eNeuro

Research Article: New Research | *Cognition and Behavior*

"Leader-follower" dynamic perturbation manipulates multi-item working memory in humans

https://doi.org/10.1523/ENEURO.0472-22.2023

Cite as: eNeuro 2023; 10.1523/ENEURO.0472-22.2023

Received: 21 November 2022
Revised: 22 October 2023
Accepted: 25 October 2023

"Leader-follower" dynamic perturbation manipulates multi-item working memory in humans

Qiaoli Huang[1-4, *], Minghao Luo[1-3], Yuanyuan Mi[5], Huan Luo[1-3, *]

[1] School of Psychological and Cognitive Sciences, Peking University
[2] PKU-IDG/McGovern Institute for Brain Research, Peking University
[3] Beijing Key Laboratory of Behavior and Mental Health, Peking University
[4] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[5] Department of Psychology, School of Social Sciences, Tsinghua University

Running title: Modulating WM using dynamic perturbation

Address for Correspondence:

* Qiaoli Huang
Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
Stephanstraße 1A, 04103, Leipzig, Germany
qiaolihuang0818@gmail.com

* Huan Luo
School of Psychological and Cognitive Sciences, Peking University
PKU-IDG/McGovern Institute for Brain Science, Peking University
52 Haidian Road, Beijing, 100087, China
huan.luo@pku.edu.cn

Abstract

Manipulating working memory (WM) is a central yet challenging question. Previous studies posit that WM items with varied memory strengths reactivate at different latencies, supporting a time-based mechanism. Motivated by this view, here we developed a purely bottom-up, "Leader-Follower" behavioral approach to manipulate WM in humans. Specifically, task-irrelevant, flickering color discs that are bound to each of the memorized items are presented during the delay period, and the ongoing luminance sequences of the color discs follow a "Leader-Follower" relationship, i.e., hundreds-of-millisecond temporal lag. We show that this dynamic behavioral approach leads to better memory performance for the item associated with the temporally advanced luminance sequence ("Leader") than that with the temporally lagged luminance sequence ("Follower"), yet with limited effectiveness. Taken together, our findings constitute evidence for the essential role of temporal dynamics in WM operation and offer a promising, non-invasive WM manipulation approach.

Significance Statement

Working memory (WM) is known to be the “sketchpad of conscious thought’’ that allows us to temporally hold and manipulate limited amounts of information to guide future behavior. A major challenge in the WM field concerns how multiple items could be simultaneously retained while not be confused with each other. Previous work advocates a time-based mechanism, with the item with stronger strength firing at earlier latency than that with weaker memory. Motivated by the time-based view, here we developed a novel behavioral approach, namely the “Leader-follower” dynamic perturbation, to alter WM performance in humans. Our findings constitute new evidence for a time-based WM mechanism and offers a brand-new behavioral approach to directly manipulate WM, but with the need for replication.

Introduction

Manipulating working memory (WM) is an important yet challenging question, answering which would also provide crucial causal evidence for the WM neural mechanism. WM information is posited to undergo reactivation or refreshing to overcome memory decay during the delay period (Curtis & D'Esposito, 2003; Vogel & Machizawa, 2004), a process that facilitates memory storage via short-term neural plasticity (STP) principles (Miller et al., 2018; Mongillo et al., 2008; Wang et al., 2006). When multiple items are retained, previous models suggest that the item-specific reactivations compete with each other over time (Oberauer & Lewandowsky, 2008, 2011), wherein individual item fires at varied phases according to their respective memory strength (Lisman & Idiart, 1995; Lisman & Jensen, 2013). The item with stronger memory strength, given its higher neural excitability, fires at an earlier latency, while the less excitable item reactivates relatively late (Bahramisharif et al., 2018; Huang et al., 2018, 2021; Siegel et al., 2009), enabling the transformation of memory strengths into neural activities with varied latencies. Hence, a potential yet unexplored WM manipulation approach is to alter the temporal relationship between item-specific reactivations during retention so that their relative memory performance could be modified.

Previous research on noninvasive WM modulation in humans has highlighted several approaches, such as frequency-specific transcranial magnetic stimulation (TMS) and transcranial Alternating Current Stimulation (tACS) (Beynel et al., 2019; Hoy et al., 2015; Sauseng et al., 2009). Moreover, presentation of a retro-cue could prioritize recalling performance via top-down attentional modulations (Griffin & Nobre, 2003; Landman et al., 2003; Myers et al., 2017; Oberauer & Hein, 2012). Recently, we developed a purely bottom-up, behavioral "dynamic perturbation" approach to interfere with the multi-item neural dynamics of sequence WM (Li et al., 2021). Notably, this approach draws upon many theoretical models and empirical findings. First, color features, even task-irrelevant, tend to be automatically bound to memorized items, i.e., object-based WM (Huang et al., 2018; Johnson et al., 2008; Li et al., 2021; Luck et al., 1997). Accordingly, presentation of color discs that are attached to memorize items could possibly reactivate and even modify memories. Second, although WM information has been posited to be stored in an active or activity-silent manner (Curtis & D'Esposito, 2003; Goldman-Rakic, 1995; Miller et al., 2018; Rose et al., 2016; Wolff et al., 2017), memory manipulation still relies on active states to drive STP-based modifications of synaptic efficacies (Barbosa et al., 2020; Masse et al., 2019, 2020). This idea is akin to the reconsolidation process in long-term emotional memories, whereby the stored information is rendered labile after being retrieved so that new information could be incorporated into and modify old memories (Agren et al., 2012; Lane et al., 2015; Schiller et al., 2010). Finally, flickering color discs have been found to be able to tag item-specific neural reactivations (Huang et al., 2018). Therefore, altering the temporal relationship between luminance sequences of color discs that are linked to each memorized item would presumably perturb the multi-item reactivation profiles to manipulate their memory performances. These points motivate the "dynamic perturbation'" approach developed in our previous study, wherein we demonstrate that temporally synchronized luminance sequences disrupt the recency effect while temporally independent luminance sequences keep the recency intact (Li et al., 2021). Nevertheless, the recency effect is just a behavioral index for sequence WM, and there still lacks an efficient bottom-up, behavioral approach to modulate multi-item WM performance at a general level.

Here we developed a new "Leader-Follower" approach for WM manipulation when participants temporarily hold two or three items simultaneously. We introduced a temporal lag at hundreds of milliseconds based on previous findings (Bahramisharif et al., 2018; Herweg et al., 2020; Huang et al., 2018; Lisman & Idiart, 1995; Mi et al., 2017; Mongillo et al., 2008), to the luminance sequences of flickering color discs during retention. Specifically, one luminance sequence ("Leader", although a randomly generated white noise that does not contain any regularities, always precedes another sequence ("Follower") by certain temporal lag. We hypothesize that the item bound to the "Leader" luminance sequence reactivates earlier than that with the "Follower" sequence and therefore has better memory performance. Four behavioral experiments on 120 participants provided modest evidence supporting that the item associated with the temporally advanced luminance sequence turns out to have better memory performance than that modulated by temporally lagged luminance sequence. Taken together, our results not only offer a new bottom-up, behavioral approach to manipulating WM performance, but also constitute new evidence supporting the critical role of temporally sequenced reactivations in multi-item WM.

Methods

Participants

One hundred and thirty-one participants (50 males, age ranging from 17 to 25 years) took part in five experiments. Two in Experiment 1, two in Experiment 2, three in Experiment 3, and four participants in Experiment were removed due to their extreme memory performance (beyond $2.5 * \sigma$), or not finishing the whole experiment, resulting in 30 participants for each experiment. An a-priori power analysis run in G-Power (Faul et al., 2009) revealed that to obtain an effect of Cohen's $d = 0.55$ for a two-sided paired sample t-test with a power of 0.8, 28 participants needs to be collected. The expected effect size of interest for a difference in normalized target probability between the "leader" and the "follower" condition was derived based on a pretest on 25 subjects, using a similar paradigm as in Experiment 1. All the participants had normal or corrected-to-normal vision with no history of neurological disorders. They were naïve to the purpose of the experiments, and have provided written informed consent prior to the start of the experiment. All experiments were carried out in accordance with the Declaration of Helsinki and have been approved by the Research Ethics Committee at Peking University.

Stimuli and tasks

Participants sat in a dark room, in front of a Display++ monitor with 100 Hz refresh rate and a resolution of 1920 * 1080, and their head stabilized on a chin rest. Participants performed a multi-item working memory task. At the beginning of trial, multiple bars ($0.56^\circ \times 1.67^\circ$ visual angle; two bars in Experiment 1&2, three bars in Experiment 3&4) were simultaneously presented at different locations of the screen, with different colors. Participants were instructed to memorize the orientations of the bars, and their colors (Experiment 1&3) or their spatial locations (Experiment 2&4). During memory maintenance, colors discs flickered for 5 s, and participants should perform a central fixation task by monitoring an abrupt luminance change of the central fixation cross. Finally, participants needed to rotate a horizontal test bar by pressing corresponding keys to one instructed memorized orientation as precise as possible, without time limit. The luminance of flickering disc was randomly generated (ranging from 0 cd/m^2 to 15 cd/m^2) and then was tailored to have equal power at all frequencies by normalizing the amplitudes of its Fourier components before applying an inverse Fourier transform separately for red and blue color. The colors and the spatial locations of the bars and discs were carefully balanced across trials to eliminate possible color-specific or spatial-specific effect. Participants should complete 192 trials in total in Experiment 1&2, which took about 1 hour, and 162 trials in total in Experiment 3&4, which also took about 1 hour.

