

Ghosts in the machine: memory interference from the previous trial

Charalampos Papadimitriou, Afreen Ferdoash, and Lawrence H. Snyder

Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri

Submitted 28 May 2014; accepted in final form 30 October 2014

Papadimitriou C, Ferdoash A, Snyder LH. Ghosts in the machine: memory interference from the previous trial. *J Neurophysiol* 113: 567–577, 2015. First published November 5, 2014; doi:10.1152/jn.00402.2014.—Previous memoranda can interfere with the memorization or storage of new information, a concept known as proactive interference. Studies of proactive interference typically use categorical memoranda and match-to-sample tasks with categorical measures such as the proportion of correct to incorrect responses. In this study we instead train five macaques in a spatial memory task with continuous memoranda and responses, allowing us to more finely probe working memory circuits. We first ask whether the memoranda from the previous trial result in proactive interference in an oculomotor delayed response task. We then characterize the spatial and temporal profile of this interference and ask whether this profile can be predicted by an attractor network model of working memory. We find that memory in the current trial shows a bias toward the location of the memorandum of the previous trial. The magnitude of this bias increases with the duration of the memory period within which it is measured. Our simulations using standard attractor network models of working memory show that these models easily replicate the spatial profile of the bias. However, unlike the behavioral findings, these attractor models show an increase in bias with the duration of the previous rather than the current memory period. To model a bias that increases with current trial duration we posit two separate memory stores, a rapidly decaying visual store that resists proactive interference effects and a sustained memory store that is susceptible to proactive interference.

spatial memory; interference; attractor model

WORKING MEMORY INVOLVES MAINTENANCE of goal-related information in an active state over a short period of time. It has been implicated in higher cognitive functions (Miyake and Shah 1999; Engle and Kane 2004; Raghubar et al. 2010; Unterrainer and Owen 2006) such as planning (Cohen 1996; Altgassen et al. 2007; Unterrainer and Owen 2006), decision-making (Payne et al. 1990; Payne et al. 1992), and language comprehension (Daneman and Carpenter 1980; Just and Carpenter 1992; Daneman and Merikle 1996), and hence an understanding of the circuitry of working memory is critical for understanding higher cognition. Spatial working memory provides an excellent model system for this purpose. The memoranda, locations in space, are continuous and well defined, as are the responses that provide a read out for these memoranda. Animals can be easily trained to perform spatial working memory tasks, and neural circuits can then be directly interrogated using various electrophysiological techniques. A number of models for working memory have been proposed (Durstewitz et al. 2000; Miyake and Shah 1999; Baddeley 2012). Here we

present findings that can constrain these models and illuminate the mechanisms by which memory is degraded.

Working memory may be degraded due to interference or distraction from task-irrelevant events that occur before or during the memory period. Prior computational and behavioral studies on memory interference have focused on the spatial working memory system because it provides continuous measures for both stimuli and responses and is well suited to studying partial memory degradation (Compte et al. 2000; Macoveanu et al. 2007; Chumbley et al. 2008). These studies focus on interference within a trial, asking how irrelevant information and distractor stimuli interfere with the memory information of a target stimulus when both are presented in the same trial.

The recurrently connected continuous attractor network (e.g., Wang 2009; Compte et al. 2000) is the premier framework for most computational studies of working memory. Continuous attractor networks have been widely successful in modeling phenomena related to working memory. Individual excitatory nodes within these networks are tuned to particular portions of the visual field. When arranged topographically, these nodes encode memory locations with a spatially localized bump of elevated activity. This bump is a stable attractor state in the system dynamics and can therefore persist indefinitely over time. Stability is achieved through a fully connected recurrent connectivity structure, with strong excitatory connections between cells representing parts of visual space that are close together and weaker connections between cells representing parts of visual space that are farther apart. The excitation is balanced by inhibitory circuits with broad connections that keep the excitatory activity in check and prevent the bump from spreading across the entire network. Continuous attractor models have been used to explore within-trial interference from distractor stimuli. The connectivity structure and bump attractor state in these models predict distractor effects with a very specific spatial profile. The models predict an attractive bias of memory toward the distractor location, with strong bias when remembered and distractor locations are close together and weaker bias when they are progressively farther apart (Compte et al. 2000). The spatial profile of within-trial interference arises because when two bumps of activity, one representing the true remembered location and one representing the distractor response, are close together, nodes between the two bumps receive the most excitatory input, and as a result the bumps merge to form a single bump at an intermediate location. When the two peaks are far apart, intermediate nodes are not excited and the global inhibition quashes the weaker bump (usually the distractor bump) before it has a chance to distort the bump representing the correct memory location. The predicted spatial pattern of within-trial interference has been

Address for reprint requests and other correspondence: C. Papadimitriou, Washington Univ. School of Medicine, Dept. of Anatomy and Neurobiology, Box 8108, 660 South Euclid Ave., St. Louis, MO 63110 (e-mail: papadimitriou.c@gmail.com).

replicated in behavioral studies (Macoveanu et al. 2007; Chumbley et al. 2008).

In addition to within-trial interference from distractors, information previously held in memory has been shown to interfere with current memory performance, a phenomenon known as proactive interference. A classic example of this effect from human literature is described by Underwood (1957), who performed experiments requiring memorization of nonsense syllable lists and found that the more sessions his subjects participated in the worse they performed due to interference from syllables memorized in prior sessions. Animal studies have identified proactive effects that degrade memory performance during categorical memory tasks (Jarvik et al. 1969; Moise 1976; Edhouse and White 1988; Dunnett and Martel 1990; see also Jonides and Nee 2006). In these studies, memorized information from the immediately preceding trial causes interference. These studies typically employ delayed-match-to-sample or recent-probes tasks with a small number (often just two) of categorical responses, and errors in these tasks are all-or-none.

In this study we looked at memory errors due to sequential, between-trial memory interference, whether and how what is remembered on one trial interferes with memory on the next trial, using a continuous oculomotor delayed response task in which subjects maintain a target location in working memory and then indicate the contents of their memory with a rapid eye movement (saccade). We found that saccades were biased toward the location of the target of the previous trial. In contrast with most previous studies, the continuous nature of our task reveals that between-trial proactive interference is graded as a function of the relative distance between the current and previous memory locations. The specific spatial profile of graded interference we identify strongly resembles within-trial errors caused by distractors, indicating that the neural mechanisms involved in between-trial proactive interference and interference due to distractors presented while a memory is already being maintained are the same or strongly related. Attractor network dynamics inherently produce errors with the same spatial structure when multiple memory traces compete within an attractor circuit (e.g., Compte et al. 2000). We therefore propose a specific mechanism for the generation of proactive interference in working memory circuits, showing that a residual ghost of activity from the previous trial competing with the current target representation in an attractor circuit can reproduce the spatial structure of the proactive bias.

We hypothesized that as the memory trace of the previous trial decays over the course of the delay period, proactive interference would decrease. Surprisingly, proactive interference instead increased asymptotically with delay length. This finding is not readily reproduced by a single sustained store. We model this effect using two stores: a biased sustained memory store (e.g., working memory) and an unbiased transient visual sensory store (e.g., something similar to iconic memory). The sustained store depends on attractor dynamics that can maintain a memory trace for long periods but as a result is susceptible to proactive interference. The visual sensory store can maintain a trace for only a few seconds, but because it is unbiased, it can protect memory representations against bias while it is available.

MATERIALS AND METHODS

All experiments were conducted with the approval of the Institutional Animal Care and Use Committee at Washington University in St. Louis.

Five male macaques were trained on a center-out memory-guided saccade task requiring them to remember peripheral target locations continuously distributed around a circle. Once the macaques became proficient at the task (>85% success rate) we recorded the end points of saccades made to the remembered target locations.

Each subject sat in a primate chair in a dark room and was head fixed securely in a straight-ahead position. Visual stimuli were projected onto a white screen 20 cm from the subject. For *monkeys C* and *W*, eye position was recorded with 0.05° resolution every 2 ms using a field coil system. For *monkeys L, D*, and *R*, an ISCAN infrared video eye-tracking system was used to record eye position. Stimuli were controlled by custom software.

