ORIGINAL ARTICLE

Dissociable functional responses along the posterior-anterior gradient of the frontal and parietal cortices revealed by parametric working memory and training

Kaiqiang Su[1] · Ziyi Huang[1] · Qianwen Li[2] · Mingxia Fan[2] · Ting Li[3] · Dazhi Yin[1,3,4]

Abstract

While the storage capacity is limited, accumulating studies have indicated that working memory (WM) can be improved by cognitive training. However, understanding how exactly the brain copes with limited WM capacity and how cognitive training optimizes the brain remains inconclusive. Given the hierarchical functional organization of WM, we hypothesized that the activation profiles along the posterior-anterior gradient of the frontal and parietal cortices characterize WM load and training effects. To test this hypothesis, we recruited 51 healthy volunteers and adopted a parametric WM paradigm and training method. In contrast to exclusively strengthening the activation of posterior areas, a broader range of activation concurrently occurred in the anterior areas to cope with increased memory load for all subjects at baseline. Moreover, there was an imbalance in the responses of the posterior and anterior areas to the same increment of 1 item at different load levels. Although a general decrease in activation after adaptive training, the changes in the posterior and anterior areas were distinct at different memory loads. Particularly, we found that the activation gradient between the posterior and anterior areas was significantly increased at load 4-back after adaptive training, and the changes were correlated with improvement in WM performance. Together, our results demonstrate a shift in the predominant role of posterior and anterior areas in the frontal and parietal cortices when approaching WM capacity limits. Additionally, the training-induced performance improvement likely benefits from the elevated neural efficiency reflected in the increased activation gradient between the posterior and anterior areas.

Keywords Activation gradient · Cognitive training · Parametric working memory · Posterior areas · Prefrontal cortex · Posterior-anterior gradient framework

Introduction

Working memory (WM) refers to a system for both information storage and manipulation, which also underpins other higher cognitive functions such as learning (Baddeley 2003; D'Esposito and Postle 2015). However, limited capacity is a core characteristic of WM with holding about 4 items at a time for adults (Cowan 2001; Buschman 2021). Mechanistically, extensive human and nonhuman primate studies have shown that WM is associated with widely distributed frontal and parietal regions, probed with a variety of characterizing approaches, such as regional activation, functional connectivity (FC), network topology, and neural dynamics (Owen et al. 2005; Chai et al. 2018; Sreenivasan and D'Esposito 2019; Eriksson et al. 2015; Braun et al. 2015; Kitzbichler et al. 2011; Zhang et al. 2022; Liang et al. 2016; Xie et al.

✉ Dazhi Yin
  dzyin@psy.ecnu.edu.cn

[1] Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), Affiliated Mental Health Center (ECNU), School of Psychology and Cognitive Science, East China Normal University, 3663 North Zhong-Shan Road, Shanghai 200062, China

[2] Shanghai Key Laboratory of Magnetic Resonance, School of Physics and Electronic Science, East China Normal University, Shanghai 200062, China

[3] Shanghai Changning Mental Health Center, Shanghai 200335, China

[4] Key Laboratory of Philosophy and Social Science of Anhui Province on Adolescent Mental Health and Crisis Intelligence Intervention, Hefei Normal University, Hefei 241002, China

2022; Palva et al. 2010). However, the understanding of neural representations of WM and how the brain copes with limited WM capacity remains inconclusive.

Nevertheless, accumulating evidence has indicated that frontal and parietal regions play distinct roles in WM (McNab and Klingberg 2008; Mackey and Curtis 2017; Katsuki and Constantinidis 2012; Murray et al. 2017), even though concurrent activation is usually observed. To better describe the functional involvement of the frontal and parietal cortices, numerous studies have used parametric load manipulation to provide a "dose–response" curve (Braver et al. 1997; Romo et al. 1999; Jaeggi et al. 2003; Jonides et al. 1997; Van Snellenberg et al. 2015; Buschman et al. 2011; Meyer et al. 2012; Lamichhane et al. 2020). Importantly, the dissociable effect of memory load on frontal and parietal activity has been presented, especially when the capacity limit is approached (Linden et al. 2003). Specifically, a monotonic increase in response to memory load was found in the dorsolateral prefrontal cortex (DLPFC) and the presupplementary motor area, while activity in the frontal eye fields and posterior parietal cortex showed a nonmonotonic response to memory load. Because previous studies have mostly focused on load-dependent responses in single brain regions, little is known about the relationship between spatial location and memory load in the frontal and parietal cortices.

Hierarchy is the organizing principle of the primate cortex in both temporal and spatial domains (Murray et al. 2014; Hasson et al. 2015). An influential theory proposes that the rostral-caudal gradient of the lateral prefrontal cortex (LPFC) is structured depending on the level of cognitive control; that is, the progressively rostral areas are committed to process increasingly abstract information (Nee and D'Esposito 2016; Badre and Nee 2018; Koechlin et al. 2003). Regarding WM, memory storage may occur at multiple levels of cortical hierarchy ranging from sensory to parietal and prefrontal cortex (PFC), reflecting the abstractness of information representations (e.g., visual features vs. categories) (Cohen et al. 1997; Lee et al. 2013; Nee and Brown 2012). Moreover, a consensus has been reached that the PFC can exert top-down control on WM storage in the parietal cortex to cope with limited memory capacity (D'Esposito and Postle 2015; Zanto et al. 2011; Freedman and Ibos 2018; Edin et al. 2009; Buschman et al. 2011). Given the hierarchical representations for both memory storage and cognitive control, a posterior-frontal gradient of functional brain organization has been raised in WM (Christophel et al. 2017). In addition, PFC plays a crucial role in organization of WM contents such as grouping into higher-level chunks, which enables decreasing WM demand (Bor et al. 2003, 2004). Although it does not directly manipulate abstractness of memory contents, load can manipulate the

cognitive control on WM storage and the organization of WM contents. Thus, WM load manipulations may potentially reveal adaptive changes in functional response along the posterior-anterior gradient of the frontal and parietal cortices, especially when approaching capacity limits.

Although the storage capacity of WM is limited, many studies have indicated that it can be improved by adaptive and extended training (Klingberg 2010; Constantinidis and Klingberg 2016; Soveri et al. 2017; Salmi et al. 2018). Cognitive training also enables us to further uncover the causal relationship between the brain and behavior. Accordingly, numerous studies have found a broad decline in activation in the frontal and parietal regions after cognitive training, reflecting an enhancement of neural efficiency (Schneiders et al. 2011; Miro-Padilla et al. 2019; Brehmer et al. 2011; Vermeij et al. 2017; Thompson et al. 2016). However, a few studies have also reported an increase in activation after training (Blacker et al. 2017; Olesen et al. 2004). Nevertheless, the mechanism of how cognitive training structures functional responses along the posterior-anterior gradient of the frontal and parietal cortices to enhance neural efficiency is still unclear.

In this study, we combined functional magnetic resonance imaging (fMRI), a parametric experimental paradigm, and cognitive training to systematically investigate activation profiles along the posterior-anterior gradient of the frontal and parietal cortices in human WM as well as the plastic changes induced by cognitive training. Specifically, we first explored load-dependent brain regions using a linear mixed-effects (LME) model. Then, the interaction analysis between spatial location (posterior, anterior) and WM load was performed for the resulting frontal and parietal regions. According to previous studies (Linden et al. 2003; Christophel et al. 2017), we defined the posterior and anterior areas of the frontal and parietal cortices with a cutoff by the Brodmann area 6 (frontal eye fields/dorsal premotor cortex), which may reconcile different WM modalities (i.e., visual, tactile, and auditory) (Wang et al. 2015; Vergara et al. 2016).

Subsequently, we investigated the effects of WM training on the functional response along the posterior-anterior gradient of the frontal and parietal cortices. By defining a summarized measure, we finally examined the effects of WM training on the activation gradient (i.e., activation difference between the posterior and anterior regions of the frontal and parietal cortices). We hypothesized that the posterior and anterior areas of the frontal and parietal cortices would show distinct responsive patterns when approaching WM capacity limits. Moreover, cognitive training may lead to an imbalanced modulation on the posterior and anterior areas of the frontal and parietal cortices to achieve enhanced neural efficiency.

Materials and methods

Participants

Fifty-nine participants were recruited from East China Normal University. Two participants were unable to complete the experiment due to technical issues with the magnetic resonance imaging (MRI) scanner. Additionally, data from six participants were excluded from the functional neuroimaging analysis due to excessive head motion. Therefore, a total of 51 participants (31 females; mean age $= 21.86 \pm 2.23$ years) were included in the final analysis. All participants were right-handed, had normal or corrected-to-normal vision, and had no history of neurological or psychiatric diseases or psychotropic medication use. The study protocol was approved by the ethics committee of East China Normal University, and each participant signed written consent prior to the experiment.

Experimental design

This study employed a randomized controlled longitudinal experimental design. All participants were to complete two MRI sessions before and after cognitive training, and each session included both resting-state and task-state fMRI. To include an active control group, participants were randomly assigned to HLT (high-load training) and LLT (low-load training) groups after the first MRI session, with the former performing adaptive n-back training and the latter performing only 1-back training (Finc et al. 2020). Each group was required to complete six training sessions as evenly as possible within two weeks, and each session consisted of 20 blocks and lasted approximately 25~30 min. The sixth training session was scheduled within 24 to 48 h prior to the second MRI session (Fig. 1A).

During the fMRI scanning task, a visuospatial n-back task was presented, in which a square appeared in sequence at 8 different spatial locations on a 3×3 grid, accompanied by a white fixation cross positioned in the center of a black screen. Each location stimulus appeared for 500 ms, and the interstimulus interval was 2500 ms (Fig. 1B). Participants were asked to determine whether the current location stimulus matched the position presented n items back, pressing the right index finger key for a match and providing no response for a mismatch. To explore how the brain responds to capacity limits, in this study, we used four different WM loads that included relatively high levels (i.e., 0-back, 2-back, 3-back, and 4-back) (Meyer et al. 2012; Linden et al. 2003). Following a block design, the stimuli were presented to participants in blocks lasting 21 s (7 trials) for the 0-back condition and 60 s (20 trials) for the conditions from 2-back to 4-back. Each block was preceded by

a 5-second fixation period and a 4-second task instruction. The order of blocks from 2-back to 4-back was randomized, with each block appearing twice and always followed and preceded by the 0-back block. There were a total of 6 blocks for the 0-back condition and 2 blocks for each of the 2-back, 3-back, and 4-back conditions. The number of target stimuli in each block accounted for 30% of all stimuli in that block. Participants were given a response window of 2500 ms and were asked to respond as accurately and quickly as possible (Jaeggi et al. 2007). The presentation of experimental stimuli and recordings of responses were established using E-prime 3.0 software.

