Research

Monkeys rely on recency of stimulus repetition when solving short-term memory tasks

John H. Wittig Jr. and Barry J. Richmond[1]

Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892-4415, USA

Seven monkeys performed variants of two short-term memory tasks that others have used to differentiate between selective and nonselective memory mechanisms. The first task was to view a list of sequentially presented images and identify whether a test matched any image from the list, but not a distractor from a preceding list. Performance was best when the test matched the most recently presented image. Response rates depended linearly on recency of repetition whether the test matched a sample from the current list or a distractor from a preceding list, suggesting nonselective memorization of all images viewed instead of just the sample images. The second task was to remember just the first image in a list selectively and ignore subsequent distractors. False alarms occurred frequently when the test matched a distractor presented near the beginning of the sequence. In a pilot experiment, response rates depended linearly on recency of repetition irrespective of whether the test matched the first image or a distractor, again suggesting nonselective memorization of all images instead of just the first image. Modification of the second task improved recognition of the first image, but did not abolish use of recency. Monkeys appear to perform nonspatial visual short-term memory tasks often (or exclusively) using a single, nonselective, memory mechanism that conveys the recency of stimulus repetition.

Old world monkeys are frequently used to examine the neural mechanisms of visual short-term memory. The prototypical assay of nonspatial short-term memory in monkeys has been delayed-match-to-sample (Miller et al. 1996; Amit et al. 2003; Schneider et al. 2009; Verrico et al. 2011). In delayed-match-to-sample, two or more stimuli are presented sequentially, separated by delays, and subjects report whether a test stimulus matches one of the sample stimuli. Two variants used to assay short-term memory in monkeys are: (1) to detect the repetition of any stimulus within the sequence (Sands and Wright 1980; Amit et al. 2003; Yakovlev et al. 2005), or (2) to detect the repetition of the first stimulus within the sequence, while ignoring, or at least suppressing a response to, repetition of intervening stimuli (Miller and Desimone 1994; Miller et al. 1996; Suzuki et al. 1997). For convenience we call the first variation Match Any, and the second variation Match First.

Both of these variations of delayed-match-to-sample have been called working memory tasks because they require remembering and manipulating the memory of recently seen stimuli (Miller and Desimone 1994; Amit et al. 2003). In Match Any all of the stimuli from the sequence must be retained in memory and then compared to the test stimulus to make a correct response. In Match First only the first stimulus must be retained. Previous studies using Match Any have shown that monkeys correctly identify sample images that were presented near the end of a sequence, that is, monkeys excel at identifying repetition of recently viewed images (Amit et al. 2003; Yakovlev et al. 2005, 2008). Previous studies using Match First reported that monkeys required additional training to correctly reject repetitions of non-first stimuli in so-called "ABBA" trials (Miller and Desimone 1994). The authors concluded that monkeys were initially solving the task using a stimulus repetition rule, and that after additional

training the monkeys learned to actively maintain just the first image using working memory.

The behavioral and/or neurophysiological responses observed during Match Any and Match First have been interpreted to mean these tasks are supported by different memory mechanisms, or processes (Miller et al. 1991; Miller and Desimone 1994; Wright 1998; Amit et al. 2003; Yakovlev et al. 2005, 2008). Definitions and models of short-term memory processes vary (Cowan 2008), but one broad distinction can be made between active and passive processes (Crowder 1969; Wright 1998; Scott et al. 2012), whereby active processes are willful and require selective attention (Watkins 1989; Baddeley 2003; Scott et al. 2012) and passive processes "can be thought of as occurring virtually automatically" (Wright 1998). Previous studies using Match First have argued that the sample is stored using an active memory process (e.g., rehearsal), based on the behavioral requirement for selective memorization and measurement of a sample-selective neurophysiological response (Miller and Desimone 1994). Previous studies using Match Any have argued that behavioral performance trends are due to the passive memory processes of interference (Wright 1998) or "spontaneous jumps of an image into or out of memory" (Yakovlev et al. 2005). Although the details and terminology used to describe passive memory processes underlying Match Any differ across studies (Wright 1998; Yakovlev et al. 2005, 2008), it is commonly assumed that the processes are different than the active memory processes that underlie Match First (Miller and Desimone 1994; Yakovlev et al. 2013).

We trained monkeys to perform both tasks using exactly the same lists of stimuli. The lists were varied in length from two to eight. Overall performance in our tasks was comparable to reports for the original tasks (Miller and Desimone 1994; Amit et al. 2003). In both Match Any and Match First, our monkeys

[1]**Corresponding author**
E-mail bjr@ln.nimh.nih.gov
Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.034181.113.

performed best when the test image matched an image that occurred late in the sequence. In Match Any, stimuli presented near the end of the sequence were most likely to be correctly identified. In Match First, we observed a similar trend in performance that suggests monkeys rely on recency of stimulus repetition to solve this task as well. It appeared as though the first image was not treated any differently than the other images in the sequence when our task conditions most closely mimicked those used in the original reports of Match First. This raises the question of whether monkeys are ever able to ignore information about the recency of stimulus repetition and selectively memorize a sample using active memory processes when solving delayed-match-to-sample.

Results

Match Any

We trained five monkeys to watch a sequence of two to eight images and report whether the last image in the sequence, the test image, matched any of the preceding samples (see Fig. 5A in Materials and Methods). The five monkeys required eight to 23 training sessions to reach criterion performance ($\geq 65\%$ correct) on the Match Any task with intermixed sequence lengths up to eight images long. After training, the monkeys were tested for 10 consecutive sessions, which provided enough data (trials per condition) to analyze the errors during the eight-image sequences. Performance was stable across the 10 consecutive sessions (ANOVA, within-subject repeated measures, subject × session: session $F = 1.1$, df $= 1$, $P = 0.30$; interaction $F = 0.2$, df $= 3$, $P = 0.88$). The mean performance on eight-image sequences was $77 \pm 3\%$ ($N = 4$, mean $\pm$ SEM; monkey B was tested with 11-image sequences instead of eight-image sequences and is not included in the mean) and errors were evenly distributed among trials in which the test matched a sample (yes is correct) or matched a distractor (no is correct) (Fig. 1A shows example data from monkey F).

To determine the source of the errors, we sorted trials in which the test matched a sample according to how recently the sample matching the test image was last presented (Fig. 1B, example for same monkey F as in Fig. 1A). Test recency is defined as the number of images before the test at which the matching sample image was last presented. Performance was best when the test image matched the most recently presented sample (index -1), and worst when the test matched the earliest sample in the sequence (index -7). The yes-response rates increased monotonically with increasing test recency. Within the trial the rates were modeled as a linear function of test recency (linear regression $r^2 = 0.55$, slope $= 5.9$, $P < 0.05$ [$P = 1 \times 10^{-14}$] [Fig. 1B, solid blue line]; the residuals at each sequence index were indistinguishable from zero, $P > 0.05$ [range $P = 0.38$ to 1.0], two-tailed t-test [Fig. 1C]). When the test image matched an image preceding the current trial, we labeled it as a preceding distractor. The monkey correctly rejected $82 \pm 2\%$ of the test images that matched a preceding distractor. The observed lower performance for samples earlier in a sequence has been interpreted to reflect working memory decay over time and/or working memory "overload" from subsequent samples (Yakovlev et al. 2005) due to retroactive interference (Wright 1998).

