Published in final edited form as:

J Exp Psychol Anim Learn Cogn. 2014 October ; 40(4): 467–476. doi:10.1037/xan0000030.

Control of Working Memory in Rhesus Monkeys (*Macaca mulatta*)

Hsiao-Wei Tu and Robert R. Hampton

Emory University

Abstract

Cognitive control is critical for efficiently using the limited resources in working memory. It is well established that humans use rehearsal to increase the probability of remembering needed information, but little is known in nonhumans, with some studies reporting the absence of active control and others subject to alternative explanations. We trained monkeys in a visual matching-to-sample paradigm with a post-sample memory cue. Monkeys either saw a remember cue that predicted the occurrence of a matching test that required memory for the sample, or a forget cue that predicted a discrimination test that did not require memory of the sample. Infrequent probe trials on which monkeys were given tests of the type not cued on that trial were used to assess whether memory was under cognitive control. Our procedures controlled for reward expectation and for the surprising nature of the probes. Monkeys matched less accurately after forget cues, while discrimination accuracy was equivalent in the two cue conditions. We also tested monkeys with lists of two consecutive sample images that shared the same cue. Again, memory for expected memory tests was superior to that on unexpected tests. Together these results show that monkeys cognitively control their working memory.

Keywords

memory control; active processing; directed forgetting; rehearsal; substitution procedure

Working memory is a system in which a relatively small amount of information is temporarily maintained for active manipulation and rapid access (Baddeley, 1986; Cowan, 2005). Because the capacity of working memory is highly constrained (Cowan, 2005; Cowan, Morey, Chen, & Bunting, 2007; Williams & Woodman, 2012), cognitive control is critical for efficiently managing limited resources for processing the most relevant information. It is well established that humans selectively rehearse appropriate information to keep working memory updated and to minimize interference from irrelevant information (Cowan, 2005; Sheard & MacLeod, 2005). For example, we may maintain the name of a new acquaintance in memory by repeating it subvocally so that we can access it immediately when necessary.

The human ability to actively control memory resources has been studied over the past few decades using different variations of the “directed forgetting” paradigm developed by Bjork, LaBerge, and LeGrand (1968). In the “item method,” participants are instructed to remember or forget individual stimuli immediately after each presentation. In the “list method,” the instruction to remember or forget occurs in the middle of a list of stimuli and applies to all the stimuli preceding the instruction. When asked in subsequent tests to report as many stimuli as possible regardless of previous instructions, subjects report more of the stimuli they were instructed to remember (e.g., Hauswald & Kissler, 2008; MacLeod, 1999; Minnema & Knowlton, 2008; Titz & Verhaeghen, 2010). Converging evidence shows that humans are able to adaptively allocate memory resources to maintain the availability of specific items in working memory (Hourihan, Ozubko, & MacLeod, 2009; Sheard & MacLeod, 2005; Williams & Woodman, 2012).

Applying these directed forgetting paradigms to nonhuman animals is challenging and has yielded controversial results (Roper & Zentall, 1993; Zentall, Roper, Kaiser, & Sherburne, 1998). As a result, studies of directed forgetting have yet to establish whether nonhuman working memory involves cognitively effortful processing, such as rehearsal, although other kinds of evidence are suggestive (Basile & Hampton, 2013).

Because we cannot directly instruct nonhuman animals to remember some stimuli and not others, we instead have to arrange contingencies of reinforcement that favor active use of memory in some conditions but not others. Typically, subjects are presented with a sample stimulus followed by either a remember cue (“R cue” hereafter), indicating that a test of memory for the sample will follow, or a forget cue (“F cue” hereafter), indicating that a memory test for the sample will not follow. On F-cued trials the memory test is either omitted (e.g., Roberts, Mazmanian, & Kraemer, 1984; Roper, Chaponis, & Blaisdell, 2005; Washburn & Astur, 1998) or a different test unrelated to memory for the given sample stimulus is substituted in place of the memory test (e.g., Grant & Barnett, 1991; Kendrick, Rilling, & Stonebraker, 1981; Maki, Olson, & Rego, 1981; Washburn & Astur, 1998). After training, probe trials are administered, in which a memory test unexpectedly follows an F cue, to assess whether memory performance is contingent on the cue given following the sample. Control of memory by the R and F cues is inferred to the extent that performance on F-cued probe trials is less accurate than that on regular R-cued trials (Roper & Zentall, 1993).

Previous studies that applied the directed forgetting paradigm to nonhuman animals were performed almost exclusively with pigeons, and the results are controversial. Alternative explanations exist for the differences in accuracy observed between R-cued trials and F-cued probes. For example, when subjects are trained using an omission procedure, in which memory tests after F cues are omitted, subjects may associate the F cue with the absence of food reward. Expectation of no reward could decrease motivation or increase the probability of behaviors that are incompatible with accurate test performance, such as turning away from the response apparatus, resulting in poorer performance on F-cued probe trials (Maki & Hegvik, 1980; Zentall, Roper, & Sherburne, 1995). When alternative rewarded tests follow F cues in substitution procedures, the interpretation of results may depend on the nature of the substituted task and the required response (Grant & Barnett, 1991; Kendrick et

al., 1981). In some designs (e.g., Grant & Barnet, 1991; Grant & Soldat, 1995), a random stimulus was presented at test after F cues and choices were rewarded indiscriminately. This may train subjects to respond non-differentially after seeing an F cue, resulting in the performance discrepancy between regular R-cued trials and F-cued probes. In other studies (e.g., Kendrick et al., 1981; Maki & Hegvik, 1980; Maki et al., 1981), the same discrimination task was always presented after F cues. Because subjects can anticipate the response required at test in such designs, F cues may trigger prospective memory for the discrimination target, and this may passively displace any items already in memory (see Washburn & Astur, 1998, for a related discussion).

In contrast to the number of studies done with pigeons, there have been few efforts to test for directed forgetting in nonhuman primates, despite the fact that similar mnemonic processing might be most likely in our close primate relatives. In several experiments, Roberts et al. (1984) did not find any evidence for active memory control in squirrel monkeys (*Saimiri sciureus*) when a large set of stimuli was used, but after reducing the stimulus set to only two images, performance on memory tests was lower after F cues than after R cues in a final test. Because an omission procedure was used in which the F cue signaled the absence of reinforcement, the inference of memory control is questionable for the reasons cited above. The difficulty in interpreting these results is highlighted by a later study using rhesus monkeys (*Macaca mulatta*; Washburn & Astur, 1998). In this study, accuracy of all tests on which reinforcement was expected was comparatively high, whether or not remembering was necessary, emphasizing the importance of controlling the expectation of reward. This last study is the most recent published investigation of directed forgetting in monkeys we know of, and the authors concluded that “no evidence was found for an active maintenance process...” Nonetheless, the conclusion of Roberts et al. (1984) that directed forgetting may occur only when memory is tested with a small set of images may have forecast more recent results contrasting memory for large and small sets of images (Basile & Hampton, 2010, 2013; Eacott, Gaffan, & Murray, 1994).

