters and drug effects. *Behavioral Neuroscience*. 1990; 104(5):655–665.
- Edhouse WV, White KG. Sources of proactive interference in animal memory. *Journal of Experimental Psychology: Animal Behavior Processes*. 1988; 14(1):56–70.
- Feredoes E, Tononi G, Postle BR. Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. *Proceedings of the National Academy of Sciences USA*. 2006; 103(51):19530–4.
- Fritz JB, Mishkin M, Saunders RC. In search of an auditory engram. *Proceedings of the National Academy of Sciences USA*. 2005; 102(26):9359–9364.
- Ghazanfar AA, Hauser MD. The auditory behaviour of primates: a neuroethological perspective. *Current Opinion in Neurobiology*. 2001; 11(6):712–720. [PubMed: 11741023]
- Gleitman H, Jung L. Retention in rats: The effect of proactive interference. *Science*. 1963; 142(3600):1683–4. [PubMed: 14075705]
- Grant DS. Proactive interference in pigeon short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes*. 1975; 1(3):207–220.
- Grant DS. Effect of sample presentation time on long-delay matching in the pigeon. *Learning and Motivation*. 1976; 7(4):580–590.
- Grant DS. Intertrial interference in rat short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes*. 1981; 7(3):217–227.
- Hampton RR, Shettleworth SJ, Westwood RP. Proactive interference, recency, and associative strength: Comparisons of black-capped chickadees and dark-eyed juncos. *Animal Learning & Behavior*. 1998; 26(4):475–485.
- Hartshorne JK. Visual working memory capacity and proactive interference. *Public Library of Science ONE*. 2008; 3(7):e2716. [PubMed: 18648493]
- Hayes KJ, Thompson R. Nonspatial delayed response to trial-unique stimuli in sophisticated chimpanzees. *Journal of Comparative and Physiological Psychology*. 1953; 46(6):499–500. [PubMed: 13109081]
- Hendrikx AJ. Short-term proactive interference revisited. *Bulletin of the Psychonomic Society*. 1986; 24(5):358–360.
- Herman LM. Interference and auditory short-term memory in the bottlenosed dolphin. *Animal Learning & Behavior*. 1975; 3(1):43–48.
- Herman LM, Bailey DR. Comparative effects of retroactive and proactive interference in motor short-term memory. *Journal of Experimental Psychology*. 1970; 86(3):407–415. [PubMed: 5491016]
- Hogan DE, Edwards CA, Zentall TR. Delayed matching in the pigeon: Interference produced by the prior delayed matching trial. *Animal Learning & Behavior*. 1981; 9(3):395–400.
- Jarrard, LE.; Moise, SL. Short-term memory in the monkey. In: Jarrard, LE., editor. *Cognitive processes of nonhuman primates*. Academic Press; New York: 1971. p. 1-24.