To study whether the monkeys were influenced by images presented in the preceding trial, we reduced the maximum sequence length from eight to four and the number of available images from 100 to five. Under these conditions, preceding-distractor images were likely to have last appeared within six sequence indices prior to the current trial, that is, within the trial or two preceding the current sequence. This provided enough data to sort trials in which the test matched a preceding distractor according to how recently the test image was last presented (Fig. 1D, red bars on white background; index -4 is the test image from the previous trial, example monkey F from Fig. 1A–C). Average performance during four-image sequences dropped from $83 \pm 4\%$ with 100 available images to $77 \pm 2\%$ with five available images ($N = 5$, example monkey F dropped from 84%
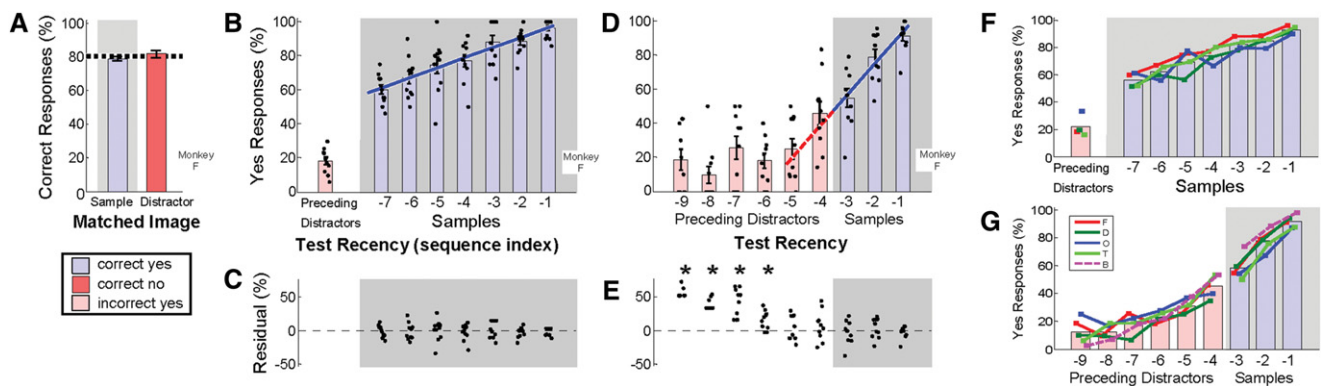
Figure 1. Match Any response rates increase linearly with recency. (A–C) Performance of one monkey ("F") on eight-image sequences (seven samples + one test) during the initial test with long sequences (10 daily sessions; 70% of trials were eight-image sequences, 100 available images). (A) Percent correct: overall (dashed line), and for trials in which the test image matched a sample (blue bar, gray background; "yes" is correct) or distractor (red bar, white background; "no" is correct). Error bars are SEM of 10 sessions. (B) Same dataset as A, but with performance reported as percent "yes" responses, and with trials further sorted by how recently a matched sample image was last presented (index -1 is the sample immediately preceding the test image, i.e., is most recent). Trials in which the test matched a distractor are aggregated together as preceding distractors. Individual session data (points) for sample indices (-1 through -7) were fit by linear regression (blue line; see Results). (C) Residuals from the fit are indistinguishable from zero for all sequence indices ($P > 0.05$, two tailed t-test). (D,E) Performance on four-image sequences (three samples + one test) during a subsequent test with shorter sequences and fewer images (10 sessions; 33% of trials were four-image sequences, five available images). In this case, trials in which the test matched a preceding distractor are also sorted by sequence index, where index -4 is the test from the previous trial and index -5 is the last sample from the previous trial. Response rates at the sample indices (-1 through -3) were fit by linear regression, which was then projected to preceding-distractor indices (-4 through -9) to predict residuals (E). (*) $P < 0.05$, two-tailed t-test. (F,G) Performance of four monkeys on the long-sequence test (F) and five monkeys on the short-sequence test (G). Performance of monkey "B" (dashed line) was ascertained after extensive testing in Match Any and Match First tasks. Lines indicate the mean of 10 sessions for each subject. Red and blue bars indicate group means at each sequence position.

to 76%). The yes-response rates (i.e., hit rates) increased monotonically with test recency in trials in which the test matched a sample (linear regression $r^2 = 0.55$, slope = 18.2, $P < 0.05$ [$P = 3.0 \times 10^{-6}$]) (Fig. 1D, solid blue line). Extrapolation of the regression line to preceding-distractor indices predicted the yes-response rate (i.e., false alarm rate) for trials in which the test matched either of the last two images from the preceding trial (at index −4 and −5 [Fig. 1D, dashed red line vs. red bars]; residuals at index −4 and −5 indistinguishable from zero by two-tailed t-test, $P > 0.05$ [$P = 0.32$ and 0.46, respectively][Fig. 1E]). The regression predicted yes-response rates for preceding distractors whether or not the previous trial resulted in reward (residuals indistinguishable from zero at index −4 and −5 by two-tailed t-test for trials following a rewarded trial [$P = 0.08$ and 0.31, respectively] and for trials following a time-out [$P = 0.45$ and 0.62]).

The monkeys appeared to use a strategy of reporting how recently a test image last appeared irrespective of whether it matched a sample or distractor. The monkeys frequently responded yes when the test image most recently appeared one or a few images ago, and they frequently responded no when the test image last appeared several (>5) images ago. The strategy yielded a high proportion of correct responses because sample images immediately preceded the test image, and distractor images were last presented at least four images prior. The continuous recency trend from samples to distractors suggests that sample images and distractor images are treated indiscriminately; therefore this strategy does not appear to depend on selective memorization of the sample images.

Response rates from each of the other four monkeys could also be fit using a linear regression on test recency. The slopes from each of these fits were indistinguishable across monkeys, therefore a single regression slope could be used to describe the group for the long sequences with 100 available images (slope range 4.9–6.8, group slope = 6.0, $P > 0.05$ [$P = 0.41$]; ANCOVA) (Fig. 1F) and for the short sequences with five available images (slope range 12.1–18.8, group slope = 16.6, $P > 0.05$ [$P = 0.43$]) (Fig. 1G). In four of five monkeys, the linear response trend for sample indices from the current trial predicted yes-response rates when the test matched either of the last two preceding-distractors from the previous trial (residuals at indices −4 and −5 indistinguishable from zero by two-tailed t-test, $P > 0.05$). The single outlier of the group (monkey T) also exhibited linear response trend with respect to test recency, although the trend from the sample indices underestimated yes-response rates at indices −4 and −5. Included among the four monkeys with a successful fit to indices −4 and −5 was one monkey (monkey B) (Fig. 1G, dashed line) that was tested after extensive exposure to Match Any (>200 sessions) and Match First (>200 sessions).

One question that arises is whether the linear response trend is related to the passage of time, or to the number of intervening images. Several events transpired between the end of one sequence and the beginning of the next: motor actions (the monkey released and reengaged the touch bar), reinforcement (reward or time-out administered), and visual stimuli (absence of images during an inter-trial-interval and the brief appearance of a "trial start" instruction target) (see Materials and Methods). Collectively, these events made the inter-

val between trials considerably longer (4–12 sec) than the interval between images within a trial (2 sec). Sorting the data by sequence index (i.e., as presented in Fig. 1G) effectively ignores the time difference between trials, yet extrapolation of the linear response trend predicts performance at indices −4 and −5 in four of five monkeys. When the same data are sorted by time of image presentation, the linear response trend underestimates performance at positions −4 and −5 in four of five monkeys (residuals significantly different from zero, $P < 0.05$). Therefore the linear trend fits the data better when the independent variable is sequence index rather than time. Overall, when performing Match Any, monkeys appear to use a strategy of reporting how recently a test image last appeared in terms of number of images prior.