We report studies in rhesus monkeys using methods that eliminate some of the alternative explanations to which early studies of nonhuman animals were subject. In Experiment 1, monkeys were trained with an R cue that predicted the occurrence of a memory test and an F cue that predicted a discrimination task that did not require memory of the sample, similar to the directed forgetting paradigm adopted in many pigeon studies (Roper & Zentall, 1993). We used manipulations that address concerns about motivational state, level of expectation of reward, and the activation of incompatible memories on F-cued trials. In Experiment 2, we presented lists of two consecutive sample images followed by one cue and tested memory for both of the studied images. This experiment tested whether the monkeys would generalize use of the R and F cues and also allowed a closer comparison to the “list method” used in human research. Both experiments test the hypothesis that monkeys use effortful and limited cognitive control to maintain sample images in working memory during the delay interval of memory tests. According to this hypothesis, monkeys actively maintain memories on R-cued trials and this effortful maintenance is rewarded when monkeys respond correctly in memory tests. In contrast, monkeys forgo active maintenance on F-cued trials because the effortful maintenance has not been reinforced in training.

Experiment 1A

The substitution procedure used in some pigeon directed forgetting research (Roper & Zentall, 1993) was used in Experiment 1A. Monkeys were trained to expect a matching-to-sample test when an R cue followed the sample and to expect a discrimination test when an F cue was presented. Only one discrimination problem was used and the specific test that followed F cues was therefore highly predictable. If monkeys actively maintain information in working memory when they expect a memory test, but do not do so when a memory test is not expected, performance on the memory tests following R cues should be more accurate than performance on the F-cued probe trials. We also evaluated whether any decrement observed after F-cued probe trials is specific to failure to maintain memory for the sample, or is rather due to a general decrement that would be present in any unexpected test. We did this by including R-cued probe trials, in which R cues were followed by an unexpected discrimination test. If discrimination performance is worse on these probe trials compared to normal F-cued trials, this would indicate that performance is impaired in any unexpected test.

Method

Subjects and apparatus—Six adult male rhesus monkeys (*Macaca mulatta*) were used. All monkeys were pair-housed, received full rations of food each day, and had *ad libitum* access to water at all times. During testing, pairs were separated by dividers such that each monkey could access only its own test equipment, but still had limited social contact with the cage mate. Monkeys had prior experience with computerized matching-to-sample tests. Monkeys were tested six days a week with a portable testing rig attached to the front of the home cage. Testing rigs consisted of a 15-inch color LCD touch-sensitive monitor (Elo TouchSystems, Menlo Park, CA) running at a resolution of 1024×768 pixels, two automatic food dispensers (Med Associates, Inc., St. Albans, VT) that delivered nutritionally balanced primate pellets (Bio-Serv, Frenchtown, NJ) into food cups below the screen, and a personal computer that controlled the experiments with custom programs written in Presentation (NeuroBehavioral Systems, Albany, CA).

Object discrimination learning—Each object discrimination problem consisted of four color clip-art images, 160 pixels high $\times$ 200 pixels wide, one of which was randomly selected as the target at the beginning of training. Trials started when a green square (100×100 pixels) appeared at the lower center of the screen. Once the subject touched it twice (FR2), the green square disappeared and the four images making up the discrimination were presented, each at a randomly determined corner of the screen. Touching (FR2) the target image was rewarded with one pellet together with an “excellent!” sound, while touching one of the other three images resulted in a “d’oh!” sound and a timeout of 5-seconds during which the screen was black. Consecutive trials were separated by an inter-trial interval of 5 seconds. Monkeys learned four discrimination problems concurrently, in sessions in which each problem occurred 40 times in a random order. Once the proportion correct was at least 0.9 for each problem simultaneously in a single session, one of the problems was randomly selected to be used in Experiment 1A.

Cued matching-to-sample—Four color clip-art images, 160 pixels high $\times$ 200 pixels wide, were used for matching. Two shapes (triangle or oval) of two colors (blue or yellow) were randomly assigned as R cues and F cues, respectively, for each monkey. Each trial began when a green square (100 $\times$ 100 pixels) appeared centered near the bottom of the screen. Touching it twice (FR2) started the study phase, in which one of the four images was pseudo-randomly selected as the sample and presented in the center of the screen. Once monkeys touched the sample twice (FR2), it disappeared. After 200 ms, an R or F cue appeared in a random location on the screen. Monkeys touched the cue twice (FR2) to enter the delay period during which the screen was black. The test phase began after the programmed memory delay had elapsed. If an R cue had appeared after the sample, all four images appeared, each randomly assigned to one corner of the screen. Selection of the sample image (FR2) shown during the study phase resulted in a food reward and an “excellent” sound. Selection of any of the three distractors resulted in a “d’oh” sound and 5-second timeout, during which the screen was black. If an F cue had appeared after the sample, the four images of the discrimination problem appeared on the screen. Choosing the target image (FR2) resulted in a food reward and an “excellent!” sound, while incorrect responses generated a “d’oh!” sound and a 5-second timeout period. Half of trials were R-cued and half F-cued. Consecutive trials were separated by an inter-trial interval of 5 seconds (Figure 1A).

Accuracy titration—The matching accuracy of each monkey on R-cued trials was titrated to between .50 and .70 by manipulating the delay interval. Reducing accuracy to this level was done to encourage memory maintenance and also to avoid ceiling effects that might mask differences in performance on R- and F- cued trials. Each session contained 160 trials, 80 with an R cue and 80 with an F cue. We started titration with a 3-second delay interval from the offset of the R or F cue to the onset of the test. If the proportion correct on R-cued trials was higher than .70 in a single session, the delay for both R-cued trials and F-cued trials was doubled. If the proportion correct on R-cued trials was lower than .50 for two consecutive sessions, the delay was decreased by a quarter. Performance on F trials was not expected to be affected by the length of delay because memory for the sample was not required. Monkeys proceeded to the next stage, probe testing, when the accuracy on R-cued trials fell between .50 and .70 for each of three consecutive sessions.