During the training sessions, the HLT group executed the adaptive visuospatial n-back task, where the load level (n-back level) increased adaptively when they achieved 80% correct responses in the block. Conversely, the n-back level decreased when they made more than 50% errors in the block. The HLT group started training from 1-back, and the n-back level they achieved was recorded for all training sessions. The LLT group only executed the 1-back task. Each participant completed 20 blocks of the n-back task during each training session, with each block consisting of $20 + n$ trials, where the number n depended on the participant's achievement on the n-back level. The training task was presented by a visuospatial n-back task program developed by Jaeggi and colleagues (Jaeggi et al., 2008) (http://brainworkshop.sourceforge.net/). Notably, the inclusion of LLT group allowed us to control the effects of repeated exposure to the task. However, there was still a potential effect that some training tasks (e.g., 2-back, 3-back) were later tested in the HLT group, but no training tasks were tested in the LLT group.
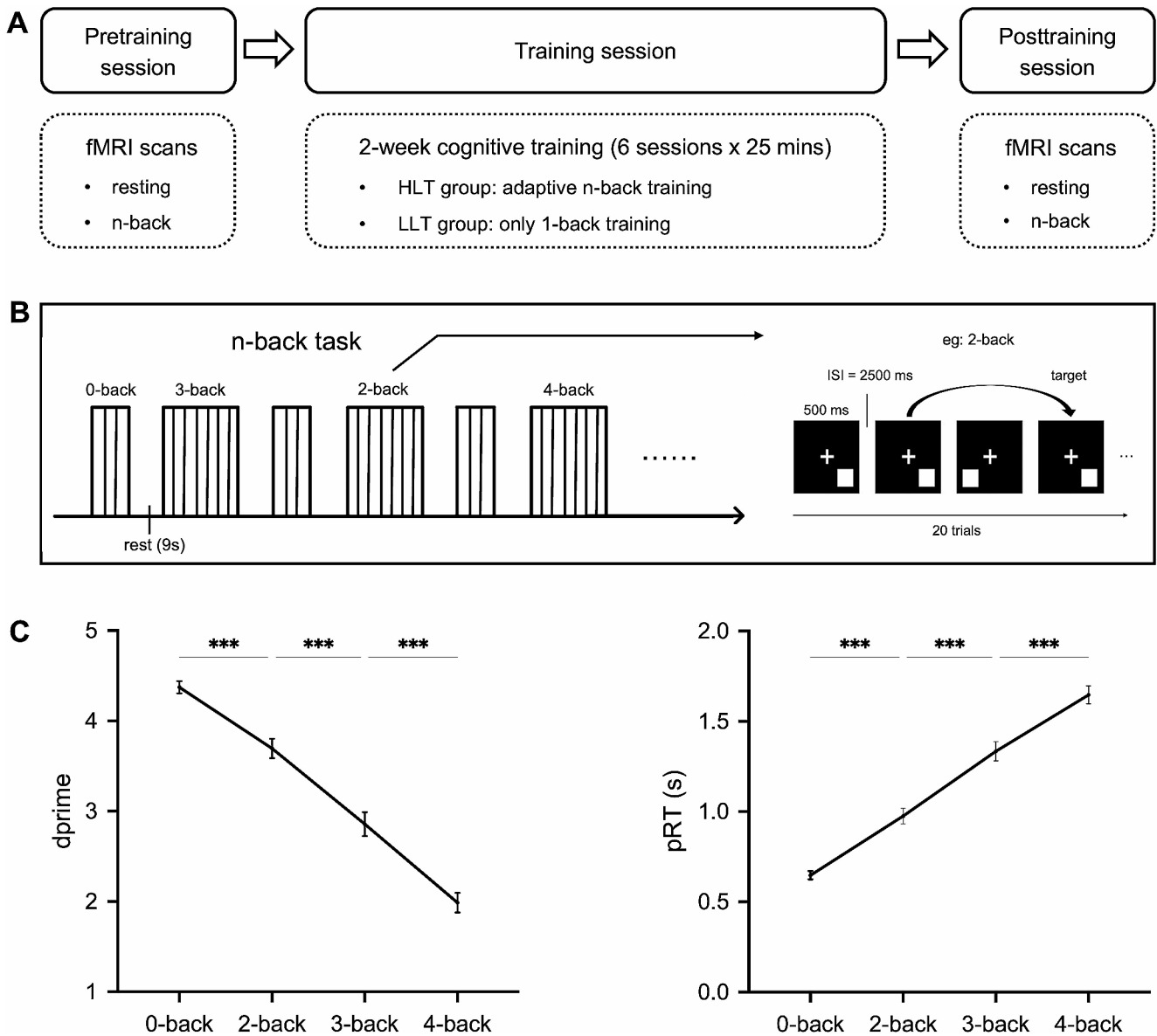
Behavioral performance measurement

To assess behavioral performance on the visuospatial n-back task, we utilized two measurements: dprime, a statistic from the signal detection theory that combines response sensitivity and response bias, and penalized reaction time (pRT), which combines accuracy and reaction time (Finc et al. 2020). For each subject, condition, and session, dprime was defined as:

$$dprime = Z(H) - Z(F) \tag{1}$$

where H represents the hit rate and F represents the false alarm rate. The calculation formulas were as follows:

$$H = \frac{\text{hits}}{\text{hits} + \text{misses}} \tag{2}$$

Fig. 1 Experimental scheme and baseline behavioral performance. (A) A randomized controlled longitudinal design with fMRI scans acquired before and after cognitive training. Participants underwent adaptive *n*-back training (HLT group) or only 1-back training (LLT group). (B) Illustration of a block-designed visuospatial *n*-back task used for fMRI. ISI = interstimulus interval. (C) Changes in behavioral performance with increasing working memory load at baseline. Error bars denote the mean standard error. HLT = high-load training, LLT = low-load training, fMRI = functional magnetic resonance imaging, pRT = penalized reaction time, and ***$p < 0.001$

$$F = \frac{\text{false alarms}}{\text{false alarms} + \text{correct rejections}} \quad (3)$$

where $Z(\text{x})$ represents the inverse of the cumulative Gaussian distribution. To obtain finite values of dprime, we employed adjusted values of 0.01 or 0.99 instead when H or F equals 0 or 1.

The pRT was defined as:

$$pRT = \frac{1}{\text{n}} \sum_{i=1} x_i \quad (4)$$

where n indicates the sum of all correct responses, false alarms, and omitted responses. x_i was obtained using the following formula:

$$x_i = \begin{cases} RT_i, & \textit{if answer was correct} \\ 2500, & \textit{if answer was incorrect} \\ 2500, & \textit{if correct response missed} \end{cases} \quad (5)$$

where RT_i denotes the correct reaction time of the response during the i-th trial, and the value of 2500 serves as a penalty for a false alarm answer or an omitted response. This

value represents the maximum possible response time during each *n*-back trial, measured in milliseconds.

Statistical analysis of behavioral performance

For the baseline, we used one-way repeated analysis of variance (ANOVA) to examine the impact of WM load on behavioral performance. Post hoc analyses were then performed to compare the differences in behavioral performance between each two WM load levels. Subsequently, we utilized an LME model to explore the linear effects of WM load on dprime and pRT (Lamichhane et al. 2020). The equation of the model was as follows:

$$
\begin{aligned}
Behav_{ij} &= B_{0j} + B_{1j}Load_{ij} + \varepsilon_{ij} \\
B_{0j} &= \alpha_{00} + u_{0j} \\
B_{1j} &= \alpha_{10} + u_{1j}
\end{aligned}
\tag{6}
$$

where $Behav_{ij}$ represents the behavioral performance scores for each WM load level i and each Participant j. The parameters α_{00} and α_{10} indicate fixed intercepts and fixed slopes, respectively, while u_{0j} and u_{1j} represent random intercepts and random slopes, respectively.

To investigate the training effects on behavior performance, we conducted a mixed ANOVA with a 2 (group: HLT, LLT) × 2 (session: pretraining, posttraining) × 4 (load: 0-back, 2-back, 3-back, 4-back) design. Post hoc analyses were then performed for the results with a significant interaction effect. The Greenhouse–Geisser correction was applied for violations of sphericity assumptions.

MRI data acquisition

Imaging data were acquired by a 3.0 Tesla MRI scanner (Siemens, Erlangen, Germany) at the Shanghai Key Laboratory of Magnetic Resonance, East China Normal University. High-resolution T1-weighted images were obtained using a magnetization-prepared rapid gradient-echo (MPRAGE) sequence with 192 sagittal slices, slice thickness = 1 mm, no gap, repetition time (TR) = 2,530 ms, echo time (TE) = 2.98 ms, flip angle (FA) = 7°, field of view (FOV) = 256 × 256 mm^2, and voxel size = 1.0 × 1.0 × 1.0 mm^3. Functional MRI scans were collected using a T2*-weighted echo-planar imaging (EPI) sequence with 33 axial slices, thickness = 3.5 mm, gap = 0.7 mm, TR = 2,000 ms, TE = 30 ms, FA = 90°, FOV = 224 × 224 mm^2, voxel size = 3.5 × 3.5 × 3.5 mm^3, and 297 contiguous volumes. The resting-state fMRI scans were also collected but were not used in the current study.

Preprocessing of functional MRI data

The fMRI data were preprocessed using Statistical Parametric Mapping 12 (SPM12; http://www.fil.ion.ucl.ac.uk/spm). After converting the raw DICOM data to the NIFTI format, the following preprocessing steps were conducted (Yin et al. 2017). The images first underwent correction for slice acquisition delay and rigid-body head movement. Excessive head motion was defined as translation > 2.5 mm or rotation > 2.5° in any direction. Subsequently, the corrected images were spatially normalized to the MNI (Montreal Neurological Institute) space using an EPI template and resampled to 3 mm isotropic voxels (Power et al. 2011; Yin et al. 2021). Finally, spatial smoothing was applied using an isotropic Gaussian kernel of 6 mm full-width at half-maximum (FWHM).

Univariate brain activation analysis

We adopted a standard general linear model (GLM) to map task-evoked brain activation. Specifically, we created three task contrasts (i.e., 2 − 0 back, 3 − 0 back, and 4 − 0 back) to obtain activation maps of different WM load levels for all subjects and all sessions. Subsequently, we used one-sample *t* tests to generate group-level activation maps for each WM load level and each session.

