

Perceptual Learning beyond Perception: Mnemonic Representation in Early Visual Cortex and Intraparietal Sulcus

Ke Jia (贾珂),^{1,2,3,4,5*} Ya Li (李雅),^{6*} Mengyuan Gong (龚梦园),^{7*} Hui Huang (黄晖),^{1,2,3,4}
Yonghui Wang (王勇慧),⁶ and Sheng Li (李晟)^{1,2,3,4}

¹School of Psychological and Cognitive Sciences, Peking University, Beijing 100871, China, ²Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100871, China, ³PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China, ⁴Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China, ⁵Department of Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom, ⁶Shaanxi Provincial Key Laboratory of Behavior and Cognitive Neuroscience, School of Psychology, Shaanxi Normal University, Xi'an 710062, China, and ⁷Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310028, China

The ability to discriminate between stimuli relies on a chain of neural operations associated with perception, memory and decision-making. Accumulating studies show learning-dependent plasticity in perception or decision-making, yet whether perceptual learning modifies mnemonic processing remains unclear. Here, we trained human participants of both sexes in an orientation discrimination task, while using functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS) to separately examine training-induced changes in working memory (WM) representation. fMRI decoding revealed orientation-specific neural patterns during the delay period in primary visual cortex (V1) before, but not after, training, whereas neurodisruption of V1 during the delay period led to behavioral deficits in both phases. In contrast, both fMRI decoding and disruptive effect of TMS showed that intraparietal sulcus (IPS) represented WM content after, but not before, training. These results suggest that training does not affect the necessity of sensory area in representing WM information, consistent with the sensory recruitment hypothesis in WM, but likely alters the coding format of the stored stimulus in this region. On the other hand, training can render WM content to be maintained in higher-order parietal areas, complementing sensory area to support more robust maintenance of information.

Key words: perceptual learning; working memory; fMRI decoding; TMS

Significance Statement

There has been accumulating progresses regarding experience-dependent plasticity in perception or decision-making, yet how perceptual experience moulds mnemonic processing of visual information remains less explored. Here, we provide novel findings that learning-dependent improvement of discriminability accompanies altered WM representation at different cortical levels. Critically, we suggest a role of training in modulating cortical locus of WM representation, providing a plausible explanation to reconcile the discrepant findings between human and animal studies regarding the recruitment of sensory or higher-order areas in WM.

Received Nov. 1, 2020; revised Mar. 22, 2021; accepted Mar. 24, 2021.

Author contributions: K.J. and S.L. designed research; K.J., Y.L., H.H., and Y.W. performed research; K.J. and Y.L. analyzed data; K.J., Y.L., M.G., and S.L. wrote the paper.

This work was supported by the National Key R&D Program of China Grant 2017YFB1002503 and by National Natural Science Foundation of China Grants 31271081, 31230029, 31800910, and 32000784.

*K.J., Y.L., and M.G. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Sheng Li at sli@pku.edu.cn.

<https://doi.org/10.1523/JNEUROSCI.2780-20.2021>

Copyright © 2021 the authors

Introduction

The ability to differentiate between similar features is essential for visual recognition in complex environment. For instance, the predators must learn to discriminate the prey items from surroundings to ensure survival. Learning and experience are known to improve the discrimination ability even in adulthood by re-organizing the brain functions and connections (Sagi and Tanne, 1994; Kourtzi and DiCarlo, 2006; Gilbert et al., 2009; Watanabe and Sasaki, 2015; Doshier and Lu, 2017; Hooks and

Chen, 2020). Previous studies have focused on how training alters perceptual encoding of the stimuli (Schoups et al., 2001; Schwartz et al., 2002; Furmanski et al., 2004; Yang and Maunsell, 2004; Yotsumoto et al., 2008; Jehee et al., 2012; Yan et al., 2014; Chen et al., 2015) or the decision-making process (Law and Gold, 2008; Kahnt et al., 2011; Kuai et al., 2013; Doshier and Lu, 2017). However, mnemonic processing also matters for discrimination judgments where the to-be-compared stimuli are often sequentially presented. In these tasks, participants are required to encode a sample item and hold it in working memory (WM) for later comparison with a test item. Yet whether and how training on these tasks modifies mnemonic processing of stimuli remain largely unclear.

The view that perceptual learning may change mnemonic processing of stimuli received support from findings of the relationship between WM and discrimination ability (Cornette et al., 2001; Brady et al., 2013; Ester et al., 2014; Zhang et al., 2016). In particular, variability of neuronal activity during WM retention is proposed as a potential indicator of the discrimination performance (Hussar and Pasternak, 2010; Qi and Constantinidis, 2015). The amount of information carried by the activity patterns during WM delay correlates with mnemonic precision (Ester et al., 2013) and performance changes as a function of WM load (Emrich et al., 2013). These findings point to the assumption that learning-dependent improvement of discriminability may be accompanied by modified WM representation of the stimuli. It has been established that multiple levels of cortical areas are recruited for representing WM information (D'Esposito, 2007; Christophel et al., 2017; Dotson et al., 2018). In particular, intraparietal sulcus (IPS) area is identified as a candidate region for mnemonic processing of the stimuli (Song and Jiang, 2006; Bettencourt and Xu, 2016; Weber et al., 2016; Lorenc et al., 2018), while the sensory recruitment account of WM suggests that primary visual area (V1) is also engaged for temporary maintenance of WM content (Pasternak and Greenlee, 2005; Ester et al., 2009; Harrison and Tong, 2009; Serences et al., 2009). Representing WM content at multiple areas could play complementary roles such that sensory areas encode precise sensory information and higher-order areas provide abstract and robust representation (D'Esposito, 2007; Christophel et al., 2017). Here, we particularly focus on these two regions of interest (ROIs) to examine learning-dependent alterations of WM representation.

To this end, we trained participants on a two-interval forced-choice (2IFC) orientation discrimination task that required temporary maintenance of the sample stimulus during a delay period. In Experiment 1, we combined functional magnetic resonance imaging (fMRI) with multivariate pattern analysis (MVPA) to investigate learning-dependent changes of WM representation in V1 and IPS. In Experiment 2, we used online repetitive transcranial magnetic stimulation (rTMS) to test the effect of training on the causal role of these two areas during WM processing. We found orientation-specific patterns during WM delay in V1 before, but not after, training, whereas V1 stimulation during the delay period impaired behavioral performance in both phases. In contrast, both fMRI decoding and TMS effect indicated that IPS represented WM content after, but not before, training. These findings suggest that perceptual learning modified mnemonic processing at different cortical levels.

Materials and Methods

Experiment 1: fMRI

Participants

Sixteen participants (nine females; age range: 18–26 years) took part in this study. The sample size is comparable to those reported in previous

work on perceptual learning (Zhang et al., 2010) or fMRI decoding of WM content using discrimination tasks (Ester et al., 2009; Gosseries et al., 2018; Lawrence et al., 2018). All participants had normal or corrected-to-normal vision, and reported being right-handed. They were naive to the aim of the study and received payment on completion of the experiment. All participants gave written informed consent and the study protocol was approved by the local ethics committee.

Stimulus and apparatus

We presented Gabor patches (Gaussian windowed sinusoidal gratings) in either upper-left or lower-right visual field with an eccentricity of 6.5° against a gray background (~35 cd/m²). The Gabor stimuli of random phase had a fixed diameter of 4°, contrast of 0.8, spatial frequency of 1.5 cycle/°. The angle of Gabor stimuli was tilted clockwise or counterclockwise from the base orientations (55° or 145°).

The stimuli were generated using Psychtoolbox 3.0 (Brainard, 1997; Pelli, 1997) for MATLAB (MathWorks). In the behavioral lab, the stimuli were presented on a Dell cathode ray tube (CRT) monitor with the size of 40 × 30 cm², resolution of 1024 × 768 and a refresh rate of 60 Hz. Gamma correction was applied to the monitor. We used a chin-rest to stabilize participants' head position and maintain the viewing distance at 90 cm. Participants were asked to make responses using a keyboard. Inside the MRI scanner, the stimuli were back-projected onto a translucent screen located inside the scanner bore (resolution, 1024 × 768; refresh rate, 60 Hz). Participants viewed the stimuli at a distance of 90 cm through a mirror placed above their eyes. An MRI-compatible response box was used for making responses.

Experimental design and statistical analysis

All participants completed four phases in this experiment, each phase consisted of multiple sessions (Fig. 1A): (1) a 2-d pretest, (2) a 6-d training, (3) a 2-d posttest I; and (4) a 2-d posttest II. Posttest I and posttest II phases were separated for around 10 d to assess the stability of training effect. Each test phase comprised of a behavioral test session (first day) and a scanning session (second day) on two consecutive days.

Behavioral tasks. We used a 2IFC, orientation discrimination task throughout the experiment. Two types of tasks, a short-delay and a long-delay task, varying in the length of delay period between the stimuli were included (Fig. 1B,C). Similar to the conventional learning regimen, we used a short-delay of 0.6 s to measure behavioral performance during the training and behavioral test sessions. To isolate memory-specific activity from fMRI signal (Harrison and Tong, 2009; Serences et al., 2009), we used a long-delay of 11.8 s during the behavioral tests and scanning sessions. Note that our WM task required holding of one orientation in WM, which differed from those manipulated WM load. This design was chosen as it is commonly used in studies on perceptual learning (Schoups et al., 2001; Jehee et al., 2012) and fMRI decoding of WM content (Harrison and Tong, 2009; Serences et al., 2009; Bettencourt and Xu, 2016).

