A direct neural signature of serial dependence in working memory

Cora Fischer[1,2], Jochen Kaiser[1,2], Christoph Bledowski[1,2*]

Affiliations

[1] Goethe University Frankfurt, Institute of Medical Psychology, Heinrich-Hoffmann-Strasse 10, 60528 Frankfurt am Main, Germany.

[2] Goethe University Frankfurt, Brain Imaging Center, Medical Faculty, 60528 Frankfurt am Main, Germany

* Corresponding authors:

Christoph Bledowski, Goethe University Frankfurt, Institute of Medical Psychology, Heinrich-Hoffmann-Strasse 10, 60528 Frankfurt am Main, Germany.

E-Mail: bledowski@em.uni-frankfurt.de

Keywords

working memory | serial dependence | neural representation | multivariate analysis | magnetoencephalography

Running Title

A direct neural signature of serial dependence

Conflicts of interest

None.

Significance Statement

Our perception is biased towards sensory input from the recent past. While this "serial dependence" has been well established in behavior, a direct neural correlate of the phenomenon has not been determined. Applying multivariate analyses to magnetoencephalographic data during a visual working memory task, we found an attraction of neural representations to previous stimuli, which was consistent with behavioral responses. This is the first evidence for a direct neural signature of serial dependence. In addition, the present task, together with the high temporal resolution of magnetoencephalography, allowed us to identify the timing of serial dependence. In line with current theoretical models, we found serial dependence to emerge at late, post-encoding stages of information processing in working memory.

Abstract

Serial dependence describes the phenomenon that current object representations are attracted to previously encoded and reported representations. While attractive biases have been observed reliably and across various modalities in behavioral reports, a direct neural correlate has not been established. Previous studies have either shown a reactivation of past information without observing a neural signal related to the bias of the current information, or a repulsive distortion of current neural representations contrasting the behavioral bias. The present study recorded neural signals with magnetoencephalography during a working memory task to identify neural correlates of serial dependence. Participants encoded and memorized two sequentially presented motion directions per trial, one of which was later retro-cued for report. Multivariate analyses provided reliable reconstructions of both motion directions. Importantly, the reconstructed directions in the current trial were attractively shifted towards the target direction of the previous trial. This neural bias mirrored the behavioral attractive bias, thus reflecting a direct neural signature of serial dependence. The use of a retro-cue task in

combination with magnetoencephalography allowed us to determine that this neural bias emerged at later, post-encoding time points. This timing suggests that serial dependence in working memory affects memorized information during read-out and reactivation processes that happen after the initial encoding. Taken together, we identified a direct neural signature of serial dependence, which occurs during later processing stages of working memory representations.

A direct neural signature of serial dependence in working memory

Our visual input can change quickly from moment to moment. Nevertheless, we represent our environment as remarkably stable and coherent across short periods of time. This coherence is thought to be supported by serial dependence (*1*), i.e., the phenomenon that a current visual object is reported as more similar to a previously encountered visual object than it actually was. This attractive bias acts most strongly between current and previous objects that are close in time and space, that are attended and that are similar to each other both with regard to the visual feature that has to be recalled and with regard to accompanying context features (*1–3*). Furthermore, serial dependence occurs at different levels of object processing including encoding (*4*), memorization (*5*) and read-out or decision (*6*, *7*).

While the cognitive aspects of serial dependence have been intensively studied (for reviews see (*8*) or (*9*)), its neural mechanisms are still largely unclear and a matter of active research. One important aspect that has been investigated on the neural level is the nature of the memory traces that impact current processing and the timing of their reactivation. Single-unit recordings from monkey prefrontal cortex combined with electroencephalography (EEG) and transcranial magnetic stimulation in humans have demonstrated that information about a previous stimulus was reactivated shortly before a new stimulus was encoded (*10*). This study found that the strength of this pre-encoding reactivation in the prefrontal cortex predicted the attractive bias of the behavioral response to the new stimulus. Moreover, subjects showed an enhanced serial bias after single-pulse transcranial magnetic stimulation of their prefrontal cortex, which further supported a causal link between the observed reactivation and serial dependence. Another study (*11*) showed that memory traces from previous trials recorded from the posterior parietal cortex (PPC) in rats reappeared during the inter-trial interval of a working memory task, were present throughout the next trial and predicted the behavioral bias.

The observed results indicate that frontal and parietal regions are involved in the storage of memory traces and their transfer into current processing.

Rather than looking for reactivation of past information, recent studies have investigated the neural activity elicited by the current stimulus itself. This approach provides a more direct measure of how the current neural representation is influenced by the recent past. Neural activity in the visual cortex measured with functional magnetic resonance imaging (fMRI) was influenced by orientations presented on a previous trial during the encoding and maintenance of a current orientation (*12*). However, as the current and previous orientations were either identical or orthogonal to each other, it remained unclear whether this neural bias reflected an attraction or repulsion in relation to the past. In a more recent fMRI study (*13*), the current neural representation of an orientation in the visual cortex was repulsed from an orientation remembered in the previous trial, which was in contrast to an observed behavioral attraction. In line with this finding, a recent magnetoencephalography (MEG) study (*14*) observed a neural repulsion of a current representation during encoding, in contrast to a behavioral attractive bias towards past information. The authors of both studies proposed that the attractive behavioral bias is driven by post-encoding processes. This view is supported by another computational model (*15*), which states that early visual representations are repulsed and the behavioral attraction emerges during post-encoding read-out of memorized information.

Until now, a direct reflection of the attractive bias of current information on the neural level has not been observed. The current study applied an inverted encoding model analysis to neural signals measured with magnetoencephalography (MEG) to reconstruct[1] the information about currently remembered stimuli. We expected that a direct neural signature of serial dependence should mirror the attractive bias in behavior by showing

[1] The term 'reconstruction' or 'reconstructed representation' refers to the reconstructed channel response, i.e. the output of the IEM analysis. It is not supposed to relate to the reconstruction of single-unit responses. It is used throughout the manuscript for brevity and readability.

an attractive shift of the reconstructed information towards the target information in the previous trial. The aforementioned models (*13*, *15*) have assumed that neural representations become attracted towards the recent past after encoding. We employed a task that allowed us to differentiate between early processes including encoding and maintenance and later, post-encoding read-out processes to inform these models and determine not only if, but also during which processing stage the neural representation becomes attracted towards the recent past. More specifically, we asked subjects to encode two motion directions and memorize them for a short delay, followed by a retro-cue that indicated which direction served as the target for an upcoming report. This retro-cue design, in combination with the high temporal resolution of MEG, enabled us to track the current representations across the whole trial to determine not only if, but also during which processing stage ~~when~~ the neural representations become attractively biased towards the direction reported in the previous trial.

Results

Behavioral Results

Ten human participants (seven females) memorized two sequentially presented visual objects per trial (S1 and S2), while their neural signals were measured via MEG. Each participant completed two sessions with a total of 1022 trials. The objects were colored dot fields and participants had to remember the motion direction of the dots. After a short delay, one of the motion directions was cued via the dot color (red or green, Fig. 1a). As expected, we observed a serial dependence towards the target of the previous trial, as the reported target motion direction of the current trial was systematically attracted to the retro-cued motion direction of the previous trial. This attraction (Fig. 1b) followed a derivative-of-Gaussian (DoG)-shaped curve with an amplitude parameter of 3.51° (bootstrapped SD: 0.479°, lower 95% of permutations between -3.51° and 2.03°, $p < .001$, $R^2 = .236$) and a w parameter of 0.024, equaling a width of 47.37° (full width at half maximum, FWHM). We ran one-sided permutation tests (n = 10) to assess whether this attraction differed from zero. In contrast, we did not observe a significant serial dependence towards the non-target of the previous trial (amplitude: -1.02°, bootstrapped SD: 0.788°, middle 95% of permutations between -1.63 and 1.49, $p = .206$, $R^2 = .014$; w parameter: 0.046, equaling 24.53° FWHM), which replicated our previous observations (*2*, *3*).

Figure 1

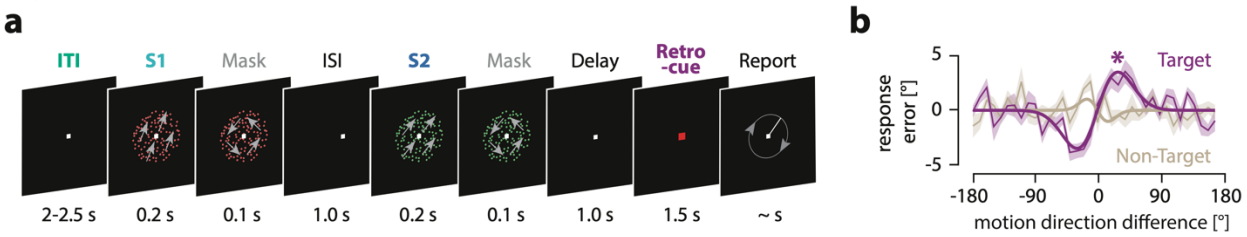

Fig. 1. Experimental paradigm and behavioral results. (A) Participants had to remember two motion directions (S1 and S2) per trial (motion directions are indicated here by gray arrows for illustration only). The stimuli were either red or green. After a short delay, a retro-cue indicated which of the two motion directions (red or green) they had to recall by adjusting a randomly oriented line. (B) Response errors in the current trial were biased towards the target direction of the previous trial (pink curve), but not towards the previous non-target direction (grey curve). The x-axis shows the motion direction difference between the previous and the current target, with positive values indicating that the previous target direction was more clockwise than the current target direction. The response error on the current trial is displayed on the y-axis, with positive values indicating a clockwise deviation from the true target direction. Therefore, shifts of the response errors to the lower left and upper right quadrant indicate an attraction whereas shifts to the upper left and lower right quadrant indicate a repulsion. The asterisk marks significant attractive serial dependence towards the previous target direction.

