Check for updates

OPEN ACCESS

EDITED BY Christine Wagner, University at Albany, United States

REVIEWED BY Yikang Liu, United Imaging Intelligence, United States Chunlin Li, Capital Medical University, China

*CORRESPONDENCE Aniol Santo-Angles ⊠ asantoangles@gmail.com

RECEIVED 16 November 2023 ACCEPTED 12 February 2024 PUBLISHED 04 March 2024

CITATION

Santo-Angles A, Temudo A, Babushkin V and Sreenivasan KK (2024) Effective connectivity of working memory performance: a DCM study of MEG data. *Front. Hum. Neurosci.* 18:1339728. doi: 10.3389/fnhum.2024.1339728

COPYRIGHT

© 2024 Santo-Angles, Temudo, Babushkin and Sreenivasan. This is an open-access article distributed under the terms of the **Creative Commons Attribution License (CC BY)**. The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Effective connectivity of working memory performance: a DCM study of MEG data

Aniol Santo-Angles^{1,2*}, Ainsley Temudo¹, Vahan Babushkin¹ and Kartik K. Sreenivasan^{1,2}

¹Division of Science and Mathematics, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, ²Center for Brain and Health, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

Visual working memory (WM) engages several nodes of a large-scale network that includes frontal, parietal, and visual regions; however, little is understood about how these regions interact to support WM behavior. In particular, it is unclear whether network dynamics during WM maintenance primarily represent feedforward or feedback connections. This question has important implications for current debates about the relative roles of frontoparietal and visual regions in WM maintenance. In the current study, we investigated the network activity supporting WM using MEG data acquired while healthy subjects performed a multi-item delayed estimation WM task. We used computational modeling of behavior to discriminate correct responses (high accuracy trials) from two different types of incorrect responses (low accuracy and swap trials), and dynamic causal modeling of MEG data to measure effective connectivity. We observed behaviorally dependent changes in effective connectivity in a brain network comprising frontoparietal and early visual areas. In comparison with high accuracy trials, frontoparietal and frontooccipital networks showed disrupted signals depending on type of behavioral error. Low accuracy trials showed disrupted feedback signals during early portions of WM maintenance and disrupted feedforward signals during later portions of maintenance delay, while swap errors showed disrupted feedback signals during the whole delay period. These results support a distributed model of WM that emphasizes the role of visual regions in WM storage and where changes in large scale network configurations can have important consequences for memory-guided behavior.

KEYWORDS

dynamic causal model (DCM), working memory, magnetoencephalography (MEG), effective connectivity, binding problem

Introduction

Neural activity supporting working memory (WM) is distributed throughout the brain (Christophel et al., 2017; Sreenivasan and D'Esposito, 2019; Mejías and Wang, 2022), motivating the view that WM may be better understood as a distributed network function, rather than being localized to specific brain regions (Lorenc and Sreenivasan, 2021; Rezayat et al., 2022), as suggested by evidence of task-dependent reorganization of large-scale brain networks during WM (e.g., Cohen and D'Esposito, 2016). This "new" perspective of WM was anticipated by Patricia Goldman-Rakic, who pioneer the study of the neurobiological underpinnings of WM (Goldman-Rakic, 1995). Goldman-Rakic and colleagues focused on the role of prefrontal cortex in higher cognitive functions, particularly dorsolateral prefrontal cortex (dlPFC) in WM, but their work also demonstrated that dlPFC was densely

connected with posterior parietal cortex, both within- and betweenhemispheres (Schwartz and Goldman-Rakic, 1984; Cavada and Goldman-Rakic, 1989). Remarkably, these two regions were also connected with a widespread and distributed network of cortical and subcortical regions supporting spatially guided behavior (Selemon and Goldman-Rakic, 1988). Despite this early evidence, the majority of subsequent investigations into WM networks, particularly in human brain imaging, have focused on interactions between prefrontal and parietal cortex based on findings that (i) these regions consistently coactivate during WM tasks (Rottschy et al., 2012, 2013; Daniel et al., 2016), (ii) modulation of frontoparietal activity by TMS/tACS impacts behavioral performance (Kessels et al., 2000; Postle et al., 2006; Polanía et al., 2012; Violante et al., 2017; Tseng et al., 2018; Biel et al., 2022), and (iii) that communication between distinct frontal and parietal sites are task-dependent and content-specific (Salazar et al., 2012; Ratcliffe et al., 2022), are modulated by WM load (Crespo-Garcia et al., 2013; Syrjälä et al., 2021), and predict individual behavioral capacity in visual WM (Palva et al., 2010). Taken together, these findings strongly support the early evidence of Goldman-Rakic and colleagues that network activity involving frontal and parietal cortices may be key for understanding WM function.

There remain several unanswered questions about how network activity supports WM. First, the majority of studies about WM network function do not investigate effective connectivity between brain regions, limiting inferences about how interregional communication supports WM. Effective connectivity refers to the directed causal influence that one neural system or region exerts over another, underscoring the dynamic and experimentdependent nature of these interactions (Friston, 2011). This concept encapsulates the notion that the strength and direction of influence between distinct brain regions can vary based on specific experimental conditions or tasks. In human imaging data, effective connectivity can be addressed using dynamic causal modeling (DCM), a Bayesian framework designed to infer, from brain activity measurements, hidden neuronal states with a neurobiological meaning, such as the context-dependent modulation of network dynamics and the differential contribution of feedforward (bottomup) and feedback (top-down) signals (David and Friston, 2003; Stephan et al., 2010; Friston et al., 2019). Previous DCM studies using fMRI with an n-back WM task reported an association between memory load and enhanced effective connectivity in frontoparietal feedforward (Dima et al., 2014; Jung et al., 2018) and feedback signals (Heinzel et al., 2017). A DCM study of local field potentials in non-human primates during a change detection task (Pinotsis et al., 2019) also reported load-dependent modulations of effective connectivity in the frontoparietal network. Furthermore, disturbances in feedback coupling were observed when the number of items surpassed cognitive capacity, underscoring the behavioral significance of large-scale effective connectivity.

Second, network studies largely focused on dorsolateral prefrontal and posterior parietal regions while ignoring the contributions of medial prefrontal regions, despite the robust activation of superior frontal areas during the delay period (Li et al., 2022), its association with spatial WM (Courtney et al., 1998; Rottschy et al., 2012, 2013) and the pattern of structural

connectivity of the superior frontal with other WM nodes in the frontoparietal networks (Briggs et al., 2020). The contributions of sensory regions have also been ignored, particularly during WM maintenance, when sensory stimulation is absent. Understanding how these regions interact with frontoparietal circuits has the potential to shed light on the debate about the role of sensory regions in WM (Xu, 2020; Lorenc and Sreenivasan, 2021), adjudicating between the sensory recruitment hypothesis (Pasternak and Greenlee, 2005; D'Esposito, 2007), which is supported by findings from human neuroimaging studies that demonstrate that visual WM contents can be decoded from visual areas (Sreenivasan et al., 2014a; Curtis and Sprague, 2021), and the countervailing view that early sensory activity during WM reflects feedback signals from parietal or frontal regions (Leavitt et al., 2017; Xu, 2020). The study of effective connectivity, particularly the contribution of feedforward/feedback signals from/to early visual areas during the delay period has the potential to make an important contribution to this debate.

Third, the behavioral relevance of WM network activity is underspecified. Behavioral relevance is often inferred from changes in network activity as a function of WM load. However, loaddependent modulations may be due to progressive deployment of neural resources in response to cognitive demands rather than WM computations per se (Palva et al., 2010). Even studies that do not rely on WM load differences do not discriminate between different types of behavioral errors that might emerge from distinct neural sources. Behavioral and computational evidence suggest that, when the integration or binding of visual features is required, behavioral errors consist a mixture of low accuracy errors (guess responses) and swap errors. The former arise from disrupted neural representations of individual features, while the latter appears as a disruption in the integration between features (Bays, 2016; Schneegans and Bays, 2017, 2019). Indeed, binding errors are associated with lesions in frontoparietal networks (Lugtmeijer et al., 2021), and feature binding has been associated with intraparietal sulcus activity, relatively independent from cognitive load (Gosseries et al., 2018) or stimulus identity (Cai et al., 2020). In an fMRI study with a delayed continuous report task (Mallett et al., 2022), successfully reconstructed location-specific representations from early visual cortex and intraparietal area in target trials. Interestingly, in swap trials, target location could not be reconstructed, but both areas successfully reconstructed the non-target stimulus on the location of subsequent swapped report, indicating that swap errors arise from incorrect responses being maintained in WM. While this evidence indicates that swap and low accuracy trials may have distinct origins at the mesocircuit scale, here we examine how large-scale network properties contribute to behavioral outcomes in WM.