Probe testing—Ten percent of the trials in each session were probes, 5% were F-cued probes and 5% R-cued probes. On F-cued probe trials the matching test was presented instead of a discrimination test, whereas on R-cued probe trials the discrimination test was presented instead of a matching test. All trials for a given monkey had the same delay length, determined by the titration procedure. The same discrimination test used in titration was used here. Regular trials were rewarded when the correct image was selected, i.e., the sample on R-cued trials and the discrimination target on F-cued trials. Probe trials were rewarded no matter which image was selected. There were 320 trials in each session, including 288 regular trials, half of which were R-cued and half were F-cued, and 32 probe trials, half R-cued and half F-cued. The same discrimination test used in titration was used here. Each monkey received five sessions, yielding 80 R-cued probe trials and 80 F-cued probe trials.

Data analysis—Memory performance on regular R-cued trials and F-cued probe trials was compared by paired *t*-tests to determine whether the R and F cues had any effect on memory for sample images. Discrimination performance on regular F-cued trials and R-cued probe trials was also compared to test whether accuracy on all unexpected tests was attenuated, whether or not memory was required. All accuracy data were transformed before analysis by taking the arcsine of the square root of proportion correct scores. This transformation is recommended to make proportions better conform to the normality assumption of parametric statistics (Kirk, 1982). All trials were programmed with the same delay interval, but monkeys responded at their own pace. Given that memory declines with increasing delay, differences in delay intervals experienced by monkeys under the two cue conditions could cause differences in accuracy. To examine this possibility and to better isolate the effect of the cues on memory, median experienced delay intervals, calculated as the interval between the offset of the sample and the registration of a test response (FR2), were also compared between regular R-cued trials and F-cued probe trials as well as between regular F-cued trials and R-cued probe trials by paired *t*-tests. For each *t* value larger than 1, we reported the standardized Cohen's d_z for within-subjects designs as suggested by Lakens (2013) as well as the 95% confidence interval for this measure of effect size calculated using a bootstrapping procedure in R (Cumming, 2012).

Results and Discussion

By the end of the titration stage, each monkey had been assigned an individually tailored delay, ranging from 8.5 seconds to 24 seconds ($M = 14.42$ sec). The average performance on regular R-cued trials in the last three sessions of all six monkeys was $.60 \pm .04$ and that on regular F-cued trials was $.99 \pm .01$.

Monkeys were significantly more accurate on regular R-cued trials, when they expected a memory test, than they were on F-cued probe trials, $t(5) = 2.88$, $p = .04$, $d_z = 1.18$, 95% CI [0.08, 2.22] (black bars in Figure 2). This discrepancy was not due to different lengths of experienced delay intervals, $t(5) = 1.33$, $p = .24$, $d_z = 0.54$, 95% CI [-0.34, 1.38], suggesting that monkeys actively maintained memory on R-cued trials. Monkeys could potentially have learned to choose indiscriminately on probe trials because all choices on probe trials were reinforced. Such learning seems unlikely, however, because probe trials only constituted 10% of the trials in each session. If the monkeys had learned to treat probe trials differently, their performance should have decreased as they learned over the five sessions. We tested for such learning by comparing accuracy in the first 20 F-cued probe trials in the first session with that in the last 20 F-cued probe trials in the last session and found no difference, $t(5) = 0.45$, $p = .67$.

In contrast to the difference found between regular and probe memory tests, performance on discrimination tests was so accurate no matter which cue occurred that there was nearly no variation for conducting *t*-tests (regular F-cued trials: $M = 1.00$, $SD = .00$; R-cued probe trials: $M = 1.00$, $SD = .01$). This ceiling level accuracy might have prevented us from measuring a decrease in discrimination ability that would be evident if performance were not at ceiling. To provide additional, and potentially more sensitive, assessments of whether monkeys were disrupted or surprised by the R-cued probe trials, we compared experienced

delay intervals of these trials to other trial types. We found no significant difference in experienced delay intervals between regular F-cued trials and R-cued probe trials, $t(5) = 0.46$, $p = .67$, indicating no disruption in performance. In addition, experienced delay intervals were significantly shorter on R-cued discrimination trials than on R-cued memory trials, $t(5) = 2.92$, $p = .03$, $d_z = 1.19$, 95% CI [0.09, 2.23], consistent with a lack of hesitation in responding on R-cued discrimination trials. Together, these results suggest that the R and F cues determined whether monkeys actively remembered the sample images during the delay interval.

Experiment 1B

The main result of Experiment 1A replicates those from some directed forgetting studies in pigeons. As was the case with those studies, the same discrimination problem always appeared after F cues, and it is possible that the F cue triggered prospective memory for the target of the upcoming discrimination problem. Activation of a memory for the target of the discrimination might displace memory for the sample seen on that trial, leading to reduced accuracy on F-cued probe trials. Such displacement can be viewed as a passive retroactive interference effect, complicating interpretation of the R cue as instructing active maintenance of memory for the sample. To reduce the probability that the F cue activates prospective memory for the target of the discrimination, we made it much more difficult for the monkeys to predict the specific discrimination problem that would follow F cues. The discrimination problem used on each particular F-cued trial was randomly selected from among ten well-learned discriminations in Experiment 1B. Because the monkeys could no longer predict which discrimination target would be appropriate for the current trial, the possibility that the F cues triggered prospective memory was greatly reduced.

Method

The same six monkeys and testing rigs used in Experiment 1A were used. All monkeys learned another two sets of four discrimination problems as described above. Dropping the discrimination problem used in Experiment 1A and another randomly selected problem resulted in the pool of ten discrimination problems used in Experiment 1B. The accuracy on R-cued matching trials was titrated again to between .50 and .70 by the same procedure as before.

Trials proceeded as in Experiment 1A, except that the discrimination task for each F-cued trial was pseudo-randomly selected from the 10 new discrimination problems. Each session contained 800 trials, half had an R cue and half had an F cue. Of the 400 trials with R cues and F cues, 360 were regular trials and 40 were probe trials. Sample images, cue types, and discrimination tasks were all counterbalanced in each session. Monkeys were tested for 2 sessions, for a total of 80 probe trials of each type. Accuracy and delay intervals were analyzed in the same way as in Experiment 1A.

Results and Discussion

After titrating the accuracy again with 10 possible discrimination problems following F cues, the delays ranged from 9 seconds to 30 seconds ($M = 21.00$ sec, $SD = 8.05$ sec). The

average performance on regular R-cued trials in the last three sessions was $.64 \pm .02$ and that on regular F-cued trials was $.99 \pm .01$.