We also estimated serial dependence between the current and the previously cued object with regard to their color congruence. Serial dependence was modulated by task-relevant color. It was stronger when the current item had the same color as the cued item of the previous trial (amplitude = 4.02°, SD = 0. 528°, lower 95% of permutations between -3.70° and 2.41°, $p < .001$, $R^2 = .167$) than when their colors differed (amplitude = 2.99°, SD = 0. 538°, lower 95% of permutations between -2.99° and 1.79°, $p = .001$, $R^2 = .106$). The amplitude difference amounted to 1.02°, $p = .038$. This replicated the context effect found in our previous study (*2*). The behavioral results thus showed that our two-item working memory task, performed during a MEG recording, produced reliably an attractive bias between target items across trials.

MEG Results

We applied an inverted encoding model (IEM) (*16–18*) to the MEG data to obtain reconstructed representations of the remembered (and currently relevant) motion directions as well as the no longer relevant target motion direction from the previous trial. If the MEG signal at a given time point carried direction-specific information, the reconstruction should show a graded response function with a peak at the given direction. If there was no direction information present, the reconstruction should be (near to) flat. To evaluate whether the reconstruction contained direction information, we aligned all reconstructed response functions to a common center direction (set to 180°), computed its fidelity and tested it at each time point (in the resolution of 100 ms) against 0 (no information present), using a cluster-based permutation test (exact test, 1-tailed) (see Methods for details).

Reconstruction of the remembered directions per epoch

First, we estimated how well we could reconstruct the motion direction of both current items throughout the current trial. For this aim we divided the current trial into three epochs including the encoding and maintenance epochs of S1 and S2 (termed S1 and S2 epochs, respectively) and the retro-cue epoch. Specifically, during the S1 epoch, we reconstructed the direction of S1, during the S2 epoch we reconstructed the direction of S2 and during the retro-cue epoch, we reconstructed the motion direction of the currently retro-cued S1 or S2 (current target). For all three epochs, we observed successful reconstructions of the corresponding motion directions with a fidelity significantly greater than zero (S1 epoch: $p < .002$, cluster extent: 300 to 1200ms, mean fidelity = 0.139; S2 epoch: $p < .002$, cluster extent: 0 to 1300ms, mean fidelity = 0.120; retro-cue epoch: $p < .002$, cluster extent: 300 to 1500ms, mean fidelity = 0.115) (Fig. 2a-b). This shows that the

information about the motion direction of S1, S2 and the retro-cued target was represented in the neural signals during the respective epochs.

For completeness, we also reconstructed the direction of S1 during the S2 epoch, as the direction of S1 was still maintained in working memory in parallel to the concurrent encoding and maintenance of the S2 direction (Fig. 3 a-b). We observed that during the S2 epoch, the direction of S1 could be successfully reconstructed with a fidelity significantly greater than zero (S1 during S2 epoch: $p = .005$, $p_{Bonferroni} = .015$, cluster extent: 100 to 700ms, mean fidelity = 0.065). However, its mean fidelity within the significant cluster was significantly smaller than the fidelities of the currently relevant direction during all other epochs, i.e. S1 during S1 epoch, S2 during S2 epoch and target during the retro-cue epoch (all p-values below .005).

Furthermore, we tried to reconstruct the motion direction that was not retro-cued during the retro-cue epoch, i.e., the non-target of a given trial that became task-irrelevant after the retro-cue. As expected, the information about the non-target item could not be reconstructed in the retro-cue phase (all cluster p-values > .05).

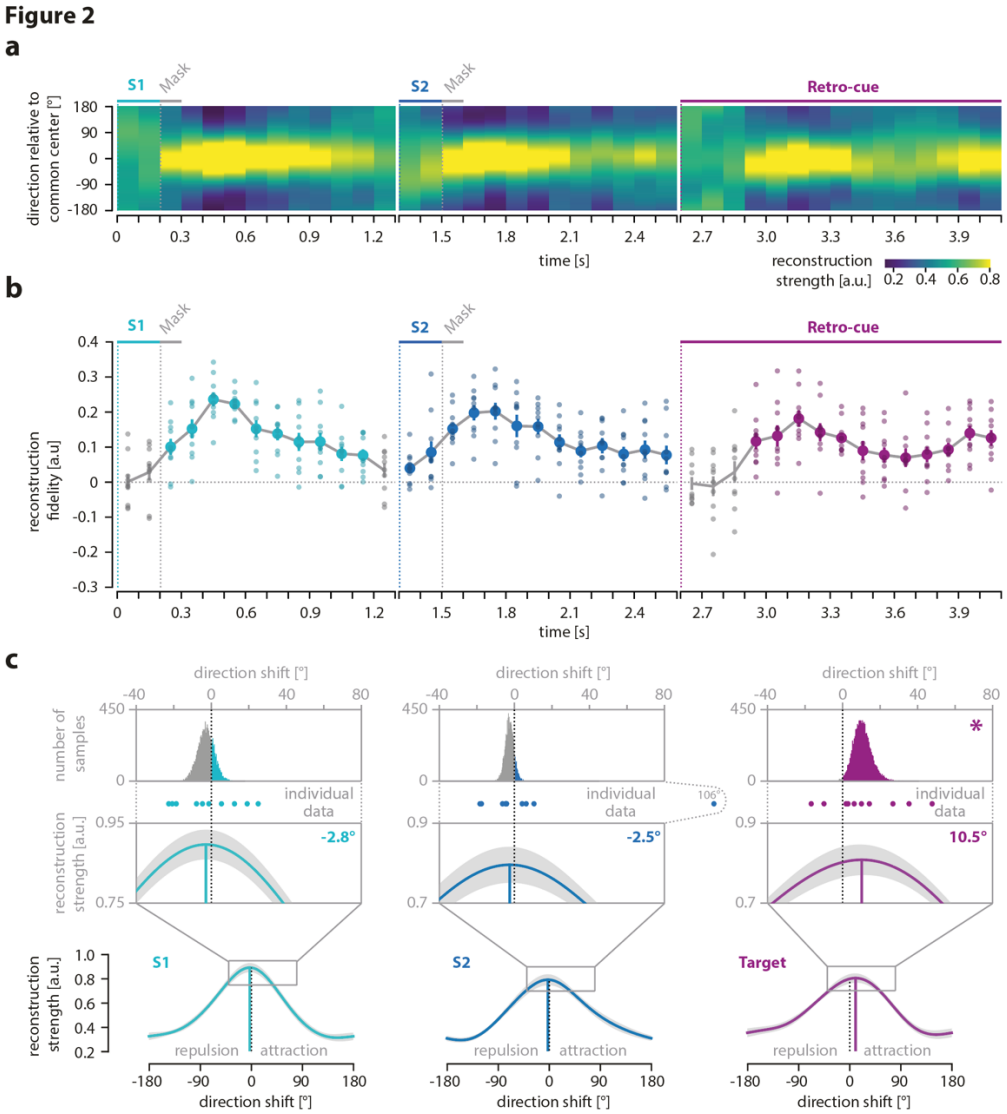

Fig. 2. Reconstruction, fidelity and systematic shifts of currently relevant motion direction. (A) The reconstructed direction of the currently relevant motion direction across the course of a trial. The left panel shows the reconstruction of the S1 direction, the middle panel the reconstruction of the S2 direction and the right panel the reconstruction of the cued direction (either S1 or S2). Higher reconstruction strength is depicted in yellow and lower strength in dark blue. Time points (x-axis) with direction information show a pronounced yellow area around the common center set to zero (y-axis) and darker areas with increasing distance from this center. (B) Fidelity of reconstructed direction of the currently relevant motion direction across the course of a trial. The left panel shows the fidelity of the S1 direction, the middle panel depicts the fidelity of the S2 direction and the right panel shows the fidelity of the cued direction (either S1 or S2). Small circles show the fidelity of each participant. Colored circles indicate time points with a fidelity significantly greater than zero (cluster-based permutation test within each epoch). (C) The reconstructed direction during

all time points with a significant reconstruction fidelity was averaged within each epoch and the maximum of the reconstructed direction was compared to the common center set to zero. A negative shift of the reconstructed mean indicates a repulsion from the target direction of the previous trial whereas a positive shift indicates an attraction. For a more detailed visualization, the range of direction shifts from -40° to +80° is magnified and reconstructed shift values for each participant are indicated by dots. The upper row shows the bootstrapping distributions (see section on MEG analysis below for details). Significant shifts are marked with an asterisk.

Attractive shift of motion direction reconstructions as a direct signature of serial dependence

The main analysis of our study aimed to identify the neural signature of serial dependence. For this aim we looked for an attractive bias of the reconstructed neural representations that mirrored the attractive behavioral bias by testing whether the mean of the significant reconstructions within the S1, S2 or retro-cue epochs showed positive shifts of their maxima from the common center of 180° towards the previous target (Fig. 2c; see Methods for details). We found that the shift of the mean reconstruction of the first and second item did not differ from 0° for the S1 epoch (-2.83°, $p = .743$, bootstrapping test, $p_{FDR} = .870$, 95% of permutations between -11.60° and 6.00°) or for the S2 epoch (-2.50°, $p = .870$, bootstrapping test, $p_{FDR} = .870$, 95% of permutations between -7.57° and 2.11°), respectively. In contrast, for the retro-cue epoch we found that the mean reconstruction of the current target was shifted towards the previous target (10.22°, $p = .009$, bootstrapping test, $p_{FDR} = .026$, 95% of permutations between 1.64° and 21.97°; Fig. 2c). In addition, there was an attractive shift of the S1 direction towards the target of the previous trial during the S2 epoch (11.69°, $p < .001$, bootstrapping test, 95% of permutations between 4.96° and 31.06°; Fig. 3c). Thus, our MEG data provided evidence for an attractive distortion of current neural representations toward the target information in the previous trial, thereby revealing a direct neural signature of serial dependence.