Behavioral Task

Each trial began with presentation of a central fixation target, and subjects were required to maintain fixation (within 4°) until it disappeared. After the subject acquired fixation, a peripheral target was presented for 150 ms at a fixed eccentricity (between 10 and 15° depending on the subject) while the subject continued to fixate. In most experiments, the target location was randomly selected in each trial from 360 locations spaced 1° apart. A delay period between 0 and 6 s (depending on the experiment and subject) followed target presentation after which the central fixation target disappeared, and the subject was required to make a saccade to the remembered peripheral target location. Saccadic responses within 6 to 10° (depending on the animal) of visual angle of the target were accepted as correct. The peripheral target reappeared 300 ms after the subject's response, and the subject was rewarded for making a corrective saccade to the target and maintaining eye position within 6° of the target for 300 ms. An intertrial interval (ITI) period between 2 and 6.5 s (depending on the experiment and subject) followed the corrective response. Correct trials were rewarded with delivery of water. Memory period errors occurred when fixation was broken before the central fixation target disappeared. These failure trials were tallied but excluded from all analyses except analysis of the error distribution. Because memory information may not have been properly encoded in failure trials, success trials following a failure trial were also excluded when considering the effect of previous target location.

Data Analyses

We measured the response error in each trial. Compared with visually guided saccades, subjects make systematic and variable errors while making saccades to remembered target locations. Due to the systematic errors, saccades to targets in the upper visual field tend to be hypermetric, saccades to targets in the lower visual hemifield tend to be hypometric, and saccades to targets on the horizontal meridian tend to be upward. These systematic errors have been shown to be influenced only by the early part (~400 to 800 ms) of memory delay, suggesting that memory processes are not major sources of these errors. However, variable error in memory tasks is influenced by delay period length over several seconds and as a result can provide information about memory decay (White and Sparks 1986; White et al. 1994; Gnadt et al. 1991). We therefore excluded systematic error effects and focused specifically on variable response error in our analyses.

Saccade angular directions were obtained during the 100- to 300-ms interval following the first saccade to the target. We calculated saccadic response error as the difference between the saccade direction and target direction in each trial. Systematic error, which was relatively constant over the duration of the

experiment, was removed from saccadic error to obtain residual variable response error. When many repetitions of discrete target locations were used, the systematic error was computed as the mean saccade endpoint for each target location. The corresponding mean was then subtracted from the saccade endpoint on each individual trial to obtain the nonsystematic error. For continuous target locations, the systematic error was computed by spatially low-pass filtering the saccade endpoints, expressed as a function of target location, using the MATLAB loess smooth function. The resulting estimate of systematic error was then subtracted from each individual saccade endpoint to obtain the nonsystematic error, which we call residual error. For each trial we also calculated the previous trial target direction relative to the current trial target direction by taking the difference between previous and current target directions. We then fit these data to the Gabor function in Eq. 1:

$$y = \text{height} \times e^{-(\text{width} \times x)^2} \times \sin(\text{width} \times x) \quad (1)$$

where y represents the residual error on each trial and x represents the relative direction of the previous trial's target. When reporting bias effect sizes, we use the peak-to-peak distance of Eq. 1, which is equivalent to $0.793 \times \text{height}$.

Continuous Attractor Network Simulations

For attractor network simulations, we used a continuous attractor network as described in Compte et al. (2000). The network consisted of 1,024 excitatory nodes, each tuned to a location in space, and 256 inhibitory nodes. Each node modeled an excitatory or inhibitory neuron. The network code is available at http://eye-hand.wustl.edu/supplemental/SpatialWMNet_Published_EyeHand.zip, and parameters and simulation paradigm can be found in the included parameters.ini file.

We simulated pairs of oculomotor delayed response trials, with each trial including a target presentation and a memory period. Each simulation began with 250 ms of spontaneous firing before the start of the first trial. A target stimulus was then presented at a spatial location for 250 ms in the form of a current injection of 70 pA to the 100 model neurons most closely tuned to that spatial location. After the memory period (500 to 3,000 ms, depending on the experiment), a stop signal (140 pA to all excitatory neurons for 100 ms) representing the end of the first trial turned off the sustained activity in the network. After an ITI of 150 ms, the next trial in the simulation began and a target stimulus with identical amplitude and duration was presented at a different location in the network. The network was read out after the second memory period ended. The population activity of the network was first smoothed across time (50 ms) and spatial location (30 adjacent neurons). A Von Mises function was then fit to the time slice at the end of the second trial's memory period, and the memory was read out as the center of the Von Mises function.

Two Store Model of Short-Term Memory

We modeled the temporal aspect of the bias effect using a model with a quickly decaying visual sensory store and a sustained working memory store. The target location estimate in the visual store was assumed to be veridical while the sustained store's estimate was taken to have a Gabor-like spatial profile. Each store's activity was modeled as a decaying exponential. To read out a target location at a given time t , each store's activity at t , was first normalized by the activity during stimulus presentation. The target location was then determined by the equation

$$T_E(t) = T_V w(t) + T_M [1 - w(t)] \quad (2)$$

where $T_E(t)$ is the estimated target location, T_V is the true target location held in the visual store, and T_M is the biased target location

held in the sustained store and determined by a Gabor function with 8° peak-to-peak based on behavioral data shown in RESULTS (see Fig. 5). The quantity $w(t)$ is the ratio of normalized activity in the visual sensory store $V(t) = V_0 e^{-t/\tau_v}$ and sustained memory store $M(t) = M_0 e^{-t/\tau_m}$ at time t :

$$w(t) = \frac{e^{-t/\tau_v}}{e^{-t/\tau_m}} \quad (3)$$

At time $t = 0$, $w(t) = 1$, and $T_E(t)$ is entirely determined from the content of the visual store and is therefore unbiased. As activity in the visual store decays, $w(t) \rightarrow 0$ and $[1 - w(t)] \rightarrow 1$, shifting the dependence of $T_E(t)$ to the biased memory of the sustained store.

Based on the large or infinite decay time constants of attractor models and our observations using neural recordings of spatial working memory circuits, we conservatively set the sustained store's decay time constant to be $\tau_m = 15$ s; the model results were essentially identical for any value > 10 s. To reproduce the time course shown in RESULTS (see Fig. 5), we set $\tau_v = 1.7$ s.

RESULTS

Attractive Bias Toward Previous Trial Target Direction

We measured errors in saccade responses in an oculomotor delayed response task as a function of the distance of the previous target from the current target. Error was defined as the angle between the target and the saccade made to the memorized target location. We subtracted systematic errors related to the current target direction, leaving only the residual error (see MATERIALS AND METHODS). We then tested whether this residual error was related to the location of the previous target, relative to the location of the current target. Figure 1A shows the data from one animal, averaged in 30° bins and fit to a Gabor function. When the target in the preceding trial was clockwise from the target in the current trial, mean residual error in saccade responses was also clockwise, that is, towards the previous target. When the target in the preceding trial was counterclockwise from the target in the current trial, mean residual error was counterclockwise, again towards the previous target. Thus we found that memory-guided saccade responses are biased in the direction of the target in the previous trial.

This was true for all five monkeys tested. In each case, Gabor fits to both the raw and binned data were highly significant, accounting for 2–7% of the variance of the raw data and 76–96% of the variance of the binned data points. The peak attractive influence of the previous target occurred when the previous target was between 41 and 78° from the current target.

Bias is unimodal. The small bias toward the previous trial direction could be due to small bias in each and every trial or due to a large error that occurs only rarely, e.g., an occasional response directed to the previous target location instead of the current location. To examine these possibilities we viewed the distribution of errors for trials in which the previous target was between $+35$ and $+85^\circ$ from the current target (Fig. 2). Saccades directed to the previous target would therefore appear at around $+60^\circ$, resulting in a bimodal distribution of errors. The distribution is instead unimodal, with a mode at $+1.29^\circ$ and a mean of $+2.21^\circ$, supporting the view that there is a small systematic bias toward the previous target on every trial.