The performance on regular R-cued trials was significantly higher than that on F-cued probe trials, $t(5) = 3.20$, $p = .02$, $d_z = 1.31$, 95% CI [0.15, 2.40] (gray bars in Figure 2), and the experienced delay intervals on F-cued probe trials and regular R-cued trials did not differ, $t(5) = 0.40$, $p = .71$. As in Experiment 1A, accuracy on the first 20 and last 20 F-cued probe trials did not differ significantly, $t(5) = 0.74$, $p = .49$, indicating that monkeys did not learn to respond indiscriminately on probe trials. In contrast to the difference in accuracy between R and F-cued memory trials, accuracy on R-cued probe trials remained as high as that of regular F-cued trials (regular F-cued trials: $M = .97$, $SD = .06$; R-cued probe trials: $M = 1.00$, $SD = .01$). As in Experiment 1A, there was no significant difference in experienced delay intervals between regular F-cued trials and R-cued probe trials, $t(5) = 1.88$, $p = .12$, $d_z = 0.77$, 95% CI [-0.18, 1.66], and monkeys were significantly faster in responding on R-cued discrimination trials than on R-cued memory trials, $t(5) = 2.61$, $p = .05$, $d_z = 1.06$, 95% CI [0.01, 2.06], consistent with a lack of hesitation in responding on R-cued discrimination trials. These results, collected under conditions that prevent prospective memory for discrimination targets, reproduce the findings from Experiment 1A indicating active control of memory on R-cued trials.

Experiment 1C

The disruption in memory accuracy following F cues observed in Experiments 1A and 1B provides strong evidence for the ability of monkeys to actively regulate the contents of working memory. However, monkeys performed almost perfectly on each discrimination problem, earning a reward nearly every trial, whereas their memory accuracy averaged only .64. This discrepancy may lead to differences in expectation of food rewards after R and F cues. If monkeys expect a food reward after seeing an F cue but then receive a memory test that has a much lower probability of reward, they may be frustrated or less motivated, providing an alternative non-mnemonic explanation for the lower accuracy on F-cued probe trials. This concern was addressed in Experiment 1C by yoking the reinforcement on regular F-cued trials with that on regular R-cued trials to equate the expectation of food reward.

Method

The same six monkeys and titrated delays were used. No additional training was needed. Trials proceeded as in Experiment 1B, but the probability of reinforcement on regular F-cued trials was yoked with that on regular R-cued trials. Because monkeys were nearly 100% correct on F-cued discrimination trials, this was accomplished by calculating the proportion of R-cued trials rewarded in the previous session and rewarding the same proportion of correct discrimination responses in the current session. Monkeys ran one session of 400 R-cued matching trials and 400 F-cued discrimination trials before we started yoking. The performance on these R-cued trials was calculated and used for yoking the F-cued trials in the first session of the current experiment. All settings, including number of discrimination tasks used and trial numbers, remained the same as in Experiment 1B. Probe trials were implemented in the first yoked session, and two sessions were obtained from

each monkey. The accuracy and response latencies of the resulting 80 probe trials of each type were compared in the same way as in Experiments 1A and 1B.

Results and Discussion

Monkeys were again more accurate on R-cue trials than on F-cue probe trials, even after the probability of reinforcement on R and F trials was equated, $t(5) = 2.77$, $p = .04$, $d_z = 1.13$, 95% CI [0.05, 2.15] (open bars in Figure 2). Further analysis showed that this difference was not caused by a difference in experienced delay interval, $t(5) = 0.06$, $p = .96$. As in Experiments 1A and 1B, performance on F-cued probe trials did not decrease significantly between the first and last 20 trials, $t(5) = 1.36$, $p = .23$, $d_z = 0.56$, 95% CI [-0.33, 1.40], indicating that the monkeys did not learn during that probe trials were differently rewarded than regular trials. Performance in the discrimination tests was high, and equivalent, on F-cued and R-cued trials (regular F-cued trials: $M = 1.00$, $SD = .00$; R-cued probe trials: $M = 1.00$, $SD = .01$). Again, experienced delay intervals on regular F-cued trials and on R-cued probe trials did not differ, $t(5) = 0.57$, $p = .59$, suggesting that there was no disruption in responding to R-cued probe trials. In addition, experienced delay interval on R-cued discrimination trials was significantly faster than those on R-cued memory trials, $t(5) = 2.86$, $p = .04$, $d_z = 1.17$, 95% CI [0.07, 2.20], consistent with a lack of hesitation in responding on R-cued discrimination trials.

It might take some experience with the yoking procedure for the monkeys to learn about the new reinforcement rate for the discrimination tests. If this were the case, we might not expect to see an effect of yoking until the second session of testing. To evaluate whether a yoking effect might emerge only after experience, accuracies in the first and second session of testing were compared. The performance in the first and the second yoked sessions did not differ significantly on regular R-cued trials ($M = .64$ in the 1st session, $M = .66$ in the 2nd session), $t(5) = 0.83$, $p = .44$, regular F-cued trials ($M = .99$, $SD = .01$ in the 1st session, $M = 1.00$, $SD = .00$ in the 2nd session), R-cued probe trials ($M = 1.00$, $SD = .00$ in the 1st session, $M = .99$, $SD = .01$ in the 2nd session), or F-cued probe trials ($M = .46$ in the 1st session, $M = .50$ in the 2nd session), $t(5) = 1.13$, $p = .31$, $d_z = 0.46$, 95% CI [-0.41, 1.29]. These results indicate that the poorer memory performance on F-cued probe trials is not due to lower motivation or lower expectation of reward.

Taken together, Experiments 1A, 1B and 1C demonstrate active control of memory by monkeys and eliminate confounding factors that commonly contaminate studies of active memory control in nonhuman animals. Specifically, our data suggest that monkeys are able to follow the R and F cues and actively maintain memory when it is necessary to keep the most current and relevant information available in working memory.

Experiment 2

The ability to actively maintain working memory was demonstrated in Experiment 1. While this establishes a parallel with the item method used in studies of human working memory, human subjects are often presented with lists of stimuli followed by an instruction to remember or forget in a list method. To determine whether the monkeys would generalize the R and F cues to a modified test situation and to provide a closer parallel to the variety of

methods used in humans, we presented two consecutive sample images followed by a single cue and tested memory for both of the studied images. If monkeys generalize use of the R and F cues from the previous experiments to the current tests with short lists of two images, we should continue to observe superior accuracy on R-cued trials compared to F-cued probe trials.

Method

Subjects and apparatus—The same six monkeys and testing equipment were used.