Experiment 1

In each trial, after a 0.5 fixation period, two bars in red and blue colors were presented at 3˚ visual angle above and below the fixation for 2 s. The orientations of the two bars were chosen randomly, with a difference of at least 10˚. The colors and spatial locations of the two bars were counterbalanced across trials. Participants were instructed to memorize the orientations and colors of the bars. After a blank interval (0.6 ~ 1 s), two discs(3˚ in radius) with the same colors as the two memorized bars were presented at the left or right side of the fixation (7˚ in eccentricity) for 5 s. The colors and spatial locations of the two discs were counterbalanced across trials. Crucially, the luminance of the two color discs was continuously modulated according to two 5 s temporal sequences ranging from dark (0 cd/m^2) to bright (15 cd/m^2). Specifically, in each trial, a 5 s temporal sequence was first randomly generated (“Leader” sequence), and then we shifted the Leader sequence 200 ms rightward and moved the final 200 ms segment of the Leader sequence to the beginning to generate a new sequence (“Follower” sequence). Note that the luminance sequences were generated anew in each trial, and it was quite hard to differentiate between Leader and Follower sequence. Throughout the 5 s maintenance period, participants performed a central fixation task by continuously monitoring an abrupt luminance change of the central fixation cross, while simultaneously holding the two bars. The fixation task is used to eliminate the effect of attentional bias. After finishing the fixation task, a horizontal test bar in red or blue color was presented to instruct participants to recall the red or blue bar’s orientation, and rotate the test bar to the target orientation as precise as possible.

Experiment 2

Experiment 2 had the same stimuli and similar paradigm as Experiment 1. The only difference was that, instead of requiring participants to memorize two bars’ orientations and their colors, we asked participants to memorize two bars’ orientations and spatial locations. Specifically, after finishing the fixation task, a retrospective cue (‘upper’ or ‘lower’ character) was presented for 1 s to instruct participants to recall the orientation at the upper or lower location. Then, a horizontal bar in white color was presented, and participants should rotate it to the instructed memorized orientation. Therefore, in Experiment 2, color information was totally task-irrelevant.

Experiment 3

Experiment 3 was a three-item memory task, and had similar task as Experiment 1. In each trial, three bars in red, blue and green colors were presented at the same eccentricity to the fixation (3˚ visual angle) for 3 s. The orientations of the three bars were chosen randomly, with a difference of at least 10˚ between any two orientations. The colors and spatial locations of the three bars were randomized. Participants were instructed to memorize the orientations and colors of the bars. After a blank interval (0.6 ~ 1 s), three discs (3˚ in radius) with the same colors as the three memorized bars were presented to at 7˚ eccentricity to the fixation for 5 s. Disc and bar with the same color were presented in the same direction of the fixation, but different spatial locations. Similarly, the luminance of the three color discs were continuously modulated according to three 5 s temporal sequences ranging from dark (0 cd/m^2) to bright (15 cd/m^2). Specifically, in each trial, a 5 s temporal sequence was first randomly generated ("Leader" sequence), and then we shifted it 150 ms rightward to generate $Follower_{1st}$ sequence. Similarly, we shifted the $Follower_{1st}$ sequence 150 ms rightward to generate $Follower_{2nd}$ sequence. Therefore, even though the three sequences were presented simultaneously, their temporal relationship showed that Leader lead $Follower_{1st}$ 150 ms, $Follower_{1st}$ lead $Follower_{2nd}$ 150 ms, and Leader lead $Follower_{2nd}$ 300 ms. After finishing the fixation task, a horizontal bar in red, blue or green color was presented to instruct participants to recall the red, blue or green bar's orientation, and rotate to the target orientation as precise as possible.

Experiment 4

Experiment 4 had the same stimuli and similar paradigm as Experiment 3, except that instead of requiring participants to memorize three bars' orientations and their colors, we asked participants to memorize three bars' orientations and their spatial locations. Specifically, after finishing the fixation task, a retrospective cue ('left', 'middle' or 'right' character) was presented for 1 s to instruct participants to recall the orientation at the left, middle or right location (horizontal direction). Then, a horizontal bar in white color was presented, and participants were asked to rotate it to the instructed memory orientation. Therefore, as Experiment 2, color information was also totally task-irrelevant in Experiment 4.

Data analysis

To quantify the memory performance for each item, a probabilistic mixture model (Bays et al., 2009) was applied to fit behavioral performance. Specifically, the mixture model simultaneously characterizes the contribution of the memory for target item, non-target item and random guess to the final report. Specifically, this model calculates probability of correctly reporting the feature value of the target item, with some variability, the probability of mistakenly reporting the feature value of one of the other, non-target items held in memory with the same variability, and the probability of generating a random response unrelated to either target or non-target items. In the present study, we focused on target probability, because it represents the memory accuracy for the target and has been widely used to quantify memory performance (Gorgoraptis et al., 2011; Li et al., 2021; Van Ede et al., 2018). Moreover, considering that the target probability is not normally distributed, we performed an empirical logit transformation: $\mathrm{logit}(p) = \ln((p + 1/2n)/(1 - p + 1/2n))$, where p is target probability and n is the number of observations transformation (de Smith, 2018). The normalized target probabilities were used for further statistical tests in all the experiments. In addition, memory precision was estimated by calculating the reciprocal of the circular standard deviation of response error (the circular difference between the reported orientation and the true target orientation).

Data and associated code are available in OSF (https://osf.io/cpvdk/).

Statistics

Classical frequentist statistics, e.g., repeated ANOVA and paired t-test, were applied to test experimental effect. Considering there are three conditions in Experiment 3&4, Holm correction was applied for post-hoc analysis.

Apart from classical frequentist statistics, we also implemented Bayesian statistics using JASP (0.16.4.0). Specifically, for paired t-test, we provided Bayes Factor, BF_{10}, which quantifies how many times the observed data are more likely under the alternative hypothesis that postulates the presence of the experimental effect (e.g., the perturbation effect) than under null hypothesis, while for repeated ANOVA, we reported the inclusion Bayes Factor, BF_{incl}, which reflects the evidence for all models with a particular experimental effect, compared to all models without that particular effect. A Bayes factor greater than 1 can be interpreted as evidence against the null, at which one convention is that a Bayes factor greater than 3 can be considered as "substantial" evidence against the null, and vice versa (a Bayes factor smaller than 1/3 indicates substantial evidence in favor of the null-model) (Wetzels et al., 2011). Bayesian post hoc tests were applied in Experiment 3&4. We reported the uncorrected Bayes factor, i.e., $BF_{10,U}$, and posterior odds, which have been corrected for multiple testing by fixing to 0.5 the prior probability that the null hypothesis holds across all comparisons (Westfall et al., 1997).

Results

"Leader-follower" dynamic perturbation modulates two-item memory performance (Experiment 1)

Thirty participants performed a two-item memory task in Experiment 1 (Fig. 1A). In each trial, two bars were simultaneously presented at the upper and lower locations, and participants needed to memorize both orientations and colors of the two bars over a 5 s delay period while performing a central fixation task. During the recalling phase, participants adjusted the orientation of a probe bar to match that of the memorized bar having the same color as the probe. Crucially, during the 5 s delay period, two task-irrelevant discs with the same colors as one of the memorized bars – one red and one blue – were bilaterally presented, and their luminance was continuously changing according to two 5 s temporal sequences (Fig. 1B). The two luminance sequences were designed to have a specific temporal relationship, with their cross-correlation coefficient peaking at 200 ms lag (Fig. 1C). Specifically, one sequence randomly generated per trial ("Leader" sequence) would be used to generate the other by introducing a 200 ms lag ("Follower" sequence). In other words, to generate two random sequences with a fixed time lag, we temporally shifted one sequence ("Leader") rightward by 200 msec to generate the "Follower" sequence. Moreover, to ensure their simultaneous occurrence, we cut the last 200 ms segment of the "Follower" sequence and shift it to its beginning so that the "Leader" and "Follower" sequences still have a fixed circular temporal lag. Finally, the color, spatial location, and "Leader-Follower" conditions were counterbalanced across trials.

All trials were then categorized based on whether the luminance sequence of the corresponding disc during the delay period (i.e., one with the same color as the probe) was a "Leader" or "Follower" sequence, regardless of its color or location. For instance, when recalling the orientation of a red bar held in memory, this trial would be labeled according to whether the luminance sequence of the red disc was a "Leader" or "Follower" sequence. Similarly, when retrieving the orientation of the blue bar, the trial condition would be determined by the blue disc, i.e., Leader or Follower.