Linear mixed-effects model analysis of brain activation

Based on task-evoked activation maps, we further utilized the LME model to explore load-dependent brain regions (Lamichhane et al. 2020). The equation of the model was as follows:

$$
\begin{aligned}
activation_{ij} &= B_{0j} + B_{1j}Load_{ij} + \varepsilon_{ij} \\
B_{0j} &= \alpha_{00} + u_{0j} \\
B_{1j} &= \alpha_{10} + u_{1j}
\end{aligned}
\tag{7}
$$

where $activation_{ij}$ denotes the activation value of each voxel for each WM load level i and each Participant j. The parameters α_{00} and α_{10} indicate a fixed intercept and fixed slope, while u_{0j} and u_{1j} represent a random intercept and random slope. We applied an AlphaSim corrected threshold of $p < 0.05$, with a voxelwise threshold of $p < 0.001$ and a minimum cluster size determined using 1000 Monte Carlo simulations.

Analysis of WM load effects on the activation along the posterior-anterior gradient of the frontal and parietal cortices at baseline.

Previous studies mainly focused on load-dependent brain responses in single regions, with limited consideration for

the variations across different spatial locations. Taking into account the presence of a posterior-frontal gradient of the functional brain organization in WM (Christophel et al. 2017), we defined the posterior and anterior areas of the frontal and parietal cortices with a cutoff by the Brodmann area 6 (frontal eye fields/dorsal premotor cortex). Then, we conducted a two-way ANOVA with the factors of spatial (posterior, anterior) and load (2-back, 3-back, 4-back). To ensure statistical rigor and avoid double dipping (Krieges-korte et al. 2009; Minamoto et al. 2015), we focused on the main effect of spatial location and the interaction effect. In addition, to explore how the posterior and anterior areas of the frontal and parietal cortices separately respond to the increase in memory load, we examined the changes in brain activation with the same increment of 1 WM item (i.e., from 2-back to 3-back and from 3-back to 4-back) for the posterior and anterior areas.

Analysis of WM training effects on the activation along the posterior-anterior gradient of the frontal and parietal cortices.

To investigate the effects of WM training, we conducted a mixed ANOVA on the activation with the factors of group (HLT, LLT), session (pretraining, posttraining), spatial (posterior, anterior), and load (2-back, 3-back, 4-back). The Greenhouse—Geisser correction was applied to violations of the sphericity assumptions. Post hoc analyses were performed using t tests.

Analysis of WM training effects on the activation gradient between the posterior and anterior areas in the frontal and parietal cortices

For defining a summarized measure, we quantified an activation gradient by calculating the activation difference between the posterior and anterior regions in the frontal and parietal cortices. Then, we performed a mixed ANOVA on the activation gradient with the factors of load (2-back, 3-back, 4-back), group (HLT, LLT), and session (pretraining, posttraining). The Greenhouse—Geisser correction was applied to violations of the sphericity assumptions. Post hoc analyses were performed using t tests.

Exploring the influence of baseline activation gradient on WM performance after adaptive training

To further test whether the people who had high activation gradient to begin with achieve higher levels of WM performance after adaptive training, we first divided the HLT group into two subgroups based on the median of baseline activation gradient. Then, two-sample t tests were performed

to examine the differences in posttraining behavioral performance between the two subgroups.

Statistical analysis of baseline differences between the two training groups in behavioral performance, activation, and activation gradient

To verify no baseline differences between the two training groups, we conducted separate group × load ANOVAs for the behavioral performance, mean activation of posterior and anterior areas, and the activation gradient at baseline. If neither the main group effect nor the group × load interaction were significant, it would indicate that the results were not influenced by preexisting group differences.

Correlation analyses

To explore the predictability of individual differences in WM levels, we performed correlation analyses between the baseline behavioral performance and the WM level achieved at the last training session in the HLT group. To investigate the relationship between the brain and behavior, we conducted correlation analyses to assess the relationship between the brain activation and behavioral performance, as well as the correlation between the activation gradient and behavioral performance.

Results

Baseline behavioral performance

The dprime and pRT of all participants showed a decrease and an increase with the increasing memory load, respectively. Through one-way repeated ANOVA, we found significant main effects of load on dprime ($F_{3,141} = 106.20$, $p < 0.001$) and pRT ($F_{3,130} = 153.50$, $p < 0.001$). Post hoc analyses showed significant differences between memory load levels ($p < 0.001$). Through the LME modeling, we further revealed a negative linear effect of WM load on dprime (slope = -0.779, $t(202) = -16.34$, $p < 0.001$) and a positive linear effect of WM load on pRT (slope = 0.336, $t(202) = 18.72$, $p < 0.001$) (Fig. 1C). These behavioral results are consistent with previous findings (Jaeggi et al. 2003, 2007; Miro-Padilla et al. 2019).

Prediction of the level of WM attainment after adaptive training

During adaptive WM training, we observed that the HLT group showed continuous improvement and reached a mean n-back level of 6.92 at the last training session (S6)

(Fig. 2A). Furthermore, we found a significant positive correlation ($r = 0.552$, $p = 0.004$) between the mean dprime across all levels of WM load at baseline and the mean n-back level achieved at the last training session for the HLT group (Fig. 2B). No significant correlations were observed between the pRT and the mean n-back level at the last training session. These findings suggest that individuals with higher baseline WM performance can achieve higher levels of WM following adaptive training.

Changes in behavioral performance after WM training

Through mixed ANOVA, we found that the main effect of session was significant (for dprime: $F_{1,49} = 16.58$; $p < 0.001$; for pRT: $F_{1,49} = 18.90$; $p < 0.001$) and the main effect of group was marginally significant (for dprime: $F_{1,49} = 3.48$; $p = 0.068$; for pRT: $F_{1,49} = 2.89$; $p = 0.096$). Furthermore, the interaction load × group × session was significant for both dprime ($F_{3,147} = 6.11$, $p < 0.001$) and pRT ($F_{3,147} = 6.47$, $p < 0.001$). Post hoc analyses showed a significant increase in dprime, mainly at 3-back ($t(25) = 3.60$, $p = 0.001$) and 4-back ($t(25) = 5.32$, $p < 0.001$) for the HLT group. The results for pRT aligned with those of dprime (Fig. 2C and D). However, the behavioral changes were not obvious for the LLT group (Fig. 2E and F). In addition, no significant main effect of group or the interaction of group × load were observed for either dprime or pRT at baseline (Fig. S1). These findings indicate that adaptive training improves WM performance, mainly at high loads.

Brain activation maps for different levels of WM load

Through group-level analysis, we found similar activation patterns for the three task contrasts (i.e., $2 - 0$ back, $3 - 0$ back, and $4 - 0$ back) with different levels of WM load, mainly involving the bilateral frontal and parietal regions (Fig. S2). This result further supports that WM is associated with distributed frontoparietal regions.

Load-dependent brain regions are distributed across the frontal and parietal cortices

Through LME modeling, we found that load-dependent regions were mainly located in the right lateral frontoparietal cortex, including the right anterior-LPFC, mid-DLPFC, predorsal premotor cortex (pre-PMd), PMd, superior parietal lobule (SPL), inferior parietal lobule (IPL), left PMd and SPL (Fig. 3A; Table 1). By overlap analysis, we found that the load-dependent regions, including the bilateral PMd and posterior parietal regions, overlapped well with

the group-level activation maps for all WM load levels. In contrast, the load-dependent regions, involving the right anterior-LPFC, mid-DLPFC, and pre-PMd, exhibited an increasing overlap with the group-level activation maps of increasing WM load (Fig. 3B). The average activation magnitude of each load-dependent region was shown in Fig. 3C. These findings suggest that WM requires a broader extent of activation in the PFC to cope with increased memory load while exclusively strengthening the activation of PMd and the posterior parietal cortex.

The posterior and anterior areas of the frontal and parietal cortices dissociably respond to the increase in memory load

Considering that there exists a posterior-frontal functional gradient in WM, we divided the load-dependent regions in the frontal and parietal cortices into anterior (including right anterior-LPFC, right mid-DLPFC, and right pre-PMd) and posterior areas (including bilateral PMd, right IPL, right SPL, and left SPL) (Fig. 4A).

Through performing a two-way repeated ANOVA with the factors of spatial (posterior, anterior) and load (2-back, 3-back, 4-back), we found that both the main effects of spatial ($F_{1,50} = 110.30$, $p < 0.001$) and load ($F_{2,98} = 16.30$, $p < 0.001$) and the interaction of spatial × load ($F_{2,100} = 4.66$, $p = 0.012$) were significant at baseline. For post-hoc analysis, we mainly focused the difference values of brain activation with the same increment of 1 WM item (i.e., from 2-back to 3-back and from 3-back to 4-back). Interestingly, we found that the increment of activation in the posterior areas was significantly higher ($t(50) = 2.97$, $p = 0.005$) than that in the anterior areas from 2-back to 3-back. However, the increment of activation in the anterior areas was higher ($t(50) = 1.79$, $p = 0.080$) than that in the posterior areas from 3-back to 4-back (Fig. 4B). These findings suggest a shift in the predominant role of posterior and anterior areas in the frontal and parietal cortices when approaching WM capacity limits.