Long-term changes or single-trial effects? The mechanisms responsible for the response bias observed could reflect persis-

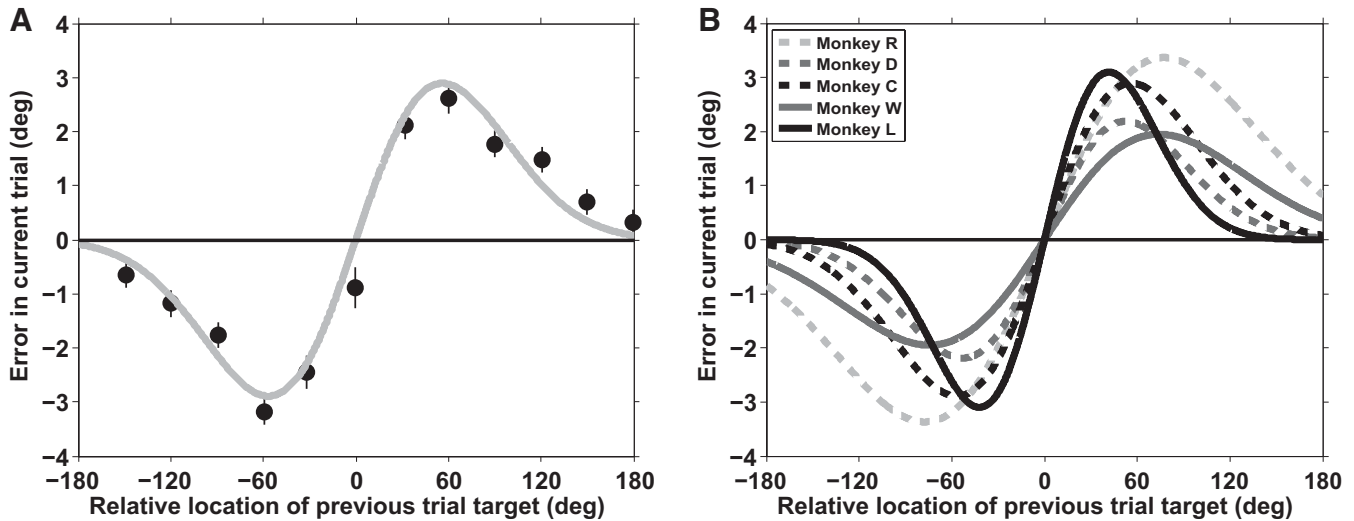


Fig. 1. Memory bias due to a prior memory. *A*: influence of previous target on the memory of current targets for *monkey C* in 7437 trials. The *x*-axis is the location of the target in the previous trial relative to the location of the target in the current trial, and the *y*-axis is the mean residual error (see MATERIALS AND METHODS). The gray line shows the Gabor fit of the data (peak-to-peak = 5.7, $P < 0.005$), and the error bars show SE. *B*: Gabor fits for 5 monkeys. Fits for each monkey were highly significant ($P < 0.005$).

tent, long-term changes that build up over time. Such changes might be beneficial for a system that commonly encounters similar information from one trial to the next. Alternatively, the response bias could be due to interference from residual information encoded during the previous trial. We first tested the temporal dynamics of the response bias by varying the length of the ITI while keeping the memory delay fixed. The response bias dropped with increasing ITI in all three animals tested (Fig. 3).

If the response bias were due to long-term persistent changes, then consecutive repetitions of target locations might increase the size of the bias (Verstynen and Sabes 2011). To determine whether the bias toward previous target locations was due to a short-term effect or to long-term plasticity, we presented targets at the same location for one to four consecutive trials, followed by a trial with a target $\pm 60^\circ$ away. We measured the response bias on the final trial of each set and

asked whether it depended on the number of consecutive repeated presentations that preceded it. Figure 4 shows data from two macaques comparing the response bias after a single presentation of the previous target vs. two or more consecutive trials using that same target. The effects were statistically indistinguishable in both monkeys. As shown, there is no significant change in bias in either subject (*monkey C*, $P > 0.33$, $n = 830$; *monkey D*, $P > 0.48$, $n = 974$). This finding together with the finding that response bias decreases with increasing ITI provides strong evidence that the bias effect is not due to long-term plasticity. Instead, the response bias is more likely due to interference from the previous trial, possibly due to persistent activity of the previous memory trace.

Effect of memory delay length on bias. We asked whether the length of the memory period would affect the bias caused by the previous trial. Simple models that employ a slowly decaying memory trace predict that the bias will decrease with

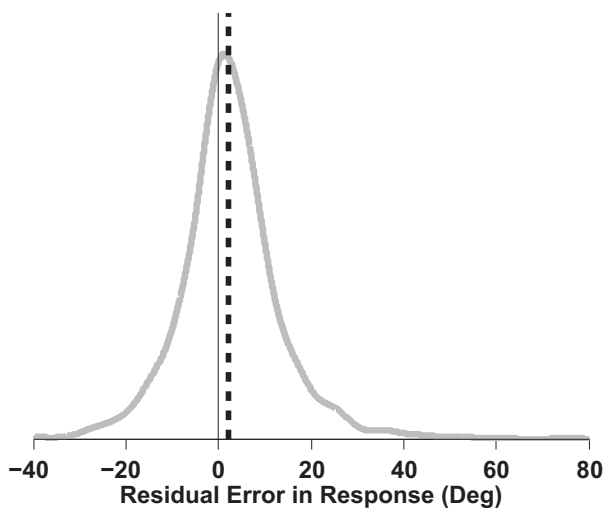


Fig. 2. Bias is unimodal. Residual error distribution for mean relative target location between $+35$ and $+85^\circ$ for all animals. Current target direction was rotated to 0° . The distribution is unimodal with a mode at $+1.29^\circ$ and a mean at 2.21° , indicating a small bias toward the previous target in each trial.

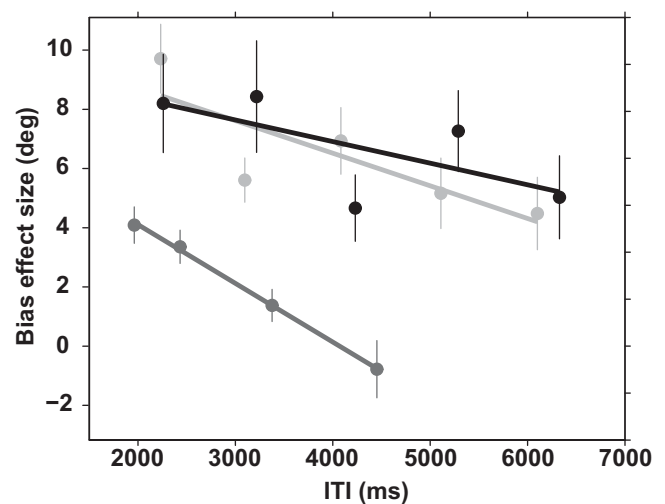


Fig. 3. Response bias as a function of intertrial interval (ITI). The response bias decreased with increasing ITI in each of the 3 animals tested. *Monkey C* (dark gray) $-2.00\%/s$, $P < 0.0005$, $n = 2,958$; *monkey R* (light gray) $-1.13\%/s$, $P = 0.09$, $n = 9,026$; *monkey D* (black) $-0.73\%/s$, $P = 0.22$, $n = 6,437$.

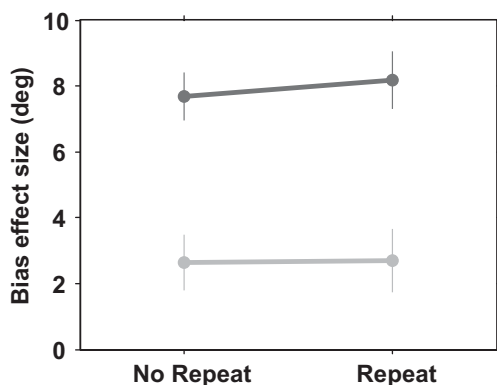


Fig. 4. Effect of target repetition on response bias. The bias in trials when a previous target location is repeated is no different from the bias in trials after 2 or more trials target location repetitions [*monkey C* (dark gray), $P > 0.33$, $n = 830$; *monkey D* (light gray), $P > 0.48$, $n = 947$]. In this experiment, targets were always presented either at the same location as or 60° away from the preceding trial's target.

longer memory periods. Longer periods lead to greater decay and hence a reduced effect on the following trial. If, on the other hand, the memory trace does not decay with time (as in an attractor network, for example), then the bias effect should remain constant with increasing delay times. Remarkably, we found instead that the bias increased asymptotically with delay length, with a mean time constant of 1.7 s (Fig. 5).