We first estimated memory precision for each item by calculating the reciprocal of circular standard deviation of response error (the circular difference between the reported orientation and the true orientation across trials) ($1 / \sigma$) (Bays et al., 2009). As shown in Fig. 1D, the "Leader" condition showed better memory performance than the "Follower" condition (Leader: mean = 1.636, s.e. = 0.100; Follower: mean = 1.483, s.e = 0.111; paired t-test, $t_{(29)} = 2.565$, $p = 0.016$, Cohen's $d = 0.468$). We then implemented the Bayesian hypothesis test and confirmed the significant memory modulation effect ($BF_{10} = 3.074$). To further assess the contribution of the memory for target item to the final report, we employed a probabilistic mixture model (Bays et al., 2009) and focused on the calculated Target probability, i.e., the proportion of responses attributed to the report of the correct target, to quantify memory performance. Moreover, to ensure normal distribution, we performed an empirical logit transformation (de Smith, 2018) on the target response probability. As shown in Fig. 1E, the "Leader" condition also showed better memory performance than the "Follower" condition (Leader: mean = 3.638, s.e. = 0.223; Follower: mean = 3.011, s.e = 0.220; paired t-test, $t_{(29)} = 2.798$, $p = 0.009$, Cohen's $d = 0.511$; Bayes factor, $BF_{10} = 4.901$) (see target probability without normalization in Extended Data Fig.1-1 and additional parameters (non-target and randomly guess probability) results in Extended Data Fig.1-2).

Taken together, consistent with our hypothesis, the "Leader-follower" dynamic perturbation during WM retention effectively modulates memory performance when participants held two items in memory, wherein the item experiencing temporal advances during retention shows better memory performance compared to the item with relative 200 ms temporal delays.

Figure 1 about here

Memory-irrelevant dynamic perturbation (Experiment 2)

In Experiment 1, the color feature was memory-relevant since participants retained both orientation and color of the two items. In Experiment 2, we examined whether the dynamic perturbation would still be effective when color is memory-irrelevant. Thirty new participants participated in Experiment 2 (Fig. 2A), wherein two bars were simultaneously presented at the upper and lower locations. Instead of memorizing colors as in Experiment 1, participants held the locations and orientations of the two bars over a 5 s delay period in memory while performing a central fixation task. During the memory test, participants were first presented with a location cue (upper or lower) based on which they adjusted a probe bar to match the memorized orientation, regardless of its color. In other words, the color feature was completely memory-irrelevant in Experiment 2. Similar to Experiment 1, the "Leader-Follower" dynamic perturbation was applied to the two colored discs during retention (Fig. 2BC).

Unfortunately, as shown in Fig. 2D, there is no significant difference between "Leader" and "Follower" condition on memory precision (Leader: mean = 1.887, s.e. = 0.628; Follower: mean = 1.920, s.e = 0.645; paired t-test, $t_{(29)} = -0.460$, $p = 0.649$, Cohen's d = -0.084; Bayes factor, $BF_{10} = 0.214$). Nevertheless, the normalized target probability showed a modulation trend (Leader: mean = 3.454, s.e. = 0.186; Follower: mean = 3.163, s.e = 0.186; paired t-test, $t_{(29)} = 1.862$, $p = 0.073$, Cohen's d = 0.340; Bayes factor, $BF_{10} = 1.012$) (Fig. 2E) (see significant memory modulation effect on target probability in Extended Data Fig 1-1). To examine the manipulation consistency between Experiment 1 and 2 in terms of the normalized target probability, we conducted a mixed-design ANOVA analysis (Experiment * Perturbation). The results reveal a significant main perturbation effect across experiments, while the main effect of Experiment and their interaction effect were non-significant (Perturbation effect: $F_{(1,58)} = 11.288$, $p = 0.001$, $\eta_p^2 = 0.163$; Experiment effect: $F_{(1,58)} = 0.004$, $p = 0.949$, $\eta_p^2 < 0.001$; Experiment * perturbation: $F_{(1,58)} = 1.520$, $p = 0.223$, $\eta_p^2 = 0.026$); this indicates a convergence of evidence from similar experimental designs. Inclusion Bayes Factor based on all models further advocates significant perturbation effect ($BF_{incl} = 15.428$), and non-significant Experiment effect ($BF_{incl} = 0.311$) and their interaction ($BF_{incl} = 0.451$).

Overall, the "Leader-Follower" dynamic perturbation still seems to modulate memory in terms of target probability when the color feature that the dynamic perturbation operates on is memory-irrelevant, but with a less stronger modulation effect than the memory-relevant perturbation (Experiment 1).

Figure 2 about here

"Leader-follower" dynamic perturbation modulates three-item memory performance (Experiment 3)

After demonstrating the limited effectiveness of the "Leader-follower" dynamic perturbation approach in the two-item memory task, we next tested the effectiveness of the approach on a three-item memory display. Thirty new participants participated in Experiment 3 (Fig. 3A), wherein they memorized both orientations and colors (red, blue, green) of three bars over a 5 s delay period. Similar to Experiment 1, during the memory test phase, participants adjusted the orientation of a probe bar to match that of the memorized bar sharing the same color. Critically, the "Leader-follower" dynamic perturbation was now applied to three task-irrelevant discs with the same colors as one of the memorized bars (red, blue, green) during the 5 s delay period, with their luminance continuously modulated by three temporally related sequences (Fig. 3B). Specifically, one sequence randomly generated in each trial ("Leader" sequence) was used to generate the other two sequences by introducing a 150 ms or 300 ms lag, corresponding to the "Follower_{1st}" and "Follower_{2nd}" sequences, respectively (Fig. 3BC). Using 150 ms and 300 ms instead of 200 ms derives from previous neural findings revealing that three-item sequence memory entails a more temporally compressed reactivation than two-item sequence memory (Huang et al., 2018). Finally, the color, spatial location, and "Leader-Follower" conditions were counterbalanced across trials.

Trials were categorized as "Leader", "Follower_{1st}", or "Follower_{2nd}" conditions, based on the corresponding luminance sequence (i.e., having the same color as the probe). As shown in Fig. 3D, the dynamic perturbation showed weak modulation on memory precision (Leader: mean = 1.153, s.e. = 0.054; Follower_{1st}: mean = 1.117, s.e = 0.069; Follower_{2nd}: mean = 1.024, s.e = 0.058; one-way repeated ANOVA; main effect of perturbation: $F_{(2,58)}$ = 2.506, p = 0.090, η_p^2 = 0.080; Bayes factor: BF_{incl} = 0.686; Post-hoc analysis; Leader vs. Follower_{1st}: $t_{(29)}$ = 0.595, p_{cor} = 0.554, Cohen's d = 0.107 (Bayesian post-hoc tests: $BF_{10,U}$ = 0.236, posterior odds = 0.138); Leader vs. Follower_{2nd}: $t_{(29)}$ = 2.167, p_{cor} = 0.103, Cohen's d = 0.390 (Bayesian post-hoc tests: $BF_{10,U}$ = 1.178, posterior odds = 0.692); Follower_{1st} vs. Follower_{2nd}: $t_{(29)}$ = 1.571, p_{cor} = 0.243, Cohen's d = 0.283 (Bayesian post-hoc tests: $BF_{10,U}$ = 0.573, posterior odds = 0.336)). Meanwhile, the normalized target probability showed a modulation trend (Leader: mean = 3.077, s.e. = 0.203; Follower_{1st}: mean = 2.599, s.e = 0.191; Follower_{2nd}: mean = 2.495, s.e = 0.184; one-way repeated ANOVA; main effect of perturbation: $F_{(2,58)}$ = 2.980, p = 0.059, η_p^2 = 0.093; Bayes factor: BF_{incl} =1.249), revealing a gradual decrease (Post-hoc analysis; Leader vs. Follower_{1st}: $t_{(29)}$ = 1.881, p_{cor} = 0.130, Cohen's d = 0.453 (Bayesian post-hoc tests: $BF_{10,U}$ = 0.781, posterior odds = 0.459); Leader vs. Follower_{2nd}: $t_{(29)}$ = 2.288, p_{cor} = 0.077, Cohen's d = 0.551 (Bayesian post-hoc tests: $BF_{10,U}$ = 1.424, posterior odds = 0.836); Follower_{1st} vs. Follower_{2nd}: $t_{(29)}$ = 0.407, p_{cor} = 0.685, Cohen's d = 0.098 (Bayesian post-hoc tests: $BF_{10,U}$ = 0.216, posterior odds = 0.127)).

Together, on a descriptive level, the Leader-Follower dynamic perturbation approach is also effective in a three-item paradigm; that is, the item associated with earlier temporal reactivations shows better memory performance compared to those endowed with relatively delayed reaction during the delay period. However, on a statistical level, the results provide a trend in the suggested direction at best.

Figure 3 about here

Memory-irrelevant dynamic perturbation in three-item memory task (Experiment 4)

Finally, we tested the memory-irrelevant dynamic perturbation approach in a three-item memory task (Experiment 4). Thirty new participants participated in the experiment (Fig. 4A), wherein they held the locations and orientations of the three bars over a 5 s delay period in memory. During the memory test phase, participants were first presented with a location cue (left, middle, or right) based on which they adjusted a probe bar to match the memorized orientation, regardless of its color. Thus, similar to Experiment 2, the color feature was completely memory-irrelevant here. Moreover, the same "Leader-Follower" dynamic perturbation as used in Experiment 3 was applied to the three colored discs during retention (Fig. 4BC).