Changes of activation in the posterior and anterior areas of the frontal and parietal cortices after WM training

Through performing a four-way mixed ANOVA with the factors including group (HLT, LLT), session (pretraining, posttraining), spatial (posterior, anterior), and load (2-back, 3-back, 4-back), we found significant main effects of group ($F_{1,49} = 7.11$, $p = 0.010$), session ($F_{1,49} = 21.57$, $p < 0.001$), load ($F_{2,98} = 19.12$, $p < 0.001$) and spatial ($F_{1,49} = 136.43$, $p < 0.001$). In addition, we found a marginally significant
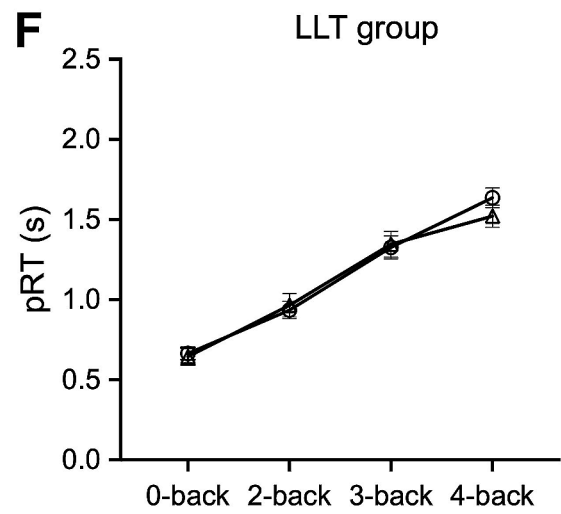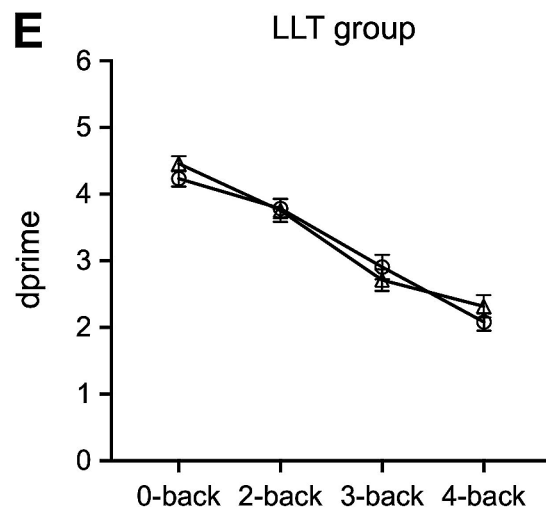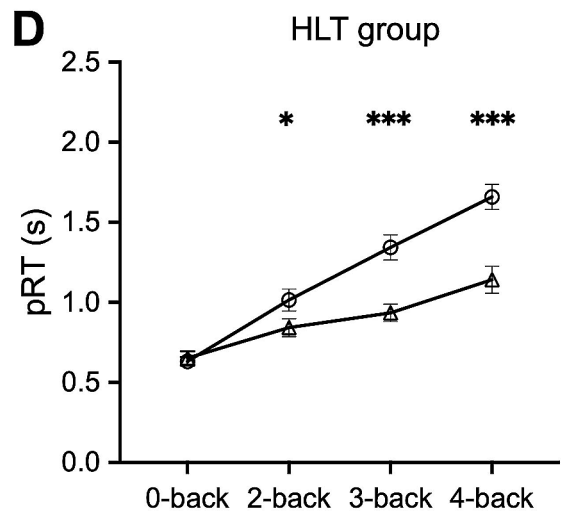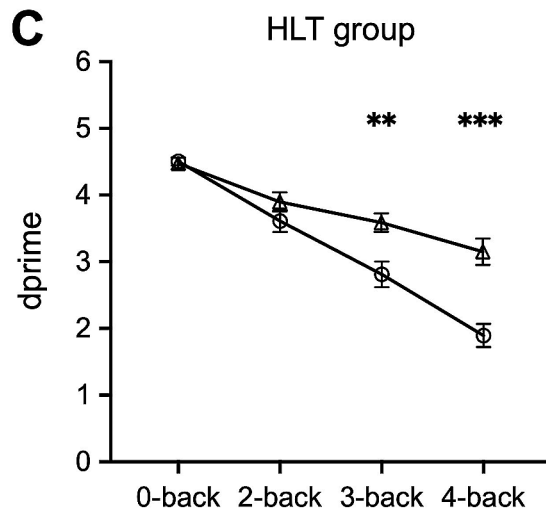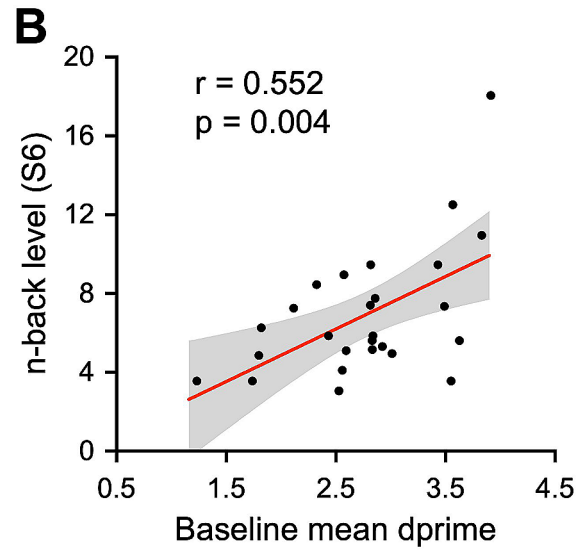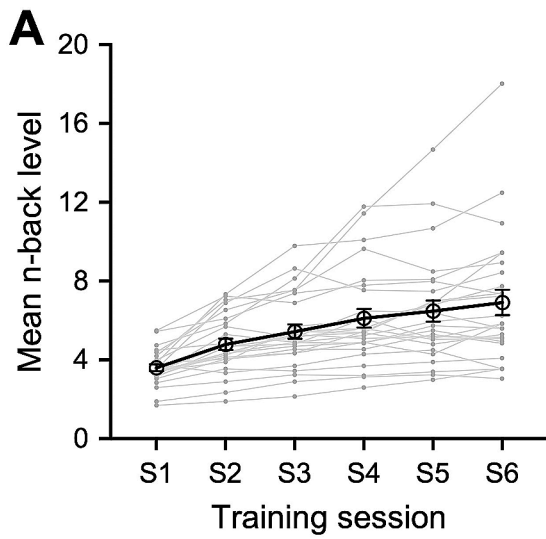
Fig. 2 Effects of WM training on behavioral performance. (**A**) A plot of the changes in *n*-back level across the adaptive training sessions. The gray lines represent the changes in the *n*-back level of individuals, and the black line denotes the changes in the mean *n*-back level. (**B**) The prediction of baseline behavioral performance to the level of WM attainment after adaptive training. (**C-F**) Changes in behavioral performance after WM training in the HLT and LLT groups. Error bars denote the mean standard error. WM = working memory, S = session, HLT = high-load training, LLT = low-load training, *p < 0.05, **p < 0.01, and ***p < 0.001

four-way interaction of group × session × load × spatial ($F_{2,97} = 2.80$, $p = 0.066$) (Table S2).

For post hoc analysis, we employed paired *t* tests to evaluate the training-induced changes of activation in the posterior and anterior areas of the frontal and parietal cortices at each load condition for each group. For the HLT group, no significant changes of activation were observed in either the posterior or anterior areas at load 2-back after training (Fig. 5A). At load 3-back, both the posterior ($t(25) = -4.44$, $p < 0.001$) and anterior ($t(25) = -3.93$, $p < 0.001$) areas exhibited a significant decrease in activation after training (Fig. 5B). At load 4-back, only the anterior areas showed a significant decrease ($t(25) = -4.49$, $p < 0.001$) in activation after training, while no significant changes of activation were observed in the posterior areas (Fig. 5C). For the LLT group, there were no significant changes of activation were found in either the posterior or anterior areas at any load condition after training (Fig. 5D-F). In addition, no significant main effect of group or the interaction of group × load were observed for activation in either the posterior or anterior areas at baseline (Fig. S3). These findings suggest that adaptive training reduces activation mainly at high loads and in the anterior areas.

Training-induced increase in the activation gradient between the posterior and anterior areas in the frontal and parietal cortices

We further defined an activation gradient as the activation difference between the posterior and anterior regions in the frontal and parietal cortices. To investigate the effect of WM training on the activation gradient, a mixed ANOVA with a 2 (session: pretraining, posttraining) × 2 (group: HLT, LLT) × 3 (load: 2-back, 3-back, 4-back) design was performed. We found a marginally significant interaction of group × session × load ($F_{2,98} = 2.80$, $p = 0.065$) (Table S3).

For post hoc analysis, a paired *t* test was employed to evaluate the training-induced changes in activation gradient at each load condition for each group. For the HLT group, no significant changes in activation gradient were observed at loads 2-back and 3-back after training. However, a significant increase in activation gradient was observed at load 4-back ($t(50) = 2.69$, $p = 0.010$) after training (Fig. 6A). In

contrast, the LLT group showed no significant differences in activation gradient at any WM load level (Fig. 6B). In addition, there were no significant main effect of group or the group × load interaction in the ANOVAs for the baseline activation gradient (Fig. S3). These results indicate that adaptive WM training selectively enhances the activation gradient at high WM load.

In the HLT group, we further found a significant negative correlation ($r = -0.494$, $p = 0.010$) between the changes in activation gradient and the changes in pRT at load 4-back (Fig. 6C). However, it was not the case in the LLT group. These findings imply that changes in the activation gradient at high load could predict improvements in behavioral performance after adaptive training. Additionally, no significant differences in posttraining behavioral performance were observed between the two subgroups of HLT defined by baseline activation gradient. This suggests that baseline activation gradient may not play a critical role in the WM performance after adaptive training, although training can optimize posterior-anterior functioning (Fig. S4).

Discussion

To reveal the neural basis of WM and its training effects is a fundamental problem in cognitive neuroscience. In addition to the abundant electrophysiological studies on specific brain areas in animals (Fuster and Alexander 1971; Buschman et al. 2011; Xie et al. 2022), noninvasive neuroimaging evidence from humans consistently showed that WM is related to a wide range of frontal and parietal regions (Eriksson et al. 2015; Owen et al. 2005). Further considering that a posterior-frontal gradient of the functional organization has been suggested in WM (Christophel et al. 2017), in this study, we mainly focused on how the activation profiles along the posterior-anterior gradient of the frontal and parietal cortices responds to increasing memory load as well as the modulation induced by cognitive training. Thus, this work not only enabled us to uncover gradient functional responses of the frontal and parietal cortices to the limited WM capacity but also provided a causal relationship between the brain and behavior.