In the previous description, we assume that it is the delay length of the previous trial that is the key factor affecting bias. In the experiments of Fig. 5, memory delay was held constant within a block of trials, that is, every trial had the same delay length. Therefore, the dependence of the bias on delay length could reflect the length of the memory delay on either the previous or the current trial. Holding a location in memory for a longer duration on the previous trial might result in a greater bias on the subsequent trial. Alternatively, the bias might increase with time within the current trial, independent of the length of the previous trial. We found the latter to be the case.

We randomly interleaved short (0.8 s) and long (3.2 s) delay lengths within a block of trials and then split the data into four sets: 1) short-delay trials followed by long-delay trials, 2) long-delay trials followed by long-delay trials, 3) long-delay trials followed by short-delay trials, and 4) short-delay trials followed by short-delay trials. Figure 6, *left*, shows a representation of trial lengths, and Fig. 6, *right*, shows the corresponding mean response bias, averaged over three animals. When the current trial delay length was long, peak response bias was large regardless of the length of the previous delay (short-long, 10.2° ; long-long, 10.1°). When the current delay was short, peak response bias was small, again regardless of the previous delay (long-short, 4.4° ; short-short, 4.5°). An ANOVA (two factors, previous and current delay length; three repeated measures) indicates that only the main effect of the current delay is significant [$F(1,11) = 21.38$; $P < 0.002$]. The same result was also obtained for each individual animal. Thus the bias from the previous trial grows over time within the current trial and is unaffected by whether the previous target was held in memory for a long or short time.

Neither saccade execution nor memory maintenance is necessary for bias. In most of our experiments, a visual target was presented, the location of that target was maintained in mem-

ory, and a saccade was directed to the remembered location. To determine whether either memory maintenance or saccade execution is necessary to produce spatial bias on the subsequent trial, standard memory-guided saccade trials were interleaved with memory-only trials and saccade-only trials. Memory-only trials differed from standard memory trials in that the subject was required to continue fixation at the end of the memory period rather than making a saccadic response. Saccade-only trials differed from standard memory trials in that no memory cue was presented and the subject performed a visually guided saccade at the end of the delay period. Biases occurred in standard memory trials that followed either a memory-only or a saccade-only trial, with effect sizes of 7.1 and 6.5° , respectively (fit $P < 0.005$ in both cases; data from 383 and 376 trials, respectively, performed by *monkey C*). This indicates that neither a memory period nor a saccade is required in the previous trial to produce an attractive bias.

A second experiment confirmed these results. In the random saccade experiment, standard memory-guided saccade trials were interleaved with trials that were identical to the standard memory trials up until the time of the go cue (fixation offset). At this time, a second "random" target appeared, and the animal was required to saccade to this new location rather than to the memorized location. The new location was statistically independent of the memorized location. We found that a standard memory-guided trial that followed a random saccade trial showed a bias toward both the previous trial's memory location (bias = 5.5 ; fit $P < 0.005$; Fig. 7, *top left*) and the previous trial's saccade location (bias = 5.6 , fit $P < 0.005$; Fig. 7 *bottom right*). This experiment shows that biases can be driven by either memorizing a target without actually moving to it or by moving to a target without having previously memorized it.

There are at least two possible interpretations of these results. First, memorizing a target location and executing a saccade may each, by themselves, be sufficient to bias a memory-guided saccade on a subsequent trial. Alternatively, it may be that merely presenting a visual stimulus is sufficient to generate a bias. There are two arguments for the latter inter-

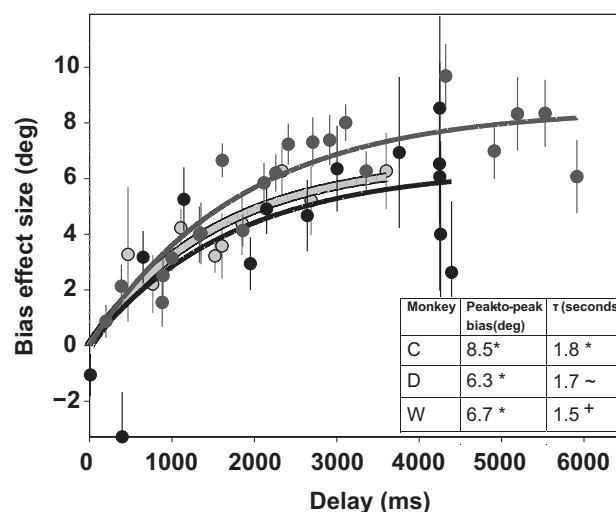
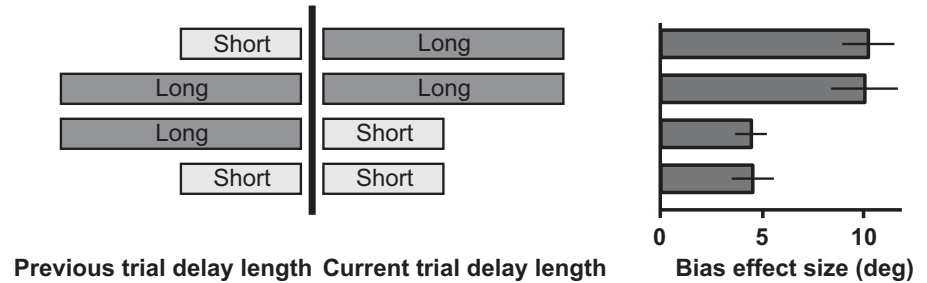


Fig. 5. Bias as a function of memory delay. Effect of increasing delay period on peak-to-peak bias for *monkey C* (dark gray, $P < 0.0001$), *monkey D* (black, $P < 0.002$), and *monkey W* (light gray, $P < 0.02$). Inset: fit coefficients for each monkey (* $P < 0.002$, + $P = 0.06$, ~ $P = 0.28$).

Fig. 6. Comparison of bias as a function of previous and current delay. Conditions from *top to bottom*: 1) short-delay trials followed by long-delay trials (10.2°), 2) long-delay trials followed by long-delay trials (10.1°), 3) long-delay trials followed by short delay trials (4.4°), and 4) short-delay trials followed by short-delay trials (4.5°). Only the main effect of the current delay is significant [$F(1,11) = 21.38$; $P < 0.002$].



pretation. First, the magnitude of the effects seen in the random-saccade, memory-only, and saccade-only trials is comparable to the magnitude of the effects seen in the standard memory trials. If memory maintenance and saccade execution contributed independently and equally to the bias effects, then we would expect that the magnitude of effects on standard memory trials would be twice as large as the effects in the other trial types. Second, previous studies using distractors show attractive bias toward irrelevant (distractor) stimuli that subjects ignore while performing a standard spatial memory task (Chumbley et al. 2008; Macoveanu et al. 2007). Taken together, the evidence suggests that neither memory maintenance nor saccade execution is necessary to produce bias and that instead it is the presentation of a visual stimulus that produces interference on a subsequent memory-guided saccade trial.

