Decoding auditory working memory content from EEG aftereffects of auditory-cortical TMS

Işıl Uluç[1,2], Mohammad Daneshzand[1,2], Mainak Jas[1,2], Parker Kotlarz[1,2], Kaisu Lankinen[1,2], Jennifer L. Fiedler[1], Fahimeh Mamashli[1,2], Netri Pajankar[1], Tori Turpin[1], Lucia Navarro de Lara[1,2], Padmavathi Sundaram[1,2], Tommi Raij[1,2], Aapo Nummenmaa[1,2], Jyrki Ahveninen[1,2]

[1]Athinoula A. Martinos Center for Biomedical Imaging, Dept. of Radiology, Massachusetts General Hospital, Charlestown, MA, USA

[2]Harvard Medical School, Boston, MA, USA

Abstract

Working memory (WM), short term maintenance of information for goal directed behavior, is essential to human cognition. Identifying the neural mechanisms supporting WM is a focal point of neuroscientific research. One prominent theory hypothesizes that WM content is carried in a dynamic fashion, involving an "activity-silent" brain state based on synaptic facilitation. Information carried in such activity-silent brain states could be decodable from content-specific changes in responses to unrelated "impulse stimuli". A potential method for such impulses is single-pulse transcranial magnetic stimulation (TMS) with its focal, precise nature. Here, we tested the activity-silent model by combining TMS/EEG and multivariate pattern analysis (MVPA) with a non-conceptual auditory WM task that employed parametric ripple sound stimuli and a retro-cue design. Our MVPA employed between-subject cross-validation and robust non-parametric permutation testing. The decoding accuracy of WM content significantly increased after a single pulse TMS was delivered to the posterior superior temporal cortex during WM maintenance. Our results are compatible with the theory that WM maintenance involves brain states which are effectively "activity-silent" relative to other intrinsic processes visible in the EEG signal. Single-pulse TMS combined with MVPA could provide a powerful way to decode information content of "activity-silent" brain states involved in WM.

Introduction

Working memory (WM), the brain system that enables maintenance and processing of recent information, plays an essential role in our daily functions. Its exploration holds the key to a deeper understanding of human cognition. As a result, the mechanisms and brain areas underlying WM maintenance have been prominent topics of study for neuroscience research. This research resulted in seemingly contradictory results that led to a long-standing controversy. On one hand, the prevailing hypothesis suggests that information is maintained through persistent firing in the prefrontal cortex (PFC). Conversely, an alternative theory posits broader distribution of activity across the brain, facilitated by dynamic coding via hidden states (Constantinidis and Klingberg, 2016; D'Esposito and Postle, 2015; Stokes, 2015). Much of this research has been conducted in the visual modality only, leaving some of the most ecologically relevant aspects of WM relatively underexplored. One such aspect is auditory WM, the subsystem of WM that enables temporary storage and manipulation of sounds and verbal information, such as spoken instructions, conversations, or music.

Initially, WM maintenance was linked to persistent activity of a subset of prefrontal neurons that respond to the incoming stimulus and remain activated even after they have vanished (Constantinidis et al., 2002; Funahashi et al., 1989; Goldman-Rakic, 1995). However, subsequent human neuroimaging studies suggested that the content of visual WM can be decoded from signal-change patterns in sensory and posterior brain areas where persistent activity is not present during WM maintenance (Christophel and Haynes, 2014; Harrison and Tong, 2009; Postle et al., 2003). This is largely in contrast to findings in non-human primates (NHP), where no WM related persistent activity is found during the maintenance period in these areas (Lemus et al., 2009; Sreenivasan et al., 2014). The identification of content-specific persistent firing patterns at the sensory level has proven challenging in NHP studies of auditory WM as well (Bigelow and Poremba, 2014; Ng et al., 2014; Scott et al., 2014). Human studies, on the other hand, have managed to decode auditory WM content from fMRI signals (Ahveninen et al., 2023; Czoschke et al., 2021; Kumar et al., 2016; Uluç et al., 2018) as well as intracranial EEG signals (Uluç et al., 2023)

from auditory cortices. The diverse and contrasting findings have inspired the development of a family of alternative theories suggesting that WM is maintained in a more distributed and dynamic fashion than initially believed (Christophel et al., 2017; Miller et al., 2018; Rose, 2020).