Through LME modeling, we noticed that the load-dependent regions were mainly located in the right lateral frontoparietal cortex, although WM activated bilateral frontoparietal regions. Early studies indicated that the load-dependent regions involved in bilateral frontoparietal regions for verbal and visual WM tasks (Linden et al. 2003; Braver et al. 1997; Jonides et al. 1997; Jaeggi et al. 2003). However, converging evidence suggests that both spatial WM and spatial selective attention are supported by a right-hemisphere dominant network of frontal and parietal
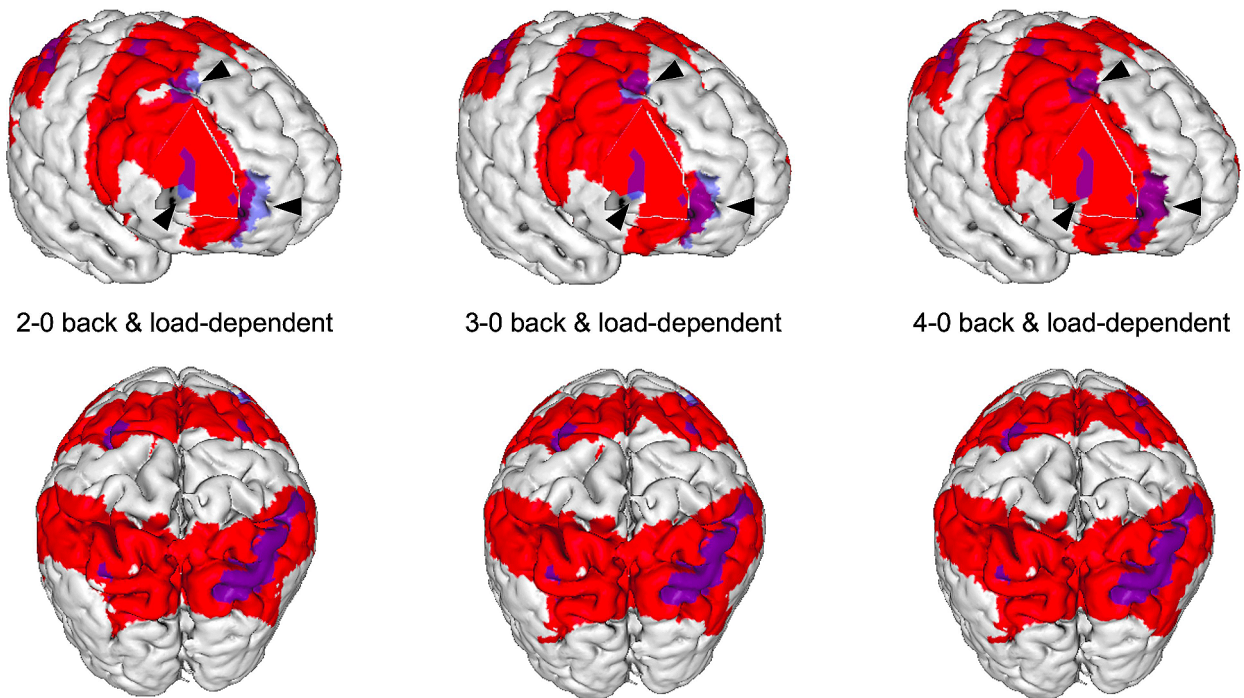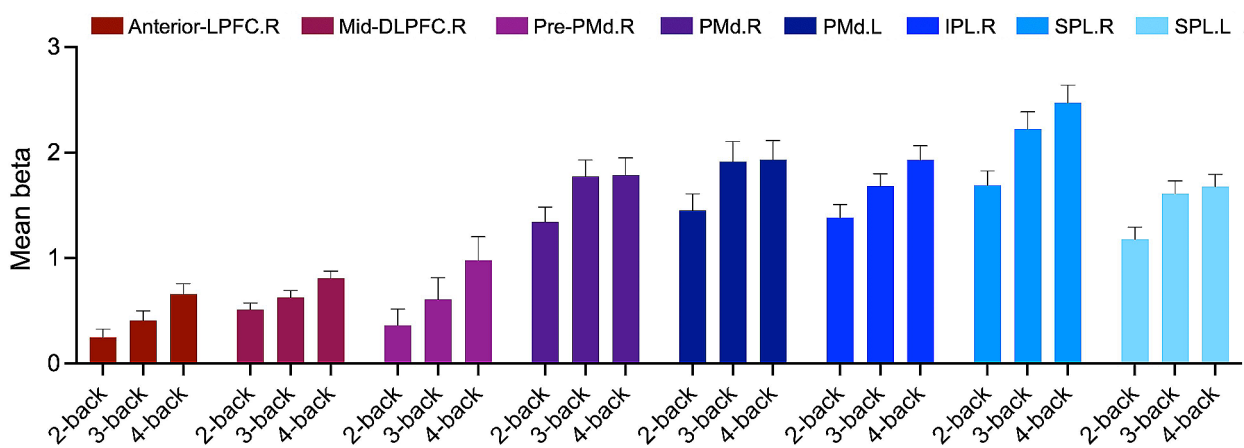
A Load-dependent activations
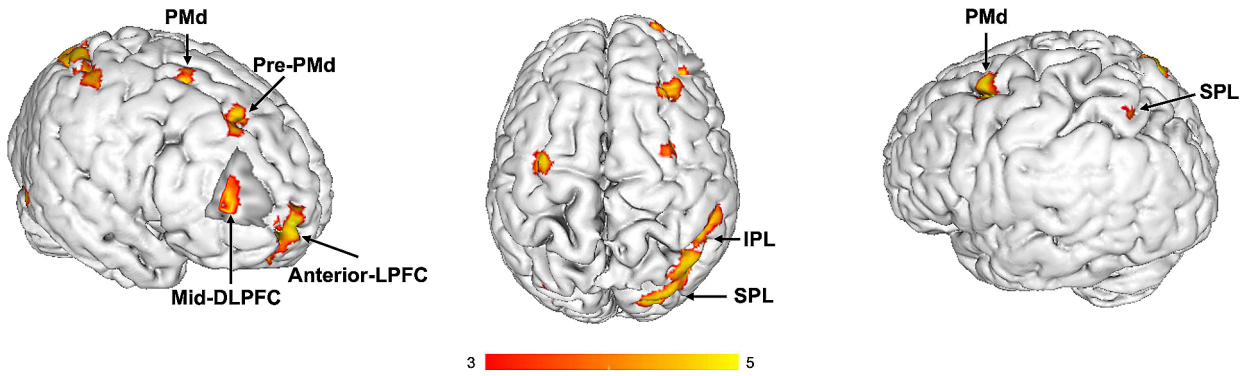
B Overlap of group-level n-back activations and load-dependent activations

2-0 back & load-dependent   3-0 back & load-dependent   4-0 back & load-dependent

C

Fig. 3 Load-dependent brain regions distributed across the frontal and parietal cortices. (**A**) Load-dependent brain regions identified using a linear mixed-effects model. The regions were mainly located in the right frontoparietal cortex. (**B**) An overlap analysis between the group-level activation patterns with different levels of WM load and the load-dependent brain regions. Red refers to group-level activation areas, and blue indicates load-dependent regions. (**C**) The average activation magnitude of each identified load-dependent region at different levels of WM load. Error bars denote the mean standard error. WM=working memory, LPFC=lateral prefrontal cortex, DLPFC=dorsolateral prefrontal cortex, PMd=dorsal premotor cortex, IPL=inferior parietal lobule, and SPL=superior parietal lobule

regions (Awh and Jonides 2001). Previous studies also demonstrated that spatial WM showed stronger activation in the right frontoparietal cortex than nonspatial WM (D'Esposito et al. 1998; Owen et al. 2005). Moreover, executive demand may improve right lateralization in the frontal cortex for spatial WM (Wager and Smith 2003). Thus, we speculate that the right lateralization of load-dependent responses is probably task modality specific.

Interestingly, we found that the posterior and anterior regions of the frontal and parietal cortices showed distinct responsive patterns with the same increment of 1 item. In other words, there was an imbalance in the responses of the posterior and anterior areas to the same increment of 1 item at different load levels. This is likely attributed to the fact that the responses of posterior areas become saturated when approaching the capacity limits. Previous studies suggested that posterior cortex plays a key role in limited capacity of visual short-term memory (Xu and Chun 2006; Todd and Marois 2004). According to Cowan's model (Cowan 1988), this capacity-limited state is explained as the focus of attention, in which there can be about 3 or 4 items (Zimmer 2008; D'Esposito and Postle 2015). Furthermore, Buschman and colleagues found that frontal cortex likely exerted top–down influences on parietal neurons to partially overcome capacity limitation (Buschman et al. 2011). Theoretically, computational models also support the main priority of DLPFC in WM which is to boost parietal memory capacity (Edin et al. 2009; Roggeman et al. 2014). In line with previous findings, our results further suggest a shift in the predominant role of posterior and anterior areas in the frontal and parietal cortices when approaching WM capacity limits.

Moreover, we found that adaptive training induced decline in activation mainly occurred at high loads and in the anterior areas. Substantial studies have reported a wide decline in brain activation after cognitive training, which is commonly interpreted as enhanced neural efficiency (Thompson et al. 2016; Brehmer et al. 2011; Miro-Padilla et al. 2019; Bastian et al. 2022). At the level of functional networks, Thompson et al. found that training-induced reduction in activation predominantly occurred in executive control network rather than dorsal attention network (Thompson et al. 2016). While the dorsal attention network

mainly includes posterior parietal cortex and caudal frontal cortex, the core of executive control network is located in DLPFC. It seems that adaptive training preferentially modulates the activity of the anterior areas in the frontal and parietal cortices, especially at high loads.

From the perspective of network topology, prior work (Finc et al. 2020) indicated that WM training can result in a more segregated network organization, suggesting an enhancement of functional efficiency. In addition, a shift of degree centrality from frontal to parietal areas has been also reported after cognitive training (Langer et al. 2013). Actually, our proposed activation gradient between the posterior and anterior areas in the frontal and parietal cortices could be considered as a mesoscale measure, to bridge macroscale brain network and microscale individual regions. Accordingly, we found a significant increase in activation gradient at load 4-back after adaptive training. In particular, a significant correlation between the changes in activation gradient and the changes in behavioral performance was further observed in the HLT group. This suggests that the reshaping of functional responses along the posterior-anterior gradient of the frontal and parietal cortices might be a more comprehensive manifestation of training-induced improvement in neural efficiency.

Notably, early studies proposed a dorsal/ventral PFC distinction related to load/task demands (Rypma and D'Esposito 1999; Rypma et al. 1999). However, empirical evidence also suggested that use of more complex tasks such as *n*-back may activate only dorsal PFC (D'Esposito et al. 1998), and the activation of ventral PFC likely reflected the material-specific verbal processes (Rypma et al. 1999). Consistently, accumulated studies, including ours, have found that the dorsal PFC is preferentially activated in response to increased memory load (Linden et al. 2003; Edin et al. 2009; Lamichhane et al. 2020). Interestingly, this study further revealed that the load-dependent regions were distributed along the posterior-anterior axis of the frontal and parietal cortices, and the posterior and anterior areas had a dissociable response to increased memory load and cognitive training. Therefore, our findings not only support that complex WM tasks may activate dorsal rather than ventral PFC, but also provide new insights into WM load and training effects from a posterior-anterior gradient perspective.