In the current study, we examined how WM performance depends on network activity using a combination of data-driven approaches to identify key nodes in a brain-wide WM network and model-driven approaches to infer directional influences and distinguish between different types of behavioral errors. We analyzed MEG data acquired while healthy subjects performed a delayed estimation task for location, and found that this task engaged nodes in lateral and medial frontal cortex, as well as parietal and occipital cortices. We used DCM to measure effective connectivity between these nodes and found that a key element of this network was feedforward connections between visual and frontoparietal regions. Finally, discriminating between trials with high accuracy responses, low accuracy responses, and swap errors allowed us to describe how differential configurations of this network are associated with different types of behavioral errors.

Methods

Sample

We invited 30 healthy adults from the NYUAD community to participate in this study, recruited from a larger pool who had previously participated in a separate behavioral study with the same task. In order to ensure that we had sufficient error trials for our analyses, we only invited subjects who had symmetric swaps (see below) on at least 5% of trials in the behavioral study. The final sample included in our analyses was 26 subjects [age: 24.33 \pm 4.57, range (19–37); four females, 22 males], after discarding three subjects due to poor performance (fewer than 15 correct responses in one or more conditions of interest), and another due to technical challenges transforming MEG data from sensor space to source space. All subjects were right-handed with normal or corrected-to-normal vision, and provided informed written consent in accordance with procedures approved by NYUAD's IRB.

Task design

Stimuli displays, timing, and responses were controlled and recorded using the Psychophysics Toolbox (Brainard, 1997) in MATLAB (The MathWorks, Inc., Natick MA, USA). We assessed WM behavior using a delayed continuous report task (Figure 1A). A fixation cross of 0.2 degrees of visual angle (DoV) was presented in the center of the screen during baseline, stimulus presentation, and memory delays. Each trial began with a 1 s fixation period, after which subjects saw three colored discs (each 0.55 DoV, presented at an eccentricity of 4.5 DoV) for 0.2 s. Subjects were asked to hold the colors and locations of all three discs over a blank 2 s memory delay. At the end of the memory delay, instead of asking subjects to report one of the items, we cued subjects to report the location of each of the three items sequentially (in a random order) by presenting a disc at fixation whose color indicated the color of the item subjects were to report. Subjects adjusted the position of a white circle of the same size and of the same eccentricity as the disc using an MEG-compatible response dial and locked in their report using a button box (Current Designs, Philadelphia PA, USA). Each location report was self-paced. Following all three reports, subjects were shown a 1 s feedback screen that indicated the presented and reported locations on that trial. The trial concluded with a selfpaced intertrial interval (ITI)-that is, the next trial began when subjects pressed a button. Subjects completed 500 \pm 75 (400-720) trials, divided into blocks of 80 trials (seven subjects) or 100 trials (19 subjects). Two subjects completed the experiment in two sessions 1 and 4 days apart; the remaining subjects completed the experiment in a single session.

The memory items were always presented in one hemifield in a given block of trials, and left and right hemifield blocks were presented in a random order. The color of the memory items on each trial were chosen from 180 color equally spaced segments drawn from a circle in CIELAB color space (radius 59°, centered at L = 54, a = 18, and b = -8) (CIE, 2004), with a minimum gap of 15 color segments between memory items on a given trial to avoid errors due to the color similarity of the memory items. Memory locations were randomly sampled from an isoeccentric (semi)circular space (depending on the hemifield block) with a minimum gap of 15° of polar angle. We additionally excluded locations that were 10° of polar angle away from the vertical meridian to avoid well-documented oblique effects (Furmanski and Engel, 2000).

Stimulus displays were projected onto a screen 85 cm away from the subject's head. Subjects were instructed to remain still during the MEG recordings and to avoid blinking or moving their eyes during the cue and delay screens. Subjects were free to move their eyes during the response screens and were encouraged to use the ITI as well as breaks between blocks to blink.

Behavioral analysis

To label trials based on subjects' behavioral performance, we fit a probabilistic mixture model with three components (target, nontarget, and guess responses) to the distribution of errors using the C016_fit function (currently called mixtureFit, bayslab.com) (Bays et al., 2009; Schneegans and Bays, 2016):

$$p\left(\hat{\theta}\right) = (1 - \beta - \gamma)\phi_{\sigma}\left(\hat{\theta} - \theta\right) + \gamma \frac{1}{2\pi} + \beta \frac{1}{m} \sum_{i}^{m} \phi_{\sigma}\left(\hat{\theta} - \theta_{i}^{*}\right),$$

where θ is the target location (in radians), $\hat{\theta}$ is the reported location, γ is the probability of random guess, β is the probability of swap, θ_i^* are the locations of non-targets with i = 1, ..., m and m = 2, and ϕ_{σ} is the circular normal distribution (Von Mises) with zero mean and σ standard deviation. Maximum likelihood estimates of parameters σ , γ , and β were obtained at subject level. Then, trialwise posterior probabilities that responses were drawn from each of the three mixture components were used to define trial types (see Figure 1C in Bays et al., 2009 and Figure 2 in Schneegans and Bays, 2016). High accuracy (correct) responses were trials with a target probability above 90% in all 3 reports, swaps errors were trials with a probability of non-target response above 70% in at least one report, and low accuracy trials comprised the remaining trials. The more lenient cut-off criterion for swap trials was motivated by the tendency of the model to underestimate swap frequency, while the conservative cut-off in correct trials in all three reports minimized the likelihood that correct trials contain lucky random guesses. Note that the label "low accuracy" does not mean that all three reports were inaccurate or guess responses; it simply indicates that subjects did not report all three items with a high degree of accuracy.



וספכנוסוז. בעכ, פווזע אוגעמו כסונפא, ור-ס, וותמסמופנמו געוכעג, אורס, ווועעופ ווסוומו עצועג, סרס, געספוסו ווסוז

MEG and MRI acquisition and preprocessing

Before MEG acquisition, each subject's head shape was digitized using a Polhemus dual source handheld FastSCAN-II. MEG data was recorded continuously using a 208-channel axial gradiometer Yokogawa system (Kanazawa Institute of Technology, Kanazawa, Japan) with a sampling rate of 1,000 Hz and an online lowpass filter of 200 Hz. The continuous MEG data was first noisereduced using eight magnetometer reference channels located away from the participant's head and using the Time Shifted Principle Component Analysis (TSPCA, block width of 5,000 and 30 shifts) as implemented in MEG160 software (Yokogawa Electric Corporation and Eagle Technology Corporation, Tokyo, Japan). To remove eye blink and heartbeat artifacts, an independent component analysis was performed with fast ICA method in MNE-Python (Ablin et al., 2018). All subsequent preprocessing steps were performed using FieldTrip (Oostenveld et al., 2011). Data was epoched from 1s pre-stimulus onset to 2.7s post-stimulus onset (including baseline, stimulus presentation, memory delay, and the first 0.5 s of response), demeaned, and low pass filtered at 140 Hz. In order to clean the data, we automatically rejected trials and channels with high variance using ft_artifact_zvalue.m function with the following settings: jumps artifacts threshold z =20, muscle artifacts threshold z = 5 and blink artifacts threshold z = 10. Additionally, we manually inspected individual epochs to remove trials that were contaminated with jump, muscle, and eye blink artifacts. Rejected channels were reconstructed using the average of all neighbors. Preprocessing resulted in an average of 54 \pm 37 (13–200) removed trials [accounting for the 11 \pm 6% of total trials (2–29)], and 2 ± 1 (1–10) rejected channels (out of 208) per participant. The average number of trials available for analysis was 446 \pm 69 (357–707): 106 \pm 94 (19–439) (23 \pm 17%) high accuracy trials, 213 ± 53 (80–328) (48 ± 11%) low accuracy trials, and 126 ± 46 (46–250) (29 \pm 10%) swap trials. The experimental setup included trials with stimuli presented in either the left or right hemifield. To analyze all trials without canceling out lateralized neural signals, we took the following approach. For trials with left hemifield presentation, we mirrored (reflected) the MEG sensors across the midline, ensuring that MEG sensors on the left (right) side aligned with the hemisphere contralateral (ipsilateral) to the stimulus presentation (as depicted in Figure 2).

In a subset of eight subjects [age: 23.16 \pm 3.54, range (20–31); two females, six males], we used a high-resolution anatomical



MRI volume collected during a separate session to constrain the transformation from sensor to source space. The anatomical scan was acquired on a Siemens Magnetom Prisma 3T MRI scanner using a 6.5 min MPRAGE-3D T1-weighted, gradient-echo acquisition sequence (TR: 2,400 ms, TE = 2.22 ms, flip angle: 8 degrees, voxel size: 0.8 mm³, 208 slices, FOV: 256 mm).