Match First

For Match First, the monkeys watched sequences of two to eight images drawn from the same image set as was used in Match Any. However, Match First required that the monkeys report whether the last image in the sequence, the test image, matched the first image in the sequence (see Fig. 5A in Materials and Methods). When the test matched an image that appeared between the sample and test, we labeled it as an intervening distractor. When the test image matched an image preceding the current trial, we labeled it as a preceding distractor. In Match First, the test image always matched a previously presented image: In half of the trials, the test matched the sample, and in the other half of the trials, the test matched either an intervening or preceding distractor.

We carried out a pilot Match First experiment with one monkey that had minimal prior exposure to Match Any (10 sessions compared to at least 100 for all other monkeys). Performance increased slowly over the course of 125 sessions, reaching an asymptotic level of 73% correct for a six-image sequence, a performance significantly better than chance (i.e., 50% correct, $\chi^2 = 101$, df = 1, $P < 0.05$). Performance was 81% correct when the test matched a sample and 65% correct when the test matched any distractor (Fig. 2A). In some trials the test matched the last intervening distractor in the sequence, such that the image appeared twice consecutively (i.e., the sequence "ABCDEE," where each letter represents a different image in the sequence), a condition that
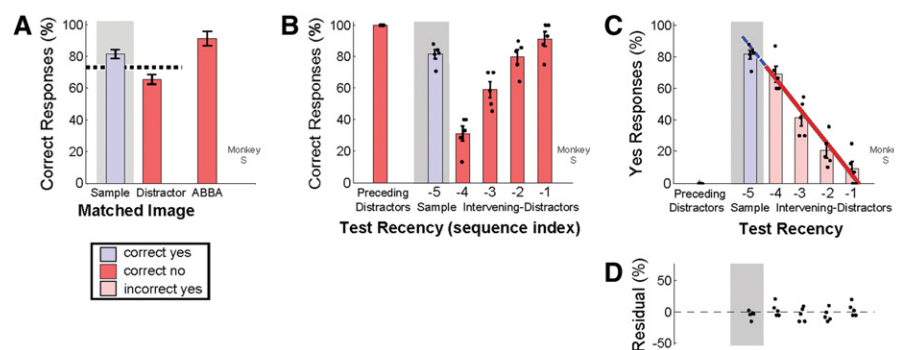
Figure 2. Pilot Match First response rates depend linearly on recency. Performance of one monkey ("S") on six-image sequences (one sample + four distractors + one test) during a pilot Match First experiment (15 daily sessions; 10% of trials were six-image sequence, 100 available images, all images presented for 1 sec; for trials where the test matched a distractor image, 5% were preceding distractors, and the remaining were equally distributed among intervening-distractor indices). (*A*) Percent correct: overall (dashed line), and for trials in which the test matched the sample, any distractor, or just the last distractor in the sequence (the ABBA-like condition, which in our task would be an ABCDEE trial). (*B–D*) Same dataset as *A*, but with trials sorted by test recency and performance reported as percent correct (*B*) or percent yes responses (*C*). Session data (points in *B–D*) are computed from three consecutive testing days for sufficient numbers of trials to estimate performance at each index. (*C*) Linear regression of response rates at the intervening-distractor indices (−1 to −4, red line) is projected to the sample index (−5, dashed blue line) to compute the predicted residuals (*D*).

mimicked the consecutive "B" images in the ABBA task of Miller and Desimone (1994). The monkey correctly rejected 91% of the test images that matched the last intervening distractor in the sequence, a performance measure that has been interpreted to reflect the use of working memory to remember the sample image while ignoring the distractors (Miller and Desimone 1994).

We examined how the rate of correct responses varied with respect to test recency during the pilot experiment (Fig. 2B). The monkey correctly rejected all trials in which the test image matched a preceding distractor (100% correct, "ABCDEF" sequences), and correctly responded yes most often when the test matched the sample (81% correct at index −5, "ABCDEA" sequences). The performance when the test matched an intervening distractor varied, with poorest performance for intervening distractors that occurred early in the sequence (index −4, "ABCDEB"), and best performance for those that occurred late in the sequence (index −1, "ABCDEE," the ABBA-like trials).

We examined the data with performance reported as the rate of yes responses rather than the rate of correct responses. A yes response for a matching distractor is a false alarm. The yes-response rates when the test matched an intervening distractor (false alarms) followed a negative linear dependence on test recency (linear regression $r^2 = 0.82$, slope = −20.1, $P < 0.05$ [$P = 3.5 \times 10^{-8}$]) (Fig. 2C, solid red line). Extrapolating the line to the first image in the sequence (dashed blue line) accurately predicted the yes-response rate for sample images (hits; residuals indistinguishable from zero, two-tailed t-test, $P > 0.05$ [$P = 0.26$]) (Fig. 2D). Thus, it appears that the linear trend in response rate as a function of test recency accounts for both the distribution of false alarms and the proportion of correctly identified sample images.

Until we saw this result we had assumed the monkey's high overall performance meant it was performing the task as intended, by selectively remembering the first image but not remembering the distractor images. Initially, we had been reassured by the monkey's ability to correctly reject a test image that matched the last intervening distractor in the sequence, because this was the criterion previously applied to monkeys performing the ABBA working memory task (Miller and Desimone 1994). In solving Match First, the monkey appeared to respond yes when the test matched an image that had been presented in the neighborhood of the first image, irrespective of whether it was a sample or distractor; we call this the neighborhood strategy.

We modified the Match First task so that monkeys trained after the pilot monkey would selectively remember the first stimulus. We increased the duration of the first image from 500 msec to 2 sec to enhance its salience, which makes the timing different than the timing used in the original ABBA task (Miller and Desimone 1994), where all stimuli were presented for 500 msec. In trials where the test matched a distractor, we increased the likelihood that the test image would match the second image to emphasize the distinction between the first and second images of a sequence; under this condition the test matched the second image in about half of the distractor trials. Even with these changes, five monkeys used the neighborhood strategy early in training (in the same manner as was seen in the pilot experiment). However, performance exceeded the maximum levels from the pilot experiment in 24–74 training sessions (average 42 ± 10 sessions to attain ≥70% correct on six-image sequences). The monkeys were then tested with intermixed sequence lengths of four, six, and eight images for 50 consecutive sessions. We increased the number of test sessions relative to Match Any to gather enough data to estimate error rates for all distractor indices in the sequences. The mean performance on eight-image sequences during this test was 75 ± 2% ($N = 5$) and errors were evenly distributed among trials in which the test matched the sample, or matched a distractor (Fig. 3A, example monkey F from Fig. 1A–E).