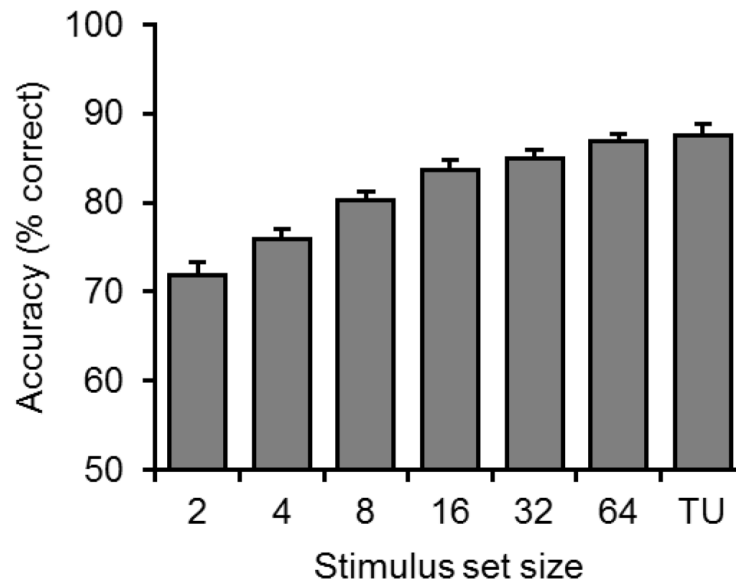
- Kane MJ, Engle RW. Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2000; 26(2):336–358.
- Keppel G, Underwood BJ. Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning & Verbal Behavior*. 1962; 1(3):153–161.
- Kojima S. Auditory short-term memory in the Japanese monkey. *International Journal of Neuroscience*. 1985; 25(3–4):255–262. [PubMed: 3980158]
- Loess H, Waugh NC. Short-term memory and intertrial interval. *Journal of Verbal Learning & Verbal Behavior*. 1967; 6(4):455–460.
- Maki WS, Moe JC, Bierley CM. Short-term memory for stimuli, responses, and reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes*. 1977; 3(2):156–177.
- Makovski T, Jiang YV. Proactive interference from items previously stored in visual working memory. *Memory and Cognition*. 2008; 36(1):43–52. [PubMed: 18323061]
- Mason M, Wilson M. Temporal differentiation and recognition memory for visual stimuli in rhesus monkeys. *Journal of Experimental Psychology*. 1974; 103(3):383–390. [PubMed: 4448951]
- Mecklinger A, Weber K, Gunter TC, Engle RW. Dissociable brain mechanisms for inhibitory control: Effects of interference content and working memory capacity. *Cognitive Brain Research*. 2003; 18(1):26–38. [PubMed: 14659494]
- Medin DL. Proactive interference in monkeys: Delay and intersample interval effects are noncomparable. *Animal Learning & Behavior*. 1980; 8(4):553–560.
- Meudell PR. Effects of length of retention interval on proactive interference in short-term visual memory. *Journal of Experimental Psychology: Human Learning and Memory*. 1977; 3(3):264–269.
- Mishkin M, Delacour J. An analysis of short-term visual memory in the monkey. *Journal of Experimental Psychology: Animal Behavior Processes*. 1975; 1(4):326–334. [PubMed: 811754]
- Moise SL. Proactive effects of stimuli, delays, and response position during delayed matching from sample. *Animal Learning & Behavior*. 1976; 4(1-A):37–40.
- Monsell S. Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*. 1978; 10(4):465–501.
- Murray EA, Mishkin M. Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *The Journal of Neuroscience*. 1998; 18(16):6568–6582. [PubMed: 9698344]
- Overman WH Jr, Doty RW. Prolonged visual memory in macaques and man. *Neuroscience*. 1980; 5(11):1825–31. [PubMed: 7432624]
- Postle BR, Berger JS, Goldstein JH, Curtis CE, D’Esposito M. Behavioral and neurophysiological correlates of episodic coding, proactive interference, and list length effects in a running span verbal working memory task. *Cognitive, Affective, & Behavioral Neuroscience*. 2001; 1(1):10–21.
- Reynolds TJ, Medin DL. Stimulus interaction and between-trials proactive interference in monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*. 1981; 7(4):334–347. [PubMed: 7288368]
- Roberts WA. Distribution of trials and intertrial retention in delayed matching to sample with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*. 1980; 6(3):217–237. [PubMed: 7391750]
- Roberts WA, Kraemer PJ. Some observations of the effects of intertrial interval and delay on delayed matching to sample in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*. 1982; 8(4):342–353. [PubMed: 7175446]
- Roitblat HL, Harley HE. Spatial delayed matching-to-sample performance by rats: Learning, memory, and proactive interference. *Journal of Experimental Psychology: Animal Behavior Processes*. 1988; 14(1):71–82.
- Roth PL. Missing data: a conceptual review for applied psychologists. *Personnel Psychology*. 1994; 47(3):537–560.
- Ruusuvirta T. Proactive interference of a sequence of tones in a two-tone pitch comparison task. *Psychonomic Bulletin & Review*. 2000; 7(2):327–331. [PubMed: 10909141]