In the short-delay task, each trial began with a central fixation of a black dot shown for 0.6 s. In the long-delay task, each trial began with a central fixation dot that was white for 0.2 s and then turned into black for 3.8 s. The change of color was designed to remind participants of the trial onset. Participants were instructed to press a button once they saw the white dot. In both tasks, the sample and test Gabor stimuli were then sequentially presented for 0.2 s each, separated by a delay period (short-delay task: 0.6 s; long-delay task: 11.8 s). Participants were asked to report whether the test Gabor was tilted clockwise or counter-clockwise relative to the sample stimulus. A uniformly distributed jitter ($\pm 5^\circ$) was added to the base orientations (i.e., 55° or 145°) to encourage perceptual comparison between two Gabors in each trial, rather than direct retrieval of a constant stimulus template.

Staircase procedure. To equate task difficulty across different conditions throughout the experiment, we used adaptive staircase method (3-down-1-up, 15 reversals, step size of 0.5°) that converges to 79.4% accuracy in the orientation discrimination tasks. We adjusted the angle difference between the sample and test stimuli independently for each condition. The threshold in each run was determined by the mean angle difference of the last eight reversals.

Behavioral test sessions (first day of pretest, posttest I, and posttest II). This session included both the short-delay and long-delay orientation discrimination tasks. The short-delay task consisted of four experimental conditions (2 stimulus orientations \times 2 stimulus locations) to assess the effect of learning and the learning specificity for orientation and location. Participants started with 16 practice trials (4 trials per condition) using a fixed angle difference (10°) and then completed 12 staircase runs (~ 65 trials per run, three staircases per condition in random order). For the first run of each condition, the starting angle difference was 8° with a step size of 0.5° . For the subsequent staircase runs, the starting value was the threshold of corresponding condition in the preceding run. Participants' performance in each session was quantified using the averaged thresholds across three staircases for each condition.

To keep consistency with the trial sequence in the scanning session, the long-delay task consisted of two stimulus conditions (i.e., $\sim 55^\circ$ or $\sim 145^\circ$) shown only at the trained location. Participants began with 20 practice trials (10 trials per condition, fixed angle difference: 10°) and then completed one run of randomly interleaved staircases (~ 65 trials for each condition). The starting angle difference was 8° with a step size of 0.5° . We quantified the performance using the threshold for each condition. No feedback on correctness was provided in any of these test sessions.

Scanning sessions (second day of the pretest, posttest I, and posttest II). Participants completed six runs of a long-delay task (16 trials per run, eight trials for each orientation in randomized order). Each run began with an 8-s fixation. Trials were separated by a 10- or 12-s interval to allow fMRI signals to return to baseline. We measured performance with staircase procedure. The starting value was the threshold inherited from the preceding behavioral session in the corresponding test phase. In addition to the discrimination task, each participant completed a retinotopic mapping scan (6 min 20 s), a localizer scan (5 min 36 s) and an anatomic image scan (see ROI definition for details). No feedback on the correctness was provided in the scanning sessions.

Training sessions. Participants were trained on an orientation discrimination task with Gabors presented at the same orientation and location throughout training. In each session, participants performed 16 runs of short-delay task. We measured performance with staircase procedure. For the first run of the first session, the starting angle difference was 8° with a step size of 0.5° . For the subsequent staircase runs, the starting value was the threshold from the preceding run. Training locations (i.e., upper-left or lower-right) and orientation (i.e., 55° or 145°) were counterbalanced across participants. In addition, we provided auditory feedback on incorrect trials. We trained participants for 6 d, resulting in a total of ~ 6200 trials.

Behavioral data analysis. To validate the training effect, we used paired *t* test to compare the discrimination thresholds between the first and last sessions of the training phase. To examine the effect of training on discrimination performance in the test phases, we calculated a mean percent improvement [MPI = (pretest threshold – posttest threshold)/pretest threshold \times 100%; Xiao et al., 2008], separately for each posttest phase. For the short-delay task, we applied a three-way repeated-measures ANOVA (2 stimulus orientations \times 2 stimulus locations \times 2 posttest phases) on MPI.

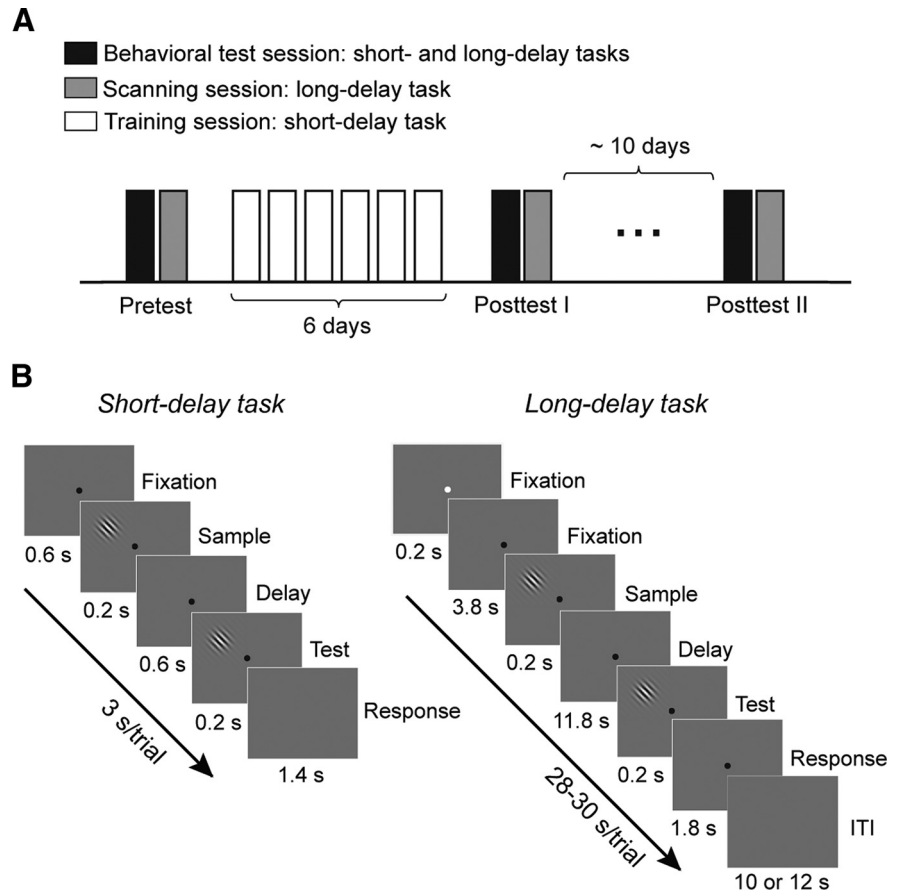


Figure 1. Procedure and tasks for Experiment 1. **A**, Participants completed four phases in Experiment 1, including (1) a 2-d pretest, (2) a 6-d training, (3) a 2-d posttest I and (4) a 2-d posttest II. Each test phase comprised of a behavioral and a scanning session, as separately denoted by black and gray bars. The first session of each test phase (behavioral session) included both short-delay and long-delay tasks. The second session of each phase (scanning session) included the long-delay task to address neural activity during WM delay. Training phase used short-delay task only. **B**, Trial sequence of short-delay (left) and long-delay (right) orientation discrimination tasks. Participants viewed two sequentially presented stimuli and reported whether the test stimulus was tilted clockwise or counterclockwise relative to the sample stimulus in both tasks.

For the long-delay task, we applied a 2-way repeated-measures ANOVA (2 stimulus orientations \times 2 posttest phases) on MPI because of the presence of stimulus solely at the trained location.

MRI data acquisition and preprocessing

Imaging data were acquired on a Siemens 3T Prisma scanner located at Peking University. All imaging data were acquired with a 20-channel head coil. For each participant, anatomic images were acquired using MP-RAGE T1-weighted sequence (TR = 2530 ms, TE = 2.98 ms, FOV = $256 \times 224 \text{ mm}^2$, flip angle: 7° , resolution $0.5 \times 0.5 \times 1 \text{ mm}^3$, number of slices: 192, slice thickness: 1 mm, slice orientation: sagittal). Functional scans were acquired using echo planar imaging (EPI) sequence (TR = 2000 ms, TE = 30 ms, FOV = $224 \times 224 \text{ mm}^2$, flip angle: 90° , matrix: 64×64 , resolution $3.5 \times 3.5 \times 3.5 \text{ mm}^3$, gap = 0.7 mm, number of slices: 33).

Each participant's anatomic image was segmented into gray and white matter using FreeSurfer (<http://surfer.nmr.mgh.harvard.edu/>). We performed the cortical reconstruction of the segmented images in BrainVoyager QX software (Brain Innovation). For the functional images, we discarded the first four volumes at the beginning of each run to ensure that the longitudinal magnetization reached steady state. The functional data were processed with slice-timing correction, head motion correction, temporal filtering (three cycles), and removal of linear trends in BrainVoyager QX. Within each scanning session, the functional data were aligned to the first volume of the first run and co-registered to the anatomic image obtained in the same session. Between scanning sessions, all anatomic images were aligned to the participant's

own anatomic data acquired in their first session and transformed to the Talairach space. The functional data in the Talairach space were resampled into $3 \times 3 \times 3 \text{ mm}^3$ resolution.