Cued matching-to-sample—The four images used on matching trials in Experiment 1 were always grouped together as a “quad” at test, with one of them randomly chosen as the sample as before. Four new color clip-art images, 160 pixels high $\times$ 200 pixels wide, were added and grouped together to make a second “quad.” Thus, one image from each quad could be randomly selected on each trial to create a two-image list of samples for that trial. Trials progressed as described in Experiment 1, except that after touching (FR2) the first sample image, it disappeared and another sample image appeared in the same location after 200 ms. Monkeys had to touch (FR2) the second sample image before receiving the R or F cue and the delay interval. At test, if an R cue had followed the two sample images, two matching tests were presented one-by-one in the same order as the corresponding sample images had been presented in the study phase. If an F cue had followed the two sample images, two of the ten pre-trained discrimination tests were randomly selected and presented one after the other (Figure 1B). Correct responses were immediately rewarded with a food pellet and an “excellent” sound. Reinforcement on F-cued trials was not yoked in this experiment. Incorrect responses resulted in a “d’oh” sound but no timeout period. The timeout period was omitted in order to keep the delay interval between the offset of each sample image and the onset of the corresponding test constant, even when there were incorrect responses. Each session contained 192 trials, half of which were R-cued and the other half F-cued. The selection of samples for the two matching tests, the order in which these two samples appeared in a trial, and the discrimination problems used were counterbalanced in each session. The same symbols were used as R cues and F cues for each individual.

Accuracy titration—Because there were now two, rather than just one, to-be-remembered images, performance on the R-cued trials was re-titrated by the same procedure described in Experiment 1. All trials initially had a 3-second delay interval from the offset of the sample to the onset of the corresponding test. Monkeys were moved to the next stage when the average proportion correct of the two matching trials after R cues fell between .50 and .70 for three consecutive sessions.

Probe testing—Each session consisted of 256 R-cued trials and 256 F-cued trials, among which 16 R-cued and 16 F-cued trials were selected to be probes. On F-cued probe trials, both discriminations were replaced by matching tests in the order corresponding to sample presentations, whereas on R-cued probe trials, both matching tests were replaced by discrimination tests. Regular trials were rewarded when the correct image was selected,

whereas probe trials were rewarded no matter which image was selected. Each monkey was tested for 8 sessions, yielding 128 probe trials of each type.

Data analysis—The effects of cues and presentation order on accuracy and experienced delay interval were both analyzed by repeated-measures ANOVAs. All accuracy data were transformed by taking the arcsine of the square root of correct proportions prior to analysis (Kirk, 1982). Experienced delay intervals were defined as the interval from the offset of each sample to the registration of a response (FR2) to the corresponding test. Median delay intervals were analyzed. Significant main effects were further analyzed by planned paired t -tests. For each t value larger than 1, we reported the standardized Cohen's d_z for within-subjects designs as suggested by Lakens (2013) as well as the 95% confidence interval for this measure of effect size calculated using a bootstrapping procedure in R (Cumming, 2012).

Results and Discussion

Accuracy titration—Titrated delays ranged from 6 seconds to 12 seconds ($M = 8.50$ sec, $SD = 2.26$ sec). These delays are significantly shorter than those used in Experiments 1B and 1C, indicating that it is more difficult for the monkeys to remember two images than one, $t(5) = 3.76$, $p = .01$, $d_z = 1.54$, 95% CI [0.28, 2.73]. Accuracy in the last three sessions of titration averaged $.62 \pm .04$ on the first test of regular R-cued trials and $.55 \pm .04$ on the second test of regular R-cued trials. Monkeys were more accurate in first tests than in second tests, $t(5) = 3.96$, $p = .01$, $d_z = 1.62$, 95% CI [0.33, 2.85], averaged across the last three sessions.

Probe testing—When the cue was followed by matching tests, there was a significant main effect of cue, $F(1, 5) = 15.30$, $MSE = 0.01$, $p = .01$, $\eta_p^2 = 0.75$. Further analysis showed that both sample images were remembered better following R cues than following F cues: first tests: $t(5) = 3.99$, $p = .01$, $d_z = 1.63$, 95% CI [0.34, 2.86]; second tests: $t(5) = 3.43$, $p = .02$, $d_z = 1.40$, 95% CI [0.21, 2.53] (Figure 3), and this discrepancy was not caused by differences in experienced delay intervals $F(1, 5) = 1.14$, $MSE = 52667.09$, $p = .34$, $\eta_p^2 = 0.19$. The performance on the first 20 F-cued probe trials did not differ significantly from that on the last 20 F-cued probe trials, $t(5) = 0.79$, $p = .47$, indicating the monkeys did not learn to respond indiscriminately on probe trials. Performance on discrimination trials was almost always at ceiling without variation for conducting statistical analysis regardless of which cue preceded the test (after R cues: $M = .98$, $SD = .02$; after F cues: $M = 1.00$, $SD = .00$), and no significant difference was found in experienced delay intervals after F or R cues, $F(1, 5) = 0.12$, $p = .74$. These results reinforce the findings from the preceding experiments and extend them to memory for multiple sample images. As in Experiment 1, we compared the response latencies on R-cued discrimination trials. Latencies for regular F-cued trials and R-cued probe trials did not differ either for first tests, $t(5) = 0.61$, $p = .57$, or second test, $t(5) = 0.22$, $p = .84$, indicating that performance was not disrupted on R-cued discrimination trials. In addition, response latencies for R-cued probe trials were significantly shorter than those for regular R-cued trials on the first tests, $t(5) = 9.33$, $p < .0001$, $d_z = 3.81$, 95% CI [1.39, 6.22], and not significantly different than those for regular R-cued trials on the second

tests, $t(5) = 1.91$, $p = .11$, $d_z = 0.78$, 95% CI $[-0.17, 1.68]$, indicating that the monkeys were not hesitant on R-cued probe trials.