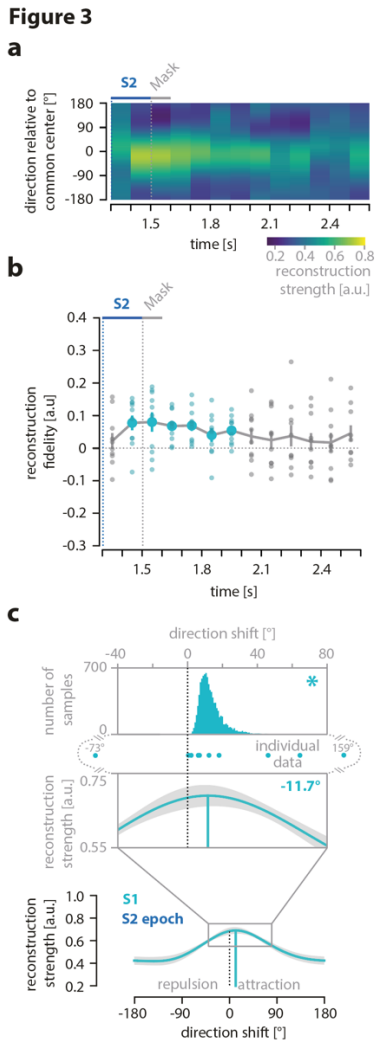

Fig. 3. Reconstruction, fidelity and systematic shifts of S1 direction during the S2 epoch. (A) The reconstructed direction of the S1 motion direction during the S2 epoch. (B) Fidelity of reconstructed S1 motion direction across the course of the S2 epoch. Small circles show the fidelity values for each participant. Colored circles indicate time points with a fidelity significantly greater than zero. (C) The reconstructed direction for all time points with a significant reconstruction fidelity was averaged within the S2 epoch to indicate attractive and repulsive shifts in relation to the target of the previous trials. Significant shifts are marked with an asterisk. For details, refer to Figure 2.

Correlation between neural representation and behavioral response

To assess whether the observed item reconstructions from the MEG signals were behaviorally relevant, we correlated them with the subjects' behavioral responses.

Specifically, we calculated the circular correlation between the maximum of the stimulus direction reconstructions and the response error for each time point and epoch separately across *single trials*, i.e., we tested whether the shift/deviation of the neural response from the true value, e.g., in the clockwise direction, was accompanied by a corresponding shift/deviation of the behavioral response.

Indeed, Fig. 4 shows that-the reconstructed shift of the retro-cued target item during the retro-cue phase predicted the upcoming behavior between 400 to 800ms after the onset of the retro-cue ($p = .003$, permutation test, $p_{Bonferroni} = .009$, mean r = .030). In contrast, there were no significant correlations for any other time points during the S1 and S2 epochs (cluster-based permutation test (exact test, 1-tailed). While the maximum of the S1 reconstruction also showed a significant shift towards the previous target, there were no significant correlations of the S1 reconstruction with the response error during the S2 epoch (no clusters with a p-value < .05).

Figure 4

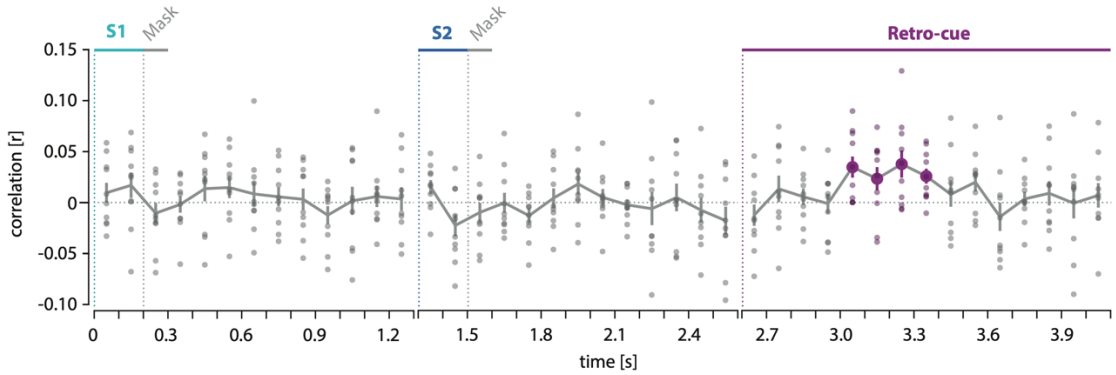

Fig. 4. Correlation between single-trial shift of reconstruction and upcoming response. For each time point, the shift of the maximum of the reconstructed direction in a single trial was correlated with the response deviation from the target direction. Small circles show the r-values of each participant. Time points where the shift of the reconstructed direction significantly correlated with the upcoming response deviation are indicated by colored circles (cluster-based permutation test within each epoch).

The presented single trial-wise correlation between the neural shift and behavioral shift did not directly incorporate the behavioral bias towards the previous trial. Thus, in order to relate the neural attractive shift and the behavioral indicators of serial dependence more

directly, we performed an additional correlation analysis on aggregated data on the *between-subject* level. In detail, we correlated the individual neuronal shift during the retro-cue epoch with the two individual parameter fits of the behavior shift, i.e., amplitude (a) and tuning width (w). In line with the correlation analysis on single-trial level, the analysis on the *between-subject* level also revealed a significant correlation between the w parameter of serial dependence and the neural shifts for the retro-cue epoch. Details of this correlation analysis are displayed in the supporting information "Correlation between the neural shift and individual serial dependence parameters" and Fig. S1.

Reactivation of the previous target direction

Finally, we used our IEM analysis approach to replicate previous results that have shown a reactivation of memory traces from the previous trial before the stimulus presentation of the current trial (*10*). Specifically, we reconstructed the motion direction of the target from the previous trial during the one second of the inter-trial interval (ITI) immediately preceding the current trial (Fig. 5). Consistent with previous findings from monkey single cell and human EEG recordings (*10*), we observed that the information about the previous target was reactivated just prior to the onset of the next trial. Direction reconstructions showed a fidelity significantly greater than zero during a time period at the end of the ITI ($p = .019$, cluster extent: -200 to 100ms (relative to S1 onset), mean fidelity = 0.069). Importantly, we also tracked the information of the previous target throughout the whole current trial to test whether it was reactivated also during the processing of the current items. Interestingly, we found that the target from the previous trial was reactivated once more, but at a later time period during the current trial, i.e., when the retro-cue indicated which item of the current trial was the target for the upcoming response ($p = .040$, cluster extent: 2700 to 2900 ms relative to trial onset, mean fidelity = 0.057). In contrast, we found no evidence of previous target reactivations during the encoding or maintenance of current items (no clusters with a p-value < .05).

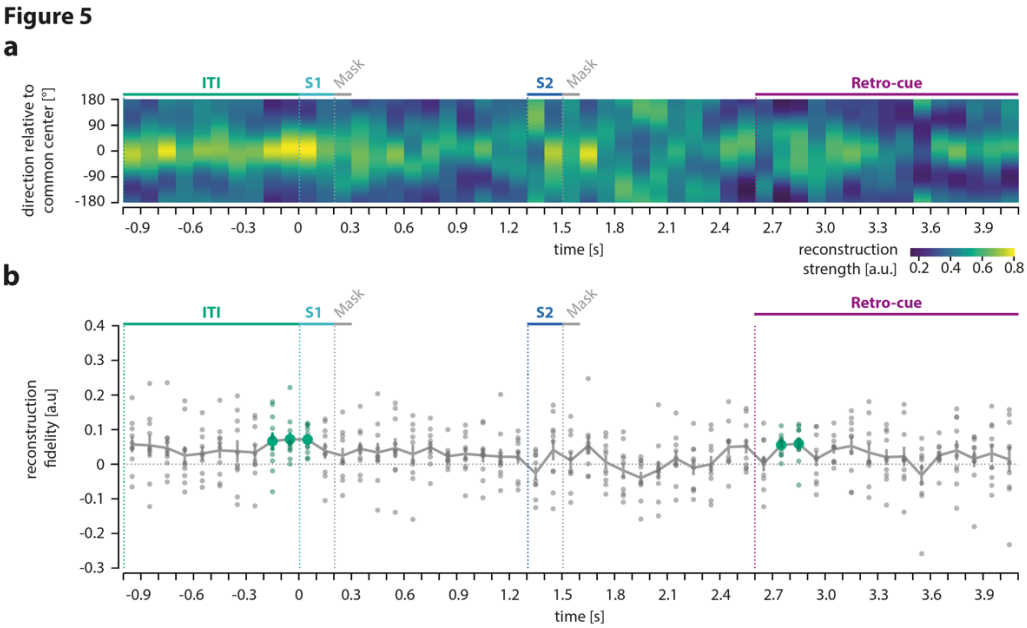

Fig. 5. Reconstruction and fidelity of previous target direction. (A) The target direction of the previous trial was reconstructed throughout the last second of the inter-trial interval (ITI) and the whole following trial. (B) Fidelity of reconstructed direction of the previous target. Small circles show the fidelity of each participant. Colored circles indicate time points with a fidelity significantly greater than zero. For details, refer to Figure 2.

Additionally, in order to test whether a re-activation of the previous target could explain the neural attractive shift observed in the current trial, we reconstructed the motion direction of the previous target from MEG data recorded during the different epochs of the current trial. Specifically, we trained the encoding model either on the presented stimuli in the S1 and S2 epochs or on the currently relevant item during the retro-cue epoch, respectively, and tried to reconstruct the target direction of the previous trial within each epoch. We observed that there was no cross-reconstruction between the current and the previous representation during the retro-cue epoch, which indicates that the observed attractive shift of the current target reconstruction following the retro-cue was not driven by a reactivation of the previous target in the same neural code. However, during the beginning of the S2 epoch, we observed a significant cross-reconstruction, i.e. previous target reactivation co-occurred with the reconstruction of S2 while it was presented to the subjects. The complete analysis and results are described in the supporting information

"Cross-validation reveals different codes for reactivation of previous information and reconstructed biases of current information" and Fig. S2.

Eye movement control

To control for influences of eye movements on the MEG data, we recorded horizontal and vertical electrooculograms to identify blinks and eye-movement-related independent components of the MEG data that were removed during our MEG preprocessing (see Methods). In addition, we also recorded continuous gaze position with a higher spatial resolution using an eye-tracker. We used the eye-tracking data to estimate whether there was a systematic relationship between eye-movements and the motion direction of currently processed items in the S1, S2 or retro-cue-phases. We found that gaze directions were systematically related to presentation and memorization of the stimulus direction (S1 and S2 epoch) and its cue-based retrieval (retro-cue-phase). They also varied considerably between subjects. Most importantly, gaze directions were not systematically related to the MEG data, thus ruling out that they had driven our MEG results. The details of the eye movement control analysis are displayed in the supporting information "Eye movement control" and Fig. S3.