Selection of sources of interest

In order to study effective connectivity, we performed a timefrequency analysis in source space to identify the network of neural sources actively involved in WM, comprising frontal, parietal, and occipital regions. First, we localized oscillatory sources using the beamformer technique (Jaiswal et al., 2020; Westner et al., 2022). The head model template in MNI space ("standard_bem") (Oostenveld et al., 2003) was transformed into a subject-specific head shape to compute the volume conduction model using a realistically shaped single shell approximation from FieldTrip's ft_prepare_headmodel.m function (Nolte, 2003). Next, we used ft_prepare_sourcemodel.m function to create a source model with a 1.25 cm isotropic 3D grid and with 1,499 sources within the head model. Finally, source and head models were used to create a forward model (ft_prepare_leadfield.m). For the subset of eight subjects with MRI data, the head model was computed using the structural T1 image, and the source model in standard space was spatially normalized into subject-specific structural image with a non-linear transformation (warp). As a result of this transformation, forward model positions were equivalent across subjects in normalized space. To compute the inverse solution, the cross-spectral density (CSD) matrix was calculated using discrete prolate spheroidal sequences ("dpss" taper in "mtmfft" method in ft_freqanalysis.m function) with all valid trials using baseline (-0.75 to -0.25 s) and delay period (0.2-2.2 s) epochs. The resulting spatial filter was used to localize sources in the frequency domain (computing power) using a beamformer dipole analysis with dynamic imaging of coherent sources ("dics" method in ft_sourceanalysis.m function) for frequencies from 2 to 100 Hz in steps of 2 Hz, with a smoothing of \pm 2Hz, and a 5% regularization. For each trial type, source activity (power) during the delay period was computed in 31 time windows of 0.5 s, in steps of 0.05 s, and baseline-corrected (relative change method) with the trial-averaged baseline epoch of 0.5 s (-0.75 to -0.25 s, equally spaced from the beginning and end of the 1s baseline). Finally, to normalize spatial maps across subjects, we transformed subject-specific maps of power from the head coordinate system to the MNI coordinate system using the inverse of the transformation matrix that was computed when generating the forward model.

Sources of interest for the DCM analysis were selected using a combined theory- and data-driven approach. Based on previous literature (Sreenivasan and D'Esposito, 2019), we predefined regions-of-interest (ROI) using the Brainnetome Atlas (Fan et al., 2016): superior frontal gyrus (SFG, areas 1-14), middle frontal gyrus (MFG, 15-26), intraparietal sulcus (IPS, 127-142), and early visual cortex (EVC, 203-206). Then, we used the time frequency matrices of power (31 time windows by 50 frequency bins) to select the sources of interest within each ROI. For each source, we performed a massive univariate one-way analysis of variance (ANOVA) with trial type as a factor with three levels (high accuracy, low accuracy, and swap trials), getting as output a time frequency matrix of F-statistics. Then, for each source, we summed the resulting uncorrected F-statistics over time and frequency, and selected the source with highest collapsed F-statistic within each ROI. All ROIs were considered separately for contralateral and ipsilateral hemispheres, except in SFG. There, given the proximity of the sources to the midline, we selected the source with the highest F-statistic bilaterally. Sources coordinates were (Figure 1B): SFG $[F_{\text{collapsed}} = 3,576; \text{MNI:} (8, 18, 68)], \text{MFG} [contralateral: <math>F_{\text{collapsed}}$ = 3,399, (-43, 30, 30); ipsilateral: $F_{\text{collapsed}}$ = 3,760, (45, 30, 30)], IPS [contralateral: $F_{\text{collapsed}} = 5,487, (-43, -58, 43)$; ipsilateral: $F_{\text{collapsed}} = 4,129, (58, -33, 55)$] and EVC [contralateral: $F_{\text{collapsed}}$ = 3,108, (-18, -95, 5); ipsilateral: $F_{\text{collapsed}} =$ 3,161, (20, -95, -8)]. Frontal and parietal nodes fall inside of a recent metaanalysis of fMRI studies about delay activity in WM (Li et al., 2022), while visual nodes fall in the occipital pole, from where previous studies decoded memory contents during delay period (Serences et al., 2009). Timecourses of selected sources of interest were reconstructed using a linear constrained minimum variance beamformer ("lcmv" method in FieldTrip) (Veen et al., 1997) with a fixed orientation projected along the first PCA axis.

DCM model

We used a DCM with cross-spectral density as data feature (range 2–100 Hz in steps of 1 Hz), as implemented in SPM12 (Friston et al., 2012; Moran et al., 2013), with the canonical microcircuit model (CMM) based on the 4-population conductance-based neural mass model (Moran et al., 2011; Bastos et al., 2012). The model comprises two components: the neuronal model characterizes the dynamics of hidden neural states,

$$\dot{x} = f(x, u, \theta) + \Gamma_{x}$$

and the observation model describes how neural activity gives rise to observed data,

$$y = h(C, x) + \epsilon$$

where *x* is the hidden neural state, *u* is the inputs to the neuronal population, θ are the model parameters, Γ_x are the stochastic noise terms in neuronal model, *y* is the observed MEG data, *h* is the

function mapping neural states to observed data, involving a lead matrix *C*, and ϵ are stochastic noise terms in the observation model.

In the conductance-based model used in the current study, neural dynamics are parametrized as the summed active and passive currents across the membrane (Pereira et al., 2021). Each neuronal population was modeled based on the Morris-Lecar model,

$$C\dot{V} = \sum_{k} g_k \left(V_K - V \right) + u + \Gamma_V$$

where C is the membrane capacitance, V is the membrane potential, g_k is the conductance of channel k, V_K is the reversal potential for channel k, u is the applied input current, Γ_V is the stochastic term modeling Gaussian noise. k comprises passive leak current, and active currents of excitatory (Na⁺) and inhibitory (Cl-) ion flow mediated by fast AMPA and GABAA receptors, depending on the type of neuronal population. Hence, membrane potential V and conductance of channels g_k comprise the hidden states x of the model in each neuronal population. Please refer to Pereira et al. (2021) for a full description of the mathematics of conductance based models, and see Moran et al. (2013) for a broad overview of neural mass models. Within each source (brain region), four neuronal populations are modeled: spiny stellate cells, superficial pyramidal cells, inhibitory interneurons, and deep pyramidal cells (Bastos et al., 2012). Neuronal populations within each source were connected as follows: (i) excitatory connections between excitatory spiny stellate and superficial pyramidal cells and from superficial pyramidal to deep pyramidal cells, (ii) recurrent connections between excitatory populations and the interlaminar population of inhibitory interneurons, and (iii) inhibitory selfconnections of each population, modeling synaptic gain (Friston et al., 2019).

The connectivity between sources, the seven brain regions described above, was modeled through the matrices A $[A \{1\}, A \{2\}]$, the intrinsic connectivity matrices representing the strength of feedforward $A\{1\}$ and feedback $A\{2\}$ connections between different neural sources (see model architecture below). Types of connections were defined based on the specific populations they projected to. Feedforward connections originated from superficial pyramidal in the projecting source to spiny stellate in the target source, feedback connections originated from deep pyramidal in the projecting source to both inhibitory interneurons and superficial pyramidal cells in the target source, and lateral (interhemispheric) connection comprised both feedforward and feedback connections. Since our focus of interest was not the comparison between different model architectures of extrinsic connectivity (inference on model architecture), but rather the performance-dependent modulation of effective connectivity (inference on model parameters), we defined a model architecture with the canonical functional hierarchy. Building upon research on the functional organization of brain networks (Bastos et al., 2015; Michalareas et al., 2016; Nee and D'Esposito, 2016; Gratton C. et al., 2018; Gratton G. et al., 2018; Marek and Dosenbach, 2018) and previous DCM studies about WM (Dima et al., 2014; Heinzel et al., 2017), prefrontal cortex was placed at the top and visual cortex at the bottom of the hierarchy, designating posterior-to-anterior connections as feedforward, anterior-to-posterior connections as feedback, and interhemispheric connections between regions at the same level of the hierarchy as lateral connections (Figure 1C). Finally, the experimental effects of interests, the baseline corrected estimates of effective connectivity during delay period, were modeled with the matrix B, comprising all the connections defined in matrices A: 12 feedforward, 12 feedback, and six lateral connections (Figure 1C). Note that there is no circularity in using the same data to define ROIs and to study their interactions with DCM, since the purpose of DCM is to test hypothesis about the mechanisms underlying the experimental effects observed with conventional analysis (Stephan et al., 2010). Massive univariate statistical analyses (e.g., statistical parametric maps) to define ROIS are standard practice in several DCM studies (Moran et al., 2011; Dima et al., 2014; Auksztulewicz and Friston, 2015; Heinzel et al., 2017; Jung et al., 2018; Friston et al., 2019; Adams et al., 2021).