ROI definition and fMRI data analysis

Definition of V1. Participants viewed rotating wedges that created traveling waves of neural activity. We identified V1 boundaries using standard phase-encoded method (Serenio et al., 1995; Engel et al., 1997). In a separate localizer run, we mapped two location-specific areas in V1, corresponding to the stimulus locations from the orientation discrimination task (i.e., upper-left and lower-right). In each trial, a Gabor patch (55° or 145°) was presented at one of the locations for 2 s. The intertrial interval (ITI) was either 2 or 4 s. The location was randomized across 32 trials. Participants were asked to detect a subtle change of orientation. For each participant and each functional localizer, we computed each voxel's response using a general linear model (GLM) comprised of two regressors, one for each stimulus location. Contrasts comparing stimulus in one location to the other led to positive responses in V1 ROI contralateral to the stimulus location. We selected 40 voxels with top-ranked β estimates for each stimulus location, the exact number of voxels was determined by the minimal number of voxels across participants and V1 ROIs. This voxel selection regime controlled for potential biases in classification accuracy because of varying number of voxels across locations and participants.

Definition of IPS. We selected IPS ROIs that were functionally defined by the delay period activity, within the anatomically constrained regions. The delay period activity was primarily assumed to reflect WM storage, while it may also relate to other control-related processes (Christophel et al., 2017; Sreenivasan and D'Esposito, 2019). In particular, after applying anatomic segmentation in FreeSurfer, we used the automated ROI labels from Destrieux atlas (Destrieux et al., 2010) to transform the identified IPS into Talairach space. For each participant, we conducted a GLM analysis that modeled the WM-related activity after the sample stimulus (i.e., delay period) and the baseline activity after the test stimulus (i.e., ITI). The resulting β estimates that indicated statistically significant increases of delay period activity ($p < 0.05$) were used for voxel selection within anatomically-defined ROIs (Xu, 2007). Because of the more prominent delay period activity in left IPS, we defined IPS based on hemisphere (i.e., left and right IPS) to accord with previous studies that had similar observation of left-lateralized delay activity in IPS (Christophel et al., 2012; Albers et al., 2013; Ester et al., 2015). We selected 250 voxels with top-ranked β estimates in each hemisphere for further analysis. The exact number of voxels was determined by the minimal number of voxels across participants and ROIs.

Univariate analysis. We assessed whether training changes the overall BOLD response during WM delay. For each participant and each run, we first extracted z-normalized response amplitude of each voxel in the predefined ROIs (V1 and IPS). Then, we took the trial-averaging BOLD response between 0 and 26 s time locked to trial onset, separately for each experimental condition during the test phase. Because of our primary focus on the sustained activity during the delay period, we averaged the response across 12 and 14 s after the trial onset (7th and 8th TRs) that were uncontaminated by the test stimulus presentation (16 s after the trial onset, 9th TR). For each of the ROIs, we then applied two-way repeated-measures ANOVA (2 stimulus orientations \times 3 test phases) on the delay activity to assess how training influenced WM-related activity.

MVPA. We used the MVPA to decode the stimulus orientation during the delay period in V1 and IPS. For each participant and each test phase, we extracted z-normalized BOLD responses between 12 and 14 s after the trial onset (7th and 8th TRs) and used the average of the two data points in each trial to represent the delay period activity. By training the classifier to discriminate between two orientations using LIBSVM (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>), we calculated the classification accuracy with a leave-one-run-out cross-validation scheme that divided the data set into training (five runs) and testing data (one run). This procedure was repeated for six times until each run was tested once (Kamitani and Tong, 2006). The classification accuracy was averaged across the folds, separately for each test phase. To evaluate whether the classification accuracy exceeded the chance level, we performed the

permutation test (see below, Permutation test). To assess how training influenced WM representation, we performed one-way repeated-measures ANOVA (3 test phases) on the classification accuracy for each brain area.

To validate that the orientation decoding in V1 reflected WM content, rather than residual sensory information, we conducted control analyses by applying MVPA on the neural activity during ITI (24 and 26 s after the trial onset, 13th and 14th TRs). The selected time window for ITI was supposed to contain the same amount of sensory information as that for the delay period (i.e., 8 and 10 s after the onset of the sample and test stimulus, respectively). To compare the classification accuracy between these two time periods (delay vs ITI), we conducted a three-way repeated-measures ANOVA (2 V1 ROIs \times 2 time periods \times 3 test phases). Further, to directly compare learning-dependent changes of decoding performance between V1 and IPS, we calculated the difference of classification accuracy between the pretest and two posttest phases (i.e., posttest I – pretest, posttest II – pretest), separately for V1 and IPS. A two-way repeated-measures ANOVA (2 ROIs \times 2 posttest phases) were conducted to assess the training effect between brain areas.

Permutation test. We evaluated the statistical significance of MVPA results using the permutation tests (Stelzer et al., 2013; Allefeld et al., 2016). In particular, for each scanning session (pretest, posttest I and posttest II) and each brain area, we took the leave-one-run-out cross-validation approach, in which we shuffled the trial labels in the training data and calculated the classification accuracy on the test data. We obtained the classification accuracy that averaged across folds for each participant. Then, we averaged the accuracies across participants to obtain a mean value. This procedure was repeated for 5000 times to compute a group-level null distribution, consistent with the method used in previous studies (Chen et al., 2011; Cocchi et al., 2017; Roth et al., 2018; Henderson and Serences, 2019). We obtained the p values by calculating the proportion of random samples that exceeds the observed value (i.e., mean classification accuracy from real data). We applied false discovery rate (FDR) method (Benjamini and Hochberg, 1995) to correct p values for multiple comparisons across predefined ROIs and test phases.

Experiment 2: TMS

Participants

Twenty-three participants (12 females; age range: 19–26 years) were recruited for this study. Three of them did not participate TMS sessions after fMRI scanning because of the lack of elevated delay period activity in IPS (for details, see below, Definition of ROIs). The sample size was comparable to those reported in previous WM-related TMS studies (Zanto et al., 2014; Zokaei et al., 2014). All participants were neurologically intact, had normal or corrected-to-normal vision, and reported being right-handed. All participants gave the informed consent and the study protocol was approved by the local ethics committee.

Stimulus and apparatus

Identical stimuli were used as that in Experiment 1. The stimuli were presented against a gray background ($\sim 15 \text{ cd/m}^2$) on a CRT monitor (refresh rate: 60 Hz) for both behavioral and fMRI experiments. In the TMS lab, the stimuli were displayed on a gray background ($\sim 19 \text{ cd/m}^2$) on a CRT monitor (refresh rate: 100 Hz).

Experimental design and statistical analysis

The TMS experiment consisted of four phases (Fig. 2A): (1) a scan for defining ROIs (V1 and IPS) per participant, (2) a 2-d pretest, (3) a 6-d training, (4) a 2-d posttest. Pretest and posttest were completed 1 d before and after the training phase, respectively. The pretest and posttest phases consisted of a behavioral session (first day) and a TMS session (second day).

Scanning session. To guide precise stimulation of the target regions in TMS sessions, each participant completed a V1 localizer scan (two runs) and an IPS localizer scan (one run), in addition to an anatomic image scan and a retinotopic mapping scan (for details, see below, Definition of ROIs).

Behavioral test sessions (first day of the pretest and posttest). Participants performed a long delay orientation discrimination task that was similar to that used in Experiment 1. In brief, the total duration of each trial was fixed to 7 s. Each trial began with a central fixation of a black dot shown for 0.6 s on a gray background. The sample and test Gabor stimuli were then sequentially presented for 0.2 s each, separated by a 4-s delay. A blank screen was then shown for 2 s. Participants were asked to report whether the test Gabor was tilted clockwise or counter-clockwise relative to the sample stimulus within 1.5 s. Notable changes were made to the task design for several practical concerns. First, we shortened the delay period from 11.8 to 4 s during online TMS stimulation. In Experiment 1, the sluggish BOLD signals require a long delay to isolate WM-related activity, which is not necessary for assessing the neurodisruptive effects of TMS. Second, taking into considerations of the stimulation site on the left hemisphere (see rTMS protocol), we presented the stimulus at one of the stimulus locations used in Experiment 1 (i.e., lower-right visual field) that corresponds to the left hemisphere. Third, we used one orientation (55°) as we aimed to compare the TMS effects on discrimination performance between test phases (i.e., pretest vs posttest) and between stimulation conditions (i.e., V1, IPS, and sham), rather than between two orientations (i.e., the trained and untrained orientations). Participants started with 40 practice trials (a fixed angle difference: 10°) and then completed two to three runs of the main task using the staircase procedure (~65 trials per run), identical to that used in the behavioral test sessions of Experiment 1. No feedback on correctness was provided in these sessions except practice trials.

Training sessions. Participants were trained to discriminate the orientation around 55° presented at the lower-right visual field. On each session, they performed 16 staircase runs, using the same protocol as that in the training sessions of Experiment 1. We provided auditory feedback on incorrect trials. We trained participants for 6 d, resulting in a total of ~6200 trials.

TMS sessions (second day of the pretest and posttest). In an orientation discrimination task (Fig. 2B), we used a fixed angle difference determined by the threshold from the behavioral test session in corresponding phase for each participant. Participants started with 40 practice trials and then completed three runs (80 trials per run). In separate runs, the magnetic stimulation was delivered to one of the three stimulation conditions (V1, IPS, and sham) during the delay period. The order of the stimulation conditions was counterbalanced across participants. No feedback on correctness was provided in the TMS sessions.

TMS and MRI parameters

MRI data acquisition and preprocessing. Imaging data were acquired on a 3T GE MEDICAL SYSTEMS scanner located at Peking University using an eight-channel head coil. For each participant, anatomic images were acquired using T1-weight sequence (TR = 6.656 ms, TE = 2.92 ms, FOV = 256 × 256 mm², flip angle: 90°, resolution 1 × 1 × 1 mm³, number of slices: 192, slice thickness: 1 mm, slice orientation: sagittal). Functional scans were acquired using EPI sequence (TR = 2000 ms, TE = 30 ms, FOV = 224 × 224 mm², flip angle: 90°, matrix: 64 × 64, resolution 3.5 × 3.5 × 3.5 mm³, gap = 0.7 mm, number of slices: 33). The same preprocessing procedure was applied to fMRI data as that in Experiment 1.