A recent alternative theory proposes that WM is maintained through intermittent bursts of neuronal firing and intervals of "activity-silent" short-term synaptic plasticity (STSP) (Mongillo et al., 2008). STSP, denoting transient changes in the strength of synaptic connections between neurons (Zucker and Regehr, 2002), has been more extensively studied in other domains of auditory cognition (Fritz et al., 2005; Fritz et al., 2010; Froemke et al., 2007; Jääskeläinen et al., 2007). In the context of WM, the synaptic theory proposes that item-specific activation of neurons during the encoding process leads to presynaptic accumulation of calcium, which facilitates postsynaptic connectivity. Due to this calcium buffer, which operates on the order of seconds, even sparse bursts of firing will be sufficient to maintain the "activity-silent" WM representations (Miller et al., 2018; Mongillo et al., 2008; Stokes, 2015). The network maintaining a synaptic WM trace will respond in a content-specific fashion, even if the non-specific input is completely unrelated to the maintained representation (Stokes, 2015). Although not directly measurable (Buonomano and Maass, 2009), information maintained in WM via the activity-silent synaptic mechanism should be decodable with machine learning techniques that can classify population-level neuronal responses elicited in response to impulse stimuli (Stokes, 2015).

The basic building blocks of the canonical cortical circuit that underlie WM (Kilpatrick et al., 2013; Miller et al., 2018; Vogel et al., 2022; Wang, 1999) are large pyramidal neurons whose post-synaptic processes dominate the EEG and MEG signals (Hämäläinen et al., 1993). Leveraging this neurophysiological principle, recent human EEG and MEG studies have tested the predictions of the synaptic model using task-irrelevant "impulse stimuli" to facilitate the decoding of WM content maintained in the activity-silent brain states, also called "hidden mnemonic states" (Mamashli et al., 2021; Rose et al., 2016; Wolff et al., 2015; Wolff et al., 2017; Wolff et al., 2020). One of these studies accomplished this using single-pulse transcranial magnetic stimulation (TMS), a method that uses a rapidly changing magnetic field to induce electric currents in the brain areas near the TMS coil. Perturbing the WM circuits during the (seemingly activity-silent)

maintenance period not only enhanced the decoding of representations that were at passively in the memory compared to the actively maintained content, but also helped participants better recall passively maintained items from WM (Rose et al., 2016). However, to our knowledge, this approach has so far not been tested in studies on auditory WM.

Whereas recordings of brain activity limited to observation only, TMS allows us to causally interact with the focal areas whose role WM we intend to evaluate (Muri and Nyffeler, 2008). In studies of human memory processes, TMS has been used to modulate maintenance of visual WM representations (Silvanto and Cattaneo, 2010) and to enhance neuronal plasticity in visual cortex (Kozyrev et al., 2018). A particular benefit of using single-pulse TMS as opposed to task-irrelevant sensory stimuli for probing memory-related brain states its effects are both temporally and anatomically specific (Romero et al., 2019). Thanks to recent advances in MRI-guided online navigation systems and highly focal stimulation coils, these pulses can be delivered to the area of interest at a predictable latency. This allows one to test an anatomically and temporally specific hypotheses while developing an understanding of how sensory areas might be contributing to WM. Here, we investigated whether the content of auditory WM, which is embedded in an activity silent brain state, can be decoded from cortical effects of single-pulse TMS. To this end, we used a retro-cue paradigm with amplitude-modulated ripple sounds as memory items. We recorded 64-channel, whole head EEG for two sessions during the WM task. In the active TMS session, a single TMS pulse was sent to the left posterior superior temporal cortex (pSTC) in the middle of the WM maintenance period in each trial. In a control session, to replicate the sensory and auditory effects of TMS session, we put a plastic block in between the head of the participant and the TMS coil. Our multivariate pattern analysis derived from whole-scalp EEG signals suggested that the decoding accuracy increases above chance level only directly after the TMS pulse in the TMS condition.

Methods

Participants

Initially we acquired structural and functional MRI images of subjects to be later used in recording MRI-guided TMS neuronavigation with a simultaneous 64-channel EEG from 23 healthy participants (12 women, 11 men; mean age ± standard deviation, SD, = 32±12 years). Anatomical MRI and resting motor thresholds (MT) for TMS were recorded in separate sessions. One participant was excluded from the study due to excessive movement artifacts and another due to their chance-level behavioral performance, resulting in a final cohort of 21 analyzed participants (11 women, 10 men; mean age ± SD=31±10 years) for the TMS session. As for the Control session, two participants opted not to continue the study and one dataset was rejected due to excessive noise, resulting in a sample of 18 participants (9 women, 9 men; mean age ± SD = 31 ± 11 years) analyzed for the control session. The order of active and Control sessions was counterbalanced across the remaining participants. All participants had normal or corrected to normal vision and self-reported normal hearing. All participants provided written informed consent. Participants were informed that they could withdraw at any time and given monetary compensation for each visit. The study design, protocol, and consent form were approved by the Mass General Brigham Institutional Review Board.