Model inversion and comparison

For each subject, we inverted two DCMs for each trial type (high accuracy trials, low accuracy trials, and swap trials), encoding the baseline corrected estimates of effective connectivity (matrix B) during early and late delay. All time intervals had a length of 800 ms [baseline (-900 to -100 ms), early delay (400-1,200 ms), and late delay (1,200-2,000 ms)], and were separated from stimulus presentation (0 to 200 ms) or report (from 2,200 ms) by at least 100 ms. The separation of delay period in early and late time windows was motivated by previous studies in nonhuman primates showing that the dynamic properties of WM coding were stronger during the initial portions of the memory delay (Murray et al., 2017), and the modulation of memory load showed distinct spectral profiles during the early and late portions of the memory delay (Buschman et al., 2011) as well as differences in effective connectivity in frontoparietal networks (Pinotsis et al., 2019). Models were inverted to fit the cross-spectral density through the optimization of variational free energy under Laplace approximation (Friston et al., 2007) to optimize log scaling parameters around default priors (spm_fx_cmm.m). Please see Zeidman et al. (2023) for a comprehensive technical review of the Bayesian inference scheme employed to fit DCM models.

Statistics over DCM parameters

We performed statistical inference on the parameters of the optimized DCM model using PEB (Friston et al., 2015; Zeidman et al., 2019). Compared to standard statistical inference, PEB has the advantage of taking into account the estimated uncertainty (variance) about model parameters, downweighing the contribution of less certain parameter estimates on group effects. The PEB approach is also advantageous in the context of model inversion and comparison, since the inversion of the same "full" DCM per subject/condition avoids the issue of fitting multiple models and performing Bayesian model comparison over models, which can lead to different DCMs falling into distinct local optima. PEB is a Bayesian hierarchical model over parameters with individual DCMs as a first level, and group effects on DCM parameters as a second level (see Zeidman et al., 2019 for details).

$$Y_i = \Gamma_i \left(\theta_i^{(1)} \right) + \varepsilon_i^{(1)} \tag{1}$$

$$\theta^{(1)} = X\theta^{(2)} + \varepsilon_i^{(2)} \tag{2}$$

The first level (Equation 1) represents the observed MEG data Y of the *i*-th subject modeled with the dynamic causal model Γ with parameters $\theta_i^{(1)}$ and noise term $\varepsilon_i^{(1)}$. The second level (Equation 2) shows that parameters of interest from first level $\theta^{(1)}$ are modeled at group level through a GLM with design matrix X and grouplevel parameters $\theta_i^{(2)}$, with zero-mean noise to account for betweensubject variability as random effects $\varepsilon_i^{(2)}$. In our study, parameters of interest in PEB analysis (within-subjects design matrix) were the baseline corrected estimates of extrinsic (between-source) effective connectivity (matrix B). We performed pairwise comparisons between trial types with a group-level design matrix with an intercept (column of 1's) encoding commonalities, and a regressor of interest encoding the pairwise contrast between trial types (e.g., high accuracy vs. low accuracy). We report PEB parameters with strong evidence (posterior probability, Pp > 0.95) (Kass and Raftery, 1995), and a large effect size, assessed through its predictive accuracy in a leave-one-out cross-validation (spm_dcm_loo.m). In this procedure, a PEB model was constructed with all subjects but one, predicting the excluded subject's covariates, and iterated through each participant serving as the left-out subject. The predictive accuracy was quantified as Pearson's correlation between PEB covariates and cross-validation predictions. We applied a statistical threshold of p = 0.05, uncorrected.

Data and code availability

All scripts to invert DCMs and perform PEB analysis are available online: https://github.com/asantoangles/dcm_working_ memory. Source timecourses used as input for DCM, in addition to inverted DCMs and PEB results, are available on Open Science Framework: https://osf.io/hbgu2/?view_only= 8bc743e2068d47f0889f7b556beb2068.

Results

Behavioral results

Absolute mean error was $24.4^{\circ} \pm 5.17^{\circ}$ (14.2° , 34.5°) for all trials, indicating that subjects were successfully able to complete the task. As expected by the definition of trial types, performance in high accuracy trials $[11.2^{\circ} \pm 1.9^{\circ} (7.4^{\circ}, 16^{\circ})]$ was better than low accuracy trials $[21.5^{\circ} \pm 4.9^{\circ} [(14.3^{\circ}, 33.8^{\circ});$ two-sample Wilcoxon test: W = 354, z = -6.12, p < 0.0001, effect size r = 0.85], and swap trials $[51.6^{\circ} \pm 5.2^{\circ} (43.4^{\circ}, 68^{\circ}); W = 351, z = -6.17, p < 0.0001, r = 0.85]$. Performance in low accuracy trials was significantly better than swap trials (W = 351, z = -6.17, p < 0.0001, r = 0.85). Additionally, we observed no significant differences in absolute mean error between trials with left vs. right hemifield stimulus presentation [W = 734, p = 0.8; left: $25.2^{\circ} \pm 5.7^{\circ} (13.2^{\circ}, 36.5^{\circ})$, right: $23.7^{\circ} \pm 4.9^{\circ} (13.9^{\circ}, 32.5^{\circ})$] and no difference in the proportion of left/right trials observed in distinct trial types (high

accuracy trials: 0.49/0.51, low accuracy trials: 0.5/0.5, swap trials: 0.51/0.49; all pairwise comparisons between trials with a Chi-square test with Yates continuity yielded *p*-values > 0.8). The similar performance across hemifields helped justify our decision to flip sensors across the midline for MEG analysis (see Methods).

DCM model fit

Correlation between empirical and predicted amplitude (modulus) was excellent ($r = 0.9907 \pm 0.004$). Model predictions of within-ROI power spectral density (PSD) matched empirical data in all frequencies. Between-ROIs cross-spectral density (CSD) predictions were below the empirical values, although empirical CSD differences between trial types were generally reproduced by model predictions (Supplementary Figure S1B). Coefficient of determination also showed a good model fit [$R^2 = 0.97 \pm 0.01$ (0.46–0.99)]. PEB results summarized below describe the effects with strong evidence (Pp > 0.95) an out-of-sample predictive accuracy with p < 0.05, uncorrected (see Supplementary material for a detailed description of PEB results).

PEB results

Low accuracy trials, in comparison with high accuracy trials (Figure 2A), during early delay, showed reduced feedback signals in ipsilateral frontooccipital and medial to superior frontal connections, and increased interhemispheric connections between parietal and frontal areas. During late delay, low accuracy trials showed reduced feedforward signals in ipsilateral frontoparietal and frontooccipital connections, and reduced lateral connections between early visual areas. Swap trials, in comparison with high accuracy trials (Figure 2B), during early delay, showed reduced feedback connectivity in ipsilateral frontooccipital and increased superior frontal to ipsilateral parietal. Feedback connectivity was also reduced in superior frontal to contralateral early visual cortex during the whole delay period. Low accuracy trials, in comparison with swap trials (Figure 2C), during early delay, showed reduced feedback ipsilateral frontoparietal connectivity and reduced feedforward connectivity in ipsilateral early visual cortex to superior frontal. During late delay, low accuracy trials showed increased feedback connectivity in several frontooccipital connections: ipsilateral middle frontal to early visual cortex, and superior frontal to ipsilateral and contralateral visual areas.

Discussion

In the current study, we report evidence that WM involves changes in effective connectivity across a network involving frontal, parietal, and visual areas; and demonstrated that these changes were performance-dependent. In comparison with high accuracy trials, low accuracy and swap error trials showed distinct patterns of effective connectivity. In early delay, both types of error trials showed reduced feedback connections in frontooccipital connections, but in late delay the pattern was qualitatively distinct: swap errors kept showing reduced frontooccipital connectivity, while low accuracy trials showed reduced feedforward connections emerging from early visual cortex and parietal to frontal areas. Taken together, our findings suggest that WM recruits a broad network of regions for optimal performance, and that errors in performance are associated with distinct systems-level reconfigurations of these networks, in agreement with the distributed view of WM (Lorenc and Sreenivasan, 2021; Mejías and Wang, 2022).