Definition of ROIs. For each participant, the functional data were aligned to the anatomic data in native space. Identical to our ROI definition approach in Experiment 1, we defined V1 using retinotopic mapping and standard phase-encoded method. Further, we applied a GLM

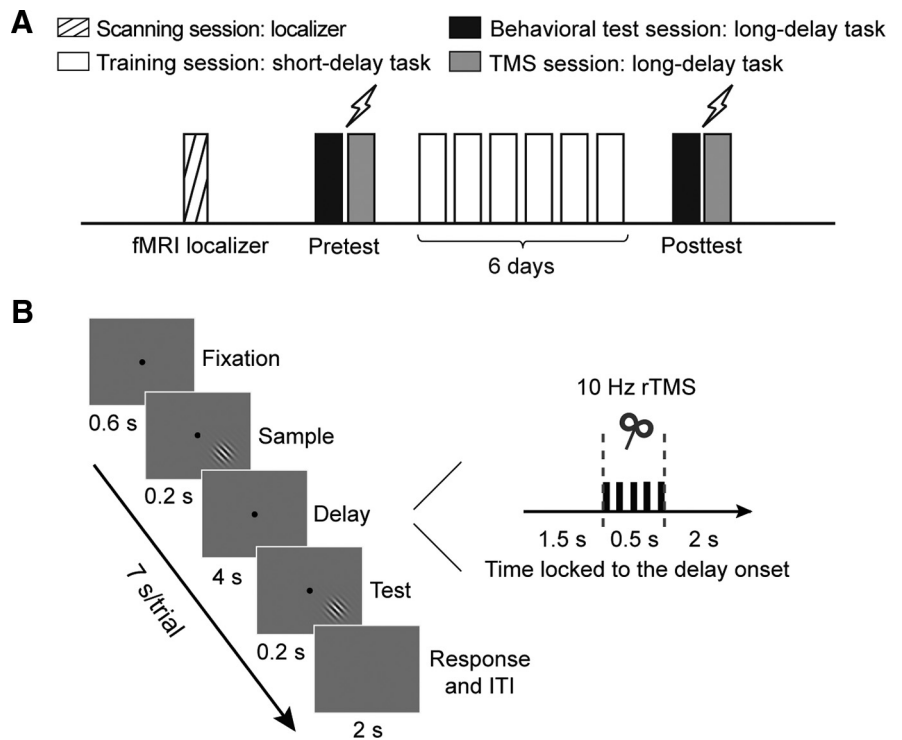


Figure 2. Procedure and tasks for Experiment 2. **A**, Participants completed four phases in Experiment 2, including (1) fMRI localizer scans for ROIs definition (V1 and IPS), (2) a 2-d pretest, (3) a 6-d training, (4) a 2-d posttest. Pretest and posttest were completed 1 d before and after the training phase. The two sessions of each test phase were behavioral and TMS sessions, respectively. Both sessions included the long-delay orientation discrimination task. Training phase included the same short-delay task as that used in Experiment 1. **B**, Trial sequence of the long-delay orientation discrimination task. Participants viewed two sequentially presented stimuli and reported whether the test stimulus was tilted clockwise or counterclockwise relative to the sample stimulus. An online 10-Hz rTMS (five pulses synchronized with 1500 ms after the offset of the sample stimulus) was delivered to one of the stimulation conditions (i.e., V1, IPS, or sham).

on data from two localizer runs to estimate each voxel's response in V1 (i.e., β estimate), allowing us to define the exact stimulus location (β : lower-right > upper-left). Further, we selected IPS voxels showing elevated delay period activity (β : delay period > ITI), while also locating within an anatomically defined IPS ROI. In addition, we included a control condition with sham TMS over the vertex. Vertex was defined as a midpoint betweeninion and nasion that was equidistant from left and right intertrachial notches. The coil was centered at the vertex with its face rotated 90° away from the scalp during stimulation. Thus, no cortical stimulation should be received during sham TMS.

rTMS protocol. To investigate the cause role of sensory and parietal areas during WM retention in discrimination tasks along with training, online rTMS was applied over V1 and IPS during the delay period. Online 10-Hz rTMS (five pulses synchronized with 1500 ms after the offset of the sample stimulus) was delivered at each stimulation site. This TMS protocol was shown to induce interference effects (Mevorach et al., 2010; Romei et al., 2010; Chang et al., 2014) and disrupt BOLD signals in the stimulated area in a concurrent fMRI-TMS study (Sack et al., 2007). We used a fixed intensity of 60% of the stimulator's maximum output for all participants that was comparable to previous studies on visual and parietal stimulation (Mevorach et al., 2010; Romei et al., 2010; Chang et al., 2014). Note that we did not use motor threshold to determine stimulation intensity for individual participant because it is not necessarily a reliable index of excitability in non-motor areas of the brain (Stewart et al., 2001; Robertson et al., 2003). Moreover, previous work showed that 10-Hz rTMS induced disruptive effects in both occipital and parietal areas (Romei et al., 2010), which were short-lived and observed only by the end of the TMS trains. Meanwhile, a concurrent EEG-TMS study found that this protocol led to progressively enhanced alpha activity during stimulation, which lasted for ~100–150 ms after the last pulse of TMS train (Thut et al., 2011). Therefore, to avoid the disruptive effect of TMS on the processing of the test stimulus following the delay period,

we added a 2-s delay by the end of the TMS trains. The TMS coil was air-cooled for >10 min after each run to prevent overheating of the coil during the experiment.

In particular, we included two target sites (i.e., left V1 and left IPS) for the following reasons. First, TMS effects on V1 for peripherally presented stimuli (>3°) are mainly restricted to the lower visual field (Kastner et al., 1998). Considering that we presented peripheral stimuli at an eccentricity of 6.5° in two stimulus locations (i.e., upper-left vs lower-right visual field) in Experiment 1, it is expected to be more effective to stimulate left V1 that responded to the lower-right visual field. Second, motivated by our IPS findings in Experiment 1 (see Results), we had a prior to stimulate the left IPS. In addition, we included a sham condition to account for nonspecific TMS effects related to variations in general behavioral state (e.g., noise, vigilance).

rTMS pulses were delivered through a MagStim Super Rapid² stimulator (The MagStim Company) in combination with a 70-mm figure-of-eight coil. Using fMRI-guided Visor Navigation System (Visor2; Advanced Neuro Technology), we separately overlaid V1 and IPS on the anatomic MR image for each participant with their centroid serving as the target site. The center of the coil was placed tangentially over these sites and a mechanical arm was used to keep the coil steady on the scalp. During V1 stimulation, the coil was held with the handle pointing right and parallel to the ground. During IPS stimulation, the coil was held with the handle pointing away ~45° along the midline (Capotosto et al., 2012; Morgan et al., 2013). The coil position in different sites was chosen based on the literature (Janssen et al., 2015) and was in real-time monitored using Visor2 throughout each session.

Behavioral analysis

Behavioral performance was quantified using the discrimination accuracy and reaction time (RT). To validate the effect of training, we used paired *t* test to compare the discrimination thresholds between the first and the last sessions of training phase. We applied two separate two-way repeated-measures ANOVAs (2 TMS conditions: active vs sham × 2 test phases: pretest and posttest) on the discrimination accuracy and RT. To parallel the comparison of fMRI decoding between two brain areas in Experiment 1, we performed two-way repeated-measures ANOVAs (2 stimulation sites: V1 vs IPS × 2 test phases: pretest and posttest) on the sham-normalized discrimination accuracy (i.e., V1 – sham, IPS – sham).

Results

Experiment 1: fMRI

Perceptual learning improves performance in short and long delay tasks

Perceptual learning improved participants' discrimination performance, as revealed by the decreased threshold from the first session (mean = 3.00°, SD = 0.81°) to the last session (mean = 2.11°, SD = 0.51°) of the training phase (paired *t* test: $t_{(15)} = 6.40$, $p < 0.001$, Cohen's $d = 1.600$; Fig. 3A).

To assess the specificity of learning effect, we applied a three-way repeated-measures ANOVA (2 stimulus orientations × 2 stimulus locations × 2 posttest phases) on MPI in the short-delay task (Fig. 3B, left panel). The results showed a main effect of stimulus orientation ($F_{(1,15)} = 6.06$, $p = 0.026$, $\eta_p^2 = 0.288$) and stimulus location ($F_{(1,15)} = 12.51$, $p = 0.003$, $\eta_p^2 = 0.455$), indicating more training-induced improvement for the trained versus untrained orientation, as well as training-induced improvement for the trained versus untrained location. None of the other main effects or interaction effects reached significance level ($ps > 0.35$). Importantly, when applying a two-way ANOVA (2 stimulus orientations × 2 posttest phases) on MPI in the long-delay task (Fig. 3B, right panel), we also observed a significant main effect of the stimulus orientation ($F_{(1,15)} = 4.80$, $p = 0.045$, $\eta_p^2 = 0.242$), suggesting that training on the short-delay task improved discrimination performance in the long-delay task. None of the other effects reached significance level ($ps > 0.80$).