Discussion

The current study aimed to identify a direct neural signature of serial dependence. Previous studies searching for neural markers of this attractive bias have investigated either the reactivation of information from the past trial in the current trial, or have tried to directly capture biases of current item representations by information from the past. The present study found evidence for both types of neural correlates. We replicated reactivations of target representations of the previous trial. More importantly, our study provides the first evidence for an attractive bias of neural representations in the current

trial towards information from the previous trial, which mirrors the behavioral effects of serial dependence and as such represents a neural signature of serial dependence.

The reactivation of past information in the current trial is crucial because an attractive bias can only occur when information from the past has left a memory trace that interacts with the processing of the current object. Barbosa, Stein and colleagues (*10*) have shown that the working memory trace of an object representation from a previous trial, measured via decoding of both neural activity in monkey prefrontal cortex and human EEG signals, was reactivated just before new visual input was presented in the current trial. This was in line with modulations of early visual evoked potentials in human EEG by past stimulus information found in numerosity tasks (*19*, *20*). Using human EEG, Bae and Luck (*21*) observed the reactivation of past stimulus information also during encoding and maintenance phases in the current trial of an orientation working memory task. In our study, we applied a similar analysis approach and tracked the reconstructed motion direction of the target item from the previous trial both during the inter-trial interval, i.e., prior to the encoding of the new stimulus, and throughout the whole current trial. Similar to the previous studies, we also found that the memory trace of the previous target was reactivated just prior to the onset of the current trial and later in the current trial after the retro-cue indicated the current target.

Reactivations of information from the previous trial may be seen as crucial but indirect neural correlates of serial dependence. A memory trace of a previously relevant representation is an important prerequisite for the emergence of serial dependence, since a current representation can only be biased by integrating the information from the past into the present trial. However, our main goal was to observe the direct effect of this past information on current representations. More specifically, we asked whether neural representations of current information are biased towards the recent past in the same way as the behavioral response. Previous studies that aimed to measure neural correlates of serial dependence directly found repulsive rather than attractive distortions of neural

representations in relation to the recent past, which were in contrast to the attractive bias in the behavioral response (*13*, *14*). The repulsive neural biases were attributed to visual adaptation during early perceptual processes. In contrast to these previous reports, the present study found attractive biases of representations in the present trial towards information from the previous trial that were consistent across the neural and behavioral level.

We propose that our combination of a retro-cue paradigm with the high temporal resolution of MEG accounts for the successful demonstration of serial dependence at the neural level in the present study. Sheehan and Serences (*13*) recorded fMRI while participants made a delayed report about a single memorized grating orientation. In contrary to the attractive behavioral bias, the orientation reconstructed from the fMRI signals measured in the visual cortex was repulsed away from the orientation presented in the previous trial. The authors proposed a computational model in which early encoding-related signals are repulsed from the recent past. The attraction observed in behavior was attributed to a later decoding process that overcomes the repulsion of sensory signals and leads to serial dependence but is not directly detectable in hemodynamic response patterns. This is consistent with the notion of an ideal observer model that combines two steps of object processing including an early encoding and a later decoding of current information to explain serial dependence (*15*). According to this model, a repulsive bias occurs early during encoding, whereas an attractive bias is due to the read-out of the current object representation in a Bayesian history-dependent manner. Previous behavioral studies have already identified read-out as a crucial processing stage for serial dependence. Fritsche and colleagues (*6*), for example, asked participants either to remember and report the orientation of a grating after a short delay (working memory task) or to compare two simultaneously presented orientations directly (perception task). They found a repulsive bias in the perception task, when the grating was still visually present and compared to another grating, whereas in the working memory task, the

current target orientation was attracted towards the orientation reported in the preceding trial. This supported an emergence of serial dependence at a late, post-perceptual, processing stage.

Similarly, Hajonides and colleagues (*14*) observed a repulsion of currently remembered orientations as decoded from MEG signals, while the behavioral recall was attracted towards the past. Participants performed a working memory task with two sequentially presented sample stimuli similar to our study. However, instead of using a temporally separated retro-cue, Hajonides and colleagues (*14*) presented a cue that indicated the target for later recall simultaneously with the second sample stimulus. As participants knew which item would be the response target as soon as the second item was presented, the read-out process for a specific item could not be disentangled from the encoding or maintenance stages. Multivariate analysis of MEG signals revealed that the reconstructed orientation information during and shortly after stimulus presentation was repulsed from the orientation recalled in the previous trial. In line with Sheehan and Serences (*13*), the authors assumed that the observed repulsion at the neural level reflected visual adaptation, whereas that the attraction of the neural representation may have occurred at later processing stages, which were not accessible by their analysis. Hajonides and colleagues (*14*) speculated that the integration of past memory traces, putatively stored in parietal areas (*11*), only happened during the behavioral response. This would mean that the neural representation of the current representation is not affected by serial dependence, but that this bias emerges as a result of integration processes during the response.

The present study provided support for the notion that the neural markers of serial dependence occur during post-perceptual phases of a working memory task, which is in agreement with existing models and behavioral findings (*13–15*). The use of a retro-cue task, which facilitates the segregation of earlier and later processing stages, in combination with motion direction reconstructions from MEG signals, revealed shifts of

the current representation towards the motion direction of the target in the previous trial. We observed such an attractive bias only at post-encoding processing stages in two time windows: First, for S1 during and after the presentation of S2 and, second, for the cued target stimulus after the presentation of the retro-cue. This demonstrates that serial dependence can be observed on the neural level as a distortion of current neural representations that mirrors the attractive bias in behavior. This direct neural signature of serial dependence emerges after encoding but prior to the response, thus speaking against the proposal that serial dependence is attributable to response-related processes (*14*).

To further underscore the relevance of the observed neural distortions, we examined whether the reconstructed neural representation predicted the behavioral response deviation. Indeed we observed a positive correlation between the deviations of the neural reconstruction and response errors on a single-trial level. Interestingly, this correlation was found during the presentation of the retro-cue, coinciding with the time point at which the neural representation of the current target stimulus was biased towards the recent past. This observation was further supported by a positive correlation between the attractive shift of the neural reconstruction during the retro-cue epoch and a behavioral measure of serial dependence on the between-subject level. In addition, we observed a reactivation of previous target information during the retro-cue presentation. The distortion of the current neural representation could thus be the result of an integration of past information. Taking these findings together, the retro-cue period emerges at the time point when different neural markers of serial dependence appear together. This suggests that the selective read-out of memorized information for an upcoming report might be a crucial process for serial dependence during working memory.

The neural representation of S1 already became attracted towards the target of the previous trial when S2 was presented. Notably, the reconstruction of the S1 direction had dropped to baseline and reappeared during the S2 epoch. This fits well with the

observation that memory content that has become neurally silent can be successfully reconstructed from the impulse response elicited by a "ping" or probe stimulus (*22*, *23*). This impulse-driven reconstruction is thought to reflect a "hidden state" of working memory, in which unattended information is stored. Therefore, the observed S1 reconstruction elicited by S2 presentation in our study likely reflects the same type of memory read-out, which might have driven the observed neural distortion. On the other hand, S1 had been maintained for a certain time when S2 was presented. As Bliss and colleagues (*5*) observed, serial dependence in behavior increases as a function of delay duration, so the distortion of the neural representation of S1 during the S2 period could also reflect an increasing bias during working memory maintenance.

Several recent studies have shown that serial dependence operates already during perception (*4*, *24*). In the current study, the neural representation of the current information became biased only after encoding. However, we did observe a reactivation of previous information as an indirect marker of SD before and during the onset of S1, replicating previous results (*10*). This early reactivation of past information could reflect a comparison of prior information with new visual input. Cicchini and colleagues (*25*) proposed a predictive coding framework, in which priors informed by past visual input are compared against incoming visual information during early encoding. This comparison results in a prediction error, which determines whether a current stimulus will be biased towards the past. In combination with the observed distortion of the neural representation, this could indicate a two-step process: New visual information is compared against past information during early encoding processes and the integration of past and current information happens during later processing stages, before the actual report is given.

How can these findings of serial dependence operating already during perception (*4*, *24*) be reconciled with other behavioral studies and our findings that pinpoint the emergence of serial dependence at post-perceptual processing stages (*5*, *6*)? One possibility is that

serial dependence does not operate at a single but rather at different levels of object processing including early perceptual and later, memory-related processes (*26*). Moreover, perception and working memory might not be as different as they appear: both functions build mental representations from sensory input, involving feedforward and feedback processing. Hence, encoding and decoding of information, including the neural computations in lower- and higher-level cortical areas, represent integral parts of both perception and working memory. Thus, in a perceptual task that requires an immediate response when visual input is still physically present, read-out of the noisy information via decoding is closely intertwined with its encoding. In working memory tasks that require encoding and memorization of several items, the read-out of a single item is typically postponed: It starts when a retro-cue or probe indicates which memorized item is task-relevant. Serial dependence that operates during the read-out stage of object processing appears to happen both early in perception and later in working memory tasks. However, as shown in this study, in order to identify the neural basis of serial dependence, a multi-item working memory task offers the advantage of separating different processing stages to isolate the time point when an object representation is actually distorted. Furthermore, while the present results indicate a particular importance of post-encoding processes for serial dependence in working memory, they do not suggest that serial dependence could not emerge during encoding. First, we cannot exclude that our data analysis was not sensitive enough to detect small shifts in the neural signals during encoding, especially as our analysis involved averaging across time points to increase the signal-to-noise ratio. Second, with the used motion direction stimuli, reconstruction strength built up slowly during the stimulus presentation, reaching its maximum only after stimulus offset. Possibly, different stimulus material with faster build-ups might help to track the neural representation during early encoding stages.