Cognitive functions emerge from a flexible reconfiguration of large-scale brain functional networks in response to task contingencies (Cohen and D'Esposito, 2016), with the integration of frontoparietal and visual networks (Spadone et al., 2015; Hearne et al., 2017). Here, we expand these previous studies by showing that large-scale network reconfigurations are performance dependent and actively involve early visual areas during WM maintenance. Our finding of reduced prefrontal feedback connectivity in error trials, both low accuracy and swaps, agrees with a DCM study of local field potentials in non-human primates during a change detection task (Pinotsis et al., 2019), where the authors reported an association between impaired behavioral performance and the break-down of prefrontal prediction signals in a network comprising prefrontal, frontal eyes field and lateral intraparietal area. We also observed reduced feedback connections between frontal nodes (middle to superior frontal in high vs. low accuracy), but most of the disrupted prefrontal feedback signals propagated to early visual cortex. Moreover, the specific connections involved depended on the type of error trials. While both low accuracy and swap error trials showed reduced feedback connections from middle frontal areas, feedback connections from superior frontal were also reduced in swap errors, but not in low accuracy errors. This pattern of performance-dependent disruption of feedback signals is consistent with the view that visual cortical engagement during WM at least partially reflects feedback signals from frontoparietal brain regions (Leavitt et al., 2017; Xu, 2020). At the same time, our observation of reduced feedforward signals in frontooccipital connections in low accuracy trials suggest that the role of visual cortex during the delay period is not solely that of a passive recipient of feedback signals, consistent with the sensory recruitment hypothesis of WM (Pasternak and Greenlee, 2005; D'Esposito, 2007; Scimeca et al., 2018).

Our findings revealed that network activity supporting WM maintenance conveys both feedforward and feedback signals involving not only frontoparietal networks (Dima et al., 2014; Heinzel et al., 2017; Jung et al., 2018), but also frontooccipital networks. Moreover, feedforward/feedback frontooccipital connections discriminated between types of errors. Swap error trials showed disrupted feedback connections to early visual areas, while low accuracy errors deployed a time-dependent pattern of effective connectivity changes of disrupted feedback (feedforward) signals in early (late) delay. This time-dependent disruption of effective connectivity in frontooccipital networks aligns with previous studies showing that the dynamic and spectral properties of WM coding differ between the early and late delay period (Buschman et al., 2011; Murray et al., 2017; Pinotsis et al., 2019). Moreover, it underscores the dynamic interplay between feedforward and feedback signals between early sensory areas

and prefrontal regions, suggesting that WM is not a static storage system but a dynamically competitive interplay between sensory and cognitive control processes during different phases of WM maintenance. Interestingly, most of these connections involve (medial) superior frontal, a region previously associated in spatial WM (Courtney et al., 1998; Rottschy et al., 2012, 2013). Indeed, our results also suggest an alternative interpretation for previous findings of functional coupling of frontomedial theta (and alpha) with posterior brain regions during WM maintenance (Payne and Kounios, 2009; Hsieh and Ranganath, 2014; Johnson et al., 2017; Riddle et al., 2020; Parto Dezfouli et al., 2021; Ratcliffe et al., 2022). Although most studies have used correlational methods, this frontomedial activation during WM has been largely interpreted as control signals to posterior brain regions where information is stored (D'Esposito and Postle, 2015). Nevertheless, we found that the direction of effective connectivity involving medial prefrontal regions during WM depends on behavioral performance: medial PFC was the source of disrupted feedback signals in swap errors, but the target of disrupted feedforward signals in low accuracy errors.

Another distinction between types of errors was observed in the interhemispheric connectivity. Low accuracy trials showed increased lateral connectivity between frontal and parietal areas, which might be interpreted as system-level compensatory responses displayed in response to the absence of appropriate communication between occipital and frontal areas. Indeed, previous studies have shown that the experimental manipulation of interparietal communication affects WM performance (Tseng et al., 2018; Grover et al., 2022). Alternatively, it is plausible that an excessive level of interhemispheric communication has detrimental effects in WM performance (Koshy et al., 2020). On the contrary, swap errors showed no disruptions in lateral connections, consistent with the notion that swaps are genuine transposition errors arising from binding issues (Schneegans and Bays, 2017, 2019) and not merely educated guesses (Pratte, 2019; Huang, 2020). Otherwise, we would not see differences between types of error trials. Moreover, the absence of disrupted frontooccipital feedforward signals in swap trials aligns with the notion that transposition errors arise from incorrect responses being confidently held in WM (Schneegans and Bays, 2019). Memory contents can be decoded from early visual cortex in correct trials (Sreenivasan et al., 2014b; Curtis and Sprague, 2021; Lorenc and Sreenivasan, 2021), as well as in swap trials (Mallett et al., 2022). However, in swap trials the decoded information corresponds to the swapped location, not the cued location, suggesting the maintenance of incorrect memory items (Mallett et al., 2022). If we assume that feedforward signals propagating from early visual cortex during delay period distribute memory contents throughout the network, the non-disruption of feedforward signals in swaps might suggest that the incorrect information hold in early visual areas are propagated through the WM network in a similar way than in correct trials.

Empirical cross-spectral density revealed that coherence between sources accounted for most differences between trial types (Supplementary Figure S1A), despite the fact that sources were selected based on local power, supporting the behavioral relevance of large-scale communication through synchronization (Roux et al., 2012; Lisman and Jensen, 2013). Although trial type differences in coherence for some connections spanned the whole range of analyzed frequencies (2–100 Hz), the most striking differences were concentrated in low frequency bands, particularly around theta (6–8 Hz) and alpha (8–12 Hz), where we observed peaks in coherence (Supplementary Figure S1). These findings agrees with recent evidence that theta frequency synchronous tACS applied across frontoparietal network improves working memory performance, possibly by enhancing coherence between distant nodes of the network (Polanía et al., 2012; Violante et al., 2017; Biel et al., 2022; Grover et al., 2022).

Here, we acknowledge and address several potential limitations in our study. First, our definition of error trials was not entirely consistent with subjects' behavior, as low accuracy or swap trials were not exclusively determined by all three responses being misplaced or involving transposed reports, suggesting that not all memory contents was maintained poorly. Nevertheless, despite this limitation, we observed robust differences in effective connectivity between trial types, underscoring the robustness and validity of our findings. Second, the scope of our results is confined by our deliberate decision to limit the model architecture to frontoparietal and frontooccipital networks. Future investigations may consider incorporating nodes such as frontotemporal regions (Braun et al., 2015; Alenazi et al., 2022), salience and cinguloopercular regions (Cai et al., 2021), and exploring their interactions with the basal ganglia (McNab and Klingberg, 2008; Voytek and Knight, 2010). Third, our source reconstruction was performed without high-resolution structural images in a subset of subjects, thereby constraining our ability to make precise claims about the spatial location of sources. To address this spatial uncertainty, we employed predefined regions of interest and subsequently selected sources of interest within these regions that exhibited the most substantial main effect of trial type. To ensure the robustness of our source selection, we performed a supplementary analysis. Source reconstruction was carried out without structural images for all subjects, and euclidean distances between selected sources with and without a subset of participants with structural images were computed. The resulting distances of $11.2 \pm 8.1 \ (0-18.4) \ (mm)$ were deemed negligible, given the spatial resolution of sourcereconstructed MEG signals. Forth, an important consideration in our study is the observed imbalance in the number of trials between different trial types. However, the implementation of the parametric empirical Bayes (PEB) approach plays a pivotal role in mitigating the potential influence of this trial number asymmetry on our results. Unlike direct comparisons of effective connectivity estimates between trial types, our approach involved contrasting baseline-corrected estimates (matrix B). By focusing on these baseline-corrected values, our intention was to minimize the impact of trial number imbalances. This method emphasizes relative changes in connectivity patterns rather than absolute values, thereby contributing to the robustness of our findings. Fifth, the study utilized data from 26 subjects, comprising four women and 22 men. This gender mismatch presents a limitation in our investigation. Despite studies that have shown small (but significant) gender differences in WM performance, with the male/female advantage depending on domain and task (Voyer

et al., 2017, 2021), there is no strong a priori reason to assume that similar behavioral performance should be supported by either the same or distinct underlying neural processes. However, gender differences in the functional dynamics supporting cognition are understudied. Preliminary evidence from task-based MEG studies in children and adolescents indicates potential sex-specific developmental effects during various cognitive tasks (Embury et al., 2019; Killanin et al., 2020, 2022; Taylor et al., 2020; Fung et al., 2021). MEG studies in adults have also shown gender differences neural responses during spatial navigation (Pu et al., 2020), face processing (Proverbio, 2021), and resting-state networks throughout the lifespan (Rempe et al., 2023; Stier et al., 2023). However, little is known about gender differences in effective connectivity, a gap we were unable to address with our current sample size. Future research should intentionally incorporate a more balanced representation of gender to elucidate potential gender-related effects on effective connectivity.

In conclusion, our study reveals performance-dependent reconfigurations in a large-scale brain network supporting WM. The network activity related to maintenance during the delay period conveys both feedforward and feedback signals, not confined solely to frontoparietal networks as depicted in the early work of Goldman-Rakic and colleagues. Instead, our findings extend this understanding to encompass frontooccipital networks, thereby emphasizing the crucial involvement of sensory areas in the dynamic processes of WM. Moreover, the disruption of these dynamics give rise to distinct types of behavioral errors. Low accuracy trials showed disrupted feedback (feedforward) signals during early (late) delay, while swap errors showed disrupted feedback signals during the whole delay period, aligning with the distributed view of WM.