Several methodological considerations were as follows. First, cognitive training inducing an increase or a decrease in activation remains controversial. While brain activation increases with improved performance after 2 weeks of training, Hempel et al. found that activation decreased at the time of consolidation of performance gains after 4 weeks (Hempel et al. 2004). Moreover, a recent study showed that WM training resulted in nonlinear changes in functional integration between multiple large-scale systems, with an

Table 1 Load-dependent brain regions

Brain region	Side	BA	MNI coordinates			Peak t score	Cluster size (voxels)
			x	y	z		
Anterior-LPFC	R	10	24	60	15	4.59	59
Mid-DLPFC	R	10/46	39	39	35	3.99	30
Pre-PMd	R	8	33	30	51	4.04	15
PMd	R	6	27	0	57	3.67	31
IPL	R	40	48	-30	45	4.59	71
SPL	R	7	15	-69	48	4.55	208
PMd	L	6	-30	-3	63	4.12	40
SPL	L	7	-30	-57	39	3.92	59

Note: Anterior-LPFC = anterior lateral prefrontal cortex; Mid-DLPF = mid-dorsolateral prefrontal cortex; Pre-PMd = predorsal premotor cortex; PMd = dorsal premotor cortex; IPL = inferior parietal lobule; SPL = superior parietal lobule; BA = Brodmann areas; MNI = Montreal Neurological Institute; L = left, and R = right
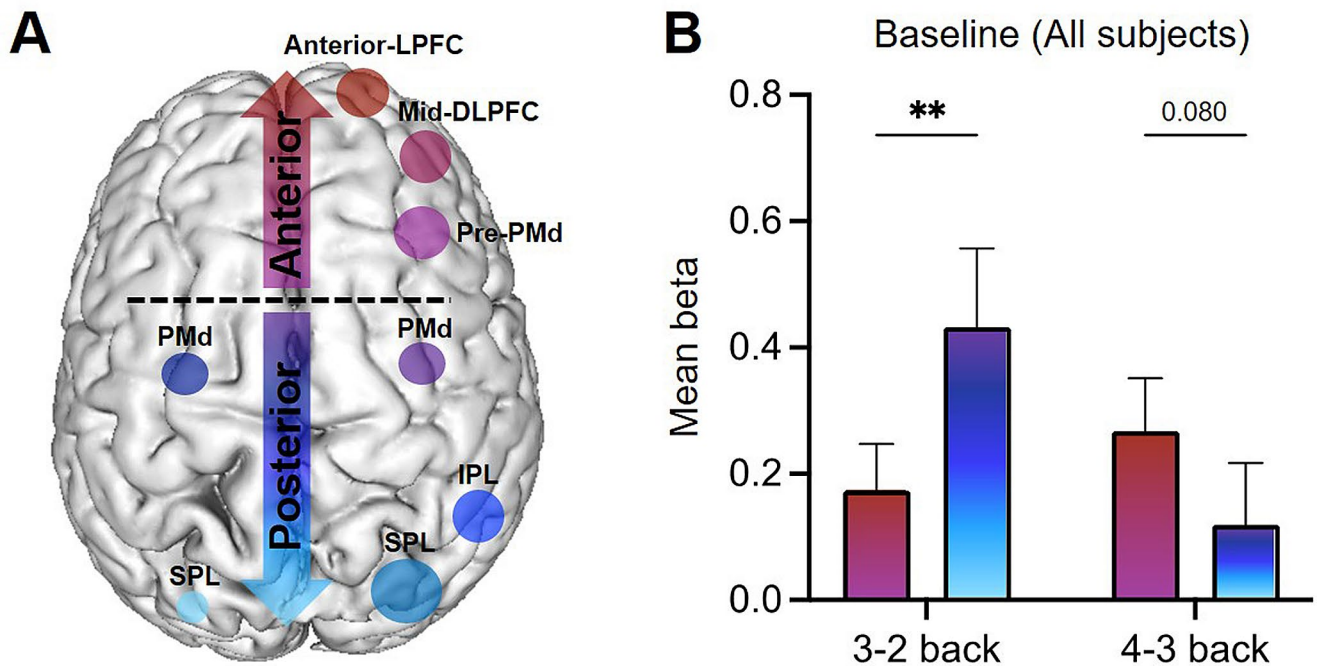
Fig. 4 The posterior and anterior areas of the frontal and parietal cortices dissociably respond to the increase in memory load at baseline. (**A**) Illustration of the load-dependent brain regions that were divided into the posterior and anterior areas. (**B**) The posterior (cold colors) and anterior (hot colors) regions showed distinct response patterns with the same increment of 1 WM item. While the increment of activation in the posterior areas was significantly higher than that in the anterior areas from 2-back to 3-back, the increment of activation in the anterior areas was higher than that in the posterior areas from 3-back to 4-back. Error bars denote the mean standard error. WM = working memory, LPFC = lateral prefrontal cortex, DLPFC = dorsolateral prefrontal cortex, PMd = dorsal premotor cortex, IPL = inferior parietal lobule, SPL = superior parietal lobule, and ** $p < 0.01$

increase in network integration at an early stage following a decrease at a later stage (Finc et al. 2020). It seems that the training duration or stage should be considered when investigating the neural plasticity related to WM training. Second, for examining the WM training effects, we found that the four-way interaction of group × session × load × spatial was marginally significant ($p = 0.066$), whereas the three-way interaction of session × load × spatial was significant ($p = 0.001$). It is likely attributed to the use of a relatively strict control group in this study. Future studies with a larger sample size might be helpful for enhancing the statistical power. Third, the adaptive training group achieved very high performance on the n-back task, with one person even reaching about 17-back. Unfortunately, we did not ask subjects to report whether and what strategies were used. In future work, this should be done as a post-test. Moreover, previous studies have revealed that encoding strategies could increase prefrontal activity unrelated to basic WM demand (Bor et al. 2003, 2004). Therefore, it is also important to dissociate functional response to increased WM load from encoding strategies.
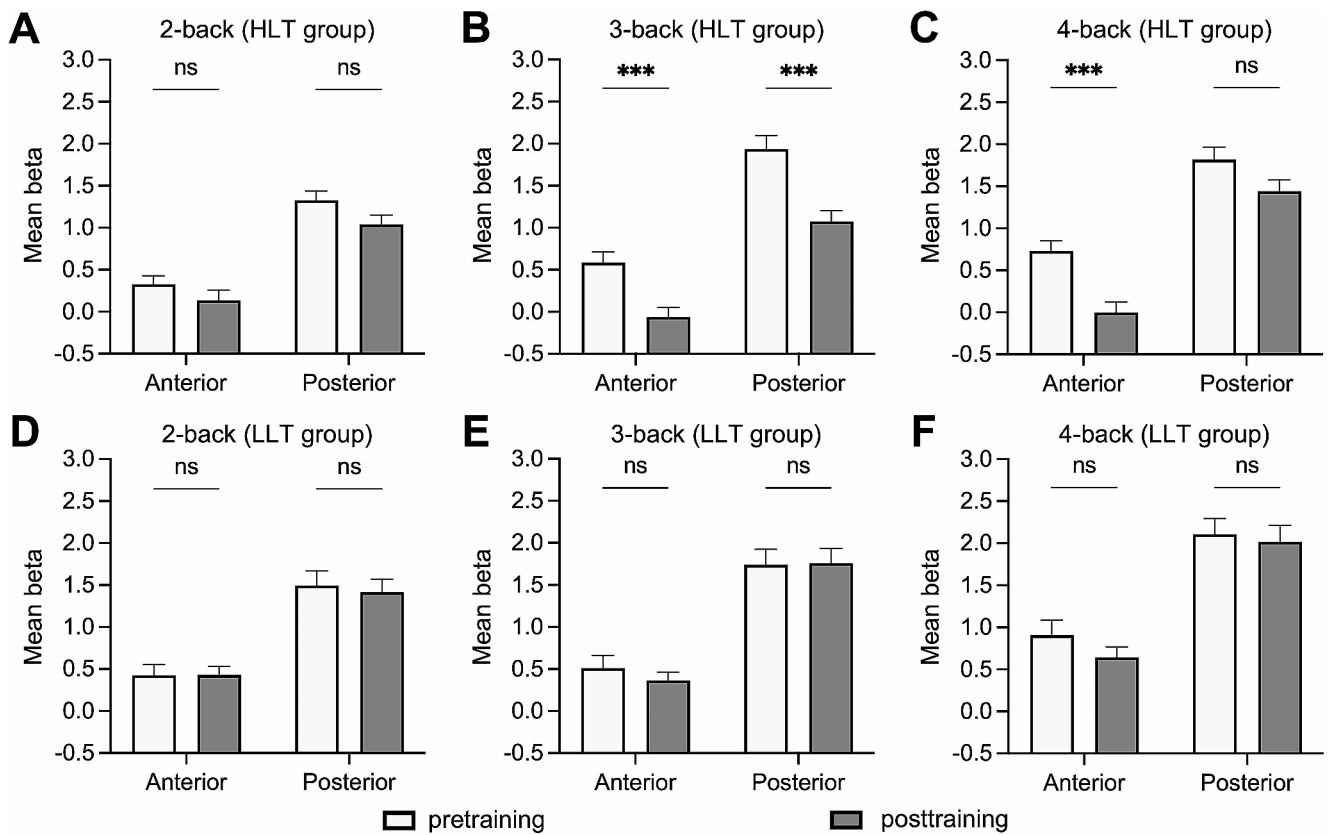
Fig. 5 Changes of activation in the posterior and anterior areas of the frontal and parietal cortices after WM training. (**A-C**) Training-induced changes of activation in the posterior and anterior areas at each load condition for the HLT group. (**D-F**) Training-induced changes of acti-vation in the posterior and anterior areas at each load condition for the LLT group. Error bars denote the mean standard error. WM = working memory, HLT = high-load training, LLT = low-load training, *** $p < 0.001$, and ns = nonsignificant
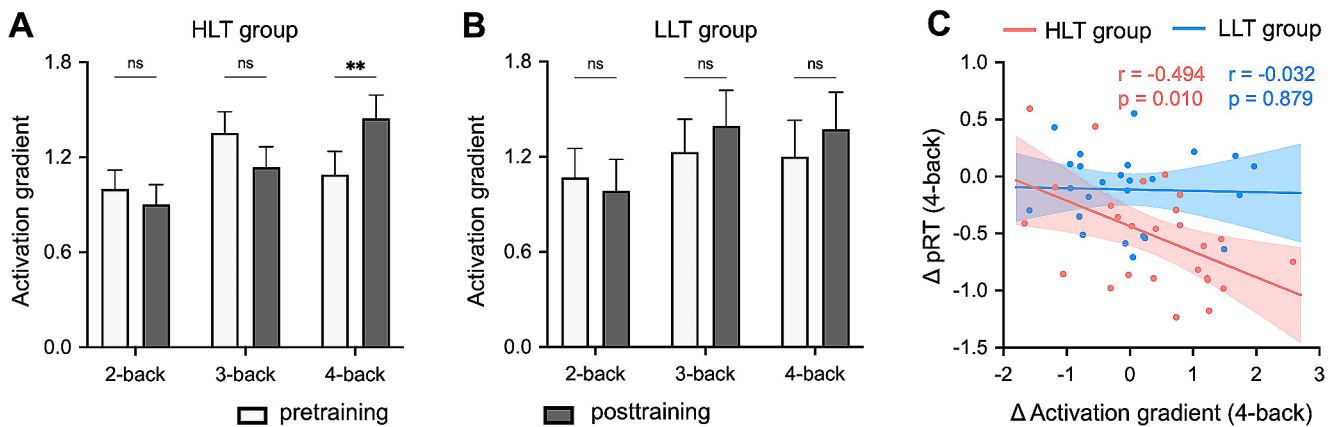
Fig. 6 WM training induced changes in the activation gradient between the posterior and anterior areas in the frontal and parietal cortices. (**A-B**) A statistical comparison of the activation gradient before and after training for the HLT and LLT groups, respectively. Error bars denote the mean standard error. (**C**) The correlation between changes in acti-vation gradient and changes in behavioral performance. WM = work-ing memory, HLT = high-load training, LLT = low-load training, ** $p < 0.01$, and ns = nonsignificant