Reaction time effects. It is possible that the spatial biases we observe are accompanied by reaction time effects. We asked whether the reaction time of a saccade to a memorized target depends on the relative location of the target on the previous trial. We removed mean reaction time as a function of current target location and computed residual reaction time as a function of previous target location. An ANOVA showed no effect of relative target location on memory-guided saccade reaction

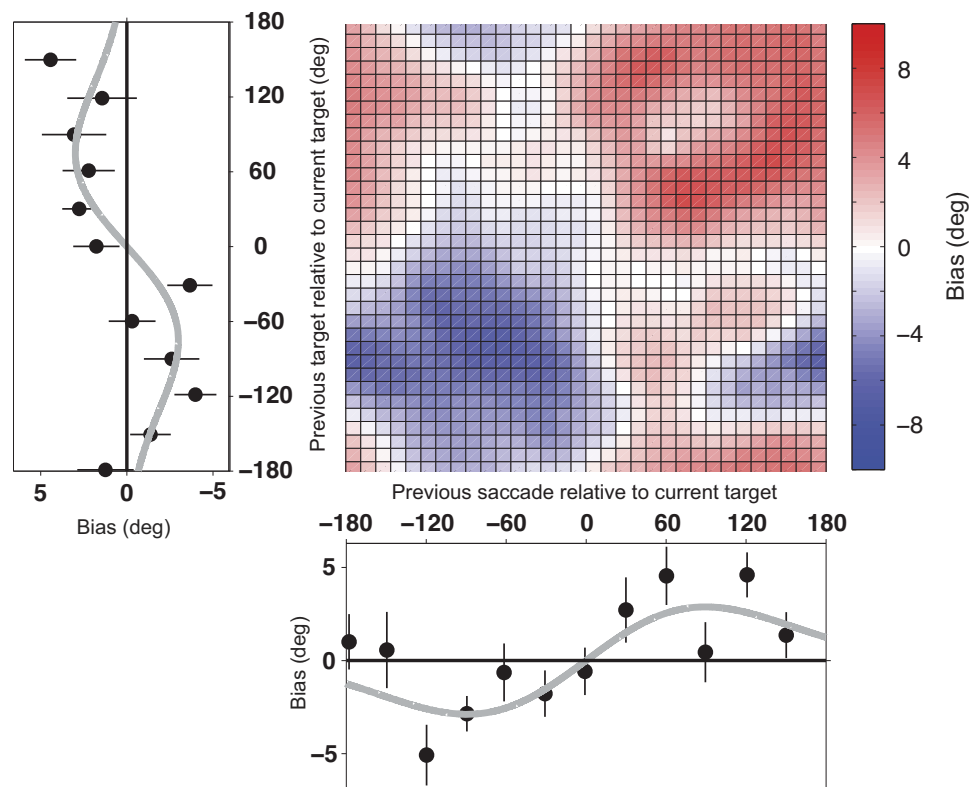
time (RT) [previous and current targets within $\pm 45^\circ$ of one another: RT = 0.26 ms; from 45 to 135° apart: RT = -0.02 ms; from 135 to 180° : RT = -0.12 ms; $F(2,32121) = 0.19$; $P = 0.83$]. Note that these same trials showed clear spatial biases (Fig. 1).

Interference in Attractor Networks

To model the data in these experiments we implemented a continuous attractor network as described in MATERIALS AND METHODS and Compte et al. (2000). In this model, memory for a particular location is maintained by strong recurrent excitatory connections between neurons with similar preferred directions. This excitatory input is counteracted by inhibitory connections between neurons with dissimilar preferred directions. The resulting bump in population activity sustains its shape over time but can drift due to stochastic fluctuations in activity.

The attractor network framework is commonly used for modeling working memory circuits. These models match many of the physiologically observed properties of the neuronal circuits believed to be involved in memory and successfully account for many of the behavioral phenomena (Durstewitz et

Fig. 7. Both previous memory location and previous saccade direction bias a subsequent memory-guided response. Standard memory-guided saccades were interleaved with trials in which a target was memorized, but then, instead of making a memory-guided saccade to that memorized target, a new target was presented and animals made a visually guided saccade to it (see text). We asked how the two processes, memorizing a previous target and making a saccade in a particular direction, affect a subsequent ("current") memory-guided saccade. The surface fit shows that the current memory-guided saccade is biased by both the previous visually guided saccade (abscissa) as well as by the previously memorized target (ordinate). The marginal plots show that the effect sizes are similar, with a previous saccade direction bias of 5.5° ($P < 0.005$) and a previous memory location bias of 5.6° ($P < 0.005$).



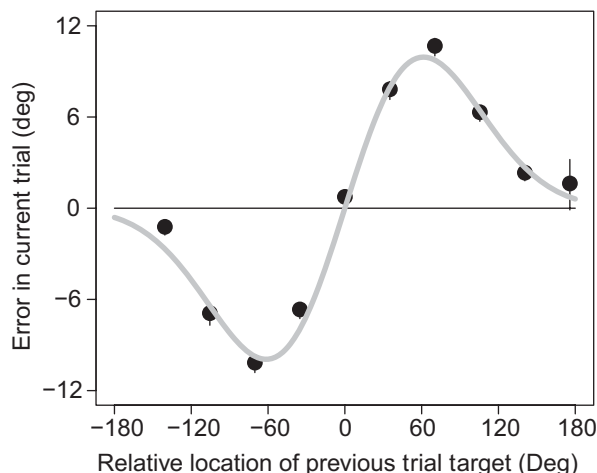


Fig. 8. Influence of previous target on the memory of current targets as predicted by the continuous attractor network model. The x-axis is the difference between the current and previous target direction and the y-axis is the mean error, as in Fig. 1. The gray line represents the Gabor fit to the model data ($P < 0.0001$) with peak-to-peak = 19.9° ($P < 0.0001$).

al. 2000; Wang 2009). These models display within-trial interference effects with a spatial profile that is reminiscent of the between-trial effects we observe in our experiments. In particular, distractors in continuous attractor models exert an attractive bias that is stronger when the distractor is close to the target location and weaker when it is farther away (Compte et al. 2000). This bias results from local positive feedback loops within the attractor network. Therefore, we chose these networks to model our between-trial interference effects, positing that the same structures that give rise to within-trial interference might also generate between-trial effects with a similar spatial profile.

The network stimulation paradigm we used in our simulations was based on our behavioral paradigm. For simplicity, each simulation consisted of just two trials, a previous and current trial. At the start of each simulation, a network of 1,024 excitatory and 256 inhibitory neurons was randomly initialized and allowed to evolve without any external input for 0.25 s. A target direction was then chosen and presented in the form of a 70-pA, 0.25-s current injection to the 100 excitatory neurons most closely tuned to that location. This was followed by a 1-s memory period during which the elevated activity was maintained in the neurons around the target location. Next, 140 pA of current was injected broadly to all excitatory neurons in the network for a duration of 0.1 s to reset the network. This reset signal displaces the network away from the memory-holding attractor state and back to a baseline state after a short time. Following an ITI period of 0.15 s during which network activity was allowed to evolve without any external stimulation, a second target stimulus with an identical profile to the first was presented at a new location. After a delay of 1 s, we fit a Von Mises function to the network activity and read out the peak as the network's response in the second trial.

The attractor network replicated the spatial aspect of the bias effect. Although the reset signal between the first and second trials causes the network to leave its attractor state, when the ITI is short there is still enough residual activity at the start of the next trial to bias the network toward the previous trial's target location. Figure 8 shows response error (the difference

between target and response locations) on the second trial as a function of the distance between the first and second target locations. Similar to the behavioral data, the simulation data show a bias toward the previous target location that is modulated by distance.

The attractor network model also reproduces the finding that bias increases with delay length. After stimulus presentation, the network only gradually settles into the attractor state. The speed at which this occurs depends on network parameters. If the first memory period is short, then the network may not reach the full attractor state. As a result, the carry over of information from the first to the second trial is reduced for short memory periods (Fig. 9).

Critically, unlike the animal data (Fig. 6), the magnitude of the response bias from the network depends on the length of the previous trial, not on the length of the current trial (Fig. 10). When the previous delay was long, response bias was large regardless of the length of the current delay (long-short, 21.3° ; long-long, 21.6°). When the previous delay was short, response bias was small, again regardless of the current delay (short-short, 14.8° ; short-long, 14.6°). An ANOVA (two factors, previous and current delay length; three repeated measures) indicated that only the main effect of the previous delay is significant [$F(1,11) = 532.81$; $P < 0.0001$]. Thus the network fails to reproduce the results of the mixed delay experiment. These results allow us to reject the hypothesis that the dynamics of a standard attractor model can solely account for the increase in bias as a function of memory period length.