Paradigm and stimuli

<u>Stimuli</u>

We employed non-conceptual, parametrically varied ripple sounds as WM items (**Fig. 1a**). The ripple-velocity pool was individualized based on each participant's pre-determined ripple-velocity discrimination thresholds. Just noticeable difference (JND) values were calculated individually for each participant using a 2-down, 1-up staircase algorithm (Ahveninen et al., 2023; Mamashli et al., 2021; Uluç et al., 2023). Based on these values, we created four auditory ripple sound stimuli, ensuring that each stimulus was positioned 1.5 JNDs apart from its closest neighbor in velocity. We used four different ripple sounds as to-be-remembered stimuli. The same four ripple stimuli

were also presented as test stimuli. The order of the memory items was pseudorandomized and participants were naïve to the number of memory items.

Paradigm

During a TMS-EEG recording, we used a retro-cue paradigm with ripple sounds as memory items. Throughout the experiment, 8 kHz low pass filtered noise was presented at 0.1 attenuation to mask the click of the TMS pulse. The trials started with a "!" presented on the screen to indicate the start of the task. It was followed by two consecutive ripple sounds with different temporal velocities that were presented for 750 ms with a 250 ms interstimulus interval. Ripple sounds were followed by a visual retro-cue indicating the to-be-remembered item, which was either a "1" to indicate that participants should remember the first item, or a "2", indicating to remember the second item. A four second maintenance period was divided into two equal parts by a TMS pulse in the middle of this period. Finally, a third test sound was presented to be compared to the cued item. The task was to determine whether the test sound was the same or different from the memorized item. Participants responded with a mouse click with their right hand. An index finger mouse click indicated the sounds were the same and the middle finger press indicated they were different. One run of the experiment consisted of 48 trials and one session consisted of 6 runs. Thus, one session had 288 runs in total. Two of the participants completed 5 runs instead of 6 due to technical issues during both active TMS and Control sessions.

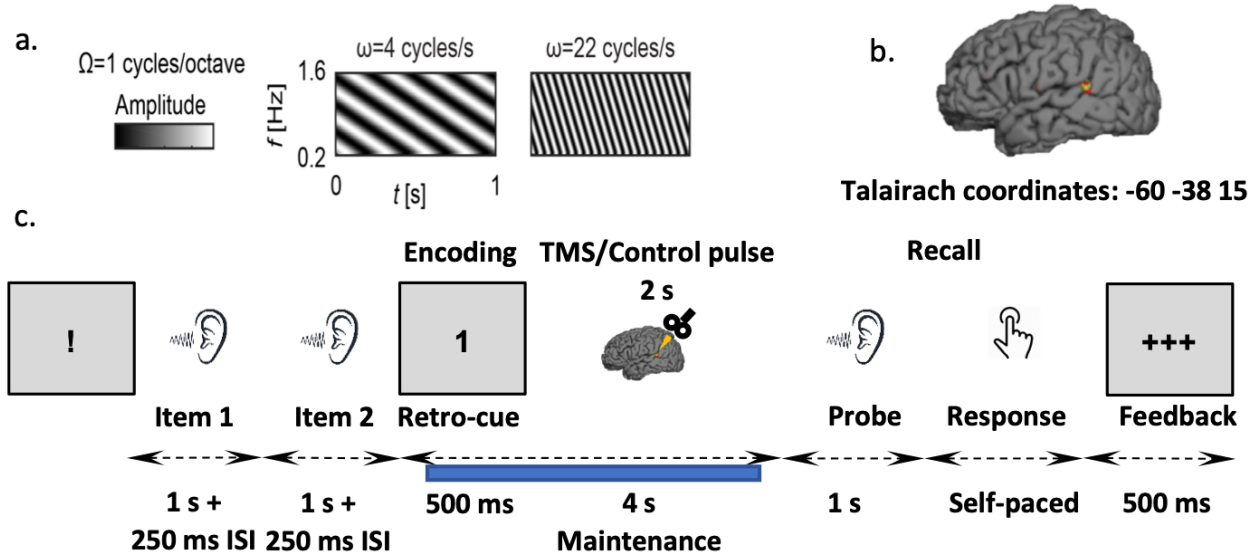

Figure 1. Task specifications. **(a).** Examples of modulation patterns for ripple sounds **(b).** The target brain area for the TMS pulse. The area is adapted from Uluç and colleagues (Uluç et al., 2018). **(c).** The auditory WM retro-cue paradigm. The timeline of events in one trial are depicted.