In summary, our study provides a direct neural signature of serial dependence that mirrors the attractive behavioral bias. This demonstrates that serial dependence in working

memory directly acts on memory representations and biases them towards the past. The attractive distortion of the current representation happened after object encoding, suggesting that read-out processes during object decoding are important for the emergence of serial dependence.

Materials and Methods

Experimental Design

Subjects

Thirteen participants were recruited from the Goethe-University Frankfurt and the Fresenius University of Applied Sciences Frankfurt. All subjects reported normal or corrected-to-normal vision and were screened for red-green-blindness. One subject participated as a pilot subject for only one MEG session, one subject was excluded after the first behavioral training session because of dependency on glasses and one subject dropped out after the first MEG session due to exhaustion because of the MEG setting. This resulted in a sample of 10 subjects (3 male), aged between 22 and 30 years (mean: 25.2 years). All subjects gave informed consent and received either a compensation of €10/h or course credit. The study was approved by the Ethics Committee of the Medical Faculty of the Goethe-University Frankfurt am Main.

Experimental Paradigm

Stimuli. Random dot patterns (RDP) were presented centrally on the screen and consisted of 200 dots colored in red (RGB: 255, 0, 0) or green (RGB: 0, 86.4105, 0) on a black background. The green was adjusted to match the luminance of the red in DKL color space and then transferred back to RGB space. The dots were presented within an invisible circular aperture which had a radius of 7.5° of visual angle. The dots had a diameter of 0.15° of visual angle and were placed randomly within the circular aperture of the RDP at stimulus onset. The dots moved with a velocity of 3.5°/s and fully coherent in a direction randomly drawn from a pool of directions between 5° and 355° spaced 10° from one another, therefore avoiding cardinal directions. Dots reaching the edge of the aperture were repositioned randomly on the edge of the opposing side of the aperture, so that dot density was kept constant throughout the presentation. Throughout the whole experiment

a white fixation square with a diagonal of 0.15° of visual angle was presented centrally on the screen, except for the cue presentation, when the fixation square changed its color to red or green to cue which item should be reported. The item was reported by adjusting a randomly oriented line to match the reported direction. The response line was white, with a width of 0.6° and a length equaling the dot field radius. The starting point of the line was the fixation square and the end point could be altered so that the line could point in all possible directions.

Procedure. The experiment consisted of a delayed-estimation task. Specifically, each trial began with the presentation of the first stimulus (S1) for 200 ms followed by a noise mask for 100 ms consisting of dots moving with 0% coherence (i.e., randomly) and of the same color as the preceding RDP. After a 1000 ms interval (ISI), the second stimulus (S2) and its noise mask were presented for 200 ms and 100 ms, respectively (Figure 1). Subjects were asked to memorized the motion direction of both RDP for a delay of 1000 ms. Then the fixation square changed its color to red or green for 1500 ms, thereby cueing which motion direction became a target item and had to be reported. Immediately after cue offset, a randomly oriented line was presented. Subjects had to report the target motion direction by rotating the line by moving an MEG-compatible track ball horizontally. No time limit was given for the response and subjects were encouraged to work as precisely as possible. After adjusting the line direction, subjects had to confirm their response by pressing the left track ball button. At the end of each trial, a fixation screen of 2000 to 2500 ms (jittered in steps of 10 ms, randomized) was presented. Subjects were instructed to fixate the fixation square throughout the whole experiment.

Between trials, 36 possible direction differences (-170°:10°:180°) were balanced so that every possible direction difference between the items of the current trial and the previous target occurred equally often. This was done with the restriction that the items of the current trial were not allowed to have the same motion direction difference to the previous target item, i.e. cued for report in the previous trial. Thereby it was ensured that the items

within a trial could not have the same motion direction. Furthermore, the congruence of color and temporal position between target items of the previous and current trial was balanced, which automatically lead to a balancing for the non-target item of the current trial, too and ensured that in each trial the two items had different colors from one another.

Every subject completed 1022 trials in two sessions on different days. Each session lasted approximately 90 minutes and was divided into 7 blocks of 73 trials with breaks in between. After the completion of each block, subjects received feedback ('You deviated by more than 30° in XX% of trials.' - German: 'Du bist in XX% der Durchgänge um mehr als 30° abgewichen.') about their performance in the previous block, displayed on the screen. Subjects were seated at a viewing distance of approximately 50 cm from the display. MATLAB software with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) was used for stimulus generation and presentation. A PROPixx projector (VPixx Technologies Inc.) with a resolution of 1920 x 1080 and running with 120 Hz refresh rate was used.

Prior to the two MEG sessions, each subject was invited for a behavioral practice session. This session consisted of 8 blocks of the task and was therefore slightly longer than the MEG experiment. We screened the participants for red-green-blindness, checked MEG and MRI contraindications, and instructed them in the behavioral task before participants started to practice the task. In addition to the feedback presented at the end of each block as in the MEG part of the study, subjects received a trialwise feedback in this first session. If their response deviated more than 30° from the cued motion direction, we presented feedback ('>30°!'). Thereby they could practice to be as precise as possible for the MEG sessions.

605 MEG recoding and preprocessing

606 **Data acquisition**

607 We recorded MEG with a whole-head MEG system (Omega 2005; VSM MedTech Ltd., Port
608 Coquitlam, Canada) with 275 axial gradiometers (271 active) at a sampling rate of 1200 Hz
609 and without online filtering. Additionally, we recorded electrocardiogram (ECG), vertical
610 and horizontal electrooculograms (EOG) as well as continuous gaze position and pupil
611 dilation of the right eye by an MEG compatible eye-tracker (Eyelink CL 1000, SR Research
612 Ltd.). Head position was continuously recorded via head localization coils placed at the
613 nasion and above both ear canal entrances using ear-plugs. The initial head position was
614 saved at the beginning of the first MEG session and presented to the participants for
615 repositioning at the beginning of the second session as well as throughout each session
616 if repositioning was necessary. Behavioral responses were recorded with an MEG-
617 compatible trackball (Current Designs Inc.), enabling subjects to continuously recall
618 motion directions.

619 **Preprocessing**

620 MEG data were preprocessed using the python-based M/EEG analysis toolbox MNE
621 (version 0.22.0; (*27*)). The MEG signal was first notch-filtered at 50Hz and up to 250Hz in
622 steps of 50Hz to remove line noise and then cut into trials from 1000 ms before S1 onset
623 until 500 ms after response onset. Trials containing SQUID jump artifacts were manually
624 identified and excluded. The remaining trials were then highpass-filtered (1Hz) and
625 decomposed via independent component analysis (ICA) with the fastica method
626 implemented in MNE. The decomposition was then applied to the notch-filtered
627 continuous data and ICs representing cardiac and ocular activity were identified via
628 correlation with ECG and EOG signals. The notch-filtered continuous data were then
629 lowpass-filtered at 25Hz and cut for the respective analysis. To analyze reactivations of the
630 previous target direction, the data were cut into one long epoch (last second of the ITI up

until the end of the retro-cue delay). For the analysis regarding the directions of the current trial, we used three epochs per trial: S1 encoding and maintenance (called S1 epoch), S2 encoding and maintenance (called S2 epoch) and retro-cue presentation and delay (called retro-cue epoch). The epochs started 200 ms before stimulus/retro-cue onset and lasted until 1500 ms after stimulus/ retro-cue onset. For the S1/S2 epochs, only 1300 ms after stimulus onset were included in the analysis as those reflected the presentation + delay time. Those epochs were baseline-corrected (-200 ms to 0 ms), the previously identified ICs reflecting cardiac and ocular activity were removed and the data resampled at 300Hz.

Statistical Analysis

Behavioral Analysis

First, we excluded trials in which the response error was at least 3 SDs higher than the subject's mean response error, or in which the response time exceeded 20 s, indicating potential attentional lapses. We also excluded the first trial of each block, as serial dependence cannot occur on those trials, and demeaned the response errors by subtracting the overall mean response error of a participant from each individual response error to remove general individual response biases independent of serial dependence.

The evaluation of serial dependence was based on individual response errors, defined as the deviation between presented and entered direction. The errors were sorted regarding the difference between the target stimulus of the current trial and the target stimulus from the previous trial as well as the relation of difference between the current item and the item of the previous trial (clockwise or counter-clockwise). The difference was computed by subtracting the direction of the current item from the direction of the item of the previous trial. Therefore, when the current item was oriented more clockwise or more counter-clockwise, this resulted in a negatively or positively signed distance, respectively.

A mean response error for a signed distance (distance*relation) deviating from 0 indicated a systematic response bias. When the sign of this systematic bias matched the sign of the distance between the directions, it indicated an attractive response bias. Conversely, an opposite sign of the systematic bias compared to the signed distance indicated a repulsive response bias. In addition, the signed response errors were sorted according to the color congruence between the two items (same vs different color). The individual mean response biases were used to evaluate the serial dependence per color congruence level. We fitted the first derivative of a Gaussian curve (DoG; e.g. (*1*), a model which is usually used to describe serial dependence. The DoG, given by

$$y = xawce^{-(wx)^2} \tag{1}$$

was fitted to the pooled mean response biases of all subjects (similar to the procedure by (*6*) per factor level, i.e. one data point per subject and distance for the respective factor level. In the DoG, x is the relative direction difference of two stimuli, a is the amplitude of the curve peak, w scales the curve width, and c is the constant $y\sqrt{2}/e^{-0.5}$. The w parameter was constrained to a value range of .01 to .1. We optimized the log likelihoods of our curve fitting using Bayesian adaptive direct search (BADS; (*28*)). BADS alternates between a series of fast, local Bayesian optimization steps and a systematic, slower exploration of a mesh grid. To estimate the variability of the parameters a and w, we bootstrapped the DoG curve fit 1,000 times, sampling the data with replacement on each iteration, and computed the standard deviation of the resulting bootstrapped distributions of a (see (*2*) for a similar procedure). To assess the significance of serial dependence on a group level, we used permutation tests, i.e. we randomly inverted the signs of each participant's mean response error, fitted a new DoG model to the pooled group data and collected the resulting amplitude parameters a in a permutation distribution. We repeated this permutation procedure 1,000 times and reported the percentage of permutations that led to equal or higher values for a than the one estimated for the empirical data as p-values. The significance level was set to $\alpha = 0.05$ (one-sided permutation test).