Finally, our findings were based on visuospatial *n*-back WM paradigm. Although proposed posterior-anterior WM gradient may reconcile distinct sensory modalities, whether these results can be generalized to other WM modalities or paradigms requires further study. Despite the same task, it

should be also noted that identified functional gradient may rely on analytical methods. For instance, prior study found a gradient of activation along the rostral-caudal axis of the LPFC for the temporal demands of cognitive control. How-ever, the mid-LPFC region that replaces the rostral LPFC

tends to be the top of the hierarchy according to the relative strength of efferent versus afferent connections (Nee and D'Esposito 2016). Fortunately, both the rostral- and mid-LPFC belong to the anterior areas following the posterior-anterior gradient defined in our current study. In addition, although both brain activation and connectivity can characterize WM load and training effects, it is difficult to uncover how the information is processed. Perhaps, using neural models such as activity flow mapping (Cole et al. 2016; Zhu et al. 2023), it is possible to parse how brain information processing is manipulated by WM load and training.

In conclusion, our results demonstrated a shift in the predominant role of the posterior and anterior areas of the frontal and parietal cortices when approaching WM capacity limits. Additionally, the performance improvement benefit from the enhanced neural efficiency that is reflected in the increased activation gradient between the posterior and anterior areas. This study suggests that gradient functional reconfiguration of the frontal and parietal cortices may form a fundamental neurophysiological mechanism for flexible WM, which also has implications for understanding the impaired WM in normal aging or neuropsychiatric disorders.

Declarations

References

Awh E, Jonides J (2001) Overlapping mechanisms of attention and spatial working memory. Trends Cogn Sci 5(3):119–126. https://doi.org/10.1016/s1364-6613(00)01593-x

Baddeley A (2003) Working memory: looking back and looking forward. Nat Rev Neurosci 4(10):829–839. https://doi.org/10.1038/nrn1201

Badre D, Nee DE (2018) Frontal Cortex and the Hierarchical Control of Behavior. Trends Cogn Sci 22(2):170–188. https://doi.org/10.1016/j.tics.2017.11.005

Bastian C, Belleville S, Udale RC, Reinhartz A, Essounni M, Strobach T (2022) Mechanisms underlying training-induced cognitive change. Nat Rev Psychol 1:30–41. https://doi.org/10.1038/s44159-021-00001-3

Blacker KJ, Negoita S, Ewen JB, Courtney SM (2017) N-back versus Complex Span Working Memory Training. J Cogn Enhanc 1(4):434–454. https://doi.org/10.1007/s41465-017-0044-1

Bor D, Duncan J, Wiseman RJ, Owen AM (2003) Encoding strategies dissociate prefrontal activity from working memory demand. Neuron 37(2):361–367. https://doi.org/10.1016/S0896-6273(02)01171-6

Bor D, Cumming N, Scott CEL, Owen AM (2004) Prefrontal cortical involvement in verbal encoding strategies. Eur J Neurosci 19(12):3365–3370. https://doi.org/10.1111/j.1460-9568.2004.03438.x

Braun U, Schafer A, Walter H, Erk S, Romanczuk-Seiferth N, Haddad L, Schweiger JI, Grimm O, Heinz A, Tost H, Meyer-Lindenberg A, Bassett DS (2015) Dynamic reconfiguration of frontal brain networks during executive cognition in humans. Proc Natl Acad Sci U S A 112(37):11678–11683. https://doi.org/10.1073/pnas.1422487112

Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC (1997) A parametric study of prefrontal cortex involvement in human working memory. NeuroImage 5(1):49–62. https://doi.org/10.1006/nimg.1996.0247

Brehmer Y, Rieckmann A, Bellander M, Westerberg H, Fischer H, Backman L (2011) Neural correlates of training-related working-memory gains in old age. NeuroImage 58(4):1110–1120. https://doi.org/10.1016/j.neuroimage.2011.06.079

Buschman TJ (2021) Balancing flexibility and interference in Working Memory. Annu Rev Vis Sci 7:367–388. https://doi.org/10.1146/annurev-vision-100419-104831

Buschman TJ, Siegel M, Roy JE, Miller EK (2011) Neural substrates of cognitive capacity limitations. Proc Natl Acad Sci U S A 108(27):11252–11255. https://doi.org/10.1073/pnas.1104666108

Chai WJ, Abd Hamid AI, Abdullah JM (2018) Working Memory from the psychological and Neurosciences perspectives: a review. Front Psychol 9:401. https://doi.org/10.3389/fpsyg.2018.00401

Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes JD (2017) The distributed nature of Working Memory. Trends Cogn Sci 21(2):111–124. https://doi.org/10.1016/j.tics.2016.12.007

Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE (1997) Temporal dynamics of brain activation during a working memory task. Nature 386(6625):604–608. https://doi.org/10.1038/386604a0

Cole MW, Ito T, Bassett DS, Schultz DH (2016) Activity flow over resting-state networks shapes cognitive task activations. Nat Neurosci 19(12):1718–1726. https://doi.org/10.1038/nn.4406

Constantinidis C, Klingberg T (2016) The neuroscience of working memory capacity and training. Nat Rev Neurosci 17(7):438–449. https://doi.org/10.1038/nrn.2016.43

Cowan N (1988) Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human

information-processing system. Psychol Bull 104(2):163–191. https://doi.org/10.1037/0033-2909.104.2.163

Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci 24(1):87–114. https://doi.org/10.1017/s0140525x01003922

D'Esposito M, Postle BR (2015) The cognitive neuroscience of working memory. Annu Rev Psychol 66:115–142. https://doi.org/10.1146/annurev-psych-010814-015031

D'Esposito M, Aguirre GK, Zarahn E, Ballard D, Shin RK, Lease J (1998) Functional MRI studies of spatial and nonspatial working memory. Brain Res Cogn Brain Res 7(1):1–13. https://doi.org/10.1016/s0926-6410(98)00004-4

Edin F, Klingberg T, Johansson P, McNab F, Tegner J, Compte A (2009) Mechanism for top-down control of working memory capacity. Proc Natl Acad Sci U S A 106(16):6802–6807. https://doi.org/10.1073/pnas.0901894106

Eriksson J, Vogel EK, Lansner A, Bergstrom F, Nyberg L (2015) Neurocognitive Architecture of Working Memory. Neuron 88(1):33–46. https://doi.org/10.1016/j.neuron.2015.09.020

Finc K, Bonna K, He X, Lydon-Staley DM, Kuhn S, Duch W, Bassett DS (2020) Dynamic reconfiguration of functional brain networks during working memory training. Nat Commun 11(1):2435. https://doi.org/10.1038/s41467-020-15631-z

Freedman DJ, Ibos G (2018) An integrative Framework for sensory, Motor, and cognitive functions of the posterior parietal cortex. Neuron 97(6):1219–1234. https://doi.org/10.1016/j.neuron.2018.01.044

Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. Science 173(3997):652–654. https://doi.org/10.1126/science.173.3997.652

Hasson U, Chen J, Honey CJ (2015) Hierarchical process memory: memory as an integral component of information processing. Trends Cogn Sci 19(6):304–313. https://doi.org/10.1016/j.tics.2015.04.006

Hempel A, Giesel FL, Garcia Caraballo NM, Amann M, Meyer H, Wustenberg T, Essig M, Schroder J (2004) Plasticity of cortical activation related to working memory during training. Am J Psychiatry 161(4):745–747. https://doi.org/10.1176/appi.ajp.161.4.745

Jaeggi SM, Seewer R, Nirkko AC, Eckstein D, Schroth G, Groner R, Gutbrod K (2003) Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: functional magnetic resonance imaging study. NeuroImage 19:210–225. https://doi.org/10.1016/s1053-8119(03)00098-3

Jaeggi SM, Buschkuehl M, Etienne A, Ozdoba C, Perrig WJ, Nirkko AC (2007) On how high performers keep cool brains in situations of cognitive overload. Cogn Affect Behav Neurosci 7(2):75–89. https://doi.org/10.3758/cabn.7.2.75

Jonides J, Schumacher EH, Smith EE, Lauber EJ, Awh E, Minoshima S, Koeppe RA (1997) Verbal working memory load affects Regional Brain activation as measured by PET. J Cogn Neurosci 9(4):462–475. https://doi.org/10.1162/jocn.1997.9.4.462

Katsuki F, Constantinidis C (2012) Unique and shared roles of the posterior parietal and dorsolateral prefrontal cortex in cognitive functions. Front Integr Neurosci 6:17. https://doi.org/10.3389/fnint.2012.00017