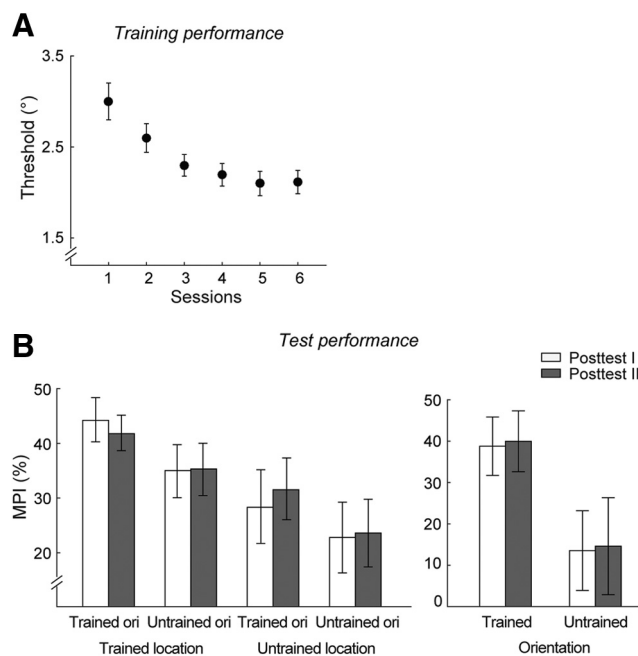


Figure 3. Behavioral results of Experiment 1. **A**, Participants' discrimination threshold decreased significantly over training sessions. **B**, Participants' performance during the posttest phases. Left, MPI in the short-delay task during the posttest phase I and phase II. MPI showed learning specificity for the trained compared with the untrained orientation presented at the trained versus untrained location in both posttest phases. Orientation was abbreviated as *ori* for the condition labels. Right, MPI in the long-delay task during the posttest phase I and phase II. MPI was significantly higher for the trained versus untrained orientation in both posttest phases. Error bars represent SEM across participants.

Perceptual learning does not change BOLD amplitude in V1 and IPS

To examine whether perceptual learning changes the response amplitude in V1 and IPS during the delay period, we used event-related analysis that compared BOLD response between the trained and untrained orientations. Figure 4 showed an example of the averaged temporal dynamics of BOLD responses in V1 and IPS from posttest I. A two-way repeated-measures ANOVA (2 stimulus orientations × 3 test phases) revealed no significant effects on the delay activity in either V1 ($ps > 0.31$) or IPS ($ps > 0.12$). Our results suggest that training did not alter overall BOLD response during the WM delay in sensory and higher-order areas.

Perceptual learning alters feature-specific WM representation in V1 and IPS

We next examined whether the feature-specific information was contained in the distributed pattern activity during the delay period and how training modulated such representation. Using MVPA (Kamitani and Tong, 2006), we decoded the stimulus orientation in V1 during the delay period, separately for each test phase (Fig. 5, left). Repeated-measures ANOVAs on the classification accuracy in V1 revealed a main effect of test phases (pretest, posttest I, and posttest II) in the contralateral ($F_{(2,30)} = 5.72$, $p = 0.008$, $\eta_p^2 = 0.276$) and the ipsilateral ($F_{(2,30)} = 5.63$, $p = 0.008$, $\eta_p^2 = 0.273$) V1. The decoding accuracies were above chance, as determined by the permutation tests (see Materials and Methods), in both V1 ROIs before training (contralateral V1: $p < 0.001$; ipsilateral V1: $p = 0.002$; FDR-corrected). In contrast, classification accuracy dropped to chance level after training for both contralateral and ipsilateral V1 ($ps > 0.788$, FDR-corrected). These results suggest that training alters the WM representation in V1. One may argue that participants learned to use

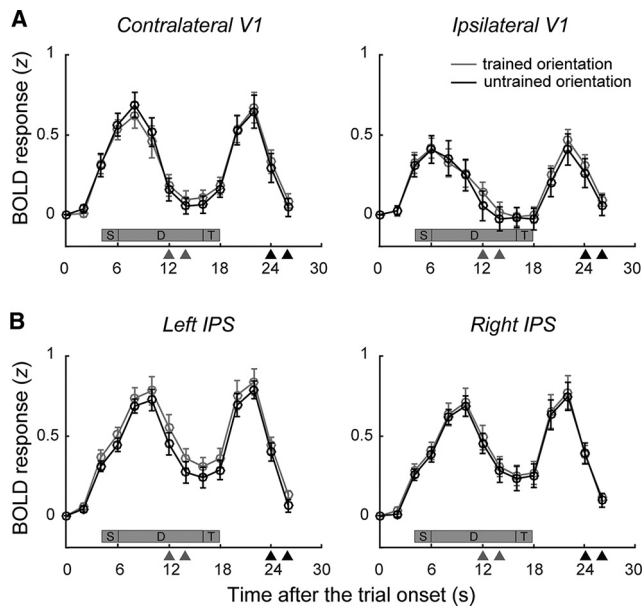


Figure 4. Time course of z-transformed BOLD activity in V1 and IPS in posttest I. **A**, Time course of BOLD activity in contralateral and ipsilateral V1. **B**, Time course of BOLD activity in left and right IPS. The gray and black triangles indicate time points selected for all subsequent analyses of the delay period and the ITI, respectively. Event labels above the x-axis are shown at the corresponding time points to represent the sample (S), delay (D), and test (T) periods, respectively. Each subplot shows the time course of BOLD response for the trained and untrained orientations. Error bars represent SEM across participants.

some verbal strategies over training that affected V1 decoding. However, the use of a verbal strategy seems unlikely to perform our fine discrimination task, where participants discriminated between small angular differences (3° – 5°) with a trial-wise jitter of orientation ($\pm 5^{\circ}$).

To ensure that selected time window for MVPA reflected WM retention and was not spuriously induced by residual effect of sensory processing, we applied the same MVPA analysis to the data from ITI, a period following the test stimulus (i.e., 8–10 s after the onset of the test stimulus). This period presumably contained comparable sensory information to that during the WM delay following the sample stimulus (i.e., 8–10 s after the onset of the sample stimulus) but without the demand of WM maintenance. This control analysis on ITI revealed chance level decoding performance in all of the test phases and V1 subregions ($ps > 0.524$) and no significant difference across test phases (repeated-measures ANOVAs: contralateral: $F_{(2,30)} = 0.02$, $p = 0.984$, $\eta_p^2 = 0.001$; ipsilateral: $F_{(2,30)} = 0.56$, $p = 0.578$, $\eta_p^2 = 0.036$). To more directly compare the classification accuracy between these two time periods (delay vs ITI), we performed a three-way repeated-measures ANOVA (2 V1 ROIs \times 2 time periods \times 3 test phases), which revealed no three-factor interaction ($F_{(2,30)} = 0.37$, $p = 0.695$, $\eta_p^2 = 0.024$), but a significant time period \times test phase interaction ($F_{(2,30)} = 8.19$, $p = 0.001$, $\eta_p^2 = 0.353$). None of the other effects reached significance ($ps > 0.056$). These results confirm that the learning-dependent changes were specific to the delay period, thus guarding against the possibility that V1 decoding of feature during WM delay was because of residual sensory processing.

To address whether high-order areas related to WM processing contained feature-specific information and how training influenced such representation, we trained the classifier to distinguish between two orientations in left and right IPS (Fig. 5, right). Repeated-measures ANOVAs revealed a main effect of

test phase (pretest, posttest I, and posttest II) in left IPS ($F_{(2,30)} = 3.56$, $p = 0.041$, $\eta_p^2 = 0.192$), but not in right IPS ($F_{(2,30)} = 0.39$, $p = 0.684$, $\eta_p^2 = 0.025$). In particular, the permutation tests revealed above-chance accuracies in left IPS after training (posttest I: $p = 0.013$; posttest II: $p = 0.001$; FDR-corrected), but not before training (pretest: $p = 0.788$, FDR-corrected). We did not observe reliable decoding in any of the test phases for right IPS ($ps > 0.165$, FDR-corrected). These findings suggest that training rendered WM information to be encoded in left IPS.

Given that multivariate fMRI results demonstrate distinct profiles of learning-dependent changes in classification performance between V1 and left IPS, we directly compared the change of decoding performance between these two brain areas. We first collapsed the results in two V1 ROIs (contralateral and ipsilateral) to obtain a single estimate, because a two-way repeated-measures ANOVA (2 V1 ROIs \times 3 test phases) on classification accuracies revealed no main effect of V1 ROIs ($F_{(1,15)} = 0.12$, $p = 0.731$, $\eta_p^2 = 0.008$) nor two-factor interaction ($F_{(2,30)} = 0.43$, $p = 0.653$, $\eta_p^2 = 0.028$). Next, to avoid potential issues relating to cross-region comparisons, we calculated the difference of classification accuracy between the pretest and posttests (i.e., posttest I – pretest, posttest II – pretest), separately for V1 and left IPS. A two-way repeated-measures ANOVA (2 ROIs \times 2 posttest phases) on the difference of classification accuracy only revealed a significant main effect of ROI ($F_{(1,15)} = 32.14$, $p < 0.001$, $\eta_p^2 = 0.682$), supporting differential effects of training on WM representation in sensory and higher-order parietal areas.

Experiment 2: TMS

Replication of the learning effect on behavior in Experiment 1

Participants were trained on a short-delay orientation discrimination task and showed decreased threshold over the training sessions (first session: mean = 2.99° , SD = 0.93° ; last session: mean = 2.22° , SD = 0.51° ; paired t test: $t_{(19)} = 5.15$, $p < 0.001$, Cohen's $d = 1.152$). Similar learning effect was also observed in the long-delay task (pretest: mean = 3.46° , SD = 1.11° ; posttest: mean = 2.53° , SD = 0.55° ; paired t test: $t_{(19)} = 3.62$, $p = 0.002$, Cohen's $d = 0.809$). These results replicated the behavioral effects observed in Experiment 1, showing that training enhanced discrimination performance in the long-delay task.