Structural MRI Data Acquisition

T1-weighted anatomical images were acquired to be used in MRI-guided TMS neuronavigation, employing a multi-echo MPRAGE pulse sequence (TR=2510 ms; 4 echoes with TEs=1.64 ms, 3.5 ms, 5.36 ms, and 7.22 ms; 176 or 208 (to cover the ears) sagittal slices with 1×1×1 mm^3 voxels, 256×256 mm^2 matrix; flip angle = 7°) (van der Kouwe et al., 2008) in a 3T Siemens Trio MRI scanner (Siemens Medical Systems, Erlangen, Germany) using a 32-channel head coil. Cortical reconstruction for each subject was generated using Freesurfer (Dale et al., 1999; Fischl et al., 1999) and read into the Localite neuronavigation system (LOCALITE GmbH, Germany) for MRI guided TMS.

TMS-EEG Data Acquisition

After a separate motor threshold (MT) visit, participants completed two counterbalanced, single-blind sessions. TMS pulses were delivered either 1) at 100% of individual MT to the posterior nonprimary auditory area left pSTC ("active TMS") or 2) 100% of MT with a 2.5 cm plastic block between the coil and scalp ("Control"). The pulses were delivered 2 s into a 4 s maintenance period. EEG, horizontal EOG, and ECG data were sampled at 25 kHz with a 64-channel EEG system (ActiChamp, Brain Products GmbH, Gilching, Germany). TMS was delivered using a figure-of-eight coil (X100 stimulator system, CB-60 coil of MagVenture, Farum, Denmark) and the 2.5 cm wide plastic block built in-house was the same shape and size of the TMS coil. During the working memory task, EEG data were recorded for 10 minutes per run across six trials.

TMS Neuronavigation

Continuous recording of the head's position and orientation relative to the TMS coil was achieved through sensors affixed to the TMS coil and the participant's head. We utilized a commercial TMS neuronavigation system (LOCALITE GmbH, Germany) with an optical camera (Polaris Spectra, Northern Digital, Inc. Waterloo, Ontario) to track the positions of head and coil, displaying this data visually in its Localite software. The participant's registration to their anatomical data, as well as the digitization of the EEG electrode locations were all done in the Localite neuronavigation software.

E-field Calculation

For each participant, we measured the TMS induced electric field (E-field) to ensure we were accurately stimulating the cortical target. We used the Boundary Element Method accelerated by Fast Multipole method (BEM-FMM) MATLAB toolbox (Makarov et al., 2018; Makarov et al., 2020),

for these calculations. We extracted the position of the TMS coil from the neuronavigation software for each participant and used it in the BEM-FMM toolbox to place the coil model at the correct location in respect to the participant head model. The high resolution and anatomically realistic head models were generated based on the T1-weighted images of each participant in the Simnibs toolbox (Saturnino et al., 2019). The model included five distinct layers of scalp, skull, cerebro-spinal fluid (CSF), grey matter, and white matter with the assumption of uniform conductivity in each layer. With the coil model/position and the detailed multi-layer head models, we then calculated the cortical E-fields on a mid grey matter surface for the participant. The detailed explanation of how BEMFMM toolbox is utilized for cortical E-field calculation can be found in (Makarov et al., 2020).

Basic EEG Preprocessing and Analysis

EEG was preprocessed using MNE Python (Gramfort et al., 2013). We used an established, rigorous preprocessing procedure (Al et al., 2023). Data were first detrended, and after the cleaning and interpolation of noisy channels, were epoched into events to exclude TMS pulse artifact. Two consecutive ICA was calculated for the epoched data. The first ICA was used to remove TMS artifacts and then the second ICA was performed to exclude the physiological artifacts. Afterwards, we applied a 60-Hz notch filtering to remove line noise and a 150 Hz low pass filter. The data was then downsampled to 1 kHz, and all epochs were visually inspected for remaining artifacts. Noisy epochs were rejected from the analysis. Signals from the occipital Iz electrode were excluded from all analyses due to excessive noise in most of the datasets.

To display the time course of brain activity during the task trials, we calculated event-related potentials (ERP) separately for each memory item during the memory period as well as a grand average for the whole trial. For memory item comparison, the cleaned data was separated into four groups according to cued memory item. Then, the separated data from all runs were concatenated and averaged across all trials for each memory item within each participant. The data were then averaged across all participants per memory item. For grand averages, the

averages were calculated across all trials irrespective of memory item. Topographical maps were calculated with a time window of 0.5 s with equal weights for all trials.