As shown in Fig. 4D, the memory precision for "Leader", "Follower_{1st}", and "Follower_{2nd}" conditions exhibited gradual decrease (Leader: mean = 1.510, s.e. = 0.091; Follower_{1st}: mean = 1.329, s.e = 0.094; Follower_{2nd}: mean = 1.170, s.e = 0.084; one-way repeated ANOVA; main effect of perturbation: $F_{(2,58)}$ =7.303 , p = 0.001, η_p^2 = 0.201; Bayes factor: BF_{incl} = 22.737; Post-hoc analysis, Leader vs. Follower_{1st}: $t_{(29)}$ = 2.033, p_{cor} = 0.093, Cohen's d = 0.368 (Bayesian post-hoc tests: $BF_{10,U}$ = 0.883, posterior odds = 0.519); Leader vs. Follower_{2nd}: $t_{(29)}$ = 3.819, p_{cor} <0.001 , Cohen's d = 0.692 (Bayesian post-hoc tests: $BF_{10,U}$ = 40.869, posterior odds = 24.007); Follower_{1st} vs. Follower_{2nd}: $t_{(29)}$ = 1.786, p_{cor} = 0.093, Cohen's d = 0.323 (Bayesian post-hoc tests: $BF_{10,U}$ = 1.197, posterior odds = 0.703)). The normalized target probability also showed significant modulation effect (Leader: mean =3.656 , s.e. = 0.217; Follower_{1st}: mean = 3.064, s.e = 0.236; Follower_{2nd}: mean = 2.630, s.e = 0.208; one-way repeated ANOVA; main effect of perturbation: $F_{(2,58)}$ = 6.435, p = 0.003, η_p^2 = 0.182; Bayes factor: BF_{incl} = 21.583; Post-hoc analysis; Leader vs. Follower_{1st}: $t_{(29)}$ = 2.062, p_{cor} = 0.087, Cohen's d = 0.490 (Bayesian post-hoc tests: $BF_{10,U}$ = 1.101, posterior odds = 0.647); Leader vs. Follower_{2nd}: $t_{(29)}$ = 3.573, p_{cor} = 0.002, Cohen's d = 0.849 (Bayesian post-hoc tests: $BF_{10,U}$ = 22.450, posterior odds = 13.187); Follower_{1st} vs. Follower_{2nd}: paired t-test, $t_{(29)}$ = 1.511, p_{cor} = 0.136, Cohen's d = 0.359 (Bayes factor: $BF_{10,U}$ = 0.614, posterior odds = 0.360)).

To examine the manipulation consistency between Experiment 3 and 4, both of which employed a three-item WM task, we conducted a mixed-design ANOVA analysis (Experiment * Perturbation) again. The results reveal a significant main perturbation effect across experiments (Perturbation effect: $F_{(2,116)} = 9.111$, $p < 0.001$, $\eta_p^2 = 0.136$ (Post-hoc analysis, Leader vs. Follower$_{1st}$: $t_{(58)} = 2.791$, $p_{cor} = 0.012$, Cohen's $d = 0.472$; Leader vs. Follower$_{2nd}$: $t_{(29)} = 4.193$, $p_{cor} < 0.001$, Cohen's $d = 0.708$; Follower$_{1st}$ vs. Follower$_{2nd}$: $t_{(29)} = 1.401$, $p_{cor} = 0.164$, Cohen's $d = 0.237$) ; Experiment effect: $F_{(1,58)} = 4.202$, $p = 0.045$, $\eta_p^2 = 0.068$; Experiment * perturbation: $F_{(2,116)} = 0.726$, $p = 0.486$, $\eta_p^2 = 0.006$), supporting the modulation effect across experiments. Inclusion Bayes Factor based on all models further advocates significant perturbation effect ($BF_{incl} = 134.346$; Post-hoc tests, Leader vs. Follower$_{1st}$: $BF_{10,U} = 3.748$, posterior odds = 2.202; Leader vs. Follower$_{2nd}$: $BF_{10,U} = 133.865$, posterior odds = 78.633; Follower$_{1st}$ vs. Follower$_{2nd}$: $BF_{10,U} = 0.435$, posterior odds = 0.256), while the main effect of Experiment ($BF_{incl} = 0.823$) and their interaction effect ($BF_{incl} = 0.376$) were non-significant. Overall, the "Leader-Follower" dynamic perturbation efficiently modulates three-item memory when the color feature that the dynamic perturbation operates on is memory-irrelevant.

To provide a possible explanation for the non-robust memory modulation effect in Experiment 2&3, we compared the memory precision between experiments, which we thought should largely reflect the task difficulty. Experiment 2 (2-item location memory) showed significant higher memory precision compared to Experiment 1(2-item color memory) (Experiment effect: $F_{(1,58)} = 5.264$, $p = 0.025$, $\eta_p^2 = 0.083$; $BF_{incl} = 2.191$), while Experiment 4 (3-item location memory) also showed significant higher memory precision than Experiment 3 (3-item color memory) (Experiment effect: $F_{(1,58)} = 7.242$, $p = 0.009$, $\eta_p^2 = 0.111$; $BF_{incl} = 5.030$). These results indicated that this purely bottom-up perturbations may only have significant effectiveness when the task is in moderate difficulty instead of too easy (2-item location memory) or too difficult (3-item color memory) to accomplish.

Figure 4 about here

Discussion

In the present study, we sought to capitalize on the "Leader-Follower" dynamic perturbation as a new behavioral manipulation mechanism to interfere with the multi-item neural dynamics and alter WM performance in humans. Four experiments on 120 participants demonstrate the effectiveness of the approach. Specifically, temporally advanced manipulation ('leader') during retention leads to better recalling performance than temporally delayed perturbation ('follower'), regardless of its relevance to the memory task. These findings, together with previous works (Barbosa et al., 2020; Li et al., 2021; Miller et al., 2018), support the substantial role of STP-based neural dynamics in mediating WM operation. Our work also offers a new bottom-up, behavioral approach to manipulating human WM. However, it is notable that memory modulation effect is not very robust across experiments and measures, which indicates that this purely bottom-up perturbation approach has limited effectiveness and needs further exploration.

There are many noninvasive approaches to altering WM performance in humans. For instance, applying TMS to relevant brain regions could modulate memory behavior (Lee & D'Esposito, 2012) and even reactivate information retained in WM (Rose et al., 2016). Oscillatory interference methods, such as rhythmic physical stimulus (Clouter et al., 2017), Repetitive TMS (rTMS) (Beynel et al., 2019; Sauseng et al., 2009), tACS with rhythmic (Hoy et al., 2015) or theta-gamma coupling (Alekseichuk et al., 2016) have also been found to efficiently impact memory performance. Here we developed a purely bottom-up, behavioral approach by presenting task-irrelevant flickering color probes during WM retention. Notably, since participants could not discriminate the temporal relationship of the luminance sequences at the perceptual level, i.e., which sequence leads and which sequence lags, the manipulation is indeed operated in an unconscious way. Moreover, the luminance sequences are randomly generated per trial, and therefore it is only their temporal relationship instead of a specific sequence that influences WM performance. Furthermore, the "Leader-Follower" dynamic perturbation aims to alter multi-item WM performance, which is different from our previous work focusing on sequence working memory (Li et al., 2021), thus offering a memory manipulation approach at a general level. Finally, distinct from the retrocue behavioral paradigm, whereby the cued item would enter the focus of attention (FoA) and get prioritized in WM (Oberauer & Hein, 2012; Öztekin et al., 2010), our method is a purely bottom-up manipulation and does not rely on top-down attentional modulations.

Crucially, our "Leader-Follower" dynamic perturbation approach draws upon accumulating findings and models advocating the central function of temporal dynamics in WM. First, multiple WM items are postulated to undergo item-by-item sequential reactivations with items of greater strength firing earlier (Lisman & Idiart, 1995; Lisman & Jensen, 2013; Oberauer & Lewandowsky, 2008, 2011), a framework that has received empirical evidence support (Axmacher et al., 2010; Burke et al., 2016; Friese et al., 2013; Heusser et al., 2016). Recently, we also demonstrate that a sequence of items is serially reactivated during the delay period, and the late item in the sequence is accompanied by better memory performance (i.e., recency effect) and earlier reactivation (Huang et al., 2018, 2021), also in line with the latency-based view. Interestingly, this latency- or time-based coding of input strength extends beyond memory findings and also occurs in perception and attention (Fiebelkorn et al., 2013, 2018; Huang & Luo, 2020; Jensen et al., 2014; Jia et al., 2017; Landau & Fries, 2012; Mo et al., 2019; Song et al., 2014). Here, we speculate that altering the early-late time relationship of neural responses indeed modifies the subsequent WM performance. Second, the time lag between luminance sequences is set also according to previous experimental findings and STP neural model, i.e., temporally compressed reactivation within 200 ms and 150 ms for two- and three-item sequences, respectively (Herweg et al., 2020; Huang et al., 2018; Li et al., 2021; Mi et al., 2017; Mongillo et al., 2008). Overall, the "dynamic perturbation" approach is motivated by previous findings, allowing us to exploit the brain's time perspective to manipulate multi-item neural dynamics and in turn alter WM performance.