When the cue was followed by matching tests, there was a significant main effect of test order on performance, $F(1, 5) = 11.45$, $MSE = 0.01$, $p = .02$, $\eta_p^2 = 0.70$, and on experienced delay interval, $F(1, 5) = 80.14$, $MSE = 361793.29$, $p < .001$, $\eta_p^2 = 0.94$, but the interaction between cue and test order was only significant in performance, $F(1, 5) = 10.16$, $MSE = 0.001$, $p = .02$, $\eta_p^2 = 0.67$, not in experienced delay interval, $F(1, 5) = 4.31$, $MSE = 22832.59$, $p = .09$, $\eta_p^2 = 0.46$. Further analysis showed that on regular R-cued trials, memory for the first image seen during study (which was also the first image tested) was significantly better than that for the second image, $t(5) = 4.93$, $p = .004$, $d_z = 2.01$, 95% CI $[0.54, 3.44]$ (Figure 3), and the experienced delay interval was significantly shorter for the first test, $t(5) = 9.13$, $p < .001$, $d_z = 3.73$, 95% CI $[1.35, 6.09]$. However, on F-cued probe trials, memory for first samples did not differ from that for second samples, $t(5) = 1.68$, $p = .15$, $d_z = 0.69$, 95% CI $[-0.24, 1.56]$ (Figure 3), while the experienced delay interval was still significantly shorter for first than for second tests, $t(5) = 8.40$, $p < .001$, $d_z = 3.43$, 95% CI $[1.22, 5.62]$. Superior memory for the first image is consistent with the well-known primacy effect in memory for lists (Basile & Hampton, 2010; Sands, Urcuioli, Wright, & Santiago, 1984; Wright, 1994). It may be that higher performance in tests with the first image reflects priority access of the first image studied to cognitive resources necessary for active memory maintenance. Such an account is consistent with our finding that this superior memory for the first image is eliminated by presentation of F cues.

While it may be that the primacy effect found on R-cued trials, but not on F-cued probe trials, is due to unequal distribution of limited cognitive resources responsible for active memory maintenance, there are other factors that might produce the pattern of results observed. First, significantly shorter experienced delay intervals for the first image could account for the higher performance on first tests after R cues. However, even though similar significant differences in experienced delay interval were found after F cues and R cues, accuracy on first and second tests differed only following R cues. This suggests that differences in experienced delay are not the main reason for the different patterns of accuracy after R and F cues. Second, taking the first test might create various types of interference that could lower performance on the second test. For example, retrieving the food reward on the first test could interfere with accuracy in the subsequent test. If retrieving a reward interfered with performance in a subsequent test, we would expect the difference in accuracy between the first and second tests to be greater on F-cued probe trials, where first responses were rewarded 100% of the time indiscriminately, than on regular R-cued trials, where first responses were correct and rewarded only about 70% of the time. But that is not what we found. Instead, we found that on regular R-cued trials, second test accuracy was significantly higher when first tests were correct than when the first tests were incorrect, $t(5) = 11.47$, $p < .001$, $d_z = 4.68$, 95% CI $[1.77, 7.60]$. Third, absence of reward following errors in first tests might cause an emotional response that would impair performance in second tests. However, such a decrement in accuracy would only be expected on regular R-cued trials but not F-cued probe trials because all first response on F-cued probe trials were rewarded. In addition, if emotional responses caused by errors and the absence of reward

were the main cause of lower accuracy in second tests, performance on the second memory tests after F cues should be as high as that after R cues when the first tests were rewarded because F-cued second tests always followed reward. However, when first memory tests were correct, monkeys performed significantly better on the second memory tests after R cues than after F cues, $t(5) = 12.05$, $p < .001$, $d_z = 4.92$, 95% CI [1.88, 7.97]. Thus, it appears that the R and F cues determined accuracy in second tests.

We would generally expect any effect of interference from first tests to be consistent, regardless of which cue was presented and whether monkeys actively rehearsed the sample images or not, because there were always two tests on each trial. In contrast we found that memory for the first image was significantly better only on R cued matching tests. While the interpretation of these results is not uncomplicated by the occurrence of two tests on every trial, the weight of the evidence favors the interpretation that the first image studied gains priority access to active memory resources and the F-cue abolishes this difference by attenuating the contribution of active memory. Nonetheless, a design that counterbalanced the order in which memory tests are given after R cues, or conducted only a single test on each trial, would provide a stronger test and should be conducted in future work.

General Discussion

In the present study, monkeys showed significantly lower performance on memory tests that followed F rather than R cues. The effect was consistent across a series of experiments designed to eliminate common confounds. Although it is still possible that the R cues reinforce or tag memory for the sample image in some way that facilitates recognition but does not necessitate active maintenance, our results show that the contents of monkeys' working memory is different following R cues and F cues even though the encoding conditions are identical in both trial types. These results parallel what has been found in human memory research (Allen & Vokey, 1998; Sheard & MacLeod, 2005; Williams & Woodman, 2012), suggesting that monkeys, like humans, may be able to strategically allocate limited working memory resources in service of current needs.

The original directed forgetting paradigm used in humans (Bjork et al., 1968) has been modified in various ways to accommodate animal training. How animals are "instructed" to forget is critical to interpretation of differences between R and F cued trials. Omitting tests following F cues may create low motivation and incompatible behaviors that confound mnemonic explanations (Maki & Hegvik, 1980; Zentall et al., 1995). Instead, a substitution procedure with a single discrimination test after F cues has been used most commonly, with inconsistent results. In studies in which no deficit on F-cued probe trials was found (Kendrick et al., 1981; Maki & Hegvik, 1980; Maki et al., 1981), it has been argued that the anticipation of making a choice might encourage animals to continue rehearsing the sample on F-cued trials, resulting in high performance on both R-cued trials and F-cued probes (Kendrick & Rilling, 1986). While possible, this is weak evidence for active memory, and at best shows poor discrimination of the conditions under which rehearsal is useful. Other studies using the substitution procedure found poorer performance on F-cued probe trials (Grant & Barnet, 1991; Roper, Kaiser, & Zentall, 1995), as we found with our monkeys in Experiment 1A. However, these results are vulnerable to the alternative explanation that

memory for the sample might be passively displaced by the prospective memory for the response required in the anticipated discrimination.

Given these concerns, it is difficult to determine conclusively whether nonverbal animals show active maintenance of memory. Roper et al. (1995) argued that true directed forgetting may only occur when the original memory for the sample is updated with new sample-independent information in working memory. In their paradigm, memory for the sample was strategically designed to be displaced by memory for the F cues, and a significantly lower accuracy for the original sample was observed on F-cued probe trials. To further investigate whether directed forgetting effects in animals is due to cognitive control of active working memory rather than passive displacement of memories with new information, we trained our monkeys on multiple discrimination tests and randomized which discrimination followed the F cue on each trial in Experiment 1B. Monkeys therefore continued to expect to have to make a choice and to be rewarded on both R- and F-cued trials, but were unlikely to prospectively remember the choice to be made in the discrimination tests. With this new variation of the substitution procedure we found a robust directed forgetting effect.