Multivariate Pattern Analysis

Decoding analysis was conducted by employing support vector machine (SVM) implemented in libsvm (Chang and Lin, 2011) and provided in the CoSMoMVPA package (http://www.cosmomvpa.org/) (Oosterhof et al., 2016) in MATLAB. To help generalize the analysis results to a larger population, we performed the classification across the subjects (Mamashli et al., 2021). Using CoSMoMVPA and Fieldtrip toolboxes in each participant, we first balanced the number of trials for each class then class-specific averages were calculated for each participant. Next, for spatial feature selection, we used principal component analysis (PCA) in MATLAB to transform the data to "virtual channels", corresponding to the eight first principal components (PC) (**Fig. 2**). To keep the dimensionality constant to allow cross-validation, we selected the number of PCs based on the average cumulative variance explained across all participants, conditions, and WM classes. The number of selected PCs refers to the point where the slope of the tangent decreased below one in a normalized plot with both dimensions scaled between zero and one. Calculating these spatial PCAs separately for each class in each subject ensured that no leakage of information occurred between participants/classes. In each task condition, including periods before and after the pulse in the TMS and Control conditions, this resulted in a two-dimensional $\left(N_{Subjects} \times 4\right) \times 8$ feature matrix that was entered into the SVM (in the TMS conditions $N_{Subjects} = 21$, in the control conditions $N_{Subjects} = 18$).

The classification was conducted as a temporal searchlight analysis (Oosterhof et al., 2016) with a 300 ms sliding window at 3 ms steps, and was done separately for the periods before and after the pulse. The classification was performed using a leave-one out cross-validation procedure: the data sets were partitioned to training and test sets such that the class-specific sub-averages of one subject used as the test set and those from the rest of the subjects as the training set. The decoding accuracy was averaged across all iterations. For each condition (active TMS, Control)

and maintenance period (before or after TMS), the analyses resulted a time series with decoding accuracies of each searchlight centroid (**Fig. 2**).

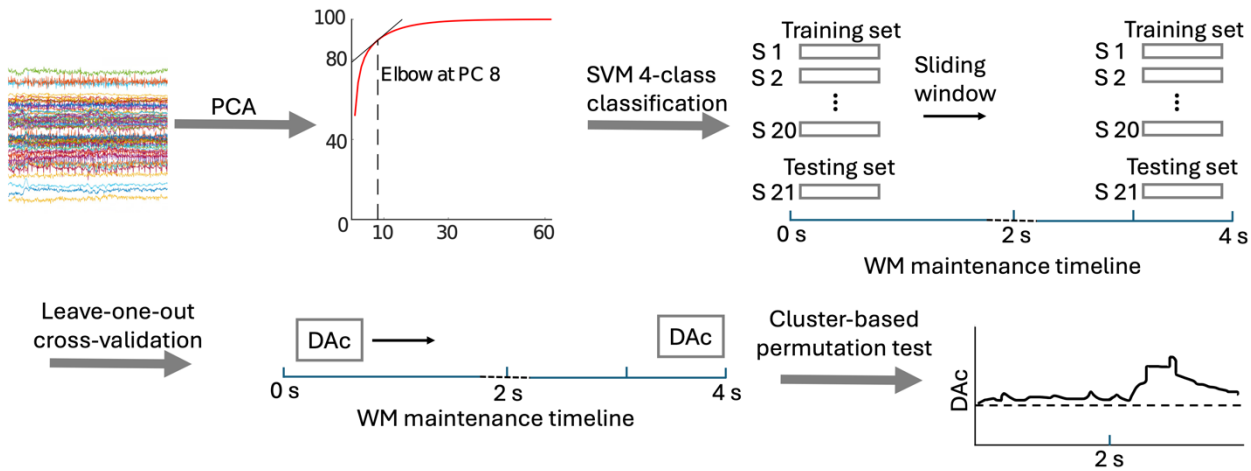

Figure 2. MVPA pipeline schematic. Preprocessed whole head EEG data was entered into a PCA for spatial feature selection. The cut-off for the PC selection (n_{PC}=8) was determined based on the elbow in the average cumulative variance curve. Then the data was entered into a searchlight analysis with between-subject 4-class SVM classification. Leave-one-out method was used for cross-validation. The analysis resulted with a decoding accuracy (DAc) time series where DAc is the value assigned to the centroid of searchlight sliding window. For statistical significance, we used maximum statistics with 500 permutations.

Statistical Significance

To address the of multiple comparisons problem, we employed a robust cluster-based maximum-statistic permutation test strategy for the between-subject decoding analyses (Mamashli et al., 2021). Statistical significance was determined using a temporal cluster-based maximum-statistic approach modified from widely used procedures in univariate analyses of ERP and MEG data (Maris and Oostenveld, 2007). We first generated 500 unique permutations of the true item-content labels of the classifier. The temporal searchlight analysis was repeated with these permuted labels to generate a distribution of decoding accuracies for each time point. For each permutation, the time series of decoding accuracies were converted to *z*-values. This was done by comparing each decoding-accuracy value to the respective permutation distribution at the

same time point. Continuous clusters with $z>1.65$ (i.e., $p<0.05$) were then identified in each permutation and the respective cluster sums of z-values were calculated. From each permutation, the largest cluster sum across all conditions was entered to the maximum-statistic null distribution. Analogously to the conventional procedure (Maris and Oostenveld, 2007), each cluster identified from the analysis with true content labels was then compared to this null distribution, to determine their statistical significance. Clusters with $p_{\text{Corrected}}<0.05$ were considered statistically significant.