Data availability statement

Source timecourses used as input for DCM, in addition to inverted DCMs and PEB results, are available on Open Science Framework: https://osf.io/hbgu2/?view_only= 8bc743e2068d47f0889f7b556beb2068.

Ethics statement

The studies involving humans were approved by New York University Abu Dhabi Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AS-A: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft. AT: Data curation, Investigation, Methodology, Project administration. VB: Data curation, Investigation. KS: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the NYU Research Enhancement Fund and the NYUAD Center for Brain and Health, funded by Tamkeen under NYUAD Research Institute grant CG012 (KS).

Acknowledgments

This research was conducted using resources provided by the High Performance Computing Cluster and the Core Technology Platform at NYUAD.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum. 2024.1339728/full#supplementary-material

References

Ablin, P., Cardoso, J.-F., and Gramfort, A. (2018). Faster independent component analysis by preconditioning with hessian approximations. *IEEE Trans. Signal Process.* 66, 4040–4049. doi: 10.1109/TSP.2018.2844203

Adams, R. A., Pinotsis, D., Tsirlis, K., Unruh, L., Mahajan, A., Horas, A. M., et al. (2021). Computational modeling of electroencephalography and functional magnetic resonance imaging paradigms indicates a consistent loss of pyramidal cell

synaptic gain in schizophrenia. *Biol. Psychiatry* 89, 1–52. doi: 10.1016/j.biopsych.2021. 02.249

Alenazi, M. F., Al-Joudi, H., Alotaibi, F., Bracewell, M., Dundon, N. M., Katshu, M. Z. U. H., et al. (2022). Spatial binding impairments in visual working memory following temporal lobectomy. *Eneuro* 9:ENEURO.0278-21.2022. doi: 10.1523/ENEURO.0278-21.2022

Auksztulewicz, R., and Friston, K. (2015). Attentional enhancement of auditory mismatch responses: a DCM/MEG study. *Cereb. Cortex* 25, 4273–4283. doi: 10.1093/cercor/bhu323

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., Friston, K. J., et al. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711. doi: 10.1016/j.neuron.2012.10.038

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., et al. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* 85, 390–401. doi: 10.1016/j.neuron.2014. 12.018

Bays, P. M. (2016). Evaluating and excluding swap errors in analogue tests of working memory. *Sci. Rep.* 6:19203. doi: 10.1038/srep19203

Bays, P. M., Catalao, R. F. G., and Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. J. Vis. 9, 1–11. doi: 10.1167/9.10.7

Biel, A. L., Sterner, E., Röll, L., and Sauseng, P. (2022). Modulating verbal working memory with fronto-parietal transcranial electric stimulation at theta frequency: does it work? *Eur. J. Neurosci.* 55, 405–425. doi: 10.1111/ejn.15563

Brainard, D. H. (1997). The psychophysics toolbox. Spatial Vis. 10, 433–436. doi: 10.1163/156856897X00357

Braun, U., Schäfer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., et al. (2015). Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11678–11683. doi: 10.1073/pnas.1422487112

Briggs, R. G., Khan, A. B., Chakraborty, A. R., Abraham, C. J., Anderson, C. D., Karas, P. J., et al. (2020). Anatomy and white matter connections of the superior frontal gyrus. *Clin. Anat.* 33, 823–832. doi: 10.1002/ca.23523

Buschman, T. J., Siegel, M., Roy, J. E., and Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11252–11255. doi: 10.1073/pnas.1104666108

Cai, W., Ryali, S., Pasumarthy, R., Talasila, V., and Menon, V. (2021). Dynamic causal brain circuits during working memory and their functional controllability. *Nat. Commun.* 12:3314. doi: 10.1038/s41467-021-23509-x

Cai, Y., Fulvio, J. M., Yu, Q., Sheldon, A. D., and Postle, B. R. (2020). The role of location-context binding in nonspatial visual working memory. *eNeuro* 7, 1–14. doi: 10.1523/ENEURO.0430-20.2020

Cavada, C., and Goldman-Rakic, P.S. (1989). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J. Comp. Neurol.* 287, 422–445. doi: 10.1002/cne.902870403

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., and Haynes, J.-D. (2017). The distributed nature of working memory. *Trends Cogn. Sci.* 21, 111–124. doi: 10.1016/j.tics.2016.12.007

CIE (2004). Colorimetry, 3rd ed. Vienna: CIE Central Bureau.

Cohen, J. R., and D'Esposito, M. (2016). The segregation and integration of distinct brain networks and their relationship to cognition. *J. Neurosci.* 36, 12083–12094. doi: 10.1523/JNEUROSCI.2965-15.2016

Courtney, S. M., and Petit, L. Maisog, J.Ma., Ungerleider, L.G., Haxby, J.V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science* 279, 1347–1351. doi: 10.1126/science.279.5355.1347

Crespo-Garcia, M., Pinal, D., Cantero, J. L., Díaz, F., Zurrón, M., Atienza, M., et al. (2013). Working memory processes are mediated by local and long-range synchronization of alpha oscillations. *J. Cogn. Neurosci.* 25, 1343–1357. doi: 10.1162/jocn_a_00379

Curtis, C. E., and Sprague, T. C. (2021). Persistent activity during working memory from front to back. *Front. Neural Circuits* 15:696060. doi: 10.3389/fncir.2021.696060

Daniel, T. A., Katz, J. S., and Robinson, J. L. (2016). Delayed match-tosample in working memory: a BrainMap meta-analysis. *Biol. Psychol.* 120, 10–20. doi: 10.1016/j.biopsycho.2016.07.015

David, O., and Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage* 20, 1743–1755. doi: 10.1016/j.neuroimage.2003.07.015

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philos.* Trans. R. Soc. B: Biol. Sci. 362, 761–772. doi: 10.1098/rstb.2007.2086

D'Esposito, M., and Postle, B. R. (2015). The cognitive neuroscience of working memory. Ann. Rev. Psychol. 66, 115–142. doi: 10.1146/annurev-psych-010814-015031

Dima, D., Jogia, J., and Frangou, S. (2014). Dynamic causal modeling of load-dependent modulation of effective connectivity within the verbal working memory network. *Hum. Brain Mapp.* 35, 3025–3035. doi: 10.1002/hbm. 22382

Embury, C. M., Wiesman, A. I., Proskovec, A. L., Mills, M. S., Heinrichs-Graham, E., Wang, Y.-P., et al. (2019). Neural dynamics of verbal working memory processing in children and adolescents. *NeuroImage* 185, 191–197. doi: 10.1016/j.neuroimage.2018.10.038

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., et al. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cereb. Cortex* 26, 3508–3526. doi: 10.1093/cercor/bhw157

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W. (2007). Variational free energy and the Laplace approximation. *NeuroImage* 34, 220–234. doi: 10.1016/j.neuroimage.2006.08.035

Friston, K., Zeidman, P., and Litvak, V. (2015). Empirical bayes for DCM: a group inversion scheme. *Front. Syst. Neurosci.* 9:164. doi: 10.3389/fnsys.2015.00164

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36. doi: 10.1089/brain.2011.0008

Friston, K. J., Bastos, A., Litvak, V., Stephan, K. E., Fries, P., Moran, R. J., et al. (2012). DCM for complex-valued data: cross-spectra, coherence and phase-delays. *NeuroImage* 59, 439–455. doi: 10.1016/j.neuroimage.2011.07.048

Friston, K. J., Preller, K. H., Mathys, C., Cagnan, H., Heinzle, J., Razi, A., et al. (2019). Dynamic causal modelling revisited. *NeuroImage* 199, 730–744. doi: 10.1016/j.neuroimage.2017.02.045

Fung, M. H., Taylor, B. K., Lew, B. J., Frenzel, M. R., Eastman, J. A., Wang, Y.-P., et al. (2021). Sexually dimorphic development in the cortical oscillatory dynamics serving early visual processing. *Dev. Cogn. Neurosci.* 50:100968. doi: 10.1016/j.dcn.2021.100968

Furmanski, C. S., and Engel, S. A. (2000). An oblique effect in human primary visual cortex. *Nat. Neurosci.* 3, 535–536. doi: 10.1038/75702

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. Neuron 14, 477-485. doi: 10.1016/0896-6273(95)90304-6

Gosseries, O., Yu, Q., Larocque, J. J., Starrett, M. J., Rose, N. S., Cowan, N., et al. (2018). Parietal-occipital interactions underlying control-and representation-related processes in working memory for nonspatial visual features. *J. Neurosci.* 38, 4357–4366. doi: 10.1523/JNEUROSCI.2747-17.2018

Gratton, C., Sun, H., and Petersen, S. E. (2018). Control networks and hubs. *Psychophysiology* 55, 1–18. doi: 10.1111/psyp.13032