With the modified version of the Match First task, yes-response rates still followed a negative linear dependence on test recency when the test matched an intervening distractor (linear regression $r^2 = 0.27$, slope = −3.2, $P < 0.05$ [$P = 2.4 \times 10^{-5}$], example monkey F) (Fig. 3B). However, the response when the test matched the sample image (index −7) was now greater than that predicted by linear extrapolation, such that the residuals at the sample position were significantly different from zero (two-tailed t-test, $P < 0.05$ [$P = 7.6 \times 10^{-9}$]) (Fig. 3C). The break from linearity at the sample index suggests that the monkey selectively memorized the first image. However, the persistence of the linear trend for trials in which the test matched an intervening distractor indicates that the monkey continued to use the neighborhood strategy as well. Projection of the linear trend to the sample index accounts for 33% of the correct responses when the test matched the sample (26% yes predicted, 81% measured). Thus, modifying the Match First task by increasing the duration of the first image and the proportion of second-image distractors appears to have
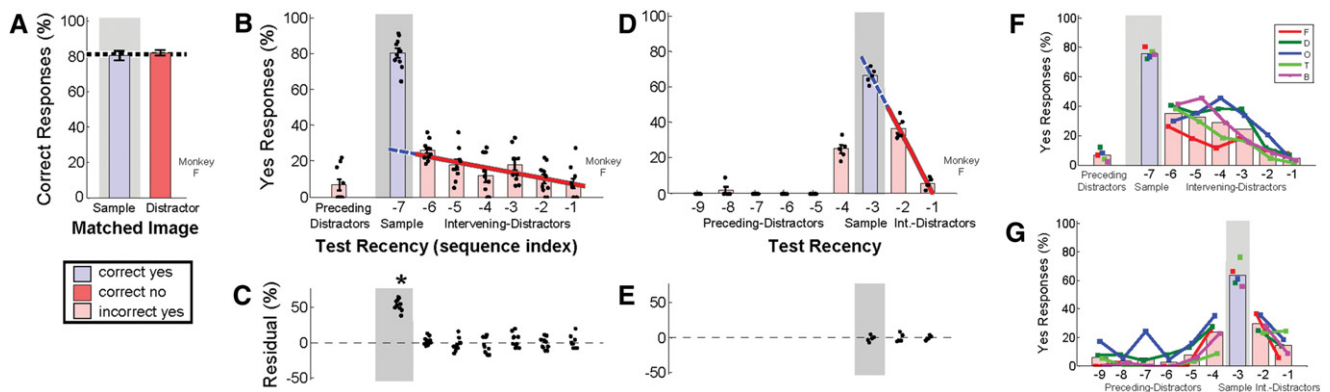
Figure 3. Match First response rates can exceed predictions from linear regression. (*A–C*) Performance on eight-image sequences (one sample + six distractors + one test) for one monkey ("F") during the initial test with long sequences (50 daily sessions; 33% of trials were eight-image sequences, 100 available images, sample image presented for 2 sec, distractor and test images for 500 msec; for trials where the test matched a distractor, 5% were preceding distractors, half of the remaining matched the second image and the other half were equally distributed among all other intervening-distractor indices). Session data (points) are computed from five consecutive testing sessions for sufficient data to estimate response rates at each index. (*D,E*) Similar measures for the same monkey on four-image sequences (one sample + two distractors + one test) during a subsequent test with shorter sequences and fewer images (20 daily sessions; 33% of trials were four-image sequences, five available images, data points computed from four consecutive sessions). (*F,G*) Performance of five monkeys on the tests in *B* and *D*.

caused the monkey to supplement the neighborhood strategy, but not to abandon it.

Next, we reduced the sequence length from eight to four and number of available images from 100 to five, so that preceding-distractor images were likely to have occurred within six sequence indices prior to the sample image. This reduced number of available images more closely resembled the six available images used in the original ABBA task (Miller and Desimone 1994). Average Match First performance during four-image sequences dropped from 89 ± 1% with 100 images to 75 ± 2% with five images ($N = 5$, example monkey F dropped from 88% to 77%). With five available images, the linear trend of response rates from intervening-distractor indices now predicted the yes-response rate at the sample index for example monkey F from Figure 3B (linear regression $r^2 = 0.93$, slope = -30.9, $P < 0.05$ [$P = 8.0 \times 10^{-6}$]; sample residuals indistinguishable from zero by two-tailed t-test $P > 0.05$ [$P = 0.59$]) (Fig. 3D,E). Error rates when the test matched a preceding distractor were highest at the index just prior to the sample (25% yes at index -4). Reducing the number of available images caused the monkey to revert to exclusive use of the neighborhood strategy, where the neighborhood includes distractor images before and after the sample image.

The yes-response rates followed a negative linear dependence on test recency when the test matched an intervening distractor for all five monkeys tested on the Match First task with eight-image sequences and 100 available images (average linear regression slope -6.4 ± 2.2, $N = 5$) (Fig. 3F). The yes-response rate when the test matched the sample was higher than predicted by extrapolation of the linear trend (sample residuals significantly different from zero for each animal, two-tailed t-test, $P < 0.05$). The neighborhood strategy accounted for 60 ± 17% of the correct responses when the test matched the sample. When the number of available images was reduced from 100 to five, three of five monkeys exclusively used the neighborhood strategy (sample residuals indistinguishable from zero in three of five monkeys), and across the group of five monkeys the neighborhood strategy accounted for 72 ± 28% of the correct responses when the test matched the sample (Fig. 3G). The largest error rates occurred when the test matched a distractor that occurred immediately before or after the sample image, such that the linear response trend for intervening-distractor indices appears to be mirrored for preceding distractors.

Control for training history: Match First without Match Any

Up to this point, all data have been reported from monkeys that were trained to perform Match Any prior to Match First. This training history differs from that of the original ABBA task, where monkeys were initially trained to perform a standard version of delayed-match-to-sample that did not include repeating distractor stimuli (Miller et al. 1993; Miller and Desimone 1994). Perhaps training history accounts for the use of the neighborhood strategy in Match First, although this would require that the monkeys change the slope of their yes-response trends from positive to negative. Above, we reported that the very first monkey trained on Match First had limited prior exposure to Match Any (10 sessions), and it was from this animal that we recognized the neighborhood effect. Observing the neighborhood effect after such a modest prior exposure to Match Any decreases the likelihood that training order was responsible for the neighborhood effect in Match First, but it does not eliminate that possibility.

To control for the order of task experience, we trained one experimentally naive monkey to perform Match First without any prior exposure to Match Any. During the initial phase of training

we mimicked the training procedure of the original ABBA task by using a variant of Match First in which the test image never matched an intervening distractor; even though intervening distractors were presented between the sample and the test, the test always matched either the sample or a preceding distractor. Otherwise the task was the same as standard Match First (100 available images, first image 2 sec in duration, distractor and test images 500 msec). Performance increased over the course of 39 sessions to >85% correct on sequences up to six images long.

We introduced trials in which the test matched an intervening distractor in three phases: first as nonreinforced catch trials (15% of all trials, 33% of nonreinforced trials) that ended without feedback, to probe whether standard DMS is solved using a neighborhood strategy (Fig. 4A, gray bars); second as reinforced trials (40% of all trials, similar to standard Match First) that ended with a liquid reward for correct no responses and a time-out for incorrect yes responses (Fig. 4C); and third as reinforced trials with the first image the same duration as all other images in the sequence to mimic the condition of the original ABBA task (six-and four-image sequences in Figure 4, E and G, respectively). In all phases of testing, yes-response rates at intervening-distractor indices were well fit by a linear regression on test recency (residuals for intervening distractors indistinguishable from zero, two-tailed t-test, $P < 0.05$) (Fig. 4B,D,F,H). In the first phase of testing, response rates at intervening-distractor indices were indistinguishable from the response rate when the test matched the sample (Fig. 4A,B), suggesting that the standard version of delayed-match-to-sample is also solved based on recency of stimulus repetition. In the second phase of testing, the response rate when the test matched the sample (index -5) was greater than what would be predicted by linear extrapolation from the intervening-distractor response rates (Fig. 4C,D). In the third phase of testing, the response to the sample image was less than or equal to what would be predicted by linear extrapolation (Fig. 4E–H). The neighborhood strategy is used during Match First whether or not a monkey is initially trained to solve Match Any. Nonetheless, a monkey capable of displaying enhanced detection of the first image will revert to exclusive use of the neighborhood strategy if all of the images are the same duration.