Two-Store Model of Short-Term Memory

To capture the dependence of the response bias on the duration of the current memory period, rather than on the duration of the previous memory period, we modified the attractor model by adding a short-term visual store. The visual store is ephemeral, discharging much more quickly than the main memory store. The visual store is fully charged by the end of the stimulus presentation, but this charge then decays with a fast time constant (Fig. 11A). This decay is too rapid to provide a conduit for contaminating the current trial with residual activity from the previous trial. The attractor network, as just

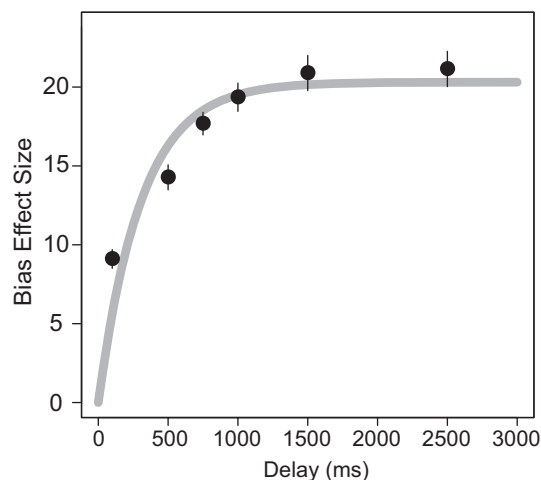
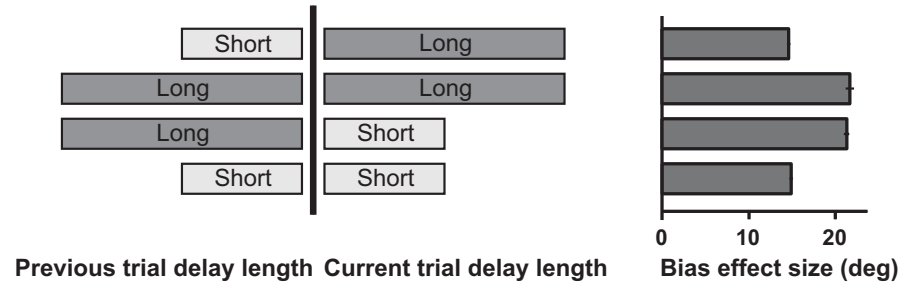


Fig. 9. Effect of increasing delay period on peak-to-peak bias predicted by the continuous attractor network model. The gray line is the exponential fit to the data ($P < 0.011$) with peak-to-peak bias = 20.3° ($P < 0.0002$) and time constant = 308 ms ($P < 0.03$).

Fig. 10. Comparison of bias as a function of previous and current delay in network simulations. Conditions from *top to bottom*: 1) short-delay trials followed by long-delay trial (14.6°), 2) long-delay trials followed by long-delay trials (21.6°), 3) long-delay trials followed by short delay trials (21.3°), and 4) short-delay trials followed by short-delay trials (14.8°). Only the main effect of the previous delay is significant [$F(1,11) = 532.81$; $P < 0.0001$].



described in the single store model, has a very long time constant, and so the stimulus can be maintained indefinitely. In the two-store model, the target location estimate is determined by the fractionally weighted estimates of the veridical visual sensory store and the biased sustained store. For visually guided and shorter memory-guided trials, the veridical visual circuit accounts for all or most of the contribution to the target estimate resulting in little bias or no bias. For longer trials, activity in the visual circuit has decayed, and therefore, the biased sustained memory circuit contributes the most strongly to the target estimate, resulting in a biased output (Fig. 11B). Thus, as the duration of the memory period increases and the activity in the visual store drops exponentially, the response bias grows as a saturating exponential (Fig. 11C). Critically, the response bias of this model, like that of our subjects, depends on the length of the current trial, not the previous trial (Fig. 11D). Parameters in this model are constrained as described in MATERIALS AND METHODS, with the visual-store time constant as the only free parameter. To reproduce the time course shown in Fig. 5, we set $\tau_v = 1.7$ s in the simulations shown in Fig. 11.

We asked whether the visual store was, like human iconic memory (see DISCUSSION), disrupted by the presentation of a visual mask. If so, then we would expect that a mask would force the system to rely on the (biased) working memory store, thereby sharply increasing the bias seen early in the memory

period. If, on the other hand, the visual store is robust, then a mask will not change the early bias. To test this, *monkey R* participated in an additional experiment in which, in half of the trials, a mask was presented 67 ms after the memory target onset. The target was presented for 50 ms, and the mask (a dense display of line-segments at random orientations, with the same color as the target) was presented for 100 ms. Memory periods were 600, 1,200, or 2,400 ms. Without a mask, bias at 600 and 1,200 ms was 41 and 44%, respectively, of the full bias effect measured at 2,400 ms (see also Fig. 5). With a mask, bias at 600 and 1,200 ms increased to 54 and 51% of the full bias, respectively. These increases were not significant ($P = 0.40$ and 0.64 respectively; $n = 7,373$), suggesting that the transient visual store in our two-store model is relatively robust to masking effects.

DISCUSSION

In this study we investigated how spatial stimuli held in working memory are influenced by memoranda from the previous trial. We found that responses are biased toward previous stimulus locations and that this bias is strongest when the previous trial target and current trial target are separated by $\sim 60^\circ$ of arc. The bias toward the previous target location did not increase with consecutive target presentations at that location, suggesting that the bias is a trial-by-trial effect perhaps

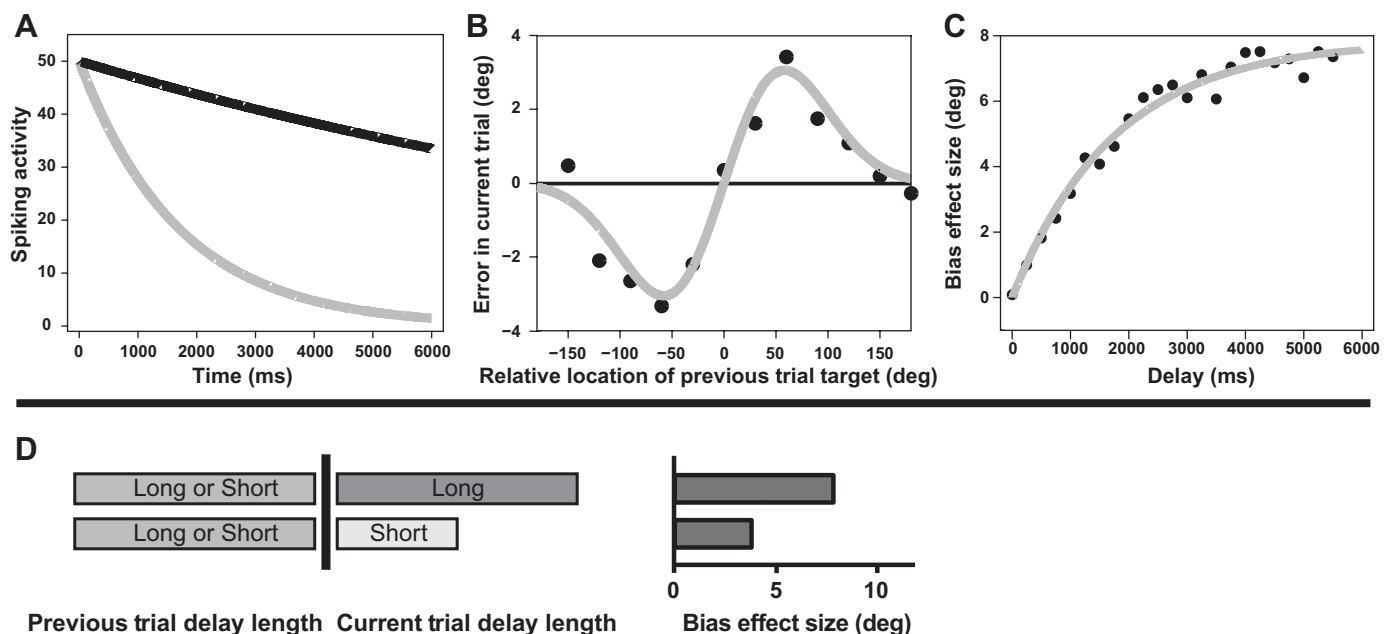


Fig. 11. Two-store model predictions. *A*: activity profiles of the visual store (gray) and the memory store (black). *B*: effect of previous trial on the memory of current target for a delay of 3s. *C*: effect of increasing delay period on the bias. *D*: predictions of the two-store model in the long-short experiment.

due to a residual memory trace rather than a long-term synaptic effect. We also found that the size of the bias increases asymptotically when increasing the memory period of the current trial. We were able to model the spatial but not the temporal aspects of the bias with a standard attractor network. To model the increase in bias with the length of the current memory period, we used a two-store model with a rapidly decaying visual store and biased sustained memory store.