The causal role of V1 and IPS in mnemonic processing over training

While our multivariate analyses in Experiment 1 suggest that perceptual learning changed the engagement of V1 and left IPS during WM retention, we further took advantage of TMS to infer causal relation between neural and behavior and examine how such relation was altered after training.

The results from Experiment 1 showed reliable orientation decoding in V1 during WM delay before, but not after, training. These results provide a seeming account that V1 became unnecessary for WM after training. We thus disrupted V1 activity during WM delay and compared the change of performance to the sham condition (Fig. 6A). A two-way repeated-measures ANOVA (2 test phases: pretest vs posttest \times 2 stimulation conditions: V1 vs sham) on discrimination accuracy revealed a significant main effect of the stimulation condition ($F_{(1,19)} = 11.05$, $p = 0.004$, $\eta_p^2 = 0.368$). None of the other effects reached significance ($ps > 0.263$). Further, TMS over V1 impaired performance both before ($t_{(19)} = -1.76$, $p = 0.047$, Cohen's $d = -0.395$, one-tailed) and after training ($t_{(19)} = -3.83$, $p = 0.001$, Cohen's $d = -0.857$, one-tailed). These results confirmed the validity of

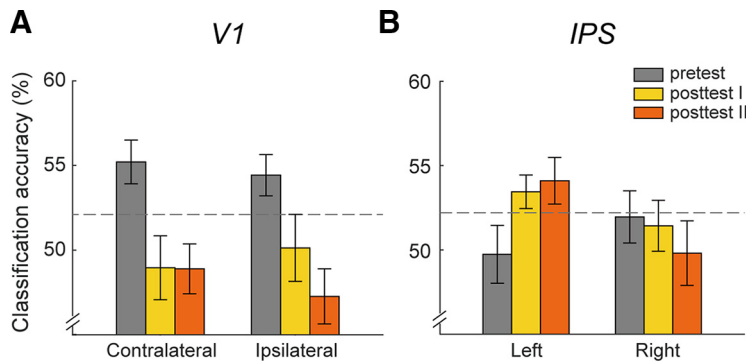


Figure 5. MVPA results of Experiment 1. **A**, Orientation decoding in contralateral and ipsilateral V1 across three test phases. Classification accuracies were above chance only for the pretest phase in both V1 ROIs. **B**, Orientation decoding in left and right IPS across three test phases. Classification accuracies were above chance only for the posttest phases in left IPS. The dashed lines denote maximal significance threshold across all conditions obtained from the permutation tests. Note that the significance thresholds were shown only for the purpose of visualization. The reported p values were obtained from the permutation tests for each ROI and each condition, and corrected for multiple comparisons using FDR method. Error bars represent SEM across participants.

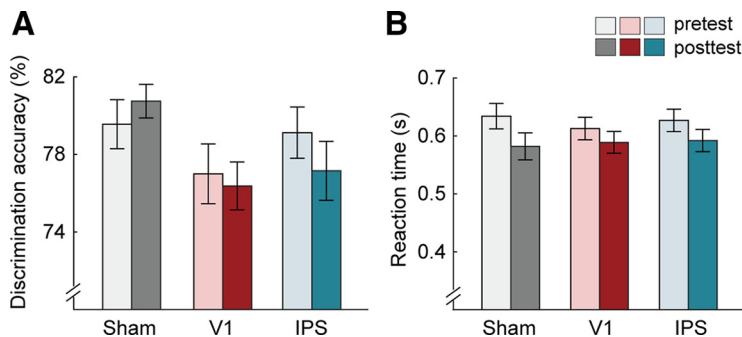


Figure 6. TMS results of Experiment 2. **A**, Discrimination accuracy at the pretest and posttest phases. From left to right, each pair of bars corresponds to behavioral performance before and after training across different TMS conditions (sham, V1, and IPS). Discrimination accuracy was significantly lower for V1 stimulation versus sham condition in both pretest and posttest phases. Discrimination accuracy was significantly lower for IPS stimulation versus sham condition only in the posttest phase. **B**, RT at the pretest and posttest phases. RTs were significantly shorter in the posttest than in the pretest phases across TMS conditions. Error bars represent SEM across participants.

TMS-specific effects on V1 and further suggest the necessary role of V1 for WM retention both before and after training.

On the contrary, the fMRI results showed reliable orientation decoding in IPS during WM delay after, but not before, training. These results predicted a learning-dependent involvement of IPS for WM maintenance. With this rationale, we disrupted IPS activity during WM delay and compared the change of performance to that in the sham condition (Fig. 6A). A two-way repeated-measures ANOVA (2 test phases: pretest vs posttest \times 2 stimulation conditions: IPS vs sham) on discrimination accuracy revealed a main effect of stimulation condition ($F_{(1,19)} = 8.38$, $p = 0.009$, $\eta_p^2 = 0.306$). Importantly, we observed a two-factor interaction effect ($F_{(1,19)} = 4.56$, $p = 0.046$, $\eta_p^2 = 0.193$). Simple effect analyses showed that stimulating IPS significantly impaired performance at the posttest phase (IPS vs sham: $t_{(19)} = -3.29$, $p = 0.002$, Cohen's $d = -0.735$, one-tailed), but not at the pretest phase (IPS vs sham: $t_{(19)} = -0.47$, $p = 0.322$, Cohen's $d = -0.105$, one-tailed). These results converged with fMRI findings and jointly suggest that perceptual learning enhanced the involvement of IPS in WM maintenance.

To parallel the cross-region comparison of the training effect on fMRI decoding, we performed a two-way repeated-measures ANOVA (2 stimulation sites: V1 vs IPS \times 2 test phases: pretest vs posttest) on the sham-normalized discrimination accuracy (see

Materials and Methods). This analysis revealed neither a main effect of stimulation sites ($F_{(1,19)} = 1.29$, $p = 0.270$, $\eta_p^2 = 0.064$) nor significant interaction ($F_{(1,19)} = 1.04$, $p = 0.321$, $\eta_p^2 = 0.052$). Although potential differences in TMS efficacy between sensory and higher-order areas may be responsible for the lack of cross-region difference, we cannot conclude that neurodisruptive effects between V1 and IPS were different. However, we should note that this result does not affect the observed learning-dependent changes of mnemonic processing in these two regions, separately.

To rule out the alternative possibility that the disruptive effect of TMS on the accuracy reflected the speed-accuracy trade-off, we performed the same ANOVA tests on RT (Fig. 6B). The results showed main effects of test phases (2 test phases: pretest vs posttest \times 2 stimulation conditions: V1 vs sham: $F_{(1,19)} = 18.11$, $p < 0.001$, $\eta_p^2 = 0.488$; 2 test phases: pretest vs posttest \times 2 stimulation conditions: IPS vs sham: $F_{(1,19)} = 14.95$, $p = 0.001$, $\eta_p^2 = 0.440$), suggesting an effect because of practice. Importantly, there were no significant interaction effects in these analyses ($ps > 0.150$), eliminating the account of speed-accuracy trade-off.

Discussion

In the present study, we provide evidence that training alters mnemonic representation of simple visual features (i.e., orientation) in a discrimination task. We focused on V1 and IPS that have been associated with WM retention for visual features (Harrison and Tong, 2009; Serences et al., 2009; Bettencourt and Xu, 2016; Weber et al., 2016). In particular, combining fMRI decoding and TMS techniques, we found orientation-specific information during WM delay that was decodable in V1 before, but not after, training; whereas the V1 stimulation led to decreased behavior performance both before and after training. In contrast, both fMRI decoding and TMS results showed that IPS represented WM content after, but not before, training. These findings thus point to learning-related changes in mnemonic representation of visual features at different cortical levels, complementing prior studies that mainly addressed learning-dependent alterations in sensory and decision-making processes.

Previous neurophysiological and neuroimaging studies have shown training-induced changes in visual cortex that presumably occurred at an early stage of encoding processes (Schoups et al., 2001; Schwartz et al., 2002; Furmanski et al., 2004; Yang and Maunsell, 2004; Yotsumoto et al., 2008; Jehee et al., 2012; Yan et al., 2014; Chen et al., 2015). Here, we examined the contribution of visual cortex to mnemonic representation. Converging pre-training results from fMRI decoding and TMS provided direct support for the theoretical hypothesis of “sensory recruitment of WM” (Pasternak and Greenlee, 2005; Harrison and Tong, 2009; Serences et al., 2009). Extending beyond these findings, we further tested the effects of training on the sensory engagement for

WM. Unexpectedly, while the posttraining TMS results suggests a learning-independent mechanism of V1 during WM maintenance, we did not find decodable WM information in V1. Of note, the null result in V1 decoding after training was paired with positive results of V1 decoding before training, using the same sets of voxels and analytical approach. It is thus unlikely that this null effect reflected lack of sensitivity in voxel selection for decoding. Therefore, we speculate that training moulds the coding format in this region that is insensitive to decoding.

Although our observation of learning-independent behavioral impairment after V1 stimulation seems at odds with the inability to decode WM content in V1 specifically after training, such discrepancy may be reconciled by a time-varying attentional modulation on sensory processing given the close relationship between WM and attention (Awh and Jonides, 2001; Gazzaley and Nobre, 2012). Itthipuripat et al. (2017) found a dominance of attentional gain modulation on sensory activity early in training, which was abolished at late phase of training. This reported change of gain modulation (i.e., early vs late phase of training) may correspond to our observation of changes in pattern difference between two features (i.e., before vs after training) in V1. In particular, the lack of gain modulation after extensive training may correspond to the absence of decodable WM content. The inability to decode memorized features, however, does not necessarily mean the absence of information. Instead of a gain mechanism, training was proposed to improve performance via noise reduction (Itthipuripat et al., 2017), to which the MVPA of fMRI data might be insensitive.