Discussion

We trained monkeys to perform two tasks, Match Any and Match First, with the intention that they would remember one or more sample images and forget or ignore distractor images. The monkeys performed well on both tasks. Examination of the distribution of correct responses revealed that the monkeys performed best when the test image matched an image that occurred late in the sequence, whether or not that image was a sample image. In Match Any, samples presented near the end of the sequence were most likely to be correctly detected, whereas samples presented near the beginning of the sequence were less likely to be correctly detected. In Match First, distractors presented near the end of the sequence were most likely to be correctly rejected, whereas distractors presented near the beginning of the sequence were less likely to be correctly rejected. In both tasks, the likelihood of a correct response increased linearly with recency of stimulus repetition.

We also analyzed the data in terms of yes-response rates, which revealed that the monkeys solved both tasks based on when a repeated stimulus was last presented, not whether the repeated stimulus had been a sample or distractor. In Match Any, yes responses were most likely when the test matched the most recent sample, the last image in the sequence, and the proportion of yes responses decreased linearly as the interval between the sample
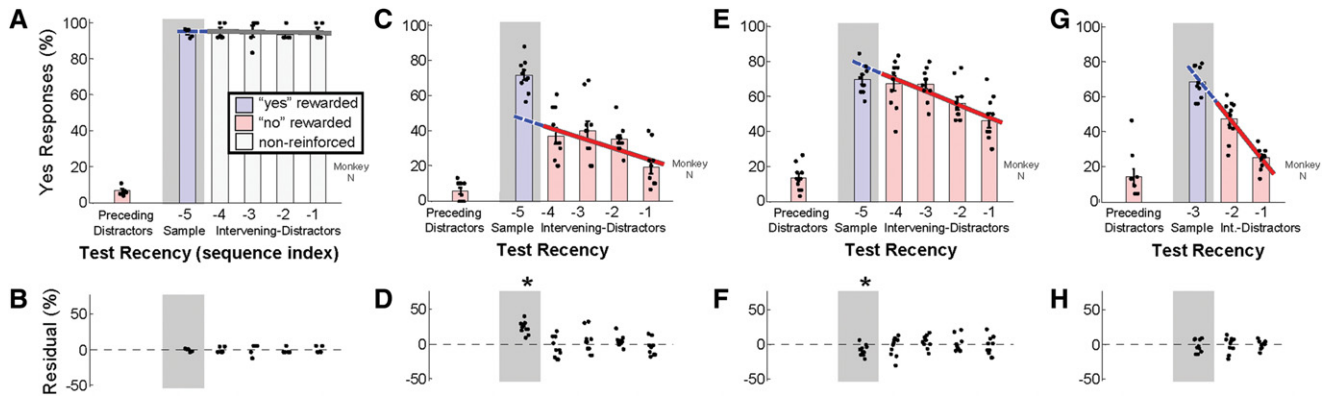
Figure 4. Match First is solved using recency without any prior exposure to Match Any. Performance of one monkey ("N") that was trained to perform Match First without any prior exposure to Match Any. Initial training consisted of sequences of up to six images (one sample + four distractors + one test), where the test image matched either the sample or a preceding distractor—although intervening distractors were present between the sample and test, the test never matched an intervening distractor. Otherwise, the task resembled Match First (100 available images, sample image presented for 2 sec, distractor and test images for 500 msec). (*A*,*B*) Performance on six-image sequences in which half the trials ended in standard reinforcement (for correct and incorrect responses) and half did not (data from five consecutive sessions after performance stabilized). In reinforced trials the test image matched the sample or a preceding distractor (blue and red bars, respectively). One-third of the nonreinforced trials were catch trials in which the test matched an intervening distractor (gray bars). Linear regression of response rates at intervening-distractor indices (−1 to −4) is projected to the sample index (−5) to predict residuals (*B*), (*) $P < 0.05$, two-tailed *t*-test. Regression $r^2 = 0.00$, slope = −0.2 is indistinguishable from zero, $P > 0.05$ ($P = 0.870$). (*C*,*D*) Performance on six-image sequences after introducing reinforced intervening-distractor trials (10 consecutive sessions after achieving criterion performance of 70% correct on a mix of four- and six-image sequences, sample 2 sec, other images 500 msec). Regression $r^2 = 0.19$, slope = −5.9, $P < 0.05$ ($P = 5.0 \times 10^{-3}$). (*E*,*F*) Performance on six-image sequences after making all images the same duration (10 consecutive sessions of only six-image sequences, all images 500 msec). Regression $r^2 = 0.34$, slope = −7.4, $P < 0.05$ ($P = 8.9 \times 10^{-5}$). (*G*,*H*) Performance after reducing the sequence length to four while keeping all images the same duration (10 consecutive sessions of only four-image sequences, all images 500 msec). Regression $r^2 = 0.63$, slope = −22.3, $P < 0.05$ ($P = 2.9 \times 10^{-5}$).

image and the test image increased. This linear trend continued when the test matched a distractor stimulus from the preceding trial, that is, stimuli for which a yes response was a false alarm. In Match First, false alarms were most likely when the test matched a distractor adjacent to the first image, and the proportion of false alarms decreased linearly as the interval between the distractor and the test image increased.

The rates of hits and false alarms in both Match Any and Match First were fit well by straight lines, with a peak at the last image for Match Any and the first image for Match First. The peak response rates when the test matched the last or first image suggest that our monkeys remember the time point at which each trial ended or started (for Match Any and Match First, respectively). Monkeys are adept at estimating the intervals between task events, including the times that trials start and end (Bayer and Glimcher 2005; Janssen and Shadlen 2005). We propose that the straight-line fits in both Match First and Match Any reflect a single underlying memory mechanism that provides a percept of how recently a test image was previously viewed, with respect to a reference time point. The difference between the behavior in the two tasks reflects the difference in the reference point: the last image in Match Any and the first image in Match First.

Our hypothesis posits that the monkeys memorize all images viewed nonselectively, irrespective of sample or distractor status, in both Match Any and Match First. The lack of an active, sample-selective, memory process for Match First suggests the use of a passive memory process, as has been proposed for Match Any by others (Wright 1998; Yakovlev et al. 2005). Interference is a passive memory process that relates the success of memorizing (or not forgetting) a particular item to the number of items stored before and after that item (Crowder 1976; Wright 1998). Within-sequence interference, where later items interfere (retroactively) with memory of earlier items, has been used to explain recency trends with both positive and negative slopes (Wright 1998), suggesting it could apply to our observations for Match Any and

Match First. Across-sequence interference, where earlier memorization interferes (proactively) with memory of later images (Wright et al. 1986), has been used to explain decreases in performance when images are drawn from a limited set that is viewed over-and-over again, which we observed in both Match Any and Match First. If the linear response trends in Match Any and Match First reflect a single underlying memory process, then that process would need to be applied nonselectively to all images viewed, and it would need to yield a percept of how recently an image was last viewed.