Results

Behavioral performance

In the active TMS session, participants' (n=21) percent correct was 79.6 ± 6.5% (mean ± SD) across the six experimental runs. In the Control session, participants performed at 78.7± 6.2% (mean ± SD) correct responses across 6 experimental runs. We also calculated the average proportion correct as a function of the ripple-velocity distance from memory item to probe. This analysis revealed a consistent relationship between participants' ability to reject non-matching Probes and the difference in ripple velocity relative to the maintained WM item in both TMS sessions (**Fig. 3**). For TMS session, when the distance was 0 (match trials), 84.5 ± 6.3% (mean ± SD), distance 1 yielded 60.4 ± 13.9% (mean ± SD), distance 2 resulted in 85.7 ± 8.6% (mean ± SD), and finally distance 4 brought 93.9 ± 4.9% (mean ± SD). In the Control session, when the distance was 0 (match trials), 84.1 ± 5.7% (mean ± SD), distance 1 yielded 58.0 ± 14.9% (mean ± SD), distance 2 resulted in 87.1 ± 10.2% (mean ± SD), and finally distance 4 brought 92.2 ± 9.0% (mean ± SD).

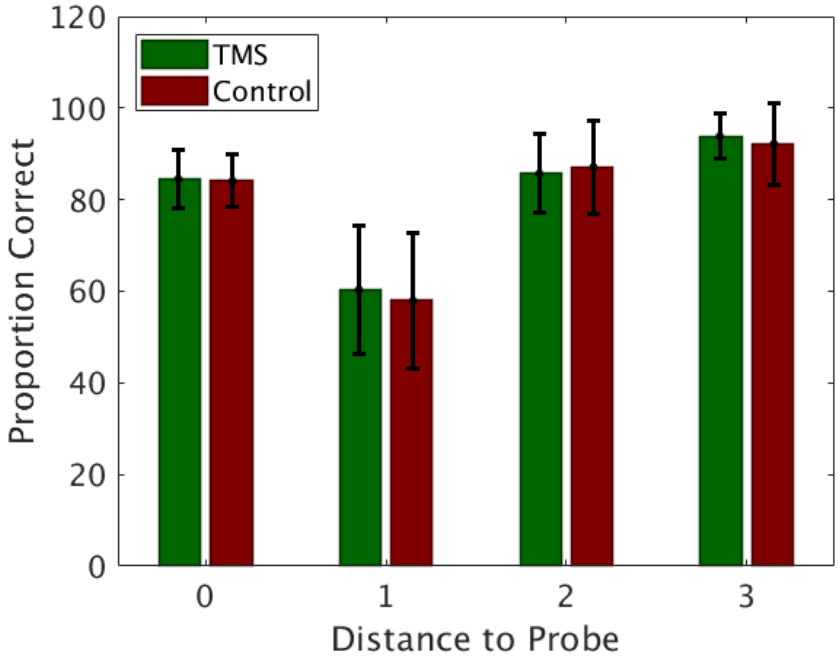

Figure 3. Proportion corrects for tor each item-probe distance. Match trial averages are depicted by distance "0".

Multivariate Pattern Analysis

We conducted a four-class between-subject classification analysis to determine whether a TMS pulse to pSTC during WM maintenance enhances the decoding of memorized content. The analysis employed a temporal searchlight approach, in which the decoding was performed based on the spatiotemporal pattern of EEG activity within a 300-ms sliding window. The statistical significance was verified through a robust between-subject cross-validation and cluster-based maximum-statistic permutation procedure. According to these analyses, the MVPA decoding accuracy for memory content rose significantly above chance level in searchlights encompassing the first few hundreds of milliseconds after the TMS pulse ($p_{Corrected}$ <0.05, cluster-based maximum-statistic permutation test; cluster sum of normalized accuracy = 137.2; **Fig. 4d**). The decoding accuracies lingered nominally above the theoretical change level at many other instances, too. However, no statistically significant decoding results were observed in any other time period in the active TMS condition or in the Control condition. Our results indicate that a TMS targeted at the non-primary auditory areas of the posterior superior temporal cortex helps decode WM content during maintenance.