MEG Analysis

Inverted encoding model analysis. To reconstruct motion directions from neural signals, we used the so-called inverted encoding model technique (*29*). We built an encoding model characterizing motion direction-selective responses in the MEG signal with the underlying assumption that the MEG signal in each sensor reflects a weighted sum of neural activity revealed by different motion directions. Our model consisted of a basis set of 18 channels spanning the room of 0-360° with centers in steps of 20°. The channels had the shape of sinusoids raised to the 18th power (adapted from (*30*) and (*31*)). This model was applied to the MEG data, which was separated in a training and a test data set. For each subject, we trained the model on one MEG session and tested it on the other one in two iterations, so that each MEG session was once training and once test data and training and test data were independent. To analyze the currently relevant motion directions, we applied this reconstruction approach to the three types of epochs (S1 epoch, S2 epoch and post-cue epoch) separately and in each epoch the currently relevant direction was the aim of the reconstruction (S1 epoch: S1 direction, S2 epoch: S2 direction, post-cue epoch: cued target direction). To increase the signal-to-noise ratio, the epoched data were averaged across time bins of 100 ms, centered on 50 ms, 150 ms, etc. after epoch onset. This led to 13 analysis time points for the S1 and S2 epoch and 15 analysis time points for the post-cue epoch. For each epoch, the directions were reconstructed for each time point separately in a time point by time point fashion. To check for reactivations of the previous target direction, we applied the same analysis on the last second of the ITI and the whole time course of the current trial, resulting in 51 analysis time points. This time, the aim of the reconstruction was the previous target direction, therefore the data were trained and tested on this direction.

In the first step, we modeled the activation of each sensor as a weighted sum of filter responses:

$$B_1 = C_1 W \tag{2}$$

B_1 is the MEG signal in each sensor at a given time point (n trials * n sensors), C_1 the modeled response of each channel function give the presented direction in each trial (n trials * n channels) and W a weight matrix representing the contribution of each channel to the measured signal in each sensor (n channels * n sensors). W was estimated by using ordinary least-squares regression:

$$\widehat{W} = (C_1^T C_1)^{-1} C_1^T B_1 \tag{3}$$

In the second step, we inverted W and applied to the independent test MEG data B_2 (n sensors * n trials) to compute the estimated channel responses $\hat{C}_2$ (n channels * n trials):

$$\hat{C}_2 = (\widehat{W}^T \widehat{W})^{-1} \widehat{W}^T B_2 \tag{4}$$

The estimated channel responses $\hat{C}_2$ were transformed from the channel space back to the direction space by weighting $\hat{C}_2$ with each channel's sensitivity from the basis set. Those reconstructed channel responses were then circularly shifted to a common center (180°). This whole reconstruction procedure was done separately per subject, session, epoch and time point and this was iterated 20 times. Across those iterations, the centers of the basis set were shifted by 1° so that the channel centers were evenly spaced in steps of 1° across the whole direction space as a result of those iterations. The single trial reconstructions were then analyzed in two different ways, described in the following paragraphs.

Shifts of the reconstructed direction. The main goal of our analysis was to examine potential shifts of the neural representation in relation to the target item of the previous trial. Therefore, the single-trial reconstructions were ordered in two groups, those with a previous target that was clockwise (CW) oriented in relation to the currently relevant item and those with a previous target that was counter-clockwise (CCW) oriented. The CCW reconstructions were flipped along the direction space. Thereby, for both CW and CCW trials, a negative deviation of the maximum of the reconstruction from 180° indicated an attraction toward the previous target, whereas a positive deviation indicated a repulsion. Those reconstruction were then first averaged within each possible motion direction and

then across them to account for different presentation numbers of the directions. This averaged reconstruction was obtained for each iteration (see above) and then later averaged across sessions. Thereby, the analysis resulted in a reconstruction per participant, epoch and time point. In the next step, the fidelity of those reconstructions was assessed per epoch and time point to identify time points with significant direction information in the MEG signal. A fidelity above 0 indicates a meaningful direction signal and was obtained by projecting the polar vector of the reconstruction at every direction angle (in steps of 1°) onto the common center (180°). The length of this projected vector indicates the information strengths of the reconstruction at each angle. The fidelity F was obtained by averaging those information strengths:

$$F = mean(r(\theta)\, cos\,(\theta)\,) \tag{5}$$

We identified the time points in each epoch with significant direction reconstruction fidelity by computing a one-sided cluster-based permutation test (as implemented in MNE) against zero. As the effect of the neural shift was expected to be very small, we averaged the reconstructions in each epoch only over those time points that showed a significant reconstruction fidelity and thus interpretable data. To examine systematic shifts, we then tested if the maximum of the reconstruction was systematically different from the common center (180°). For display purposes, we subtracted the reconstructed maximum from 180° to compute the direction shifts. A positive shift thus reflected attraction and a negative shift reflected repulsion.

Statistical inference of shifts. To assess the significance of potential shifts, we used a bootstrapping procedure similar to (*32*). To this aim, we randomly selected reconstructions from 10 participants (with replacement), then computed the mean reconstruction over the resampled participants and collected the shift. This was done on 10,000 iterations. The distribution of shifts was then used to examine if the observed shift was significantly larger than zero. A shift was significant when 95% of the resampling iterations led to a shift equal

or higher than 0°, i.e., we conducted a one-sided bootstrapping test as we expected an attraction, which was indicated by a shift larger than 0°. To obtain an indicator for the variability of the shifts, the middle 95% of shifts resulting from the described resampling procedure are reported.

Correlation between neural representation and behavioral response. In addition, we correlated the maximum of single trial reconstructions with the upcoming behavioral response to test whether shifts in behavior and in the neural signal on a single-trial level were systematically related to each other. Both behavioral and neural shifts in any direction (clockwise or counter-clockwise) in a single-trial response can originate from or be modulated by different sources like, e.g., noise in the motor response, but they can also result from attractive serial dependence that can be observed by averaging over trials. We collected the reconstructions before flipping them with regard to the previous target and removed trials where the behavioral response deviated more than three SDs from the individual response error. Then we computed the circular correlation between the deviation of the reconstructed maximum from the presented direction and the response error for each participant, session, epoch and time point. For S1 and S2 epochs, only epochs from trials where the respective item was cued as the response target were included.

Statistical inference of correlations. The obtained correlation coefficients were averaged across sessions and then time points for each epoch with a correlation significantly differing from zero were identified by computing a one-sided cluster-based permutation test (as implemented in MNE) against zero.

Reconstruction of previous target direction. We also examined whether the neural signals of the current trial contained memory traces of the previous target direction. Therefore, another IEM analysis was employed, but this time trained and tested on the target direction of the previous trial throughout the last second of the ITI and the current trial. Time points that contained significant direction information from the previous target were

identified by computing the fidelity of the reconstructions as described above and tested against zero with a one-sided cluster-based permutation test.

Multiple comparison correction. In order to correct for multiple comparisons, we applied an FDR correction (Benjamini-Hochberg) to those analyses that resulted in one p-value per epoch, i.e., the bootstrapping tests of the shifts against zero, and reported these corrected values alongside the original ones. Furthermore, for cluster-based permutation tests that were applied to each epoch individually, we multiplied p-values higher than the minimally possible ($p<.002$) for significant clusters by the number of epochs (3), thus resulting in a Bonferroni correction.

Software

All behavioral analysis were performed with MATLAB 2019a (*33*) and the following toolboxes/functions: Circular Statistics Toolbox (*34*), BADS (*28*) and EzyFit v2.44 (*35*). The MEG analysis were performed with Python 3.8 in the Anaconda distribution v.2020.11 (*36*) using spyder v.4.1.5 (*37*) and the following toolboxes: autoreject v0.2.2 (*38*), MNE v0.22.0 (*27*), numpy v1.19.2 (*39*), pandas v1.1.3 (*40*, *41*) and scipy v1.5.2 (*42*).

References

1. J. Fischer, D. Whitney, Serial dependence in visual perception. *Nature Neuroscience*. **17**, 738–743 (2014).
2. C. Fischer, S. Czoschke, B. Peters, B. Rahm, J. Kaiser, C. Bledowski, Context information supports serial dependence of multiple visual objects across memory episodes. *Nat Commun*. **11**, 1932 (2020).
3. S. Czoschke, C. Fischer, J. Beitner, J. Kaiser, C. Bledowski, Two types of serial dependence in visual working memory. *British Journal of Psychology*. **110**, 256–267 (2019).
4. Y. Murai, D. Whitney, Serial dependence revealed in history-dependent perceptual templates. *Current Biology*. **31**, 3185-3191.e3 (2021).
5. D. P. Bliss, J. J. Sun, M. D'Esposito, Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*. **7** (2017), doi:10.1038/s41598-017-15199-7.
6. M. Fritsche, P. Mostert, F. P. de Lange, Opposite Effects of Recent History on Perception and Decision. *Current Biology*. **27**, 590–595 (2017).
7. D. Pascucci, G. Mancuso, E. Santandrea, C. D. Libera, G. Plomp, L. Chelazzi, Laws of concatenated perception: Vision goes for novelty, decisions for perseverance. *PLOS Biology*. **17**, e3000144 (2019).
8. A. Kiyonaga, J. M. Scimeca, D. P. Bliss, D. Whitney, Serial Dependence across Perception, Attention, and Memory. *Trends in Cognitive Sciences*. **21**, 493–497 (2017).
9. D. Pascucci, Ö. D. Tanrikulu, A. Ozkirli, C. Houborg, G. Ceylan, P. Zerr, M. Rafiei, Á. Kristjánsson, Serial dependence in visual perception: A review. *Journal of Vision*. **23**, 9 (2023).
10. J. Barbosa, H. Stein, R. L. Martinez, A. Galan-Gadea, S. Li, J. Dalmau, K. C. S. Adam, J. Valls-Solé, C. Constantinidis, A. Compte, Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nat Neurosci*. **23**, 1016–1024 (2020).
11. A. Akrami, C. D. Kopec, M. E. Diamond, C. D. Brody, Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature*. **554**, 368–372 (2018).