Gratton, G., Cooper, P., Fabiani, M., Carter, C. S., and Karayanidis, F. (2018). Dynamics of cognitive control: theoretical bases, paradigms, and a view for the future. *Psychophysiology* 55, 1–29. doi: 10.1111/psyp.13016

Grover, S., Wen, W., Viswanathan, V., Gill, C. T., and Reinhart, R. M. G. (2022). Long-lasting, dissociable improvements in working memory and long-term memory in older adults with repetitive neuromodulation. *Nat. Neurosci.* 25, 1237–1246. doi: 10.1038/s41593-022-01132-3

Hearne, L. J., Cocchi, L., Zalesky, A., and Mattingley, J. B. (2017). Reconfiguration of brain network architectures between resting-state and complexity-dependent cognitive reasoning. *J. Neurosci.* 37, 8399–8411. doi: 10.1523/JNEUROSCI.0485-17.2017

Heinzel, S., Lorenz, R. C., Duong, Q. L., Rapp, M. A., and Deserno, L. (2017). Prefrontal-parietal effective connectivity during working memory in older adults. *Neurobiol. Aging* 57, 18–27. doi: 10.1016/j.neurobiolaging.2017.05.005

Hsieh, L. T., and Ranganath, C. (2014). Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *NeuroImage* 85, 721–729. doi: 10.1016/j.neuroimage.2013.08.003

Huang, L. (2020). Distinguishing target biases and strategic guesses in visual working memory. *Atten. Percept. Psychophys.* 82, 1258–1270. doi: 10.3758/s13414-019-01913-2

Jaiswal, A., Nenonen, J., Stenroos, M., Gramfort, A., Dalal, S. S., Westner, B. U., et al. (2020). Comparison of beamformer implementations for MEG source localization. *NeuroImage* 216, 116797–116797. doi: 10.1016/j.neuroimage.2020.116797

Johnson, E. L., Dewar, C. D., Solbakk, A. K., Endestad, T., Meling, T. R., Knight, R. T., et al. (2017). Bidirectional frontoparietal oscillatory systems support working memory. *Curr. Biol.* 27, 1829–1835.e4. doi: 10.1016/j.cub.2017.05.046

Jung, K., Friston, K. J., Pae, C., Choi, H. H., Tak, S., Choi, Y. K., et al. (2018). Effective connectivity during working memory and resting states: a DCM study. *NeuroImage* 169, 485–495. doi: 10.1016/j.neuroimage.2017.12.067

Kass, R. E., and Raftery, A. E. (1995). Bayes factors. J. Am. Stat. Assoc. 90, 773–795. doi: 10.1080/01621459.1995.10476572

Kessels, R. P. C., D'Alfonso, A. A. L., Postma, A., and De Haan, E. H. F. (2000). Spatial working memory performance after high-frequency repetitive transcranial magnetic stimulation of the left and right posterior parietal cortex in humans. *Neurosci. Lett.* 287, 68–70. doi: 10.1016/S0304-3940(00)01146-0

Killanin, A. D., Embury, C. M., Picci, G., Heinrichs-Graham, E., Wang, Y.-P., Calhoun, V. D., et al. (2022). Trauma moderates the development of the oscillatory dynamics serving working memory in a sex-specific manner. *Cereb. Cortex* 32, 5206–5215. doi: 10.1093/cercor/bhac008

Killanin, A. D., Wiesman, A. I., Heinrichs-Graham, E., Groff, B. R., Frenzel, M. R., Eastman, J. A., et al. (2020). Development and sex modulate visuospatial oscillatory dynamics in typically-developing children and adolescents. *NeuroImage* 221:117192. doi: 10.1016/j.neuroimage.2020.117192

Koshy, S. M., Wiesman, A. I., Proskovec, A. L., Embury, C. M., Schantell, M. D., Eastman, J. A., et al. (2020). Numerical working memory alters alpha-beta

oscillations and connectivity in the parietal cortices. *Hum. Brain Mapp.* 41, 3709–3719. doi: 10.1002/hbm.25043

Leavitt, M. L., Mendoza-Halliday, D., and Martinez-Trujillo, J. C. (2017). Sustained activity encoding working memories: not fully distributed. *Trends Neurosci.* 40, 328–346. doi: 10.1016/j.tins.2017.04.004

Li, X., O'Sullivan, M. J., and Mattingley, J. B. (2022). Delay activity during visual working memory: a meta-analysis of 30 fMRI experiments. *NeuroImage* 255:119204. doi: 10.1016/j.neuroimage.2022.119204

Lisman, J. E., and Jensen, O. (2013). The theta-gamma neural code. *Neuron* 77, 1002–1016. doi: 10.1016/j.neuron.2013.03.007

Lorenc, E. S., and Sreenivasan, K. K. (2021). Reframing the debate: The distributed systems view of working memory. Vis. Cogn. 29, 416-424. doi: 10.1080/13506285.2021.1899091

Lugtmeijer, S., Schneegans, S., Lammers, N. A., Geerligs, L., de Leeuw, F. E., and de Haan, R. P. C. (2021). Consequence of stroke for feature recall and binding in visual working memory. *Neurobiol. Learn. Mem.* 179, 107387–107387. doi: 10.1016/j.nlm.2021.107387

Mallett, R., Lorenc, E. S., and Lewis-Peacock, J. A. (2022). Working memory swap errors have identifiable neural representations. *J. Cogn. Neurosci.* 34, 776–786. doi: 10.1162/jocn_a_01831

Marek, S., and Dosenbach, N. U. F. (2018). The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. *Dialogues Clin. Neurosci.* 20, 133–141. doi: 10.31887/DCNS.2018.20.2/smarek

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

Mejías, J. F., and Wang, X. J. (2022). Mechanisms of distributed working memory in a large-scale network of macaque neocortex. *eLife* 11, 1–33. doi: 10.7554/eLife.72136

Michalareas, G., and Vezoli, J. van Pelt, S., Schoffelen, J.-M., Kennedy, H., Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron* 89, 384–397. doi: 10.1016/j.neuron.2015.12.018

Moran, R., Pinotsis, D. A., and Friston, K. (2013). Neural masses and fields in dynamic causal modelling. *Front. Comput. Neurosci.* 7, 1–12. doi: 10.3389/fncom.2013.00057

Moran, R. J., Stephan, K. E., Dolan, R. J., and Friston, K. J. (2011). Consistent spectral predictors for dynamic causal models of steady-state responses. *NeuroImage* 55, 1694–1708. doi: 10.1016/j.neuroimage.2011.01.012

Murray, J. D., Bernacchia, A., Roy, N. A., Constantinidis, C., Romo, R., Wang, X.-J., et al. (2017). Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 114, 394–399. doi: 10.1073/pnas.1619449114

Nee, D. E., and D'Esposito, M. (2016). The hierarchical organization of the lateral prefrontal cortex. *eLife* 5, 1–26. doi: 10.7554/eLife.12112

Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoenchephalography forward calculation in realistic volume conductors. *Phys. Med. Biol.* 48, 3637–3652. doi: 10.1088/0031-9155/48/22/002

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011:156869. doi: 10.1155/2011/156869

Oostenveld, R., Stegeman, D. F., Praamstra, P., and Van Oosterom, A. (2003). Brain symmetry and topographic analysis of lateralized event-related potentials. *Clin. Neurophysiol.* 114, 1194–1202. doi: 10.1016/S1388-2457(03)00059-2

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

Parto Dezfouli, M., Davoudi, S., Knight, R. T., Daliri, M. R., and Johnson, E. L. (2021). Prefrontal lesions disrupt oscillatory signatures of spatiotemporal integration in working memory. *Cortex* 138, 113–126. doi: 10.1016/j.cortex.2021.01.016

Pasternak, T., and Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 6, 97-107. doi: 10.1038/nrn1603

Payne, L., and Kounios, J. (2009). Coherent oscillatory networks supporting short-term memory retention. *Brain Res.* 1247, 126–132. doi: 10.1016/j.brainres.2008.09.095

Pereira, I., Frässle, S., Heinzle, J., Schöbi, D., Do, C. T., Gruber, M., et al. (2021). Conductance-based dynamic causal modeling: a mathematical review of its application to cross-power spectral densities. *NeuroImage* 245, 118662–118662. doi: 10.1016/j.neuroimage.2021.118662

Pinotsis, D. A., Buschman, T. J., and Miller, E. K. (2019). Working memory load modulates neuronal coupling. *Cereb. Cortex* 29, 1670–1681. doi: 10.1093/cercor/bhy065

Polanía, R., Nitsche, M. A., Korman, C., Batsikadze, G., and Paulus, W. (2012). The importance of timing in segregated theta phase-coupling for cognitive performance. *Curr. Biol.* 22, 1314–1318. doi: 10.1016/j.cub.2012.05.021