The response bias that we see from the previous stimulus is similar to proactive interference effects reported in prior literature. Classic examples of proactive interference in humans show long-term buildup of interference over repeated memorization sessions often spanning many days (e.g., Underwood 1957). However, there are also reports of interference effects within a single session. In particular, Dunnet and Martel (1990) report proactive interference due specifically to the immediately preceding trial (but see Wright et al. 2012). In some studies, the magnitude of interference grows asymptotically with the length of the delay period (e.g., Moise 1976; Dunnett and Martel 1990; Edhouse and White 1988), exactly as we show here (Fig. 5), although this was not commented on in the original studies. Another point of similarity between our results and previous studies of proactive interference is that interference effects are larger when the previous and current stimuli share similar properties, such as their spatial location (e.g., Wickens et al. 1963; Makovski and Jiang 2008). Most of these previous studies use categorical memoranda and measure interference effects in percent correct. We use continuous memoranda and see analogous interference effects in accuracy, suggesting that similar mechanisms are at play.

A preceding memory may cause interference by overwriting a current memory, adding nonsystematic noise, or adding a systematic bias to the current memory. Distinguishing between these possibilities based only on the proportion of correct responses can be difficult. An advantage of using continuous memoranda and responses is that we can more finely characterize the spatial and temporal structure of the interference and thereby address how previous information interferes with current memory. This in turn provides insight into the underlying circuits for working memory. In particular, we find that the interference is not caused due to overwriting by the previous memory or the introduction of nonsystematic noise but instead that previous information combines with and systematically biases current information.

This systematic bias is not due to persistent, long-term adaptation. Verstynen and Sabes (2011) describe such an effect in a visually guided reaching task. They find that attractive bias increases with repeated presentations of the same target location. This was not the case in the current study (Fig. 3). Instead, the bias from the previous trial appears to be a single-trial effect, likely due to either residual activity within the network or from short-term changes in synaptic efficacy from the previous trial.

The spatial profile of the bias effect is similar to that of distractor effects in attractor network simulations (Compte et al. 2000) arising from the recurrent connectivity structure in attractor models. We show that a standard attractor network shows proactive interference effects in the form of an attractive bias toward the previous target location and replicates the spatial aspect of the bias. The animal's bias increases as delay increases, rising with a time constant of just over 1.5 s (Fig. 5),

and the critical time interval is the duration of the memory period in the current trial (Fig. 6). Although the attractor model predicts a similar increase in bias with delay, the bias depends on the duration of the previous memory period (Figs. 9 and 10). This occurs because, with short trials, the memory trace in the attractor model does not have sufficient time to build up to the full attractor state, and as a result there is less residual activity remaining after the reset signal, leading to a weaker effect on the subsequent trial.

We asked whether saccade execution or active memory maintenance was necessary to produce bias. We show that while a visually guided saccade or memory period will, on their own, bias a subsequent memory-guided saccade, neither are necessary. Furthermore, behavioral studies with distractors show that merely presenting an irrelevant visual stimulus is sufficient to produce an attractive bias (Macoveanu et al. 2007; Chumbley et al. 2008). These findings suggest that visual target presentation, which occurs in both saccade-only and memory-only trials, is the most likely cause for the observed response bias in every case.

A large body of work in humans and nonhuman primates has shown that the reaction time of responses to visual stimuli depends on the location of previous targets. For example, reaction times in a center-out visually guided saccade task depends on where the target lies, relative to the target location on the previous trial (Dorris et al. 1999). Monkeys execute saccades more quickly to a target whose location coincides with the target of the previous trial. This may be related to the phenomenon of inhibition of return, which was first described in humans, although in humans the polarity of the effect is reversed; saccades to targets at repeat locations are slowed rather than sped up (Rafal et al. 1994; Taylor 1997; Taylor and Klein 1998.) These reaction time effects have been modeled as a consequence of transient suppression of cells involved in saccade generation (e.g., Dorris et al. 1999). The same mechanism could also lead to spatial biases in saccades. If so, then reaction time and spatial bias effects should appear in tandem. However, as described in RESULTS, we found spatial but not reaction time effects in the memory-guided saccade task. In saccade-only trials, we found reaction time slowing when the target was in the same direction as the previous trial's target but not spatial biases [RT difference (same – opposite): 35.7 ms, $P < 0.002$; peak-to-peak bias: 0.59° , $P = 0.74$; $n = 377$]. This double dissociation provides strong evidence that spatial bias and reaction time effects rely at least in part on different neuronal mechanisms.

To model a buildup in bias that depends on the current trial duration, we posit two separate memory stores, a visual sensory store and a sustained store. In this model, the saccade to the remembered target is a combination of information from the two stores. The visual sensory store decays quickly, losing all information within a few seconds after target onset. Classic human psychology studies have identified iconic memory, a short-term store that retains the visuo-spatial characteristics of stimuli for a brief interval after stimulus offset (e.g., Sperling 1960). Due to its short duration and its apparent inability to survive eye movements (Irwin 1992; Irwin and Andrews 1996), iconic memory is unlikely to show substantial interference from the previous trial and in this respect is a good match for the quickly decaying visual store posited in our model. However, most studies describe iconic memory as completely

decaying in 700 ms or less (e.g., Sperling 1960; but see Averbach and Sperling 1961). This is too short to match our data, which require a time constant closer to 1.7 s for the decay of the visual store. In addition, iconic memory is overwritten when a mask is presented following the memory stimulus. Our model predicts that, by overwriting the contents of iconic memory almost immediately after target presentation, the memory-guided saccade will rely only on the longer, biased memory store and will therefore show substantial bias even with very short memory periods. Contrary to this, we find that a mask has little effect on the early bias. Thus both the observed time constant and the effect of a visual mask suggest that iconic memory is not involved. Other visual short-term memory stores intermediate to iconic memory and working memory have been identified in behavioral (Slighte et al. 2008; Pinto et al. 2013; Vandenberg et al. 2014; see also Griffin and Nobre 2003; Makovski and Jiang 2007; Makovski et al. 2008) and neurophysiological studies (Slighte et al. 2009; see also Pasternak and Greenlee 2005; Bisley et al. 2004). These stores have longer time constants than the iconic store, with some studies showing information retention for up to 4 s.

The sustained spatial working memory store holds information for a much longer time than the visual store and is biased by residual activity from the previous trial. To produce a behavioral response, the signals from the visual sensory store and sustained memory store are combined and weighted by the ratio of their activity. Immediately after target presentation, the visual sensory store is highly active and provides most of the signal used to guide the saccade. After several seconds, this unbiased store has decayed away, and the biased memory store provides most of the signal used to guide the saccade. This model successfully replicates the observed features of memory (Fig. 11). These findings suggest that short-term memory circuits may be composed of multiple memory stores with independent decay rates. Furthermore, the visual sensory store may hold a more veridical representation of visual information than the sustained memory store and may act to shield against systematic biases arising from the recurrent attractor circuit dynamics of the sustained store.

Conclusions

We characterized proactive interference in a simple oculomotor delayed response paradigm and describe several new and significant properties. First, proactive interference is graded, not all-or-none. Second, the spatial profile of proactive interference is well modeled by attractor circuits. Third, proactive interference resembles interference from within-trial distractors, which suggests both proactive interference and distractor interference arise due to the same memory mechanisms. Fourth, proactive interference was manifest only on delayed saccade responses, not on visually guided saccades. This distinguishes the process from inhibition of return, which affects all saccades. Fifth, proactive interference increased as a function of delay length. This is significant because it implies that not one but two memory stores are in operation: a short-term store (distinct from iconic memory) that is not susceptible to proactive interference, and a long-term store that is susceptible.