- Ruusuvirta T, Astikainen P, Wikgren J. Proactive interference of differently ordered tone sequences with the accuracy and speed of two-tone frequency comparisons. *Music Perception*. 2002; 19(4): 551–563.
- Ruusuvirta T, Wikgren J, Astikainen P. Proactive interference in a two-tone pitch-comparison task without additional interfering tones. *Psychological Research*. 2008; 72(1):74–78. [PubMed: 17024429]
- Sands SF, Wright AA. Primate memory: Retention of serial list items by a rhesus monkey. *Science*. 1980; 209(4459):938–940. [PubMed: 6773143]
- Scott BH, Mishkin M, Yin P. Monkeys have a limited form of short-term memory in audition. *Proceedings of the National Academy of Sciences USA*. 2012; 109(30):12237–41.
- Stepien LC, Cordeau JP. Memory in monkeys for compound stimuli. *The American Journal of Psychology*. 1960; 73(3):388–395.
- Suzuki WA, Zola-Morgan S, Squire LR, Amaral DG. Lesions of the perirhinal and parahippocampal cortices in the monkey produce long-lasting memory impairment in the visual and tactual modalities. *The Journal of Neuroscience*. 1993; 13(6):2430–51. [PubMed: 8501516]
- Thompson RK, Herman LM. Auditory delayed discriminations by the dolphin: Nonequivalence with delayed-matching performance. *Animal Learning & Behavior*. 1981; 9(1):9–15.
- Underwood BJ. Interference and forgetting. *Psychological Review*. 1957; 64(1):49–60. [PubMed: 13408394]
- van Hest A, Steckler T. Effects of procedural parameters on response accuracy: Lessons from delayed (non-)matching procedures in animals. *Cognitive Brain Research*. 1996; 3(3–4):193–203. [PubMed: 8806022]
- Visscher KM, Kahana MJ, Sekuler R. Trial-to-trial carryover in auditory short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2009; 35(1):46–56.
- Wixted JT, Rohrer D. Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1993; 19(5):1024–1039.
- Worsham RW. Temporal discrimination factors in the delayed matching-to-sample task in monkeys. *Animal Learning & Behavior*. 1975; 3(2):93–97.
- Wright AA. Auditory and visual serial position functions obey different laws. *Psychonomic Bulletin & Review*. 1998; 5(4):564–584.
- Wright AA. Auditory list memory and interference processes in monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*. 1999; 25(3):284–296. [PubMed: 10423854]
- Wright, AA. Memory processing. In: Wasserman, EA.; Zentall, TR., editors. *Comparative cognition: Experimental explorations of animal intelligence*. Oxford University Press; New York: 2006. p. 164-185.
- Wright AA. An experimental analysis of memory processing. *Journal of the Experimental Analysis and Behavior*. 2007; 88(3):405–433.
- Wright AA, Katz JS, Ma WJ. How to be proactive about interference: lessons from animal memory. *Psychological Science*. 2012; 23(5):453–8. [PubMed: 22491142]
- Wright AA, Shyan MR, Jitsumori M. Auditory same/different concept learning by monkeys. *Animal Learning & Behavior*. 1990; 18(3):287–294.
- Wright, AA.; Urcuioli, PJ.; Sands, SF. Proactive interference in animal memory research. In: Kendrick, DF.; Rilling, M.; Denny, R., editors. *Theories of animal memory*. Erlbaum; Hillsdale, NJ: 1986. p. 101-125.
- Yonelinas AP. The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*. 2002; 46(3):441–517.
- Zentall TR, Hogan DE. Memory in the pigeon: Proactive inhibition in a delayed matching task. *Bulletin of the Psychonomic Society*. 1974; 4(2–A):109–112.