An alternative interpretation for the lack of decodable WM information in V1 after training is that training induced synaptic changes in WM storage (Mongillo et al., 2008; Christophel et al., 2017; Masse et al., 2020). Previous studies using computational modeling offered a plausible mechanism of activity-silent short-term retention of WM, where the feature-specific information can be retained in the pattern of synaptic weights even in absence of persistent delay activity (Mongillo et al., 2008; Masse et al., 2020). We thus speculate that such effect of synaptic facilitation might be sufficient for mnemonic processing of features after training without relying on pattern-level differences. Nevertheless, our current study was not sufficient to discriminate between these two possibilities (i.e., noise reduction vs synaptic changes). Further investigations with advanced neurobiological techniques are needed to clarify this issue.

Another governing assumption in WM research has been that the retention of WM information is supported by the delay period activity in the frontoparietal network (D'Esposito and Postle, 2015; Sreenivasan and D'Esposito, 2019). Here, we also showed elevated delay period activity in IPS, which was not influenced by training. Although the pattern activity during the delay period did not contain feature information (i.e., orientation) before training, as also reported in other studies (Linden et al., 2012; Riggall and Postle, 2012), combined evidence from fMRI decoding and TMS suggest that after training, IPS became engaged for feature-specific mnemonic processing. Such learning-dependent changes in IPS may contribute to the formation of a more stabilized WM representation in higher-order areas. Combined with the findings in V1, we suggest that discrimination training may alter feature-specific WM coding at different cortical levels for complementary roles (Bettencourt and Xu, 2016; Lorenc et al., 2018). It is worth noting that some previous studies showed decodable WM content in IPS without training (Bettencourt and Xu, 2016; Cai et al., 2019; Rademaker et al., 2019), differing from the lack of IPS decoding before training in

our data. This apparent discrepancy can be ascribed to the different parietal subregions defined across studies: we used the delay-period activity to define IPS, whereas other studies used retinotopic mapping procedure (Rademaker et al., 2019) or tasks measuring WM capacity (Bettencourt and Xu, 2016). This possibility was supported by the study from Bettencourt and Xu (2016), where they found decodable WM content in IPS regions sensitive to WM load, but not for IPS defined by topographic or anatomic features. In addition, the differences in the presented stimulus may also contribute to differences in fMRI decoding: we used a small, peripheral stimulus, in contrast to the large and centrally presented stimulus used in their studies (Bettencourt and Xu, 2016; Rademaker et al., 2019).

We should note that IPS may not be the unique region that shows robust delay period activity, as prefrontal cortex (PFC) also exhibited elevated delay activity that can be selective for different stimuli (Miller et al., 1996; Stokes et al., 2013; Ester et al., 2015). However, many neuroimaging studies using MVPA failed to decode WM content in this region (Christophel et al., 2012; Riggall and Postle, 2012; Lee et al., 2013; Ester et al., 2015; LaRocque et al., 2017). The absence of decodable WM content may reflect the limitations of applying MVPA techniques to PFC, as suggested by a meta-analysis (Bhandari et al., 2018) and other studies showing stimulus-specific response in PFC by means of the inverted encoding models (Ester et al., 2015). Another possibility is that PFC activity primarily represent higher order information, such as task rules, abstract representations of categories (Lee et al., 2013; Ester et al., 2015). Future studies may use other analytical approaches and neurophysiological methods to investigate learning-dependent modification in the WM representation in PFC.

Previous studies have produced mixed results regarding the involvement of V1 and IPS during WM delay, especially between human and non-human primates. Here, we point to the role of training that may explain such discrepancy. While human only need a few trials to familiarize the task, monkeys have to go through extensive training before recording their neural activity, analogous to the measurement of posttraining performance in human learning studies. In this regard, the decodable WM content found in human V1 (Harrison and Tong, 2009; Serences et al., 2009) versus weak (or absent) WM information reported in sensory areas from neurophysiological studies (Zaksas and Pasternak, 2006; Mendoza-Halliday et al., 2014) may reflect the key difference in training. Similarly, the decodable WM information contained in elevated delay period activity from neurophysiological studies (Baeg et al., 2003; Averbach and Lee, 2007) versus the mixed results of WM decoding in parietal cortex from human neuroimaging studies (Linden et al., 2012; Riggall and Postle, 2012) may also relate to the difference in training. Thus, we provide a plausible account that can reconcile the discrepant findings across species. That is, the neural locus of WM representation depends critically on training.

Interestingly, a previous study that used feature discrimination task showed a lack of TMS effect when stimulating IPS both before and after training (Chang et al., 2014). However, in that study, two features were concurrently, rather than sequentially, shown to the participants. This design did not require short-term memory retention of a particular item and thus involvement of IPS was not necessary for performing the task. The contrast between our results and the findings in Chang et al. (2014) supported the critical dependence on parietal cortex to afford stable WM representation after training, specifically when the task included mnemonic processing of the stimuli (Postle, 2006; Xu, 2017).

In summary, discrimination training influenced the mnemonic processing of visual features in sensory and higher-order areas. Although the sensory engagement for WM is relatively independent of training, training may alter the coding format of WM content in this region. In contrast, the recruitment of higher-order parietal areas for WM representation depends on training, potentially contributing to a more stabilized representation along with improved discriminability. The feature-specific WM representation at different cortical areas may serve complementary roles to support learning-related brain plasticity throughout multiple cortical hierarchy (Watanabe and Sasaki, 2015; Doshier and Lu, 2017; Maniglia and Seitz, 2018).

References

- Albers AM, Kok P, Toni I, Dijkerman HC, De Lange FP (2013) Shared representations for working memory and mental imagery in early visual cortex. *Curr Biol* 23:1427–1431.
- Allefeld C, Gørgen K, Haynes JD (2016) Valid population inference for information-based imaging: from the second-level t-test to prevalence inference. *Neuroimage* 141:378–392.
- Averbeck BB, Lee D (2007) Prefrontal neural correlates of memory for sequences. *J Neurosci* 27:2204–2211.
- Awh E, Jonides J (2001) Overlapping mechanisms of attention and spatial working memory. *Trends Cogn Sci* 5:119–126.
- Baeg EH, Kim YB, Huh K, Mook-Jung I, Kim HT, Jung MW (2003) Dynamics of population code for working memory in the prefrontal cortex. *Neuron* 40:177–188.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 57:289–300.
- Bettencourt KC, Xu Y (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat Neurosci* 19:150–157.
- Bhandari A, Gagne C, Badre D (2018) Just above chance: is it harder to decode information from prefrontal cortex hemodynamic activity patterns? *J Cogn Neurosci* 30:1473–1498.
- Brady TF, Konkle T, Gill J, Oliva A, Alvarez GA (2013) Visual long-term memory has the same limit on fidelity as visual working memory. *Psychol Sci* 24:981–990.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Cai Y, Sheldon AD, Yu Q, Postle BR (2019) Overlapping and distinct contributions of stimulus location and of spatial context to nonspatial visual short-term memory. *J Neurophysiol* 121:1222–1231.
- Capotosto P, Corbetta M, Romani GL, Babiloni C (2012) Electrophysiological correlates of stimulus-driven reorienting deficits after interference with right parietal cortex during a spatial attention task: a TMS-EEG study. *J Cogn Neurosci* 24:2363–2371.
- Chang DHF, Mevorach C, Kourtzi Z, Welchman AE (2014) Training transfers the limits on perception from parietal to ventral cortex. *Curr Biol* 24:2445–2450.
- Chen N, Bi T, Zhou T, Li S, Liu Z, Fang F (2015) Sharpened cortical tuning and enhanced cortico-cortical communication contribute to the long-term neural mechanisms of visual motion perceptual learning. *Neuroimage* 115:17–29.
- Chen Y, Namburi P, Elliott LT, Heinze J, Soon CS, Chee MWL, Haynes JD (2011) Cortical surface-based searchlight decoding. *Neuroimage* 56:582–592.
- Christophel TB, Hebart MN, Haynes JD (2012) Decoding the contents of visual short-term memory from human visual and parietal cortex. *J Neurosci* 32:12983–12989.
- Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes JD (2017) The distributed nature of working memory. *Trends Cogn Sci* 21:111–124.
- Cocchi L, Yang Z, Zalesky A, Stelzer J, Hearne LJ, Gollo LL, Mattingley JB (2017) Neural decoding of visual stimuli varies with fluctuations in global network efficiency. *Hum Brain Mapp* 38:3069–3080.
- Cornette L, Dupont P, Bormans G, Mortelmans L, Orban GA (2001) Separate neural correlates for the mnemonic components of successive discrimination and working memory tasks. *Cereb Cortex* 11:59–72.
- D'Esposito M (2007) From cognitive to neural models of working memory. *Philos Trans R Soc B Biol Sci* 362:761–772.
- D'Esposito M, Postle BR (2015) The cognitive neuroscience of working memory. *Annu Rev Psychol* 66:115–142.
- Destrieux C, Fischl B, Dale A, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* 53:1–15.
- Doshier B, Lu ZL (2017) Visual perceptual learning and models. *Annu Rev Vis Sci* 3:343–363.
- Dotson NM, Hoffman SJ, Goodell B, Gray CM (2018) Feature-based visual short-term memory is widely distributed and hierarchically organized. *Neuron* 99:215–226.
- Emrich SM, Riggall AC, Larocque JJ, Postle BR (2013) Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *J Neurosci* 33:6516–6523.
- Engel SA, Wandell BA, Glover GG (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex* 7:181–192.
- Ester EF, Serences JT, Awh E (2009) Spatially global representations in human primary visual cortex during working memory maintenance. *J Neurosci* 29:15258–15265.
- Ester EF, Anderson DE, Serences JT, Awh E (2013) A neural measure of precision in visual working memory. *J Cogn Neurosci* 25:754–761.
- Ester EF, Ho TC, Brown SD, Serences JT (2014) Variability in visual working memory ability limits the efficiency of perceptual decision making. *J Vis* 14:2.
- Ester EF, Sprague TC, Serences JT (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* 87:893–905.
- Furmanski CS, Schluppeck D, Engel SA (2004) Learning strengthens the response of primary visual cortex to simple patterns. *Curr Biol* 14:573–578.
- Gazzaley A, Nobre AC (2012) Top-down modulation: bridging selective attention and working memory. *Trends Cogn Sci* 16:129–135.
- Gilbert CD, Li W, Piech V (2009) Perceptual learning and adult cortical plasticity. *J Physiol* 587:2743–2751.
- Gosseries O, Yu Q, Larocque JJ, Starrett MJ, Rose NS, Cowan N, Postle BR (2018) Parietal-occipital interactions underlying control- and representation-related processes in working memory for nonspatial visual features. *J Neurosci* 38:4357–4366.
- Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458:632–635.
- Henderson M, Serences JT (2019) Human frontoparietal cortex represents behaviorally relevant target status based on abstract object features. *J Neurophysiol* 121:1410–1427.
- Hooks BM, Chen C (2020) Circuitry underlying experience-dependent plasticity in the mouse visual system. *Neuron* 106:21–36.
- Hussar C, Pasternak T (2010) Trial-to-trial variability of the prefrontal neurons reveals the nature of their engagement in a motion discrimination task. *Proc Natl Acad Sci USA* 107:21842–21847.
- Itthipuripat S, Cha K, Byers A, Serences JT (2017) Two different mechanisms support selective attention at different phases of training. *PLoS Biol* 15:e2001724.
- Janssen AM, Oostendorp TF, Stegeman DF (2015) The coil orientation dependency of the electric field induced by TMS for M1 and other brain areas. *J Neuroeng Rehabil* 12:1–13.
- Jehee JFM, Ling S, Swisher JD, Bergen RV, Tong F (2012) Perceptual learning selectively refines orientation representations in early visual cortex. *J Neurosci* 32:16747–16753.
- Kahnt T, Grueschow M, Speck O, Haynes JD (2011) Perceptual learning and decision-making in human medial frontal cortex. *Neuron* 70:549–559.
- Kamitani Y, Tong F (2006) Decoding seen and attended motion directions from activity in the human visual cortex. *Curr Biol* 16:1096–1102.
- Kastner S, Demmer I, Ziemann U (1998) Transient visual field defects induced by transcranial magnetic stimulation over human occipital pole. *Exp Brain Res* 118:19–26.
- Kourtzi Z, DiCarlo JJ (2006) Learning and neural plasticity in visual object recognition. *Curr Opin Neurobiol* 16:152–158.
- Kuai S, Levi D, Kourtzi Z (2013) Learning optimizes decision templates in the human visual cortex. *Curr Biol* 23:1799–1804.
- LaRocque JJ, Riggall AC, Emrich SM, Postle BR (2017) Within-category decoding of information in different attentional states in short-term memory. *Cereb Cortex* 27:4881–4890.