In addition to substituting F-cued tests with multiple discrimination problems, there are several features of our design that may differ from those used with pigeons. Typically, before the inclusion of probe trials, animals are trained to associate the R cue with a memory test and the F cue with another test that does not require memory. The performance on R-cued and F-cued trials may be high after training, but it does not indicate successful learning of the cues. Because remembering is not particularly effortful on these trials, subjects may simply ignore the post-sample cue and try to remember on every trial to increase the probability of getting a reward. We modified our paradigm in at least three ways to emphasize the importance of the cues during training and encourage monkeys to actively remember only when necessary. First, the location of cue presentation changed on each trial, and the monkeys had to visually search for the cue before touching it to proceed. This reduced the possibility of inattentive responses and strengthened the association between each cue and its corresponding test. Second, the retention interval was titrated to bring accuracy into a moderate range. Because actively keeping memory available is an effortful process, it is possible that monkeys would not adopt this strategy unless it is beneficial to do so. When the delay interval becomes longer, memory for the sample naturally weakens, and active maintenance becomes necessary to increase the chance of food reward compared to passively waiting for the test. Third, the small image set used on R-cued trials created substantial interference that reduced the possibility of using familiarity for matching and may have enhanced the necessity of rehearsal (Basile & Hampton, 2013; Eacott et al., 1994; Roberts et al., 1984). A major property of human working memory is that it is vulnerable to competing cognitive demand (Logie, 1986; Phillips & Christie, 1977). In monkeys, memory for images drawn from a small image set is especially vulnerable to competing cognitive demand (Basile & Hampton, 2013), suggesting that such memories are actively, rather than passively maintained, a property consistent with the memory control effects we report here.

Experimental designs other than directed forgetting paradigms have been developed to study the active control of working memory in nonhuman primates. For example, distractors, such as a sample-independent stimulus (Artchakov et al., 2009; Takeda, Naya, Fujimichi,

Takeuchi, & Miyashita, 2005), a motor response (Washburn & Astur, 1998), or a categorization task (Basile & Hampton, 2013), were inserted during the retention interval to compete with memory for the sample image for the limited cognitive resources in working memory. Whether active rehearsal was observed seems to depend on the extent to which distractor tasks increase cognitive load. List learning is another common method, in which primacy is often argued to reflect active rehearsal-like process of items in memory (Basile & Hampton, 2010; Cook, Wright, & Sands, 1991), but alternative interpretations exist (Sands et al., 1984).

Our findings of significantly higher performance on the first matching tests than the second ones *only* after R cues but not after F cues coincide with the results of human studies. For example, clear primacy and recency effects were reported when participants were asked to recall items in the to-be-remembered list, but only a slight primacy and no recency effect was found when recalling items in the to-be-forgotten list (Pastötter, Kliegl, & Bäuml, 2012; Sheard & MacLeod, 2005). Sheard & MacLeod (2005) further concluded that selective rehearsal may be the main explanation for directed forgetting. Although other factors may have also contributed to the “primacy-like” pattern we observed on regular R-cued trials, this first attempt to incorporate cues in a list learning paradigm offers an opportunity to study active memory control in nonverbal animals in a way that parallels the paradigms used in human research, facilitating cross-species comparisons of memory processes.

Here we eliminated memory-irrelevant factors that confound many previous reports of memory control in nonhumans, including effects of reward expectancy and potential interference from prospective memory. The case for memory control was further supported by extension of the effects of the R and F cues to multiple stimuli in lists. Given the fact that matching was more often rewarded in the presence of the R cue than the F cue, this pattern of behavior would at one time have been described as better stimulus control of matching behavior by the R cue than the F cue. While such a description can correctly predict performance and well capture the contingencies present in our experimental design, it says little about the cognitive processes that might be responsible for the control of memory retention by the R and F cues. We do interpret our results as showing that matching performance is under differential stimulus control by the R and F cues, but additionally suggest that it is not just “performance” that is under the control of these cues, but a memory process that facilitates accurate matching. Future work will evaluate whether this memory process is an active one, vulnerable to competing cognitive demand, as we have found to be the case with some, but not all, memories (Basile and Hampton, 2013). The nonhuman primate model of cognitive control of memory established here provides a basis for further investigations of the underlying mechanisms by which brains control memory.

Acknowledgments

This work was supported by National Institute of Mental Health Grant No. R01MH082819 and by National Science Foundation Grant Nos. BCS-0745573 and IOS-1146316. This project was funded by the National Center for Research Resources P51RR165 and is currently supported by the Office of Research Infrastructure Programs/OD P51OD11132. We thank Emily K. Brown, Jessica A. Joiner, and Tara A. Dove-VanWormer for assistance with testing animals and animal care.

References

- Allen, SW.; Vokey, JR. Directed forgetting and rehearsal on direct and indirect memory tests. In: Golding, JM.; MacLeod, CM., editors. *Intentional Forgetting: Interdisciplinary Approaches*. Mahwah, NJ: Lawrence Erlbaum Associates; 1998. p. 173-195.
- Artchakov D, Tikhonravov D, Ma Y, Neuvonen T, Linnankoski I, Carlson S. Distracters impair and create working memory-related neuronal activity in the prefrontal cortex. *Cerebral Cortex*. 2009; 19:2680–2689. [PubMed: 19329569]
- Baddeley, AD. *Working Memory*. Oxford: Oxford University Press; 1986.
- Basile BM, Hampton RR. Rhesus monkeys (*Macaca mulatta*) show robust primacy and recency in memory for lists from small, but not large, image sets. *Behavioural Processes*. 2010; 83:183–190. [PubMed: 20035843]
- Basile BM, Hampton RR. Dissociation of active working memory and passive recognition in rhesus monkeys. *Cognition*. 2013; 126:391–396. [PubMed: 23287420]
- Bjork RA, LaBerge D, LeGrand R. The modification of short-term memory through instructions to forget. *Psychonomic Science*. 1968; 10:55–56.
- Cook RG, Wright AA, Sands SF. Interstimulus interval and viewing time effects in monkey list memory. *Animal Learning and Behavior*. 1991; 19:153–163.
- Cowan, N. *Working Memory Capacity*. New York, NY: Psychology Press; 2005.
- Cowan, N.; Morey, CC.; Chen, Z.; Bunting, M. What do estimates of working memory capacity tell us?. In: Osaka, N.; Logie, RH.; D'Esposito, M., editors. *The Cognitive Neuroscience of Working Memory*. New York, NY: Oxford University Press; 2007. p. 43-58.
- Cumming, G. *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York, NY: Routledge; 2012.
- Eacott MJ, Gaffan D, Murray EA. Preserved recognition memory for small sets, and impaired stimulus identification for large sets, following rhinal cortex ablations in monkeys. *European Journal of Neuroscience*. 1994; 6:1466–1478. [PubMed: 8000570]
- Grant DS, Barnet RC. Irrelevance of sample stimuli and directed forgetting in pigeons. *Journal of the Experimental Analysis of Behavior*. 1991; 55:97–108. [PubMed: 16812631]
- Grant DS, Soldat AS. A postsample cue to forget does initiate an active forgetting process in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*. 1995; 21:218–228.
- Hauswald A, Kissler J. Directed forgetting of complex pictures in an item method paradigm. *Memory*. 2008; 16:797–809. [PubMed: 18608977]
- Hourihaan KL, Ozubko JD, MacLeod CM. Directed forgetting of visual symbols: Evidence for nonverbal selective rehearsal. *Memory & Cognition*. 2009; 37:1059–1068. [PubMed: 19933451]
- Kendrick, DF.; Rilling, ME. AIM: A theory of active and inactive memory. In: Kendrick, DF.; Rilling, ME.; Denny, MR., editors. *Theories of Animal Memory*. Hillsdale, NJ: Erlbaum; 1986. p. 129-152.
- Kendrick DF, Rilling M, Stonebraker TB. Stimulus control of delayed matching in pigeons: Directed forgetting. *Journal of the Experimental Analysis of Behavior*. 1981; 36:241–251. [PubMed: 16812243]
- Kirk, RE. *Experimental Design: Procedures for the Behavioral Sciences*. 4. Thousand Oaks, CA: SAGE Publications, Inc; 1982.
- Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*. 2013; 4:1–12. [PubMed: 23382719]
- Logie RH. Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*. 1986; 37:229–247.
- MacLeod CM. The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review*. 1999; 6:123–129. [PubMed: 12199306]
- Maki WS, Hegvik DK. Directed forgetting in pigeons. *Animal Learning & Behavior*. 1980; 8:567–574.
- Maki WS, Olson D, Rego S. Directed forgetting in pigeons: Analysis of cue functions. *Animal Learning & Behavior*. 1981; 9:189–195.