An alternative interpretation is that different memory processes support Match Any and Match First, with Match Any depending on recency of repetition and Match First depending on selective memorization of the sample image (Miller and Desimone 1994; Yakovlev et al. 2013). In Match First, the monkeys correctly responded yes most often when the test matched the sample image, with a "halo" of false alarms around the ordinal position of the sample that extended to distractors. The peak in yes-response rates when the test matched the sample argues that the monkeys "understood" the task rule. If the monkeys were selectively remembering the sample image, then the halo of errors would reflect imperfect selectivity when storing and/or maintaining the sample image. Although our data do not refute the selective memory hypothesis, based on parsimony we favor the interpretation that all images are memorized indiscriminately (i.e., no special storage mechanism for sample images), and that response rates reflect the proximity between the time the trial started and the recency of the test image (i.e., an estimate of "first-ness"). Our favored interpretation accounts for the linear response trends in both Match First and Match Any using a single memory percept, recency of repetition.

By manipulating conditions in the Match First task, we were able to coax the monkeys into selectively responding to the sample image in a way that cannot be explained solely with recency of repetition. This enhanced response to the first image relative to

the projected linear trend was seen only when we made three modifications: increased the duration of the first image to four times that of the distractors, increased the likelihood that the test matched a distractor neighboring the first image, and used a large stimulus set. Each of these modifications enhanced the saliency of the sample, which should counteract passive interference. Therefore, the enhanced response to the first image may reflect selective protection from passive interference mechanisms. Alternatively, the enhanced response might reflect selective memorization of the sample image via active rehearsal, making it akin to human working memory (Logie 2011; Baddeley 2012). Selectively remembering a sample while ignoring distractors is a common requirement in working memory tasks for humans, and humans appear to be very good at doing this (Jonides and Nee 2006; Ecker et al. 2010). However, the best the monkeys ever did was to use a mix of selective and nonselective memory. When experimental conditions were most favorable for selective memorization, it accounted for a small, but nonnegligible, fraction of the correctly identified sample images. Therefore, even though monkeys appear to be capable of selectively memorizing the sample image, response decisions are often (or exclusively) made based on the recency of stimulus repetition.

Our results lead to an interpretation of how monkeys perform Match First that is at odds with previous studies. Match Any and Match First resemble, but were not identical to, tasks used to study nonspatial short-term memory in monkeys in the past: the delayed-match-to-multiple-sample task (Amit et al. 2003; Yakovlev et al. 2005) and the ABBA task (Miller and Desimone 1994; Miller et al. 1996; Suzuki et al. 1997). In the original tasks, images were sequentially presented on a video screen and monkeys had to hold a touch-bar until a sample image was repeated. Early release was a false alarm and late release was a miss. In the terminology presented here, to complete a trial correctly the monkeys had to respond no to every image before the sample (by continuing to hold the bar) and then respond yes to the sample. In our variants of the tasks, monkeys were cued to make a response to a single test image, that is, the monkeys had to make a yes–no decision only once per trial. Overall performance and the proportion of hits and false alarms with Match Any and Match First were similar to reports with the original tasks. For example, in the delayed-match-to-multiple-sample task: Overall performance of two monkeys for sequences between four and eight images long was between 75% and 85% correct (Yakovlev et al. 2005), performance appeared to us to depend linearly on test recency, and false alarms were most likely when an image matched a distractor from the immediately preceding trial. In the ABBA task, overall performance of two monkeys on just the ABBA trials was >85% correct in one study (Miller and Desimone 1994) and the average performance of two monkeys from a subsequent study (only one of which performed ABBA) was 76% correct, with false alarm errors (18% of all trials) more common than misses (6%) (Suzuki et al. 1997). These ABBA performance levels are similar to those we observed in Match First when all images were the same duration or the stimulus set was small, both of which were conditions used in the original ABBA task. Thus, the monkeys performing the original short-term memory tasks were likely to have been using the same strategies we observe here.

Neurophysiological recordings in monkeys show that some stimulus-selective neurons in the inferior temporal cortex encode the relative familiarity of a stimulus (Fahy et al. 1993; Li et al. 1993; Kaliukhovich and Vogels 2011). Perhaps this neural response supports our monkeys' ability to recognize when a test image was last viewed. There is also activity in a subset of neurons in the monkey inferior temporal cortex and prefrontal cortex that appears to encode whether a test image matches a sample and not a distractor (Miller and Desimone 1994; Miller

et al. 1996). Perhaps such sample-selective neuronal responses support our monkeys' ability to selectively remember the first image. However, the neurophysiological recordings were made under conditions where we see no evidence of a selective memorization in our experiments (i.e., all images the same duration and a reduced stimulus set). One possibility is that neurons are sample selective even though the monkey is not. Another possibility is that the sample-selective neurophysiological response reflects the same process that causes a peak in yes-response rates when the test matches the sample. To gain an understating of how sample-selective neuronal activity is related to the behavior reported here will almost certainly require recording neurons while the recency of intervening distractors is varied parametrically.

Measurements of behavioral performance in nonspatial delayed-match-to-sample tasks have been interpreted as showing that monkeys remember the sample using working memory, that is, they use selective memorization (Miller and Desimone 1994; Amit et al. 2003; Schneider et al. 2009; Verrico et al. 2011). However, this interpretation has been made even when the delayed-match-to-sample task does not include intervening distractors (Schneider et al. 2009; Verrico et al. 2011), which are critical for distinguishing selective memorization from the neighborhood strategy. Even tasks that included intervening distractors often did not repeat those images as the test image (Miller et al. 1993; Woloszyn and Sheinberg 2009). When we trained one monkey to perform a version of Match First where intervening distractors were included in each trial but not repeated as the test image, the monkey seemed to be reporting (with 95% accuracy) whether or not the test image had been presented recently (cf. Fig. 4A). The original version of the ABBA task (Miller and Desimone 1994; Miller et al. 1996) used intervening-distractor images that were repeated one or two images later (so called "ABBA" or "ABCBA" trials, where the second "A" was only presented after correct rejection of the second "B"). Performance was reported to be lower in ABBA trials early in training, and the authors concluded that their monkeys had been solving delayed-match-to-sample based on recency of stimulus repetition. Once performance increased to 85% correct on ABBA trials it was assumed that the monkeys were using working memory to maintain just the "A" while suppressing automatic or passive memories of the intervening distractors. We only became aware of how the monkeys were really solving our task when we systematically evaluated performance in terms of yes-response rates for all possible stimulus positions.

Our results raise the possibility that the linear response trends reported above in both Match Any and Match First are supported by a single memory mechanism that reflects how recently a stimulus was last viewed. This may be akin to the mechanism underlying familiarity judgments in humans, a judgment that is thought to reflect a continuous index of memory strength (Yonelinas 2002), even over short time intervals (Oberauer 2008). Our results suggest that monkeys performing nonspatial visual short-term memory tasks are incapable of ignoring information about recency, and have little or no capacity for the active memory processes that humans willfully employ when storing information in working memory.