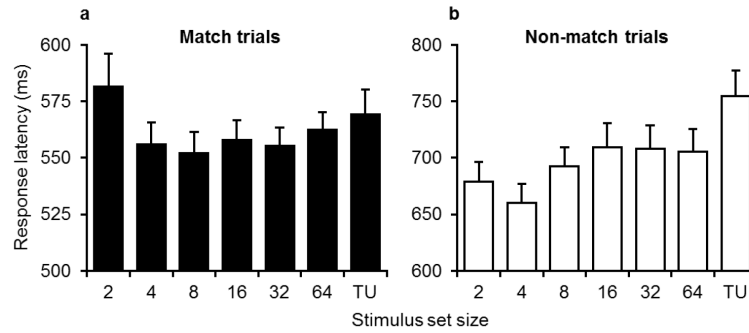


**Fig. 1.** Average number of trials not completed (out of 128) as a function of stimulus set size. Early quitting was more frequently observed during sessions using the smallest stimulus set size (two). TU = trial unique. Error bars indicate the standard error of the mean.



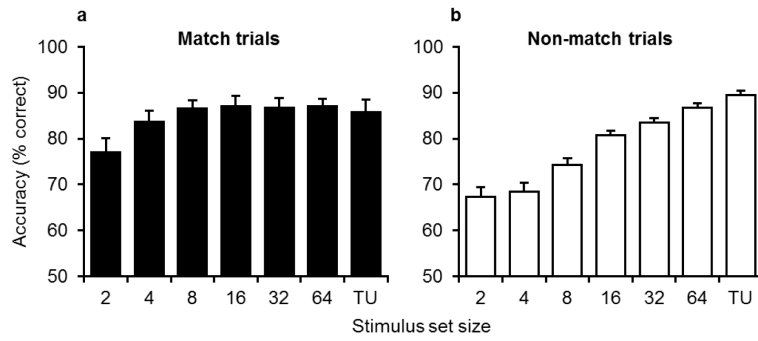
**Fig. 2.**

Overall accuracy improves as a function of stimulus set size. Accuracy for trial-unique sessions was significantly greater than sessions using stimulus set sizes of 32 or smaller. TU = trial unique. Error bars indicate the standard error of the mean.



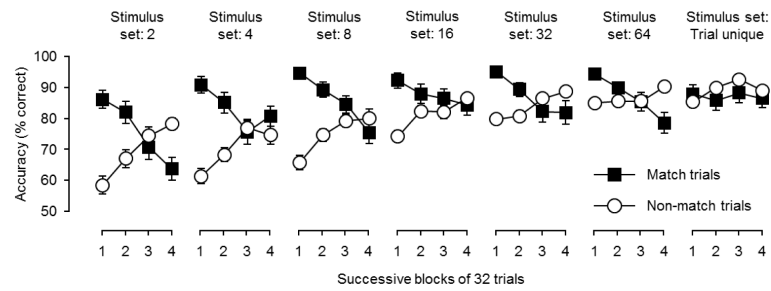
**Fig. 3.**

Response latency for match and non-match trials as a function of stimulus set size. **a** Response latency for match trials was significantly slower during sessions using the smallest set size (two), suggesting increased processing time for correct “match” responses under relatively high PI conditions. **b** By contrast, erroneous button presses on non-match trials were significantly slower for trial unique conditions, suggesting that errors are committed more quickly under high PI conditions. TU = trial unique. Error bars indicate the standard error of the mean.