Kitzbichler MG, Henson RN, Smith ML, Nathan PJ, Bullmore ET (2011) Cognitive effort drives workspace configuration of human brain functional networks. J Neurosci 31(22):8259–8270. https://doi.org/10.1523/JNEUROSCI.0440-11.2011

Klingberg T (2010) Training and plasticity of working memory. Trends Cogn Sci 14(7):317–324. https://doi.org/10.1016/j.tics.2010.05.002

Koechlin E, Ody C, Kouneiher F (2003) The architecture of cognitive control in the human prefrontal cortex. Science 302(5648):1181–1185. https://doi.org/10.1126/science.1088545

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12(5):535–540. https://doi.org/10.1038/nn.2303

Lamichhane B, Westbrook A, Cole MW, Braver TS (2020) Exploring brain-behavior relationships in the N-back task. NeuroImage 212:116683. https://doi.org/10.1016/j.neuroimage.2020.116683

Langer N, von Bastian CC, Wirz H, Oberauer K, Jancke L (2013) The effects of working memory training on functional brain network efficiency. Cortex 49(9):2424–2438. https://doi.org/10.1016/j.cortex.2013.01.008

Lee SH, Kravitz DJ, Baker CI (2013) Goal-dependent dissociation of visual and prefrontal cortices during working memory. Nat Neurosci 16(8):997–U935. https://doi.org/10.1038/nn.3452

Liang X, Zou Q, He Y, Yang Y (2016) Topologically reorganized Connectivity Architecture of Default-Mode, Executive-Control, and salience networks across Working Memory Task loads. Cereb Cortex 26(4):1501–1511. https://doi.org/10.1093/cercor/bhu316

Linden DE, Bittner RA, Muckli L, Waltz JA, Kriegeskorte N, Goebel R, Singer W, Munk MH (2003) Cortical capacity constraints for visual working memory: dissociation of fMRI load effects in a fronto-parietal network. NeuroImage 20(3):1518–1530. https://doi.org/10.1016/j.neuroimage.2003.07.021

Mackey WE, Curtis CE (2017) Distinct contributions by frontal and parietal cortices support working memory. Sci Rep 7(1):6188. https://doi.org/10.1038/s41598-017-06293-x

McNab F, Klingberg T (2008) Prefrontal cortex and basal ganglia control access to working memory. Nat Neurosci 11(1):103–107. https://doi.org/10.1038/nn2024

Meyer ML, Spunt RP, Berkman ET, Taylor SE, Lieberman MD (2012) Evidence for social working memory from a parametric functional MRI study. Proc Natl Acad Sci U S A 109(6):1883–1888. https://doi.org/10.1073/pnas.1121077109

Minamoto T, Yaoi K, Osaka M, Osaka N (2015) The rostral prefrontal cortex underlies individual differences in working memory capacity: an approach from the hierarchical model of the cognitive control. Cortex 71:277–290. https://doi.org/10.1016/j.cortex.2015.07.025

Miro-Padilla A, Bueicheku E, Ventura-Campos N, Flores-Compan MJ, Parcet MA, Avila C (2019) Long-term brain effects of N-back training: an fMRI study. Brain Imaging Behav 13(4):1115–1127. https://doi.org/10.1007/s11682-018-9925-x

Murray JD, Bernacchia A, Freedman DJ, Romo R, Wallis JD, Cai X, Padoa-Schioppa C, Pasternak T, Seo H, Lee D, Wang XJ (2014) A hierarchy of intrinsic timescales across primate cortex. Nat Neurosci 17(12):1661–1663. https://doi.org/10.1038/nn.3862

Murray JD, Jaramillo J, Wang XJ (2017) Working memory and decision-making in a Frontoparietal Circuit Model. J Neurosci 37(50):12167–12186. https://doi.org/10.1523/JNEUROSCI.0343-17.2017

Nee DE, Brown JW (2012) Rostral-caudal gradients of abstraction revealed by multi-variate pattern analysis of working memory. NeuroImage 63(3):1285–1294. https://doi.org/10.1016/j.neuroimage.2012.08.034

Nee DE, D'Esposito M (2016) The hierarchical organization of the lateral prefrontal cortex. Elife 5. https://doi.org/10.7554/eLife.12112

Olesen PJ, Westerberg H, Klingberg T (2004) Increased prefrontal and parietal activity after training of working memory. Nat Neurosci 7(1):75–79. https://doi.org/10.1038/nn1165

Owen AM, McMillan KM, Laird AR, Bullmore E (2005) N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum Brain Mapp 25(1):46–59. https://doi.org/10.1002/hbm.20131

Palva JM, Monto S, Kulashekhar S, Palva S (2010) Neuronal synchrony reveals working memory networks and predicts individual memory capacity. Proc Natl Acad Sci U S A 107(16):7580–7585. https://doi.org/10.1073/pnas.0913113107

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE (2011) Functional network organization of the human brain. Neuron 72(4):665–678. https://doi.org/10.1016/j.neuron.2011.09.006

Roggeman C, Klingberg T, Feenstra HE, Compte A, Almeida R (2014) Trade-off between capacity and precision in visuospatial working memory. J Cogn Neurosci 26(2):211–222. https://doi.org/10.1162/jocn_a_00485

Romo R, Brody CD, Hernandez A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. Nature 399(6735):470–473. https://doi.org/10.1038/20939

Rypma B, D'Esposito M (1999) The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. Proc Natl Acad Sci U S A 96(11):6558–6563. https://doi.org/10.1073/pnas.96.11.6558

Rypma B, Prabhakaran V, Desmond JE, Glover GH, Gabrieli JDE (1999) Load-dependent roles of frontal brain regions in the maintenance of working memory. NeuroImage 9(2):216–226. https://doi.org/10.1006/nimg.1998.0404

Salmi J, Nyberg L, Laine M (2018) Working memory training mostly engages general-purpose large-scale networks for learning. Neurosci Biobehav Rev 93:108–122. https://doi.org/10.1016/j.neubiorev.2018.03.019

Schneiders JA, Opitz B, Krick CM, Mecklinger A (2011) Separating intra-modal and across-modal training effects in visual working memory: an fMRI investigation. Cereb Cortex 21(11):2555–2564. https://doi.org/10.1093/cercor/bhr037

Soveri A, Antfolk J, Karlsson L, Salo B, Laine M (2017) Working memory training revisited: a multi-level meta-analysis of n-back training studies. Psychon Bull Rev 24(4):1077–1096. https://doi.org/10.3758/s13423-016-1217-0

Sreenivasan KK, D'Esposito M (2019) The what, where and how of delay activity. Nat Rev Neurosci 20(8):466–481. https://doi.org/10.1038/s41583-019-0176-7

Susanne M., Jaeggi Martin, Buschkuehl John, Jonides Walter J., Perrig (2008) Improving fluid intelligence with training on working memory Proceedings of the National Academy of Sciences 105(19) 6829-6833 10.1073/pnas.0801268105

Thompson TW, Waskom ML, Gabrieli JD (2016) Intensive Working Memory Training produces functional changes in large-scale Frontoparietal Networks. J Cogn Neurosci 28(4):575–588. https://doi.org/10.1162/jocn_a_00916

Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 428(6984):751–754. https://doi.org/10.1038/nature02466

Van Snellenberg JX, Slifstein M, Read C, Weber J, Thompson JL, Wager TD, Shohamy D, Abi-Dargham A, Smith EE (2015) Dynamic shifts in brain network activation during supracapacity working memory task performance. Hum Brain Mapp 36(4):1245–1264. https://doi.org/10.1002/hbm.22699

Vergara J, Rivera N, Rossi-Pool R, Romo R (2016) A neural Parametric Code for storing information of more than one sensory modality in Working Memory. Neuron 89(1):54–62. https://doi.org/10.1016/j.neuron.2015.11.026

Vermeij A, Kessels RPC, Heskamp L, Simons EMF, Dautzenberg PLJ, Claassen J (2017) Prefrontal activation may predict working-memory training gain in normal aging and mild cognitive impairment. Brain Imaging Behav 11(1):141–154. https://doi.org/10.1007/s11682-016-9508-7

Wager TD, Smith EE (2003) Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci 3(4):255–274. https://doi.org/10.3758/cabn.3.4.255

Wang LP, Li XX, Hsiao SS, Lenz FA, Bodner M, Zhou YD, Fuster JM (2015) Differential roles of delay-period neural activity in the monkey dorsolateral prefrontal cortex in visual-haptic crossmodal working memory. Proc Natl Acad Sci U S A 112(2):E214–E219. https://doi.org/10.1073/pnas.1410130112

Xie Y, Hu P, Li J, Chen J, Song W, Wang XJ, Yang T, Dehaene S, Tang S, Min B, Wang L (2022) Geometry of sequence working memory in macaque prefrontal cortex. Science 375(6581):632–639. https://doi.org/10.1126/science.abm0204

Xu Y, Chun MM (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. Nature 440(7080):91–95. https://doi.org/10.1038/nature04262

Yin D, Liu W, Zeljic K, Lv Q, Wang Z, You M, Men W, Fan M, Cheng W, Wang Z (2017) Failure in cognitive suppression of negative affect in adolescents with generalized anxiety disorder. Sci Rep 7(1):6583. https://doi.org/10.1038/s41598-017-07063-5

Yin DZ, Wang XF, Zhang XY, Yu QR, Wei Y, Cai Q, Fan MX, Li L (2021) Dissociable plasticity of visual-motor system in functional specialization and flexibility in expert table tennis players. Brain Struct Funct 226(6):1973–1990. https://doi.org/10.1007/s00429-021-02304-w

Zanto TP, Rubens MT, Thangavel A, Gazzaley A (2011) Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. Nat Neurosci 14(5):656–661. https://doi.org/10.1038/nn.2773

Zhang H, Zhao R, Hu X, Guan S, Margulies DS, Meng C, Biswal BB (2022) Cortical connectivity gradients and local timescales during cognitive states are modulated by cognitive loads. Brain Struct Funct 227(8):2701–2712. https://doi.org/10.1007/s00429-022-02564-0

Zhu H, Huang Z, Yang Y, Su K, Fan M, Zou Y, Li T, Yin D (2023) Activity flow mapping over probabilistic functional connectivity. Hum Brain Mapp 44(2):341–361. https://doi.org/10.1002/hbm.26044

Zimmer HD (2008) Visual and spatial working memory: from boxes to networks. Neurosci Biobehav Rev 32(8):1373–1395. https://doi.org/10.1016/j.neubiorev.2008.05.016