12. E. St. John-Saaltink, P. Kok, H. C. Lau, F. P. de Lange, Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex. *Journal of Neuroscience*. **36**, 6186–6192 (2016).

13. T. C. Sheehan, J. T. Serences, Attractive serial dependence overcomes repulsive neuronal adaptation. *PLOS Biology*. **20**, e3001711 (2022).

14. J. E. Hajonides, F. Van Ede, M. G. Stokes, A. C. Nobre, N. E. Myers, Multiple and Dissociable Effects of Sensory History on Working-Memory Performance. *J. Neurosci.* **43**, 2730–2740 (2023).

15. M. Fritsche, E. Spaak, F. P. de Lange, A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. *eLife*. **9**, e55389 (2020).

16. G. J. Brouwer, D. J. Heeger, Cross-orientation suppression in human visual cortex. *Journal of Neurophysiology*. **106**, 2108–2119 (2011).

17. G. J. Brouwer, D. J. Heeger, Decoding and Reconstructing Color from Responses in Human Visual Cortex. *Journal of Neuroscience.* **29**, 13992–14003 (2009).

18. T. C. Sprague, J. T. Serences, Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience*. **16**, 1879–1887 (2013).

19. M. Fornaciai, J. Park, Attractive Serial Dependence in the Absence of an Explicit Task. *Psychological Science*, 095679761773738 (2018).

20. M. Fornaciai, J. Park, Neural Dynamics of Serial Dependence in Numerosity Perception. *J Cogn Neurosci*. **32**, 141–154 (2020).

21. G.-Y. Bae, S. J. Luck, Reactivation of Previous Experiences in a Working Memory Task. *Psychol Sci*. **30**, 587–595 (2019).

22. M. J. Wolff, J. Jochim, E. G. Akyürek, M. G. Stokes, Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*. **20**, 864–871 (2017).

23. M. J. Wolff, J. Ding, N. E. Myers, M. G. Stokes, Revealing hidden states in visual working memory using electroencephalography. *Frontiers in Systems Neuroscience*. **9** (2015), doi:10.3389/fnsys.2015.00123.

24. M. Manassi, A. Liberman, A. Kosovicheva, K. Zhang, D. Whitney, Serial dependence in position occurs at the time of perception. *Psychonomic Bulletin & Review* (2018), doi:10.3758/s13423-018-1454-5.

25. G. M. Cicchini, A. Benedetto, D. C. Burr, Perceptual history propagates down to early levels of sensory analysis. *Current Biology*. **31**, 1245-1250.e2 (2021).

26. D. Whitney, M. Manassi, Y. Murai, Searching for serial dependencies in the brain. *PLoS Biol*. **20**, e3001788 (2022).

27. A. Gramfort, M. Luessi, E. Larson, D. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, M. Hämäläinen, MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*. **7** (2013) (available at https://www.frontiersin.org/articles/10.3389/fnins.2013.00267).

28. L. Acerbi, W. J. Ma, "Practical Bayesian Optimization for Model Fitting with Bayesian Adaptive Direct Search" in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett, Eds. (2017; http://papers.nips.cc/paper/6780-practical-bayesian-optimization-for-model-fitting-with-bayesian-adaptive-direct-search.pdf), pp. 1836–1846.

29. T. C. Sprague, K. C. S. Adam, J. J. Foster, M. Rahmati, D. W. Sutterer, V. A. Vo. *Eneuro, 5*(3).

30. E. F. Ester, T. C. Sprague, J. T. Serences, Parietal and Frontal Cortex Encode Stimulus-Specific Mnemonic Representations during Visual Working Memory. *Neuron*. **87**, 893–905 (2015).

31. T. C. Sprague, IEM-tutorial. *GitHub repository* (2016), (available at https://github.com/tommysprague/IEM-tutorial).

32. E. F. Ester, T. C. Sprague, J. T. Serences, Categorical biases in human occipitoparietal cortex. *Journal of Neuroscience*. **40.4**, 917–931 (2020).

33. MATLAB, *R2019a* (The MathWorks Inc., Natick, Massachusetts, 2019).

34. P. Berens, CircStat : A MATLAB Toolbox for Circular Statistics. *J. Stat. Soft.* **31** (2009), doi:10.18637/jss.v031.i10.

35. F. Moisy, EzyFit 2.44 (https://www.mathworks.com/matlabcentral/fileexchange/10176-ezyfit-2-44), MATLAB Central File Exchange. Retrieved April 6, 2018. (2016).

36. Anaconda Software Distribution. *Anaconda Documentation* (2020), (available at https://docs.anaconda.com/).

37. P. Raybaut, Spyder-documentation. *Available online at: pythonhosted. org* (2009).

38. M. Jas, D. A. Engemann, Y. Bekhti, F. Raimondo, A. Gramfort, Autoreject: Automated artifact rejection for MEG and EEG data (2016), doi:10.48550/ARXIV.1612.08194.

39. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T. E. Oliphant, Array programming with NumPy. *Nature*. **585**, 357–362 (2020).

40. W. McKinney, others, "Data structures for statistical computing in python" in *Proceedings of the 9th Python in Science Conference* (Austin, TX, 2010), vol. 445, pp. 51–56.

41. T. pandas development team, pandas-dev/pandas: Pandas (2020), , doi:10.5281/zenodo.3509134.

42. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*. **17**, 261–272 (2020).

Acknowledgments:

We thank Magdalena Feldmann, Nicole Huizinga, Philipp Deutsch and Susanna Wenzel for their help in data collection and Benjamin Rahm for helpful discussions.

Funding:

PhD scholarship by the German Academic Scholarship Foundation awarded to CF

Author contributions:

Conceptualization: CF, JK, CB

Methodology: CF, JK, CB

Investigation: CF

Formal analysis: CF

Visualization: CF, CB

Supervision: JK, CB

Writing—original draft: CF, CB

Writing—review & editing: CF, JK, CB

Competing interests:

Authors declare that they have no competing interests.

Data and code availability:

All code and preprocessed data used for the presented analyses are available on [to be inserted upon publication].

Supporting Information

Correlation between the neural shift and individual serial dependence parameters. In addition to the single trial-based correlation between neural and behavioral shifts reported in the main manuscript, we calculated correlations between the neural shift and individual a and w parameter obtained from DoG serial dependence fits on a between-subject level. This analysis revealed no significant correlations during the S1 and S2 epochs for neither a nor w parameter (a parameter, S1: r = .01, p = .9767; S2: r = .08, p = .8114; w parameter: S1: r = .36, p = .3096; S2: r = .30, p = .4054). While during the retro-cue epoch, there was no significant correlation with the a parameter either (r = -.35, p = .3258), we did observe a significant correlation between individual neural shifts and w parameters (r = -.70, p = .0246). This means that smaller individual w parameters indicating a broader tuning of serial dependence went along with larger neural shifts. This further demonstrates that the neural shift during the retro-cue epoch reflected behavioral serial dependence. These results are displayed in the **Figure S1**.

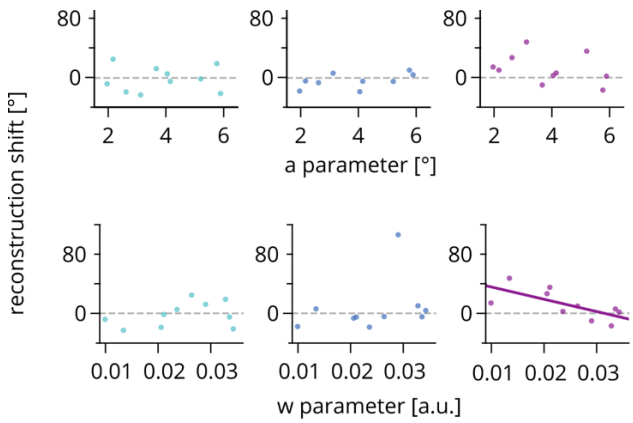

Figure S1. Relation between the individual a (upper panels) and w (lower panels) parameters of a DoG serial dependence fit (x-axis) and the neural reconstruction shift (y-axis). Colored circles show individual values for the S1 (left panels), S2 (middle panels) and retro-cue epochs (right panels). Solid lines depict significant correlations, dashed grey lines mark a reconstruction shift of 0°.

Cross-validation reveals different codes for reactivation of previous information and reconstructed biases of current information. In the retro-cue epoch, we observed both an attractive bias towards the target of the previous trial and a transient reactivation of the previous target itself. Therefore, one alternative explanation of the observed attractive bias could be a reactivation of the previous target in the same neural code as the current target that in sum lead to an attractive shift of the reconstruction. To rule out this explanation, we ran a cross-validation analysis. Therefore, we trained the encoding model on the currently relevant item (S1 during the S1 epoch, S2 during the S2 epoch and the cued item during the retro-cue epoch) and aimed to reconstruct the motion direction of the previous trial.
This analysis revealed that during the retro-cue epoch, where we observed the neural shift toward the past, as well as during the S1 epoch, there was no reactivation in the same neural code as the current representation. In contrast, a reactivation of target information from the previous trial using the same neural code could be observed during stimulus presentation in the S2 epoch. This indicated that during the S2 phase, information from the current S2 and the previous target were reconstructed in the same neural code. These results are displayed in the **Figure S2**.