We developed a "Leader-Follower" dynamic perturbation aiming to introduce a specific temporal lag in the reactivation profiles of memorized items to manipulate their memory strengths. We hypothesize that items with relatively earlier reactivation during retention would have better memory performance than that with relatively later reactivation. The manipulation is implemented by generating temporally shifted luminance sequences (i.e., Leader sequence, Follower sequence) for color discs that are bound to each memorized item during retention. Although the temporal manipulation is possibly at an unconscious level, i.e., participants could not tell which sequence advances over time, our brain is known to be indeed endowed with tremendous capabilities to calculate the temporal lag between events, from tens of milliseconds to hundreds of milliseconds. Moreover, the continuous attractor neural network model established in our previous work, by incorporating plausible biological principles, also supports that temporal lag is encoded in the system and influences memory representations (Li et al., 2021)

Retaining information in WM has traditionally been hypothesized to rely on persistent firing but computational models and recent findings propose a hidden-state WM view, i.e., items could be silently retained in STP-based synaptic weights (Huang et al., 2021; Miller et al., 2018; Mongillo et al., 2008; Rose et al., 2016; Trübutschek et al., 2019; Wolff et al., 2017), even lasting for tens of seconds long with periodical refresh (Fiebig & Lansner, 2017). Then how could we access information in this activity-silent network? Recent studies demonstrate that presenting a nonspecific impulse (i.e., PING) during retention could transiently perturb the WM network and reactivate memories (Fan et al., 2020; Huang et al., 2021; Wolff et al., 2017). This methodological advance has allowed researchers to directly access WM information and predict subsequent behavior. Here we use task-irrelevant luminance sequences to first reactivate memory information, and then apply continuous perturbation to impose temporal relationships between items to interfere with their neural dynamics and manipulate WM. This approach resembles the reconsolidation process in long-term memory, such that the stored fear memory would be rendered labile when retrieved, and new information could be inserted and modify old memories within this period (Agren et al., 2012; Lane et al., 2015; Schiller et al., 2010). Meanwhile, different from long-term memory relying on long time scales, our approach is operated at a shorter temporal scale, i.e., 100-200 ms, a critical time scale in STP-based WM operation.

Taken together, based on accumulating neural findings and theoretical models, we develop a new "Leader-Follower" dynamic perturbation behavioral approach to alter multi-item WM in humans, by presenting temporally related luminance sequences during the delay period. We demonstrate that the item associated with the 'leader' luminance sequence shows better memory performance than the item bound to the 'follower' luminance sequence. Our results suggest the essential role of neural temporal dynamics in WM operation and offer a promising, non-invasive WM manipulation approach.

Author contribution
Conceptualization: Qiaoli Huang, Yuanyuan Mi, Huan Luo.
Data curation: Qiaoli Huang, Minghao Luo.
Formal analysis: Qiaoli Huang.
Funding acquisition: Huan Luo.
Investigation: Qiaoli Huang.

Methodology: Qiaoli Huang.
Resources: Huan Luo.
Software: Qiaoli Huang.
Supervision: Huan Luo.
Validation: Qiaoli Huang.
Visualization: Qiaoli Huang, Huan Luo.
Writing – original draft: Qiaoli Huang, Huan Luo.
Writing – review & editing: Qiaoli Huang, Huan Luo.

Acknowledgments

This work was supported by the National Science and Technology Innovation 2030 Major Program 2021ZD0204103, and National Natural Science Foundation of China Grants 31930052 to H.L.

Reference

Agren, T., Engman, J., Frick, A., Björkstrand, J., Larsson, E. M., Furmark, T., & Fredrikson, M. (2012). Disruption of reconsolidation erases a fear memory trace in the human amygdala. *Science*, *337*(6101), 1550–1552. https://doi.org/10.1126/science.1223006

Alekseichuk, I., Turi, Z., Amador de Lara, G., Antal, A., & Paulus, W. (2016). Spatial Working Memory in Humans Depends on Theta and High Gamma Synchronization in the Prefrontal Cortex. *Current Biology*, *26*(12), 1513–1521. https://doi.org/10.1016/j.cub.2016.04.035

Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., & Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences*, *107*(7), 3228–3233.

Bahramisharif, A., Jensen, O., Jacobs, J., & Lisman, J. (2018). Serial representation of items during working memory maintenance at letter-selective cortical sites. *PLoS Biology*, *16*(8), 1–21. https://doi.org/10.1371/journal.pbio.2003805

Barbosa, J., Stein, H., Martinez, R. L., Galan-Gadea, A., Li, S., Dalmau, J., Adam, K., Valls-Solé, J., Constantinidis, C., & Compte, A. (2020). Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nature Neuroscience*, *23*(8), 1016–1024.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7.

Beynel, L., Davis, S., Crowell, C., Hilbig, S., Palmer, H., Brito, A., Hile, C., Lim, W., Nguyen, D., & Dannhauer, M. (2019). Site-specifics effects of online repetitive transcranial magnetic stimulation (rTMS) on working memory (WM). *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*, *12*(2), 564.

Burke, John, Jacobs, Joshua, Lega, Bradley, Kahana, Michael, & J. (2016). Slow-Theta-to-Gamma Phase-Amplitude Coupling in Human Hippocampus Supports the Formation of New Episodic Memories. *Cerebral Cortex*, *26*(1), 268–278.

Clouter, A., Shapiro, K. L., & Hanslmayr, S. (2017). Theta Phase Synchronization Is the Glue that Binds Human Associative Memory. *Current Biology*, *27*(20), 3143-3148.e6. https://doi.org/10.1016/j.cub.2017.09.001

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, *7*(9), 415–423.

de Smith, M. J. (2018). *Statistical analysis handbook*. The Winchelsea Press.

Fan, Y., Han, Q., Guo, S., & Luo, H. (2020). Dissociated Neural Representations of Content and Ordinal Structure in Auditory Sequence Memory. *SSRN Electronic Journal*, *41*(29), 6290–6303. https://doi.org/10.2139/ssrn.3651509

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160.

Fiebelkorn, I. C., Pinsk, M. A., & Kastner, S. (2018). A dynamic interplay within the frontoparietal network underlies rhythmic spatial attention. *Neuron, 99*(4), 842–853.

Fiebelkorn, I. C., Saalmann, Y. B., & Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. *Current Biology*, *23*(24), 2553–2558.

Fiebig, X. F., & Lansner, X. A. (2017). A spiking working memory model based on Hebbian short-term potentiation. *The Journal of Neuroscience*, *37*(1), 83–96. https://doi.org/10.1523/JNEUROSCI.1989-16.2017

Friese, U., Köster, M., Hassler, U., Martens, U., Trujillo-Barreto, N., & Gruber, T. (2013). Successful memory encoding is associated with increased cross-frequency coupling between frontal theta and posterior gamma oscillations in human scalp-recorded EEG. *Neuroimage*, *66*, 642–647.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, *14*(3), 477–485.

Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *Journal of Neuroscience*, *31*(23), 8502–8511. https://doi.org/10.1523/JNEUROSCI.0208-11.2011

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194.

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta oscillations in human memory. *Trends in Cognitive Sciences*, *24*(3), 208–227.

Heusser, A. C., Poeppel, D., Ezzyat, Y., & Davachi, L. (2016). Episodic sequence memory is supported by a theta-gamma phase code. *Nature Neuroscience*. https://doi.org/10.1038/nn.4374

Hoy, K. E., Bailey, N., Arnold, S., Windsor, K., John, J., Daskalakis, Z. J., & Fitzgerald, P. B. (2015). The effect of γ-tACS on working memory performance in healthy controls. *Brain and Cognition*, *101*, 51–56.

Huang, Q., Jia, J., Han, Q., & Luo, H. (2018). Fast-backward replay of sequentially memorized items in humans. *eLife*, *7*, 1–21. https://doi.org/10.7554/eLife.35164

Huang, Q., & Luo, H. (2020). Saliency-based Rhythmic Coordination of Perceptual Predictions. *Journal of Cognitive Neuroscience*, *32*(2), 201–211.

Huang, Q., Zhang, H., & Luo, H. (2021). Sequence structure organizes items in varied latent states of working memory neural network. *eLife*, *10*, 1–21. https://doi.org/10.7554/elife.67589

Jensen, O., Gips, B., Bergmann, T. O., & Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends in Neurosciences, 37*(7), 357–369.

Jia, J., Liu, L., Fang, F., & Luo, H. (2017). Sequential sampling of visual objects during sustained attention. *PLoS Biology*, *15*(6), 1–19. https://doi.org/10.1371/journal.pbio.2001903

Johnson, J. S., Hollingworth, A., & Luck, S. J. (2008). The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 41–55. https://doi.org/10.1037/0096-1523.34.1.41

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, *22*(11), 1000–1004.

Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, *43*(2), 149–164.

Lane, R. D., Ryan, L., Nadel, L., & Greenberg, L. (2015). Memory reconsolidation, emotional arousal, and the process of change in psychotherapy: New insights from brain science. *Behavioral and Brain Sciences*, *38*, 1–64. https://doi.org/10.1017/S0140525X14000041

Lee, T. G., & D'Esposito, M. (2012). The dynamic nature of top-down signals originating from prefrontal cortex: A combined fMRI-TMS study. *Journal of Neuroscience*, *32*(44), 15458–15466. https://doi.org/10.1523/JNEUROSCI.0627-12.2012

Li, J., Huang, Q., Han, Q., Mi, Y., & Luo, H. (2021). Temporally coherent perturbation of neural dynamics during retention alters human multi-item working memory. *Progress in Neurobiology*, 102023.

Lisman, J. E., & Idiart, M. A. (1995). Storage of 7+/-2 short-term memories in oscillatory subcycles. *Science*, *267*(5203), 1512–1515.

Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. *Neuron*, *77*(6), 1002–1016.

Luck, S., Vogel, J., & Edward, K. (1997). The capacity of visual working memory for features and conjuctions. *Nature*, *390*(6657), 279–281.