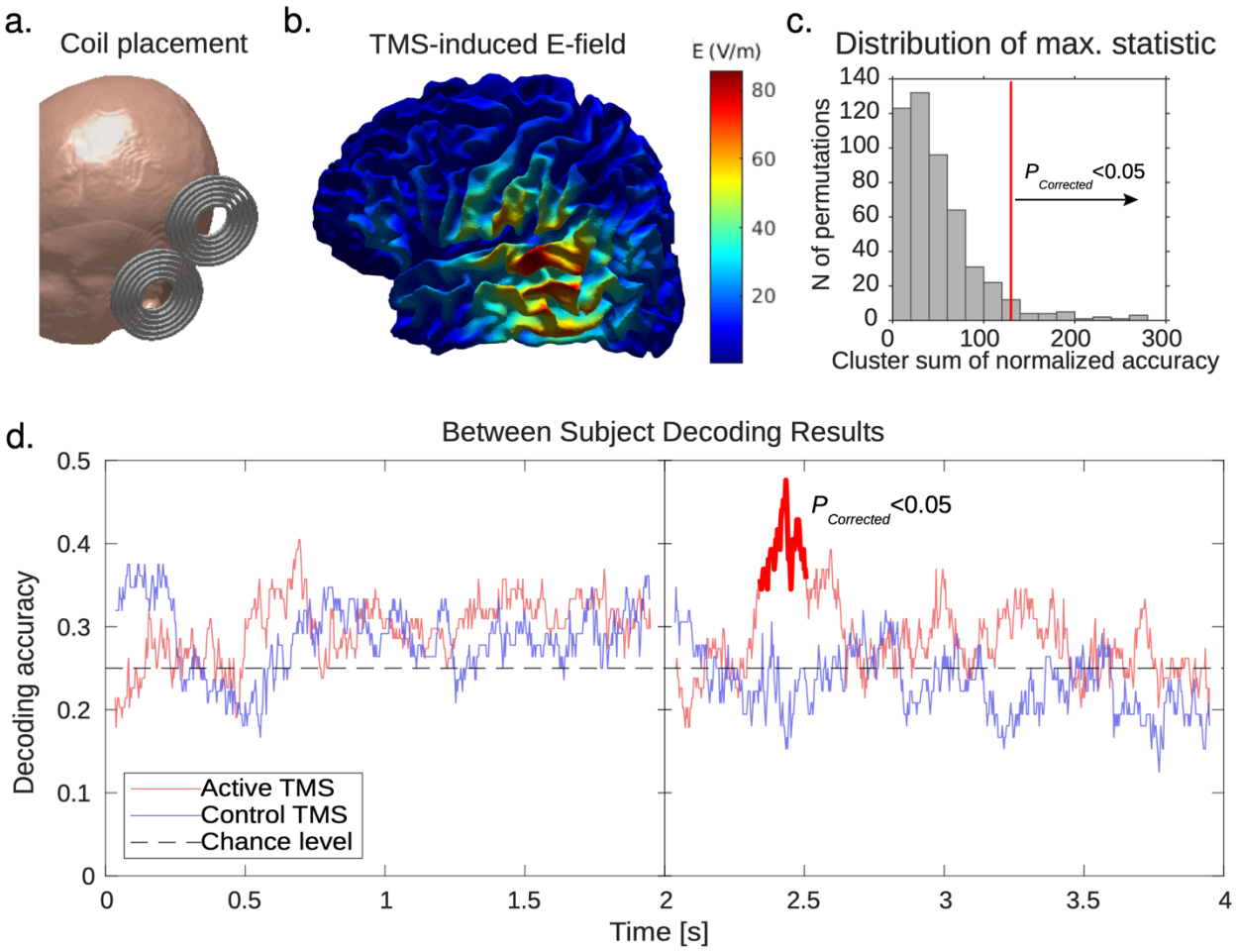

Figure 4. The results of searchlight MVPA decoding WM content from EEG. **(a)** Example coil position during active TMS session in a single participant. **(b)** Example e-field calculation in active TMS session in a single participant. **(c)** Null distribution for 500 permutations. A vertical red line illustrates the critical value for $p_{corrected}$<0.05. **(d)** Between-subject Multivariate Pattern Analysis Decoding Accuracies. Red line denotes active TMS condition. Bold red is the patch where decoding accuracy is significantly higher than chance (p<0.05). Blue line denotes Control condition.

Event Related Potentials

Figure 5 depicts the grand-average ERPs calculated for the trial period. The averages across subjects show the response to ripple sound stimuli (**Fig. 5a, b**) and the visual cue (start of **Fig. 5c**), but we found no persistent delay activity during the WM period (**Fig. 5c, d**). As a control, we also averaged the response to different cued content during WM period across subjects. No systematic differences were found between signals for different cued content during the maintenance period before or after TMS pulse.