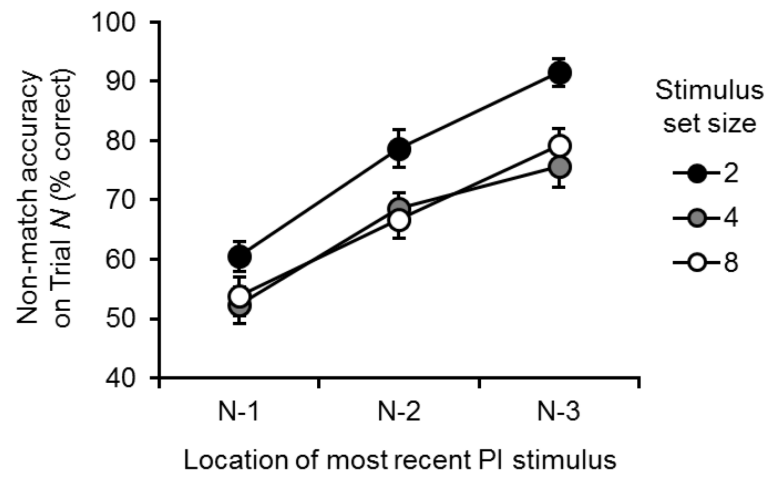
**Fig. 4.**

Accuracy for match and non-match trials as a function of stimulus set size. **a** Accuracy for match trials was relatively stable across stimulus set sizes, except that fewer correct match responses were made during sessions using the smallest sets (two and four). **b** PI associated with the smaller stimulus sets had a much larger impact on non-match accuracy, rising steadily from 67% in the two-stimulus set condition to 89% in the trial-unique condition. TU = trial unique. Error bars indicate the standard error of the mean.

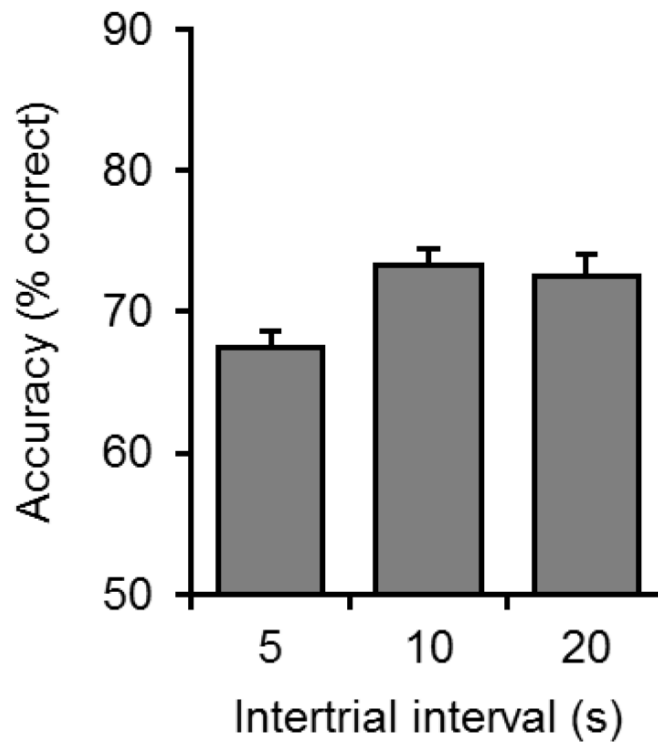


**Fig. 5.**

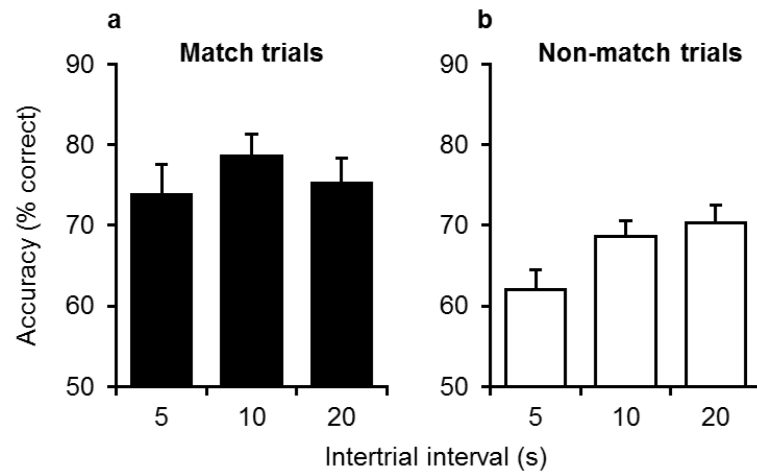
Progressive changes in accuracy by trial type for the first through the fourth quarters of the experimental session (successive blocks of 32 trials). Non-match errors became less frequent, whereas match errors became more common as the session progressed. The magnitude of this interaction diminished with increasing stimulus set size, such that no significant interaction was observed for trial-unique sessions. TU = trial unique. Error bars indicate the standard error of the mean.