Masse, N. Y., Rosen, M. C., & Freedman, D. J. (2020). Reevaluating the Role of Persistent Neural Activity in Short-Term Memory. *Trends in Cognitive Sciences*, *24*(3), 242–258. https://doi.org/10.1016/j.tics.2019.12.014

Masse, N. Y., Yang, G. R., Song, H. F., Wang, X.-J., & Freedman, D. J. (2019). Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nature Neuroscience*, 1.

Mi, Y., Katkov, M., & Tsodyks, M. (2017). Synaptic correlates of working memory capacity. *Neuron*, *93*(2), 323–330. https://doi.org/10.1016/j.neuron.2016.12.004

Miller, E. K., Lundqvist, M., & Bastos, A. M. (2018). Working Memory 2.0. In *Neuron*. https://doi.org/10.1016/j.neuron.2018.09.023

Mo, C., Lu, J., Wu, B., Jia, J., Luo, H., & Fang, F. (2019). Competing rhythmic neural representations of orientations during concurrent attention to multiple orientation features. *Nature Communications*, *10*(1), 1–10.

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*(5869), 1543–1546.

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends in Cognitive Sciences*, *21*(6), 449–461. https://doi.org/10.1016/j.tics.2017.03.010

Oberauer, K., & Hein, L. (2012). Attention to information in working memory. *Current Directions in Psychological Science*, *21*(3), 164–169.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review*, *115*(3), 544.

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: A computational implementation of the Time-Based Resource-Sharing theory. *Psychonomic Bulletin & Review*, *18*(1), 10–45.

Öztekin, I., Davachi, L., & McElree, B. (2010). Are Representations in Working Memory Distinct From Representations in Long-Term Memory? Neural Evidence in Support of a Single Store. *Psychological Science*, *21*(8), 1123–1133. https://doi.org/10.1177/0956797610376651

Rose, N. S., Larocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136–1140.

Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., Glennon, M., Gerloff, C., Birbaumer, N., & Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Current Biology*, *19*(21), 1846–1852. https://doi.org/10.1016/j.cub.2009.08.062

Schiller, D., Monfils, M. H., Raio, C. M., Johnson, D. C., Ledoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*, *463*(7277), 49–53. https://doi.org/10.1038/nature08637

Siegel, M., Warden, M. R., & Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences*, *106*(50), 21341–21346. https://doi.org/10.1073/pnas.0908193106

Song, K., Meng, M., Chen, L., Zhou, K., & Luo, H. (2014). Behavioral oscillations in attention: Rhythmic α pulses mediated through θ band. *Journal of Neuroscience*, *34*(14), 4837–4844.

Trübutschek, D., Marti, S., Ueberschär, H., & Dehaene, S. (2019). Probing the limits of activity-silent non-conscious working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(28), 14358–14367. https://doi.org/10.1073/pnas.1820730116

Van Ede, F., Chekroud, S. R., Stokes, M. G., & Nobre, A. C. (2018). Decoding the influence of anticipatory states on visual perception in the presence of temporal distractors. *Nature Communications*. https://doi.org/10.1038/s41467-018-03960-z

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751.

Wang, Y., Markram, H., Goodman, P. H., Berger, T. K., Ma, J., & Goldman-Rakic, P. S. (2006). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nature Neuroscience*, *9*(4), 534–542.

Westfall, P. H., Johnson, W. O., & Utts, J. M. (1997). A Bayesian perspective on the Bonferroni adjustment. *Biometrika*, *84*(2), 419–427.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*, *6*(3), 291–298.

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, *20*(6), 864–871. https://doi.org/10.1038/nn.4546

Figure Legends

Figure 1

Figure 1. “Leader-follower” dynamic perturbation during retention modulates two-item memory performances (Experiment 1, N = 30)
(A) “Leader-follower” dynamic perturbation paradigm. In each trial, participants were presented with two bars and memorized their orientations and colores. During the memory test, participants adjusted the orientation of a probe bar to match that of the memorized bar having the same color as the probe. During the 5 s delay period, participants performed a central fixation task, while two task-irrelevant, flickering discs having the same color as each of the memorized bars (blue and red) were presented bilaterally, with their luminances continuously modulated by two 5 s temporal sequences (Leader or Follower sequences), respectively. The color, spatial location, and “Leader-Follower” conditions were counterbalanced across trials. (B) The Leader temporal sequence was a 5 s white noise randomly generated per trial, and the Follower sequence was created by circular-shifting the Leader sequence 200 ms rightward. Note that the two sequences were presented simultaneously rather than asynchronously. (C) The Leader-Follower cross-correlation over time as a function of temporal lag, peaking at 200 ms. (D) Memory performance. Grand averaged (mean + SEM) memory precision during recalling test for Leader (purple) and Follower (turquoise) conditions, with dots denoting individual participants. (E) Same as D, but for normalized target probability. *: $p < 0.05$. For target probability without normalization see Extended Data Fig.1-1. For additional parameters (non-target and randomly guess probability) results in Extended Data Fig.1-2.

Figure 2

Figure 2. Task-irrelevant "Leader-follower" dynamic perturbation (Experiment 2, N = 30).

(A) Task-irrelevant dynamic perturbation paradigm. Experiment 2 was the same as Experiment 1, except that participants needed to memorize the orientations and locations (upper or lower) of the two bar stimuli regardless of their color features. During the memory test period, a location cue (upper or lower) was first presented, based on which participants rotated the horizontal white bar to the corresponding memorized orientation. Critically, a "Leader-follower" dynamic perturbation as in Experiment 1 was applied during the delay period, i.e., two discs of the same color as each of the memorized bars (blue and red) were presented bilaterally, with their luminances continuously modulated by a Leader or Follower sequences, respectively. (B) The Leader temporal sequence was a 5 s white noise randomly generated per trial, and the Follower sequence was created by circular-shifting the Leader sequence 200 ms rightward. The two luminance sequences were presented simultaneously rather than asynchronously. (C) The Leader-Follower cross-correlation over time as a function of temporal lag, peaking at 200 ms. (D) Memory performance. Grand averaged (mean + SEM) memory precision during recalling test for Leader (purple) and Follower (turquoise) conditions, with dots denoting individual participants. (E) Same as D, but for normalized target probability. For target probability without normalization see Extended Data Fig.1-1. For additional parameters (non-target and randomly guess probability) results in Extended Data Fig.1-2.

Figure 3

Figure 3. “Leader-follower” dynamic perturbation modulates three-item memory performance (Experiment 3, N = 30)

(A) Experiment 3 paradigm. In each trial, participants were presented with three bars and memorized their orientations and colores. During the memory test, participants adjusted the orientation of a probe bar to match that of the memorized bar having the same color as the probe. During the 5 s delay period, participants performed a central fixation task, while three task-irrelevant, flickering discs having the same color as each of the memorized bars (blue, red, green) were presented simultaneously, with their luminances continously modulated by three 5 s temporal sequences (Leader, Follower_{1st}, Follower_{2nd}), respectively. The color, spatial location, and “Leader-Follower” conditions were counterbalanced across trials. (B) The Leader temporal sequence was a 5 s white noise randomly generated per trial, and the Follower_{1st} and Follower_{2nd} sequences were created by circular-shifting the Leader sequence 150 ms and 300 ms rightward, respectively. (C) The Leader- Follower_{1st} and Leader-Follower_{2nd} cross-correlation over time as a function of temporal lag, peaking at 150 ms and 300 ms, respectively. (D) Memory performance. Grand averaged (mean + SEM) memory precision for Leader (purple), Follower_{1st} (turquoise), and Follower_{2nd} (yellow) conditions. Dots denote individual participants. (E) Same as D, but for normalized target probability. For target probability without normalization see Extended Data Fig.1-1. For additional parameters (non-target and randomly guess probability) results in Extended Data Fig.1-2.

Figure 4

Figure 4. Memory-irrelevant "Leader-follower" dynamic perturbation (Experiment 4, N = 30)
(A) Task-irrelevant dynamic perturbation paradigm. Experiment 4 was the same as Experiment 3, except that participants needed to memorize the orientations and locations (left/middle/right) of the three bar stimuli regardless of their color features. During the memory test period, a location cue was first presented, based on which participants rotated the horizontal white bar to the corresponding memorized orientation. Critically, a "Leader-follower" dynamic perturbation as in Experiment 3 was applied during the delay period, i.e., three discs of the same color as each of the memorized bars (blue, red, green) were presented simultaneously, with their luminances continuously modulated by Leader, Follower_{1st} , or Follower_{2nd} sequence, respectively. (B) The Leader temporal sequence was a 5 s white noise randomly generated per trial, and the Follower_{1st} and Follower_{2nd} sequences were created by circular-shifting the Leader sequence 150 ms and 300 ms rightward, respectively. (C) The Leader- Follower_{1st} and Leader- Follower_{2nd} cross-correlation over time as a function of temporal lag, peaking at 150 ms and 300 ms, respectively. (D) Memory performance. Grand averaged (mean + SEM) memory precision for Leader (purple), Follower_{1st} (turquoise), and Follower_{2nd} (yellow) conditions. Dots denote individual participants. (E) Same as D, but for normalized target probability. For target probability without normalization see Extended Data Fig.1-1. For additional parameters (non-target and randomly guess probability) results in Extended Data Fig.1-2.