- Law CT, Gold JJ (2008) Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat Neurosci* 11:505–513.
- Lawrence SJ, van Mourik T, Kok P, Koopmans PJ, Norris DG, de Lange FP (2018) Laminar organization of working memory signals in human visual cortex. *Curr Biol* 28:3435–3440.
- Lee SH, Kravitz DJ, Baker CI (2013) Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nat Neurosci* 16:997–999.
- Linden DEJ, Oosterhof NN, Klein C, Downing PE (2012) Mapping brain activation and information during category-specific visual working memory. *J Neurophysiol* 107:628–639.
- Lorenc ES, Sreenivasan KK, Nee DE, Vandenbroucke ARE, D'Esposito M (2018) Flexible coding of visual working memory representations during distraction. *J Neurosci* 38:5267–5276.
- Maniglia M, Seitz AR (2018) Towards a whole brain model of perceptual learning. *Curr Opin Behav Sci* 20:47–55.
- Masse NY, Rosen MC, Freedman DJ (2020) Reevaluating the role of persistent neural activity in short-term memory. *Trends Cogn Sci* 24:242–258.
- Mendoza-Halliday D, Torres S, Martinez-Trujillo JC (2014) Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat Neurosci* 17:1255–1262.
- Mevorach C, Hodsoll J, Allen H, Shalev L, Humphreys G (2010) Ignoring the elephant in the room: a neural circuit to downregulate salience. *J Neurosci* 30:6072–6079.
- Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16:5154–5167.
- Mongillo G, Barak O, Tsodyks M (2008) Synaptic theory of working memory. *Science* 319:1543–1546.
- Morgan HM, Jackson MC, Van Koningsbruggen MG, Shapiro KL, Linden DEJ (2013) Frontal and parietal theta burst TMS impairs working memory for visual-spatial conjunctions. *Brain Stimul* 6:122–129.
- Pasternak T, Greenlee MW (2005) Working memory in primate sensory systems. *Nat Rev Neurosci* 6:97–107.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Postle BR (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139:23–38.
- Qi XL, Constantinidis C (2015) Lower neuronal variability in the monkey dorsolateral prefrontal than posterior parietal cortex. *J Neurophysiol* 114:2194–2203.
- Rademaker RL, Chunharas C, Serences JT (2019) Coexisting representations of sensory and mnemonic information in human visual cortex. *Nat Neurosci* 22:1336–1344.
- Riggall AC, Postle BR (2012) The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J Neurosci* 32:12990–12998.
- Robertson EM, Théoret H, Pascual-Leone A (2003) Studies in cognition: the problems solved and created by transcranial magnetic stimulation. *J Cogn Neurosci* 15:948–960.
- Romei V, Gross J, Thut G (2010) On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *J Neurosci* 30:8692–8697.
- Roth ZN, Heeger DJ, Merriam EP (2018) Stimulus vignetting and orientation selectivity in human visual cortex. *Elife* 7:e37241.
- Sack AT, Kohler A, Bestmann S, Linden DEJ, Dechent P, Goebel R, Baudewig J (2007) Imaging the brain activity changes underlying impaired visuospatial judgments: simultaneous fMRI, TMS, and behavioral studies. *Cereb Cortex* 17:2841–2852.
- Sagi D, Tanne D (1994) Perceptual learning: learning to see. *Curr Opin Neurobiol* 4:195–199.
- Schoups A, Vogels R, Qian N, Orban G (2001) Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412:549–553.
- Schwartz S, Maquet P, Frith C (2002) Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. *Proc Natl Acad Sci USA* 99:17137–17142.
- Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. *Psychol Sci* 20:207–214.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RBH (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–893.
- Song JH, Jiang Y (2006) Visual working memory for simple and complex features: an fMRI study. *Neuroimage* 30:963–972.
- Sreenivasan KK, D'Esposito M (2019) The what, where and how of delay activity. *Nat Rev Neurosci* 20:466–481.
- Stelzer J, Chen Y, Turner R (2013) Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. *Neuroimage* 65:69–82.
- Stewart LM, Walsh V, Rothwell JC (2001) Motor and phosphene thresholds: a transcranial magnetic stimulation correlation study. *Neuropsychologia* 39:415–419.
- Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J (2013) Dynamic coding for cognitive control in prefrontal cortex. *Neuron* 78:364–375.
- Thut G, Schyns PG, Gross J (2011) Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. *Front Psychol* 2:1–10.
- Watanabe T, Sasaki Y (2015) Perceptual learning: toward a comprehensive theory. *Annu Rev Psychol* 66:197–221.
- Weber EM, Peters B, Hahn T, Bledowski C, Fiebach CJ (2016) Superior intraparietal sulcus controls the variability of visual working memory precision. *J Neurosci* 36:5623–5635.
- Xiao LQ, Zhang JY, Wang R, Klein SA, Levi DM, Yu C (2008) Complete transfer of perceptual learning across retinal locations enabled by double training. *Curr Biol* 18:1922–1926.
- Xu Y (2007) The role of the superior intraparietal sulcus in supporting visual short-term memory for multifeature objects. *J Neurosci* 27:11676–11686.
- Xu Y (2017) Reevaluating the sensory account of visual working memory storage. *Trends Cogn Sci* 21:794–815.
- Yan Y, Rasch MJ, Chen M, Xiang X, Huang M, Wu S, Li W (2014) Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nat Neurosci* 17:1380–1389.
- Yang T, Maunsell JHR (2004) The effect of perceptual learning on neuronal responses in monkey visual area V4. *J Neurosci* 24:1617–1626.
- Yotsumoto Y, Watanabe T, Sasaki Y (2008) Different dynamics of performance and brain activation in the time course of perceptual learning. *Neuron* 57:827–833.
- Zaksas D, Pasternak T (2006) Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *J Neurosci* 26:11726–11742.
- Zanto TP, Chadick JZ, Gazzaley A (2014) Anticipatory alpha phase influences visual working memory performance. *Neuroimage* 85:794–802.
- Zhang JY, Zhang GL, Xiao LQ, Klein SA, Levi DM, Yu C (2010) Rule-based learning explains visual perceptual learning and its specificity and transfer. *J Neurosci* 30:12323–12328.
- Zhang YX, Moore DR, Guiraud J, Molloy K, Yan TT, Amitay S (2016) Auditory discrimination learning: role of working memory. *PLoS One* 11:e0147320.
- Zokaei N, Manohar S, Husain M, Feredoes E (2014) Causal evidence for a privileged working memory state in early visual cortex. *J Neurosci* 34:158–162.