Materials and Methods

Subjects

Seven adult (4–10 kg) rhesus monkeys were used (five male, two female). All the experimental procedures were carried out in accordance with the ILAR Guide for the Care and Use of Laboratory Animals and approved by the Animal Care and Use Committee of the National Institute of Mental Health.

Behavioral training

For all behavioral training and testing, monkeys squatted in a primate chair inside a sound-attenuated dark room. Visual stimuli were presented on a computer video monitor in front of the monkey. Experimental control and data acquisition were performed using the REX program (Hays et al. 1982). During the testing period, the monkeys' liquid intake was controlled to ensure adequate motivation to perform the behavioral tasks while maintaining a healthy body weight. Monkeys were tested 5–7 d/wk with daily sessions 90–120 min in duration with 200–500 trials/session, depending on task parameters. There were fewer trails in sessions with longer sequence lengths.

Naive monkeys were initially trained to perform a simple sequential visual color discrimination task in which they were rewarded for releasing a bar within 1 sec when a small red square turned green. After reaching 80% correct, the task was transformed into an "A-not-A" pattern discrimination task. Images appeared and disappeared behind the red square in the center of the screen. The red square turned green while one of the images was present. If the image was the positive one, say a dog, the correct response was to release during the 1-sec duration of the green square (a yes response). If the green spot appeared with any other image, the correct answer was no and the correct response was to release the bar after the green spot disappeared. We were concerned that the monkeys might adopt a strategy of intentionally looking away from distractor images to improve performance. Therefore we interleaved different sequence lengths to make such a strategy difficult. In practice, the behavioral data indicate that monkeys viewed and remembered all images presented (see Results).

After practice, even the most delayed yes responses were easily distinguishable from no responses. For example, during the first 10 sessions of Match Any testing (cf. Fig. 1), monkey F responded yes 1381 times and no 1473 times, with 90% of the yes response times between 292 and 438 msec of the green square appearing, and 90% of the no response times between 1213 and 1601 msec of the green square appearing (i.e., 213–601 msec after the square extinguished). Correct responses were reinforced with liquid reward and incorrect responses resulted in a time-out (4–30 sec, depending on task and motivation level). A blue square appeared and remained on the screen during the reinforcement. If the monkey released the bar before the green square appeared, or 5 sec after it extinguished, the trial was considered aborted and no reward or time-out was administered. Monkeys aborted <1% of all trials (e.g., monkey F aborted 15 of 2869 trials during the first 10 sessions of Match Any testing). In this and all subsequent tasks, trial types were distributed evenly so 100% correct performance would require an equal number of yes and no responses. Some animals exhibited an inherent bias toward yes or no responses. The bias was ameliorated by adjusting the relative volume of liquid reward for each response type. The monkeys' heads were unconstrained and eye position was not tracked during the experiments reported in this manuscript.

Match Any and Match First

Five of seven monkeys were trained in the same manner. After performing above 85% correct for 3 d on the pattern discrimination task, they began Match Any training and testing (~100 sessions over 4 mo) followed by Match First training and testing (~200 sessions over 7 mo). Monkey S piloted the Match First task after only 10 sessions of preliminary Match Any training and testing, and monkey N was trained on a variant of Match First with no prior exposure to Match Any. Monkey B was tested with longer maximum sequence lengths in Match Any (11 images instead of eight, no four-image sequences). Its long-sequence data are not included, and its short-sequence data were collected after completing additional Match Any and Match First training.

Match Any: Answer yes when the test image matched any of the other images from the sequence in the current trial, and answer no otherwise (Fig. 5A). Training started with a two-image sequence, where the monkeys indicated whether the second image matched the first image. Performance of five monkeys was
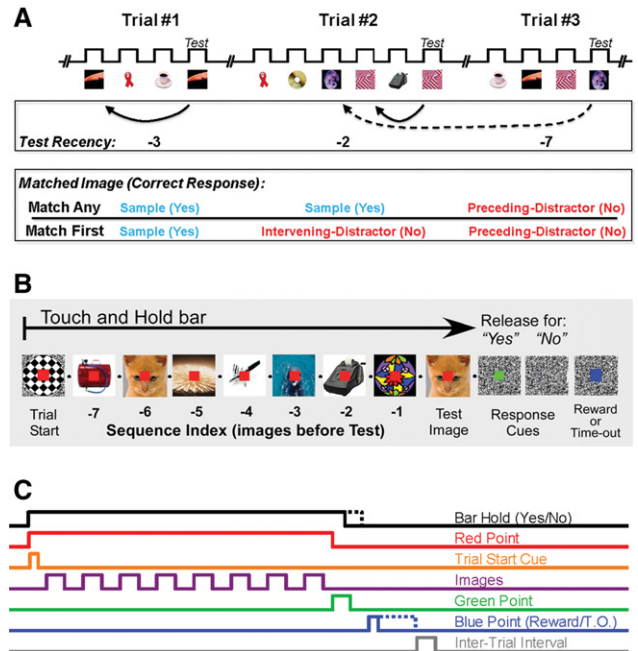
Figure 5. Match Any and Match First have the same trial structure. (*A*) Schematic example of three consecutive trials: a four-image sequence, followed by a six-image sequence, and then a four-image sequence. Trace indicates timing of image onsets and offsets, with breaks representing the interval between trials. In both tasks, the test image (last image in each sequence) always matches a previously presented image; test recency quantifies when the test image was last presented in terms of sequence index. The correct response at the end of each trial depends on whether the test matches a sample ("yes" correct) or distractor ("no" correct). In Match Any, each trial contains a series of sample images and ends with a test image, which matches either a sample from the current trial or a distractor that was last viewed in a previous trial (a preceding distractor). In Match First, only the first image is a sample and the images between the first and the test are intervening distractors. In Match First, the test image matches a sample, an intervening distractor, or a distractor that was last viewed in a previous trial. For Match Any, the correct responses to the three trials would be yes, yes, no, whereas for Match First, the correct responses to the same three trials would be yes, no, no. (*B*) Schematic of images and instruction targets during an example eight-image sequence. The trial begins when the monkey touches a bar, which triggers the presentation of a red instruction target (red square), before a trial-start instruction target (checkered circle) is presented and removed. While the red target is on, up to eight images are presented and then removed in sequence (filled bullet [•] indicates an interval with no image; total sequence duration up to 16 sec). The red target turns green to indicate that the last image viewed was the test image. A "yes" response is registered when the monkey releases the bar during the 1-sec-long green target; a "no" response is registered when bar release is delayed until after the green target disappears. Correct responses are rewarded with juice or water. Incorrect responses result in a 4-sec to 30-sec time-out. A blue target appears after a response has been made and remains on-screen throughout the feedback (whether reward or time-out). (*C*) Timing of actions, images, instructions targets, and feedback.

above chance within 13 sessions (mean 6.2). We incrementally increased the maximum sequence length to eight while randomly interleaving short sequences. After training, four monkeys were tested with 10 consecutive sessions of eight-image sequences (seven samples + one test image) randomly interleaved with shorter sequences (70% of trials were eight-image sequences). Subsequently, they were tested for five sessions on shorter sequence lengths (two, three, and four, randomly interleaved) with 100 available images and then 10 sessions with five available images. At this point the timing of the green and blue squares

relative to the test image changed: For all subsequent Match Any testing and Match First training or testing, the test image remained in the background when the red square transitioned to green, as well as during the subsequent reinforcement period (when the blue square was on the screen). Last, the monkeys were tested on just two-image sequences with either 100 or two available images (five and 10 sessions, respectively).