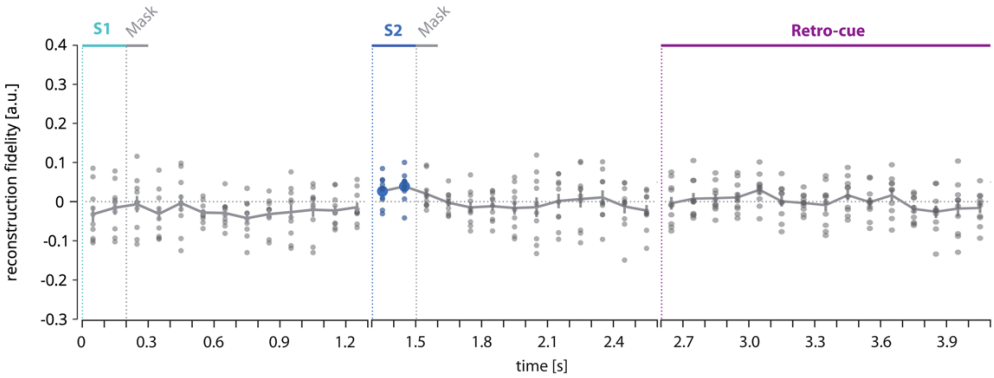

Fig. S2. Cross-validated reconstruction fidelity of previous direction, model trained on currently relevant direction. Fidelity of reconstructed direction of the previous target using a model that was trained on the currently relevant direction. The left panel shows the fidelity of the previous target during the S1 epoch, the middle panel depicts the fidelity during the S2 epoch and the right panel shows the fidelity of the cued direction (either S1 or S2). Small circles show the fidelity of each participant. Colored circles indicate time points with a fidelity significantly greater than zero (cluster-based permutation test within each epoch).

Eye Movement control. Participants were instructed to maintain fixation on a small square at the center of the screen throughout each trial and independent component analysis (ICA) was used during preprocessing to remove any eye movement artifacts. However, recent studies showed that systematic eye movements can potentially confound decoding results (e.g. 1, 2). In the present study, participants viewed dynamic stimuli (moving dot patterns) and had to remember a spatial feature (motion direction), which made systematic eye movement behavior more likely. Therefore, we thoroughly tested a) whether eye movements were associated with the presented/remembered motion direction and if so, b) whether eye movements could have driven our MEG results.

As a first step of our eye movement control analysis, we tested how much of the variance of gaze direction could be explained by the stimulus direction that was either currently presented or had to be maintained at a given moment in time. More specifically, we first transformed the x and y positions measured by the eye tracker (in relation to a baseline from 200ms until trial onset) into polar coordinates to obtain a gaze direction in degrees. This transformation was performed on the same time points as the MEG analysis, i.e., gaze directions were averaged within bins of 100ms. Then we computed the circular variance of these gaze directions separately for each specific stimulus direction and time bin per subject. This circular variance was compared to a random distribution of gaze direction variance. To obtain this distribution, we shuffled the motion direction labels randomly over trials and then computed the circular variance for each pseudo direction, which was unrelated to the direction that was actually presented in the trial. The empirical and shuffled variances were averaged across directions and then a ratio was computed between them. The resulting value reflects the extent to which the variance of the stimulus-related gaze directions was reduced compared with random gaze direction patterns. This procedure was able to reveal not only systematic gaze directions that were in accordance with the stimulus direction or the opposite direction, but also picked up all stimulus-related gaze directions, even if the relation differed across participants or time. We identified time points with a significant amount of stimulus-related gaze directions in each analyzed epoch with cluster-based permutation tests. **Fig. S3a** shows that gaze directions were systematically related to the presented or remembered motion direction at different time points during each epoch. However, stimulus-related gaze directions differed considerably between participants as indicated by the standard errors.

In the second step, we plotted the time courses of stimulus-related gaze direction separately for each subject. This is shown in **Fig. S3b**. This revealed that while some participants showed a high amount of stimulus-related gaze directions, others showed very little or no stimulus-related gaze directions at all. In **Fig. S3b,** we marked the time courses in dark red for subjects with the highest amount of direction-related gaze positions ("high") and dark green for those with the lowest amount ("low"), with gradients in between.

In the third step, we addressed the question of how the stimulus-related gaze directions related to our MEG results. To this aim, we plotted the MEG-based reconstruction fidelity of presented and remembered motion direction separately for each subject and epoch and, crucially, coded them by the amount of stimulus-related gaze directions shown by each subject. This is shown in **Fig. S3c**. Comparing the stimulus reconstruction fidelity time courses with the individual color codes derived from the gaze-direction analyses revealed that the reconstruction fidelities of each participant did not correspond to their amount of stimulus-related gaze behavior (compare **Fig. S3b** with **Fig. S3c**). This suggests that there was no systematic relationship at any time point between the fidelity measure and the amount of direction-related gaze positions.

To support this observation statistically, we calculated Spearman rank correlations between the amount of stimulus-related gaze directions and reconstruction fidelities individually for each time point of each epoch. The resulting p-values are shown in **Fig. S3d**. Even before correcting for multiple comparisons, the p-values were consistently above 0.05, except for one single time point in the retro-cue epoch. Moreover, the p-values within each epoch did not show any consistent time profile but varied considerably between time points, supporting our interpretation that MEG-based reconstruction fidelity was not driven by the amount of stimulus-related gaze directions.

In consequence, we also did not find any relationship between the attractive neural shift that was based on these reconstructed MEG signals (i.e., the main finding of our study) and stimulus-related gaze directions. Specifically, there was no correlation between each subject's rank of the strength of neural shift towards the previous target and the subject's rank of the amount of stimulus-related gaze directions in any of the epochs (S1: $p = .986$, S2: $p = .260$, retro-cue: $p = .244$).

We conclude that our MEG results were unlikely to be explained by eye movements.

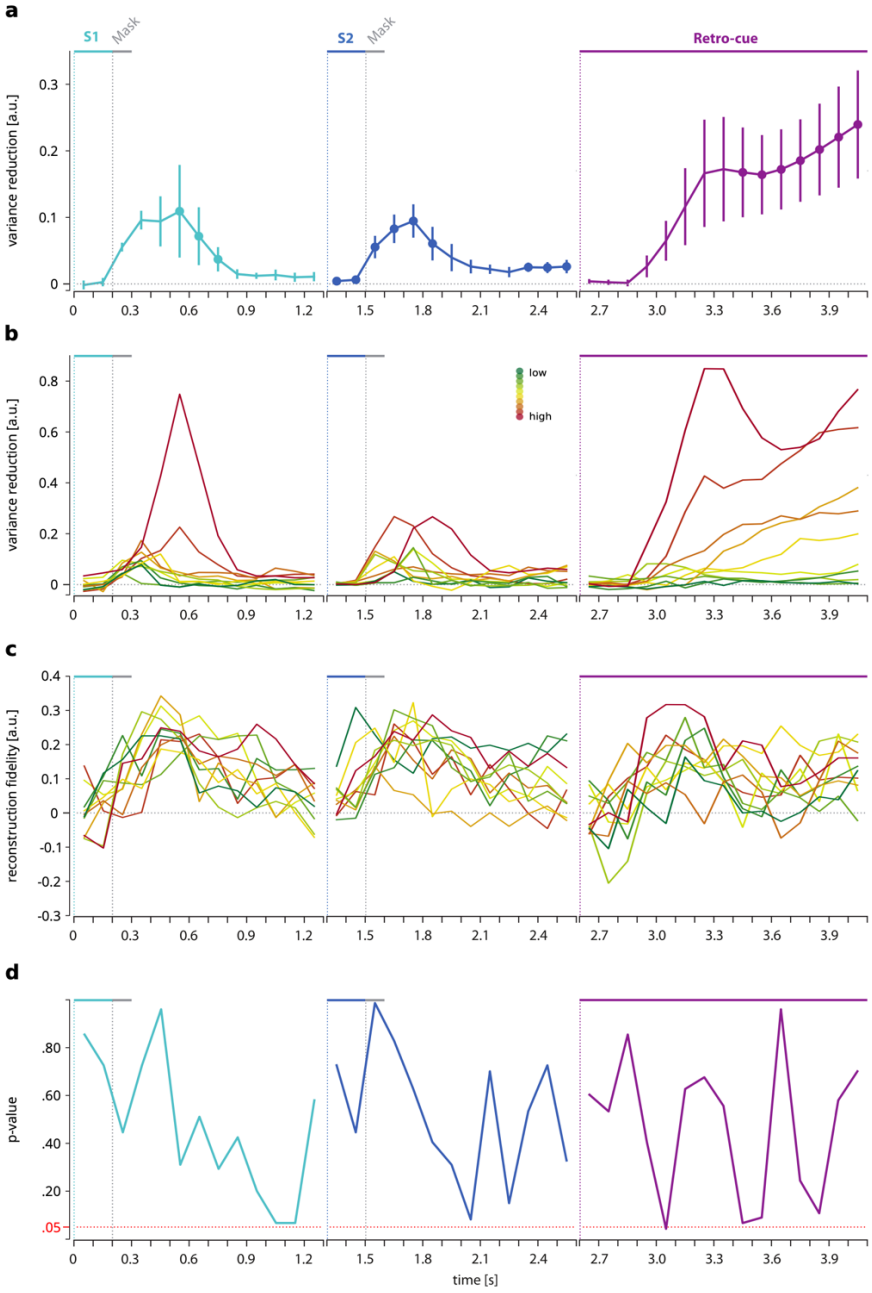

Fig. S3. Systematic eye movements and reconstruction time courses were uncorrelated. (**a**) Amount of stimulus-related gaze direction (indicated by variance reduction) during each epoch as a function of time with error bars indicating standard errors and colored circles indicating time points with a variance reduction significantly greater than zero (cluster-based permutation test within each epoch; p<.05). (**b**) Time courses of individual stimulus-related gaze directions. Dark red indicates the subject with the highest average variance reduction in a given epoch while green indicates the one with the smallest average variance reduction, with color gradients in between according to the individual rank of average variance reduction. (**c**) Time courses of individual reconstruction fidelities. Lines are colored as in (**b**), reflecting the stimulus-related gaze direction rank of a subject in a given epoch. (**d**) Time courses of p-values (uncorrected for multiple comparisons) of time point-

by-time point Spearman rank correlations between the systematic gaze directions as shown in (**b**) and the reconstruction fidelity as shown in (**c**).

SI References

1. Mostert, P., Albers, A. M., Brinkman, L., Todorova, L., Kok, P., & De Lange, F. P. Eye movement-related confounds in neural decoding of visual working memory representations. *ENeuro*, **5**(4) (2018).
2. Quax, S. C., Dijkstra, N., van Staveren, M. J., Bosch, S. E., & van Gerven, M. A. Eye movements explain decodability during perception and cued attention in MEG. *Neuroimage*, **195**, 444-453 (2019).