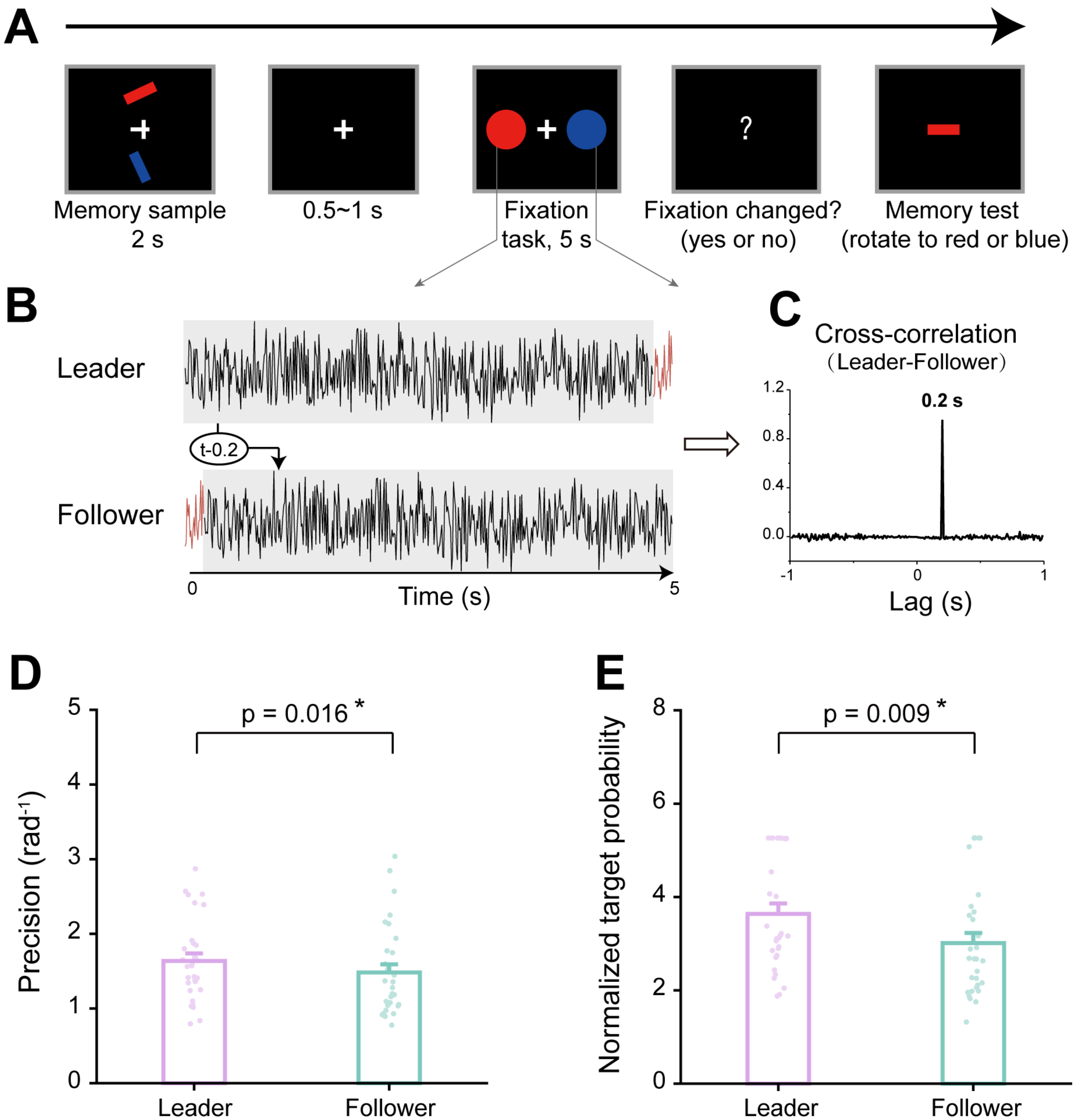
Extended Data

Figure 1-1

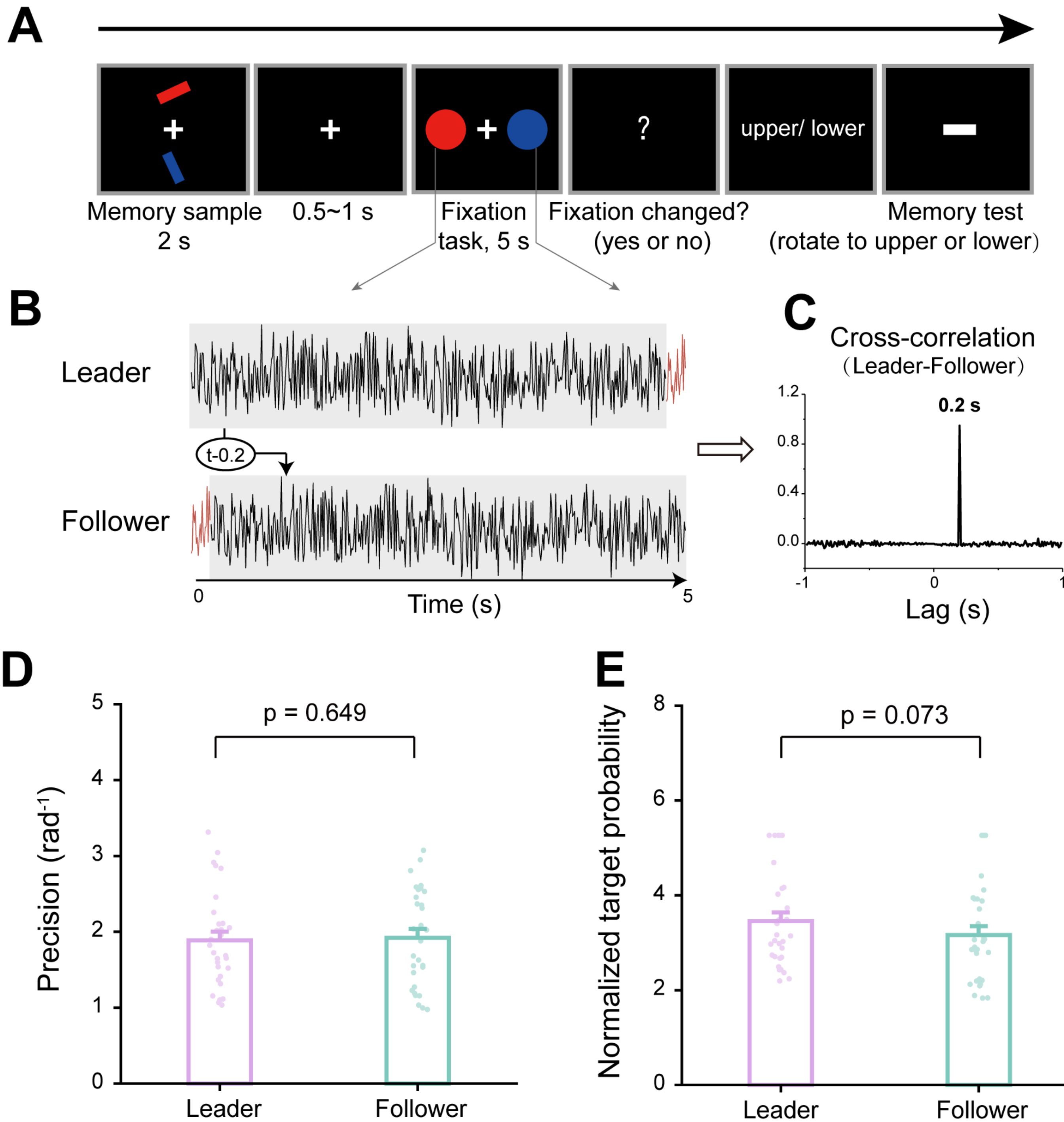
(A) Target probability for the Leader (purple), Follower (turquoise) conditions, with dots denoting individual subjects in Experiment 1. (B-D) Same as A, but for Experiment 2, Experiment 3 and Experiment 4. Correction for multiple comparisons was applied to Experiment 3&4.