GRANTS

This work was supported by National Eye Institute Grant EY-012135 (to L. H. Snyder) and National Institute of Mental Health Grant 5F31-MH-094076-04 (to C. Papadimitriou).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: C.P., A.F., and L.H.S. conception and design of research; C.P. and A.F. performed experiments; C.P. and A.F. analyzed data; C.P., A.F., and L.H.S. interpreted results of experiments; C.P. prepared figures; C.P. and A.F. drafted manuscript; C.P. and L.H.S. edited and revised manuscript; L.H.S. approved final version of manuscript.

REFERENCES

- Altgassen M, Phillips L, Kopp U, Kliegel M. Role of working memory components in planning performance of individuals with Parkinson's disease. *Neuropsychologia* 45: 2393–2397, 2007.
- Averbach E, Sperling G. Short term storage of information in vision. In: *Information Theory*, edited by Cherry C. Washington, DC: Butterworth, 1961, p. 196–221.
- Baddeley A. Working memory: theories, models, controversies. *Annu Rev Psychol* 63: 1–29, 2012.
- Bisley JW, Zaksas D, Droll JA, Pasternak T. Activity of neurons in cortical area MT during a memory for motion task. *J Neurophysiol* 91: 286–300, 2004.
- Chumbley JR, Dolan RJ, Friston KJ. Attractor models of working memory and their modulation by reward. *Biol Cybern* 98: 11–18, 2008.
- Cohen G. *Memory in the Real World* (2nd ed.). Hove, UK: Psychology, 1996.
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10: 910–923, 2000.
- Daneman M, Carpenter PA. Individual differences in working memory and reading. *J Verb Learn Verb Behav* 19: 450–466, 1980.
- Daneman M, Merikle PM. Working memory and language comprehension: a meta-analysis. *Psychon Bull Rev* 3: 422–433, 1996.
- Dorris MC, Taylor TL, Klein RM, Munoz P. Influence of previous visual stimulus or saccade on saccadic reaction times in monkey. *J Neurophysiol* 81: 2429–2436, 1999.
- Dunnett SB, Martel FL. Proactive interference effects on short-term memory in rats: I. Basic parameters and drug effects. *Behav Neurosci* 104: 655–665, 1990.
- Durstewitz D, Seamans JK, Sejnowski TJ. Neurocomputational models of working memory. *Nat Neurosci* 3, Suppl: 1184–1191, 2000.
- Edhouse WV, White KG. Sources of proactive interference in animal memory. *J Exp Psychol* 14: 56–70, 1988.
- Engle RW, Kane MJ. Executive attention, working memory capacity, and a two-factor theory of cognitive control. In: *The Psychology of Learning and Motivation*, edited by Ross BH. New York: Elsevier Academic, 2004, vol. 44, p. 145–199.
- Gnadt JW, Bracewell RM, Andersen RA. Sensorimotor transformation during eye movements to remembered visual targets. *Vision Res* 31: 693–715, 1991.
- Griffin IC, Nobre AC. Orienting attention to locations in internal representations. *J Cogn Neurosci* 15: 1176–1194, 2003.
- Irwin DE. Memory for position and identity across eye movements. *J Exp Psychol Learn Mem Cogn* 18, 307–317, 1992.
- Irwin DE, Andrews RV. Integration and accumulation of information across saccadic eye movements. In: *Attention and Performance XVI*, edited by Inui T, McClelland JL. Cambridge, MA: MIT Press, 1996, p. 125–155.
- Jarvik ME, Goldfarb TL, Carley JL. Influence of interference on delayed matching in monkeys. *J Exp Psychol* 81: 1–6, 1969.
- Jonides J, Nee DE. Brain mechanisms of proactive interference in working memory. *Neuroscience* 139: 181–193, 2006.
- Just MA, Carpenter PA. A capacity theory of comprehension: individual differences in working memory. *Psychol Rev* 1: 122–149, 1992.
- Macoveanu J, Klingberg T, Tegnér J. Neuronal firing rates account for distractor effects on mnemonic accuracy in a visuo-spatial working memory task. *Biol Cybern* 96: 407–419, 2007.

- Makovski T, Jiang Y.** Distributing versus focusing attention in visual short-term memory. *Psychon Bull Rev* 14: 1072–1078, 2007.
- Makovski T, Jiang Y.** Proactive interference from items previously stored in visual working memory. *Mem Cognit* 36: 43–52, 2008.
- Makovski T, Sussman R, Jiang YV.** Attention in visual working memory reduces interference from memory probes. *J Exp Psychol* 34: 369–380, 2008.
- Miyake A, Shah P.** *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. New York: Cambridge Univ. Press, 1999.
- Moise SL Jr.** Proactive effects of stimuli, delays, and response position during delayed matching from sample. *Anim Learn Behav* 4: 37–40, 1976.
- Pasternak T, Greenlee MW.** Working memory in primate sensory systems. *Nat Rev Neurosci* 6: 97–107, 2005.
- Payne JW, Beuman JR, Johnson EJ.** The adaptive decision maker: effort and accuracy in choice. In: *Insights in Decision Making: A Tribute to Hillel J. Einhorn*, edited by Hogarth RM. Chicago, IL: Univ. Chicago Press, 1990, p. 129–153.
- Payne JW, Bettman JR, Johnson EJ.** Behavioral decision research: a constructive processing perspective. *Annu Rev Psychol* 43: 87–131, 1992.
- Pinto Y, Sligte IG, Shapiro KL, Lamme VA.** Fragile visual short-term memory is an object-based and location-specific store. *Psychon Bull Rev* 20: 732–739, 2013.
- Rafal RD, Egly R, Rhodes D.** Effects of inhibition of return on voluntary and visually guided saccades. *Can J Exp Psychol* 48: 284–300, 1994.
- Raghubar KP, Barnes MA, Hecht SA.** Working memory and mathematics: a review of developmental, individual, difference, and cognitive approaches. *Learn Indiv Diff* 20: 110–122, 2010.
- Sligte IG, Scholte HS, Lamme VA.** Are there multiple visual short-term memory stores? *PLoS One* 3: e1699, 2008.
- Sligte IG, Scholte HS, Lamme VA.** V4 activity predicts the strength of visual short-term memory representations. *J Neurosci* 29: 7432–7438, 2009.
- Sperling G.** The information available in brief visual presentations. *Psychol Monographs Gen Appl* 74: 1–29, 1960.
- Taylor TL.** *Generating and Measuring Inhibition of Return* (PhD thesis). Halifax, Nova Scotia, Canada: Dalhousie Univ., 1997.
- Taylor TL, Klein RM.** On the causes and effects of inhibition of return. *Psychon Bull Rev* 5: 625–643, 1998.
- Underwood BJ.** Interference and forgetting. *Psychol Rev* 64: 49–60, 1957.
- Unterrainer JM, Owen AM.** Planning and problem solving: from neuropsychology to functional neuroimaging. *J Physiol (Paris)* 99: 308–317, 2006.
- Vanderbroucke ARE, Sligte IG, Barrett AB, Seth AK, Fahrenfort JJ, Lamme VA.** Accurate metacognition for visual sensory memory representations. *Psychol Sci* 25: 861–873, 2014.
- Verstynen T, Sabes PN.** How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. *J Neurosci* 31: 10050–10059, 2011.
- Wang XJ.** Attractor network models. In: *Encyclopedia of Neuroscience*, edited by Squire LR. Oxford, UK: Academic, 2009, vol. 1, p. 667–679.
- White JM, Sparks DL.** Saccades to remembered targets: a behavioral and neurophysiological study. *Inv Ophthalm Vis Sci Suppl* 27: 155, 1986.
- White JM, Sparks DL, Stanford TR.** Saccades to remembered target locations: an analysis of systematic and variable errors. *Vision Res* 34: 79–92, 1994.
- Wickens DD, Born DG, Allen CK.** Proactive inhibition and item similarity in short-term memory. *J Verb Learn Verb Behav* 2: 440–445, 1963.
- Wright AA, Katz JS, Ma WJ.** How to be proactive about interference: lessons from animal memory. *Psychol Sci* 23: 453–458, 2012.