Match First: Answer yes if the test image matched the first image, and answer no otherwise (Fig. 5A). We started training using a three-image sequence and incrementally increased the sequence length (three, four, six; single sequence length at a time) based on performance criteria. During early phases of training, an automated shaping paradigm reduced the duration of intervening-distractor images (not the first or test image) to 100 msec if performance dropped below 50% on the preceding 10 trials. After training, the monkeys were tested for 50 consecutive sessions of four-, six-, and eight-image sequences randomly interleaved within each session. Subsequently, individuals were tested in different conditions (e.g., reduced first image duration, reduced number of available images) as described in Results.

Image sets and stimulus timing

All visual stimuli were jpeg or pcx image files displayed over a black-and-white noise background on a LCD monitor (800 × 600-pixel resolution) using Presentation software (Neurobehavioral Systems). There were two types of stimuli, instruction targets and images (Fig. 5B). The instruction targets were red, green, and blue squares (10 × 10 pixels), which indicated whether a trial was currently running, prompted bar release, and indicated that a trial was complete by coinciding with reward delivery or a time-out period. A trial-start instruction target (100 × 100 pixels) was either circular or square to indicate whether a trial was Match First or Match Any. Only a single task was tested per session in the data presented here, so the shape of this target did not change between trials. Images (200 × 200 pixels) were drawn from a random mix of color photographs and clip art imagery obtained from the Internet. One thousand such images were equally divided into 10 libraries that were used repeatedly during testing and training. A single library of 100 images was used for ~10 sessions before moving to the next. The timing of instruction target presentation was identical for Match Any and Match First (Fig. 5C). For Match Any, the delay between successive images was 1 sec, and all images were presented for 1 sec. For Match First, the delay between images was 750 msec, and all images were presented for 500 msec, except for the first image, which was presented for 2 sec to enhance its salience.

Data analysis and statistics

Data were analyzed and visualized in Matlab (Mathworks). Linear regression models were confirmed to be appropriate by ANOVA analysis of the residuals from the predicted values, and multiple regression slopes were compared by ANCOVA (Zar 2010). Differences were deemed statistically significance at a $P < 0.05$ level.

Acknowledgments

References

Amit DJ, Bernacchia A, Yakovlev V. 2003. Multiple-object working memory—a model for behavioral performance. *Cereb Cortex* **13:** 435–443.

Baddeley A. 2003. Working memory: Looking back and looking forward. *Nat Rev Neurosci* **4:** 829–839.

Baddeley A. 2012. Working memory: theories, models, and controversies. *Annu Rev Psychol* **63:** 1–29.

Bayer HM, Glimcher PW. 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47:** 129–141.

Cowan N. 2008. What are the differences between long-term, short-term, and working memory? *Prog Brain Res* **169:** 323–338.

Crowder RG. 1969. Behavioral strategies in immediate memory. *J Verb Learn Verb Behav* **8:** 524–528.

Crowder RG. 1976. *Principles of learning and memory.* Erlbaum, Hillsdale, NJ.

Ecker UK, Lewandowsky S, Oberauer K, Chee AE. 2010. The components of working memory updating: An experimental decomposition and individual differences. *J Exp Psychol Learn Mem Cogn* **36:** 170–189.

Fahy FL, Riches IP, Brown MW. 1993. Neuronal activity related to visual recognition memory: Long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Exp Brain Res* **96:** 457–472.

Hays AV, Richmond BJ, Optican LM. 1982. A Unix-based multiple process system for real-time data acquisition and control. In *WESCON conference proceedings*, Vol. 2, pp. 1–10.

Janssen P, Shadlen MN. 2005. A representation of the hazard rate of elapsed time in macaque area LIP. *Nat Neurosci* **8:** 234–241.

Jonides J, Nee DE. 2006. Brain mechanisms of proactive interference in working memory. *Neuroscience* **139:** 181–193.

Kaliukhovich DA, Vogels R. 2011. Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cereb Cortex* **21:** 1547–1558.

Li L, Miller EK, Desimone R. 1993. The representation of stimulus familiarity in anterior inferior temporal cortex. *J Neurophysiol* **69:** 1918–1929.

Logie RH. 2011. The functional organization and capacity limits of working memory. *Curr Dir Psychol Sci* **20:** 240–245.

Miller EK, Desimone R. 1994. Parallel neuronal mechanisms for short-term memory. *Science* **263:** 520–522.

Miller EK, Li L, Desimone R. 1991. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* **254:** 1377–1379.

Miller EK, Li L, Desimone R. 1993. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J Neurosci* **13:** 1460–1478.

Miller EK, Erickson CA, Desimone R. 1996. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* **16:** 5154–5167.

Oberauer K. 2008. How to say no: Single- and dual-process theories of short-term recognition tested on negative probes. *J Exp Psychol Learn Mem Cogn* **34:** 439–459.

Sands SF, Wright AA. 1980. Primate memory: Retention of serial list items by a rhesus monkey. *Science* **209:** 938–940.

Schneider JS, Decamp E, Clark K, Bouquio C, Syversen T, Guilarte TR. 2009. Effects of chronic manganese exposure on working memory in non-human primates. *Brain Res* **1258:** 86–95.

Scott BH, Mishkin M, Yin P. 2012. Monkeys have a limited form of short-term memory in audition. *Proc Natl Acad Sci* **109:** 12237–12241.

Suzuki WA, Miller EK, Desimone R. 1997. Object and place memory in the macaque entorhinal cortex. *J Neurophysiol* **78:** 1062–1081.

Verrico CD, Liu S, Asafu-Adjei JK, Sampson AR, Bradberry CW, Lewis DA. 2011. Acquisition and baseline performance of working memory tasks by adolescent rhesus monkeys. *Brain Res* **1378:** 91–104.

Watkins MJ. 1989. Willful and nonwillful determinants of memory. In *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (ed. Roediger HL III, Craik FIM), pp. 59–71. Erlbaum, Hillsdale, NJ.

Woloszyn L, Sheinberg DL. 2009. Neural dynamics in inferior temporal cortex during a visual working memory task. *J Neurosci* **29:** 5494–5507.

Wright A. 1998. Auditory and visual serial position functions obey different laws. *Psychon Bull Rev* **5:** 564–584.

Wright AA, Urcuioli PJ, Sands SF. 1986. Proactive interference in animal memory research. In *Theories of animal memory* (ed. Kendrick DF, et al.), pp. 101–125. Erlbaum, Englewood Cliffs, NJ.

Yakovlev V, Bernacchia A, Orlov T, Hochstein S, Amit D. 2005. Multi-item working memory—a behavioral study. *Cereb Cortex* **15:** 602–615.

Yakovlev V, Amit DJ, Romani S, Hochstein S. 2008. Universal memory mechanism for familiarity recognition and identification. *J Neurosci* **28:** 239–248.

Yakovlev V, Amit Y, Hochstein S. 2013. It's hard to forget: Resetting memory in delay-match-to-multiple-image tasks. *Front Hum Neurosci* **7:** 765.

Yonelinas AP. 2002. The nature of recollection and familiarity: A review of 30 years of research. *J Mem Lang* **46:** 441–517.

Zar JH. 2010. *Biostatistical analysis.* Prentice Hall, Upper Saddle River, NJ.