- Minnema MT, Knowlton BJ. Directed forgetting of emotional words. *Emotion*. 2008; 8:643–652. [PubMed: 18837614]
- Pastötter B, Kliegl O, Bäuml KHT. List-method directed forgetting: The forget cue improves both encoding and retrieval of postcue information. *Memory & Cognition*. 2012; 40:861–873. [PubMed: 22588948]
- Phillips WA, Christie DFM. Interference with visualization. *Quarterly Journal of Experimental Psychology*. 1977; 29:637–650. [PubMed: 601182]
- Roberts WA, Mazmanian DS, Kraemer PJ. Directed forgetting in monkeys. *Animal Learning & Behavior*. 1984; 12:29–40.
- Roper KL, Chaponis DM, Blaisdell AP. Transfer of directed-forgetting cues across discrimination tasks with pigeons. *Psychonomic Bulletin & Review*. 2005; 12:1005–1010. [PubMed: 16615320]
- Roper KL, Kaiser DH, Zentall TR. True directed forgetting in pigeons may occur only when alternative working memory is required on forget-cue trials. *Animal Learning & Behavior*. 1995; 23:280–285.
- Roper KL, Zentall TR. Directed forgetting in animals. *Psychological Bulletin*. 1993; 113:513–532. [PubMed: 8316612]
- Sands, SF.; Urcuioli, PJ.; Wright, AA.; Santiago, HC. Serial position effects and rehearsal in primate visual memory. In: Roitbalt, HL.; Bever, TG.; Terrace, HS., editors. *Animal Cognition*. Hillsdale, NJ: Erlbaum; 1984. p. 375-388.
- Sheard, ED.; MacLeod, CM. List method directed forgetting: Return of the selective rehearsal account. In: Ohta, N.; MacLeod, CM.; Uttl, B., editors. *Dynamic Cognitive Processes*. New York, NY: Springer; 2005. p. 219-248.
- Takeda M, Naya Y, Fujimichi R, Takeuchi D, Miyashita Y. Active maintenance of associative mnemonic signal in monkey inferior temporal cortex. *Neuron*. 2005; 48:839–848. [PubMed: 16337920]
- Titz C, Verhaeghen P. Aging and directed forgetting in episodic memory: A meta-analysis. *Psychology and Aging*. 2010; 25:405–411. [PubMed: 20545424]
- Washburn DA, Astur RS. Nonverbal working memory of humans and monkeys: Rehearsal in the sketchpad? *Memory & Cognition*. 1998; 26:277–286. [PubMed: 9584435]
- Williams M, Woodman GF. Directed forgetting and directed remembering in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2012; 38:1206–1220.
- Wright AA. Primacy effects in animal memory and human nonverbal memory. *Animal Learning and Behavior*. 1994; 22:219–223.
- Zentall, TR.; Roper, KL.; Kaiser, DH.; Sherburne, LM. A critical analysis of directed-forgetting research in animals. In: Golding, JM.; MacLeod, CM., editors. *Intentional Forgetting: Interdisciplinary Approaches*. Mahwah, NJ: Lawrence Erlbaum Associates; 1998. p. 265-287.
- Zentall TR, Roper KL, Sherburne LM. Most directed forgetting in pigeons can be attributed to the absence of reinforcement on forget trials during training or to other procedural artifacts. *Journal of the Experimental Analysis of Behavior*. 1995; 63:127–137. [PubMed: 7714447]

Figure 1.

(A) The directed forgetting paradigm. The R and F cues were presented at random locations on the screen. An R cue was followed by a matching test of memory for the sample image (upper panel), while an F cue was followed by a pre-trained discrimination test (lower panel). The red circles indicate the correct responses. (B) An example of a two-sample trial. If an R cue was presented after the two sample images, two matching tests appeared in sequence after the delay interval in the order corresponding to sample presentation. If an F cue was presented after the two sample images, two discrimination tests randomly chosen from 10 pre-trained discriminations appeared one after the other following the delay interval.

Figure 2.

Accuracy of regular R-cued trials and F-cued probe trials in the three conditions in Experiment 1. An R cue predicted a matching test that required memory for the sample image; while an F cue predicted a discrimination test unrelated to memory for the sample image. Memory accuracy on F-cued probe trials was significantly poorer than that on regular R-cued trials. The results were consistent across Experiment 1 (black bars), 1B (gray bars), and 1C (open bar), showing that monkeys followed the cues and allocated their working memory accordingly. Error bars indicate standard errors.

Figure 3.

Sample images were better remembered when matching tests, not discrimination tests, were expected. In addition, when monkeys were cued to remember the sample images, the first matching tests were significantly more accurate than the second one, suggesting active rehearsal might occur during the delay interval. In contrast, when monkeys expected to receive discrimination tests on F-cued probed trials, accuracy was equally poor for both the first and second memory test. Error bars indicate standard errors.