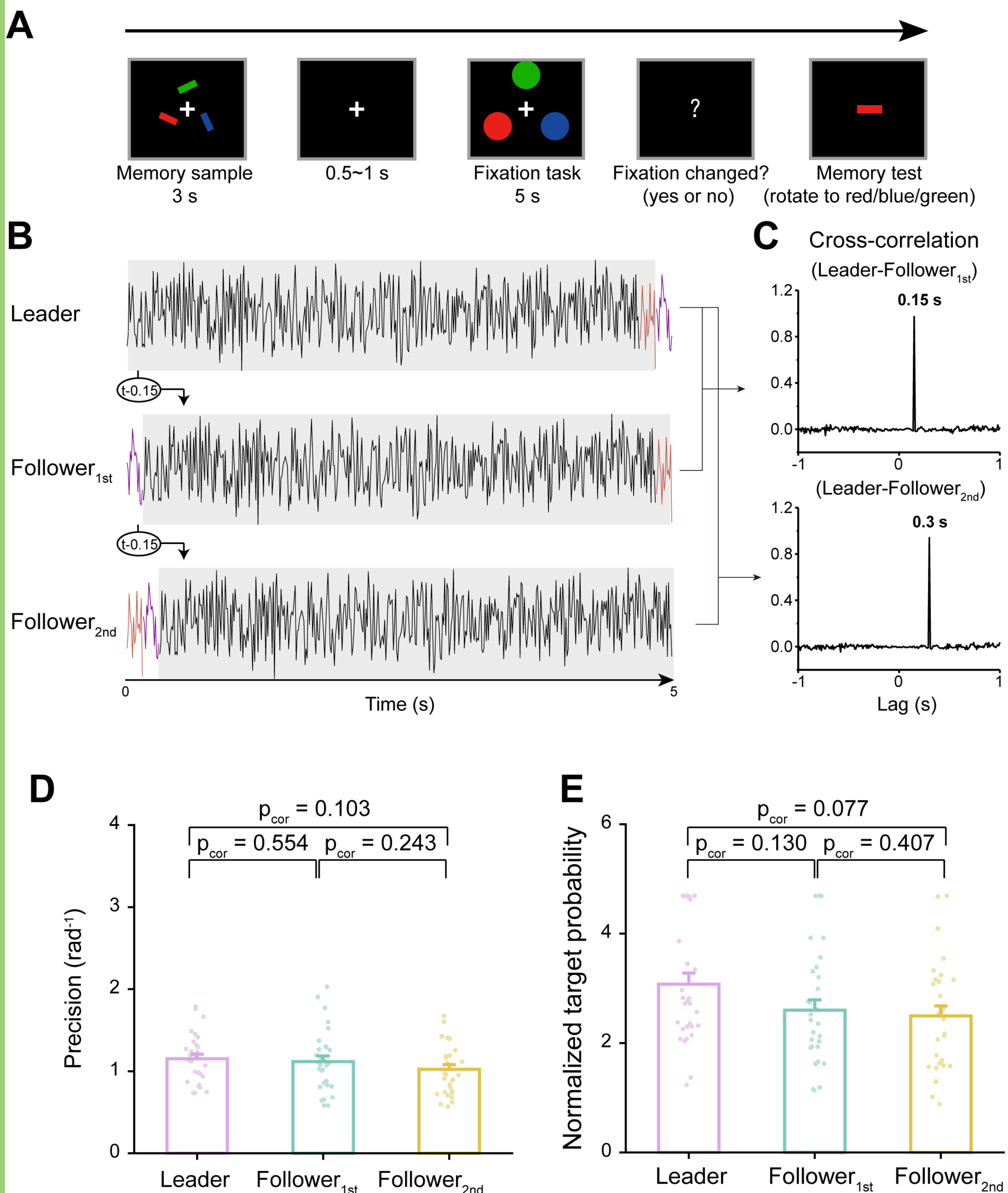
Figure 1-2

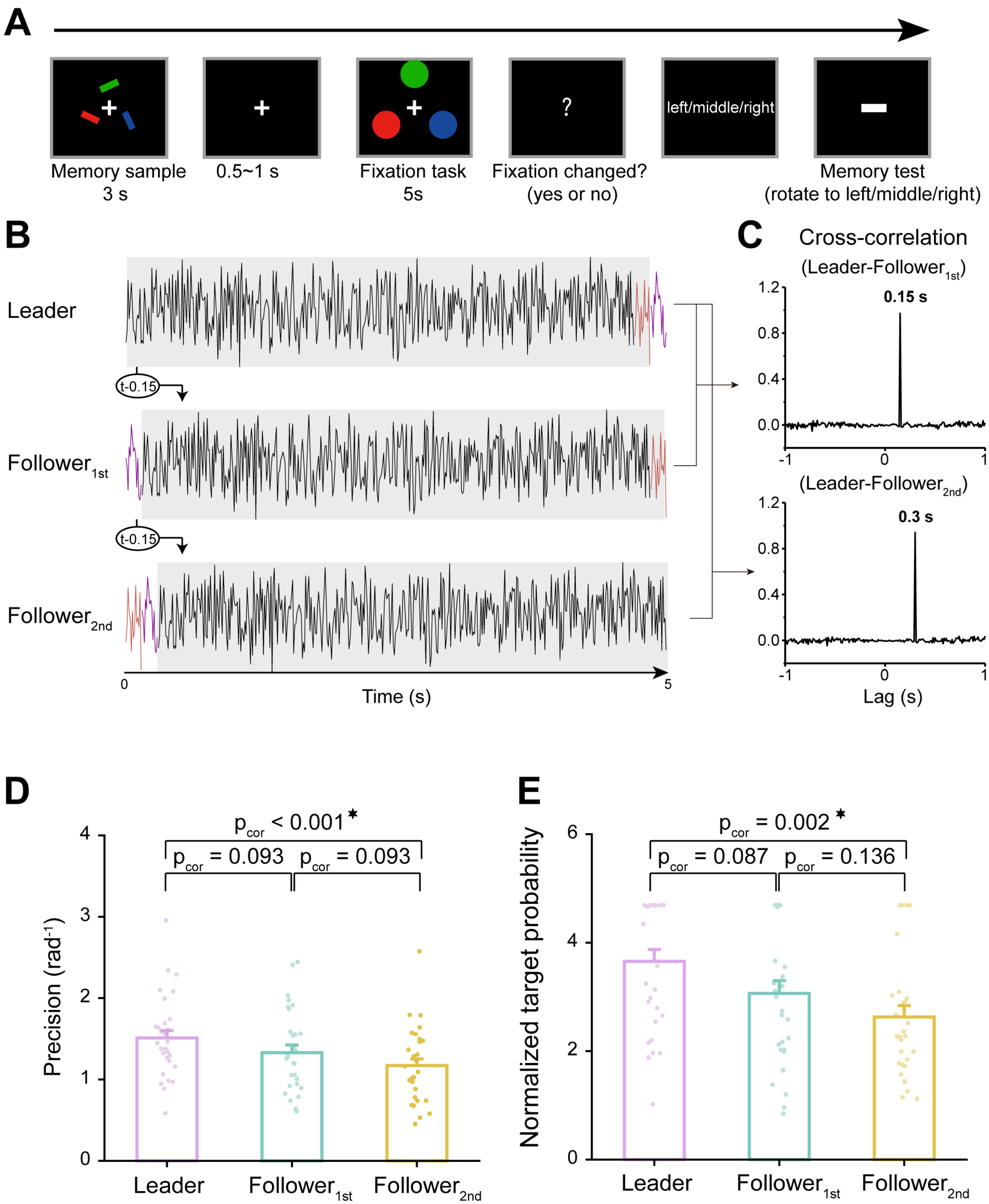
(A) Left panel: non-target probability for the Leader (purple), Follower (turquoise) conditions, with dots denoting individual subjects in Experiment 1. Right panel: random guess probability in Experiment 1. (B-D) Same as A, but for Experiment 2, Experiment 3 and Experiment 4. Correction for multiple comparisons was applied to Experiment 3&4.



A
Memory sample
2 s
0.5~1 s
Fixation
task, 5 s
?
Fixation changed?
(yes or no)
Memory test
(rotate to red or blue)
B
Leader
t-0.2
Follower
0
Time (s)
5
C
Cross-correlation
(Leader-Follower)
1.2
0.8
0.4
0.0
0.2 s
-1
0
1
Lag (s)
D
Precision (rad-1)
p = 0.016 *
5
4
3
2
1
0
Leader
Follower
E
Normalized target probability
p = 0.009 *
8
6
4
2
0
Leader
Follower


A
Memory sample
2 s
0.5~1 s
Fixation
task, 5 s
?
Fixation changed?
(yes or no)
upper/ lower
Memory test
(rotate to upper or lower)
B
Leader
t-0.2
Follower
0
Time (s)
5
C
Cross-correlation
(Leader-Follower)
1.2
0.8
0.4
0.0
0.2 s
-1
0
1
Lag (s)
D
p = 0.649
Precision (rad-1)
5
4
3
2
1
0
Leader
Follower
E
p = 0.073
Normalized target probability
8
6
4
2
0
Leader
Follower


A
Memory sample
3 s
0.5~1 s
Fixation task
5 s
Fixation changed?
(yes or no)
Memory test
(rotate to red/blue/green)
B
Leader
t-0.15
Follower$_{1st}$
t-0.15
Follower$_{2nd}$
0
5
Time (s)
C
Cross-correlation
(Leader-Follower$_{1st}$)
0.15 s
(Leader-Follower$_{2nd}$)
0.3 s
Lag (s)
D
p_{cor} = 0.103
p_{cor} = 0.554
p_{cor} = 0.243
Precision (rad^{-1})
Leader
Follower$_{1st}$
Follower$_{2nd}$
E
p_{cor} = 0.077
p_{cor} = 0.130
p_{cor} = 0.407
Normalized target probability
Leader
Follower$_{1st}$
Follower$_{2nd}$


A
Memory sample
3 s
0.5~1 s
Fixation task
5s
?
Fixation changed?
(yes or no)
left/middle/right
Memory test
(rotate to left/middle/right)
B
Leader
t-0.15
Follower1st
t-0.15
Follower2nd
0
5
Time (s)
C
Cross-correlation
(Leader-Follower1st)
0.15 s
(Leader-Follower2nd)
0.3 s
Lag (s)
D
Precision (rad-1)
pcor < 0.001*
pcor = 0.093
pcor = 0.093
Leader
Follower1st
Follower2nd
E
Normalized target probability
pcor = 0.002*
pcor = 0.087
pcor = 0.136
Leader
Follower1st
Follower2nd