**Fig. 6.** Non-match accuracy on trials for which the test stimulus had been presented on Trials  $N-1$ ,  $N-2$ , or  $N-3$ . Although overall non-match accuracy was lowest for the two-stimulus set condition (see text), accuracy on the subset of trials with recent PI included in this analysis was greater for the two-stimulus set than for the four- or eight-stimulus sets. Intertrial PI had a graded effect on non-match accuracy for all three conditions. Error bars indicate the standard error of the mean.

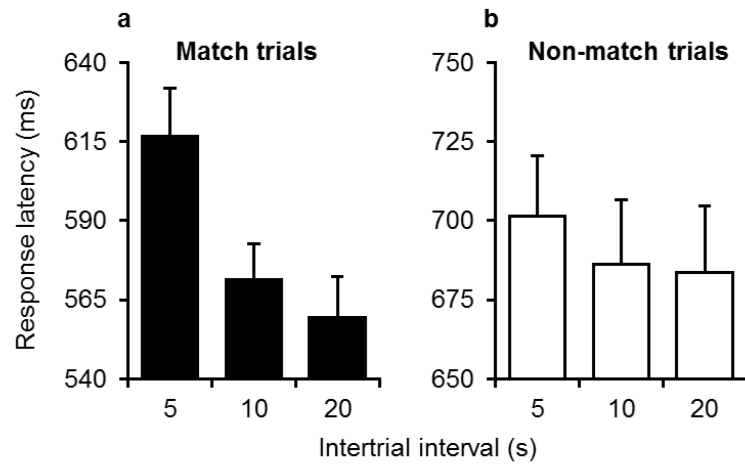


**Fig. 7.** Overall accuracy as a function of the duration of the ITI. Accuracy improved significantly when the ITI was extended from 5s to 10 s, but no further advantage was gained by increasing the ITI to 20 s. Error bars indicate the standard error of the mean.

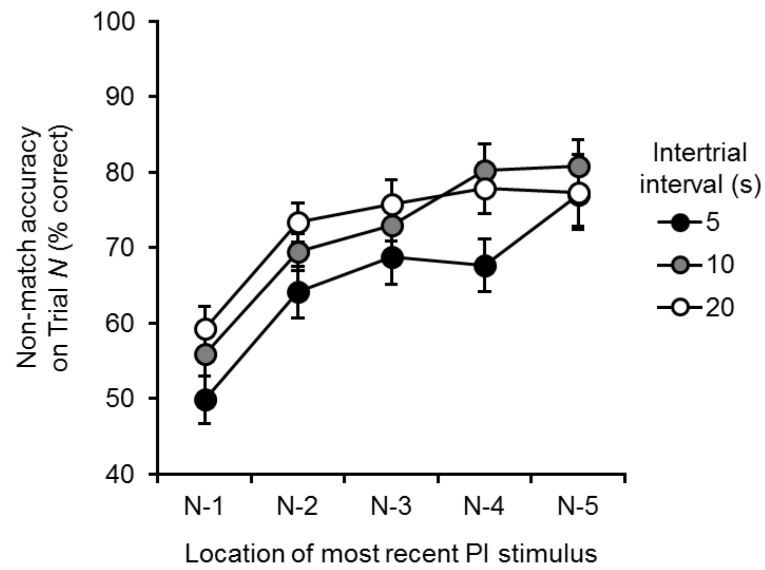


**Fig. 8.** Accuracy for match and non-match trials as a function of the duration of the ITI. **a** There was no significant effect of ITI on match accuracy. **b** However, non-match accuracy improved significantly by increasing the ITI from 5 s to 10 or 20 s. Error bars indicate the standard error of the mean.





**Fig. 9.** Response latency for match and non-match trials as a function of the duration of the ITI. **a** Correct match responses were significantly slower for sessions using the shortest ITI (five seconds). **b** No significant effect of ITI was found for erroneous responses on non-match trials. Error bars indicate the standard error of the mean.



**Fig. 10.** Intertrial PI as a function of the duration of the ITI. PI had the largest effect in the 5-s ITI condition, whereas the 10-s and 20-s ITI conditions were similar. PI had a significant influence spanning multiple trials for all three ITI conditions. Error bars indicate the standard error of the mean.