Postle, B. R., Ferrarelli, F., Hamidi, M., Feredoes, E., Massimini, M., Peterson, M., et al. (2006). (2006). Repetitive transcranial magnetic stimulation dissociates working

memory manipulation from retention functions in the prefrontal, but not posterior parietal, cortex. J. Cogn. Neurosci. 18, 1712–1722. doi: 10.1162/jocn.2006.18.10.1712

Pratte, M. S. (2019). Swap errors in spatial working memory are guesses. *Psychon. Bull. Rev.* 26, 958–966. doi: 10.3758/s13423-018-1524-8

Proverbio, A. M. (2021). Sexual dimorphism in hemispheric processing of faces in humans: a meta-analysis of 817 cases. *Soc. Cogn. Affect. Neurosci.* 16, 1023–1035. doi: 10.1093/scan/nsab043

Pu, Y., Cornwell, B. R., Cheyne, D., and Johnson, B. W. (2020). Gender differences in navigation performance are associated with differential theta and high-gamma activities in the hippocampus and parahippocampus. *Behav. Brain Res.* 391:112664. doi: 10.1016/j.bbr.2020.112664

Ratcliffe, O., Shapiro, K., and Staresina, B. P. (2022). Fronto-medial theta coordinates posterior maintenance of working memory content. *Curr. Biol.* 32, 2121–2129.e3. doi: 10.1016/j.cub.2022.03.045

Rempe, M. P., Ott, L. R., Picci, G., Penhale, S. H., Christopher-Hayes, N. J., Lew, B. J., et al. (2023). Spontaneous cortical dynamics from the first years to the golden years. *Proc. Natl. Acad. Sci. U.S.A.* 120:e2212776120. doi: 10.1073/pnas.2212776120

Rezayat, E., Clark, K., Dehaqani, M.-R. A., and Noudoost, B. (2022). Dependence of working memory on coordinated activity across brain areas. *Front. Syst. Neurosci.* 15:787316. doi: 10.3389/fnsys.2021.787316

Riddle, J., Scimeca, J. M., Cellier, D., Dhanani, S., and D'Esposito, M. (2020). Causal evidence for a role of theta and alpha oscillations in the control of working memory. *Curr. Biol.* 30, 1748–1754.e4. doi: 10.1016/j.cub.2020.02.065

Rottschy, C., Caspers, S., Roski, C., Reetz, K., Dogan, I., Schulz, J. B., et al. (2013). Differentiated parietal connectivity of frontal regions for "what" and "where" memory. *Brain Struct. Funct.* 218, 1551–1567. doi: 10.1007/s00429-012-0476-4

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., et al. (2012). Modelling neural correlates of working memory: a coordinate-based meta-analysis. *NeuroImage* 60, 830-846. doi: 10.1016/j.neuroimage.2011.11.050

Roux, F., Wibral, M., Mohr, H. M., Singer, W., and Uhlhaas, P. J. (2012). Gamma-band activity in human prefrontal cortex codes for the number of relevant items maintained in working memory. *J. Neurosci.* 32, 12411-12420. doi:10.1523/JNEUROSCI.0421-12.2012

Salazar, R. F., Dotson, N. M., Bressler, S. L., and Gray, C. M. (2012). Contentspecific fronto-parietal synchronization during visual working memory. *Science* 338, 1097–1100. doi: 10.1126/science.1224000

Schneegans, S., and Bays, P. M. (2016). No fixed item limit in visuospatial working memory. *Cortex* 83, 181–193. doi: 10.1016/j.cortex.2016.07.021

Schneegans, S., and Bays, P. M. (2017). Neural architecture for feature binding in visual working memory. J. Neurosci. 37, 3913–3925. doi: 10.1523/JNEUROSCI.3493-16.2017

Schneegans, S., and Bays, P. M. (2019). New perspectives on binding in visual working memory. Br. J. Psychol. 110, 207-244. doi: 10.1111/bjop.12345

Schwartz, M. L., and Goldman-Rakic, P. S. (1984). Callosal and intrahemispheric connectivity of the prefrontal association cortex in rhesus monkey: relation between intraparietal and principal sulcal cortex. J. Comp. Neurol. 226, 403–420. doi: 10.1002/cne.902260309

Scimeca, J. M., Kiyonaga, A., and D'Esposito, M. (2018). Reaffirming the sensory recruitment account of working memory. *Trends Cogn. Sci.* 22, 190–192. doi: 10.1016/j.tics.2017.12.007

Selemon, L., and Goldman-Rakic, P. (1988). Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: evidence for a distributed neural network subserving spatially guided behavior. J. Neurosci. 8, 4049–4068. doi: 10.1523/JNEUROSCI.08-11-04049.1988

Serences, J. T., Ester, E. F., Vogel, E. K., and Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* 20, 207–214. doi: 10.1111/j.1467-9280.2009.02276.x

Spadone, S., Della Penna, S., Sestieri, C., Betti, V., Tosoni, A., Perrucci, M. G., et al. (2015). Dynamic reorganization of human resting-state networks during visuospatial attention. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8112–8117. doi: 10.1073/pnas.1415439112

Sreenivasan, K. K., Curtis, C. E., and D'Esposito, M. (2014a). Revisiting the role of persistent neural activity during working memory. *Trends Cogn. Sci.* 18, 82–89. doi: 10.1016/j.tics.2013.12.001

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

Sreenivasan, K. K., Gratton, C., Vytlacil, J., and D'Esposito, M. (2014b). Evidence for working memory storage operations in perceptual cortex. *Cogn. Affect. Behav. Neurosci.* 14, 117–128. doi: 10.3758/s13415-013-0246-7

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E. M., Daunizeau, J., and Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *NeuroImage* 49, 3099–3109. doi: 10.1016/j.neuroimage.2009.11.015

Stier, C., Braun, C., and Focke, N. K. (2023). Adult lifespan trajectories of neuromagnetic signals and interrelations with cortical thickness. *NeuroImage* 278:120275. doi: 10.1016/j.neuroimage.2023.120275

Syrjälä, J., Basti, A., Guidotti, R., Marzetti, L., and Pizzella, V. (2021). Decoding working memory task condition using magnetoencephalography source level long-range phase coupling patterns. *J. Neural Eng.* 18:016027. doi: 10.1088/1741-2552/abcefe

Taylor, B. K., Embury, C. M., Heinrichs-Graham, E., Frenzel, M. R., Eastman, J. A., Wiesman, A. I., et al. (2020). Neural oscillatory dynamics serving abstract reasoning reveal robust sex differences in typically-developing children and adolescents. *Dev. Cogn. Neurosci.* 42:100770. doi: 10.1016/j.dcn.2020.100770

Tseng, P., Iu, K. C., and Juan, C. H. (2018). The critical role of phase difference in theta oscillation between bilateral parietal cortices for visuospatial working memory. *Sci. Rep.* 8, 1–9. doi: 10.1038/s41598-017-18449-w

Veen, B. D. V., Drongelen, W. V., Yuchtman, M., and Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–880. doi: 10.1109/10.623056

Violante, I. R., Li, L. M., Carmichael, D. W., Lorenz, R., Leech, R., Hampshire, A., et al. (2017). Externally induced frontoparietal synchronization modulates network dynamics and enhances working memory performance. *eLife* 6, 1–22. doi: 10.7554/eLife.22001

Voyer, D., Saint Aubin, J., Altman, K., and Gallant, G. (2021). Sex differences in verbal working memory: a systematic review and meta-analysis. *Psychol. Bull.* 147, 352–398. doi: 10.1037/bul0000320

Voyer, D., Voyer, S. D., and Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: a meta-analysis. *Psychon. Bull. Rev.* 24, 307–334. doi: 10.3758/s13423-016-1085-7

Voytek, B., and Knight, R. T. (2010). Prefrontal cortex and basal ganglia contributions to visual working memory. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18167–18172. doi: 10.1073/pnas.1007277107

Westner, B. U., Dalal, S. S., Gramfort, A., Litvak, V., Mosher, J. C., Oostenveld, R., et al. (2022). A unified view on beamformers for M/EEG source reconstruction. *NeuroImage* 246:118789. doi: 10.1016/j.neuroimage.2021. 118789

Xu, Y. (2020). Revisit once more the sensory storage account of visual working memory. *Vis. Cogn.* 28, 433-446. doi: 10.1080/13506285.2020. 1818659

Zeidman, P., Friston, K., and Parr, T. (2023). A primer on Variational Laplace (VL). *NeuroImage* 279:120310. doi: 10.1016/j.neuroimage.2023. 120310

Zeidman, P., Jafarian, A., Seghier, M. L., Litvak, V., Cagnan, H., Price, C. J., et al. (2019). A guide to group effective connectivity analysis, part 2: second level analysis with PEB. *NeuroImage* 200, 12–25. doi: 10.1016/j.neuroimage.2019. 06.032