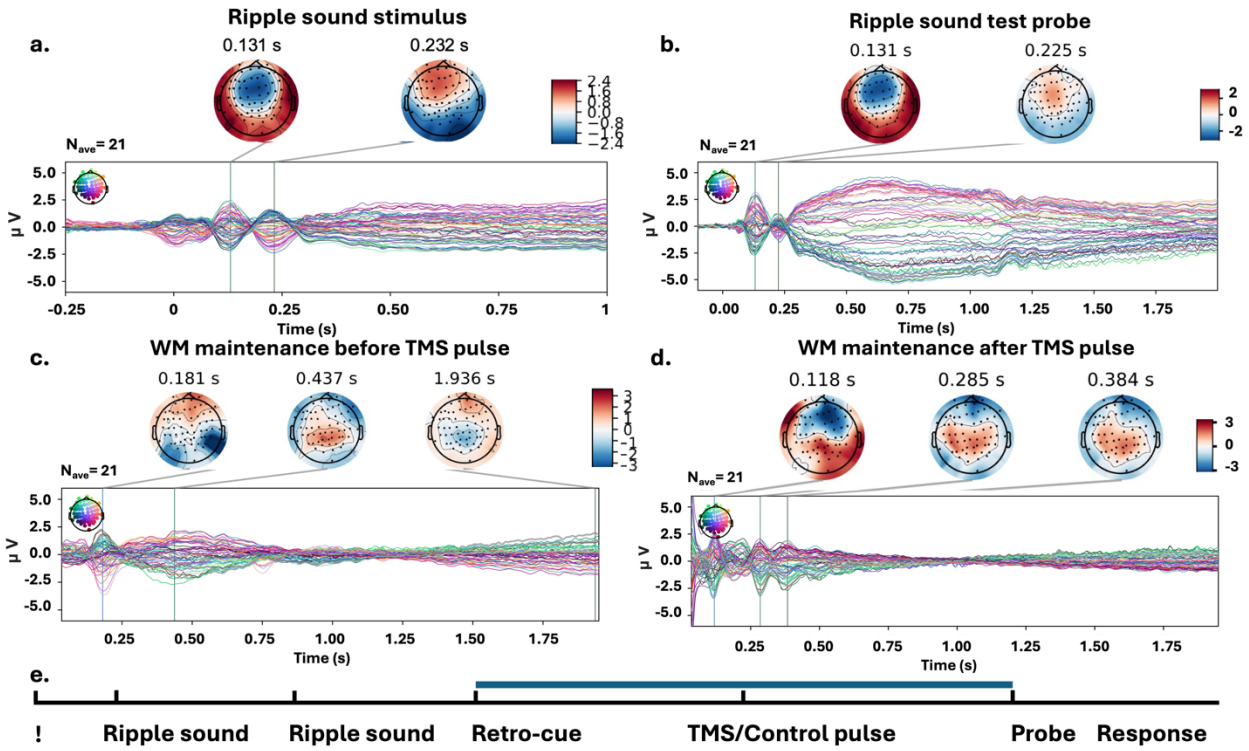

Figure 5. Topographical and time plots of ERPs generated during the trial period for the active TMS condition. Different colors in the averaged signal refer to different electrodes. **(a).** Event related potentials and topographical plots for the stimulus presentation. **(b).** Event related potentials topographical plots for the retro-cue presentation and the maintenance period until TMS pulse. **(c).** Event related potentials topographical plots for maintenance period after TMS pulse. **(d).** Event related potentials and topographical plots for the test probe. **(e).** Trial timeline. The order of events was depicted. For detailed timeline see **Fig. 1**.

Control Analysis

We performed a control decoding analysis with the same parameters as our main analysis using the task-irrelevant (i.e., "un-cued") items, which were to be forgotten after the retro-cue. This control MVPA showed no significant decoding for neither active TMS nor Control conditions for the entire WM period.

DISCUSSION

Here, we investigated auditory WM using a "perturbation approach", which combines MRI-navigated, single-pulse TMS and simultaneously recorded EEG, to probe content-specific mnemonic states otherwise obscured by the much larger EEG "background" activity. To decode WM content from EEG signals during the maintenance period, we employed a temporal-searchlight MVPA with robust between-subject cross-validation and non-parametric permutation testing. As inferred from spatiotemporal information in 300-ms sliding time windows of EEG data, the decoding accuracy of non-conceptual auditory WM content briefly rose significantly above chance level after single-pulse TMS was delivered to the non-primary auditory area left pSTC. No significant MVPA results were found before the TMS pulse or after a control pulse that was delivered to the same target area but with a 2-cm plastic block between the coil and the participant's scalp.

One possible explanation for the present finding is offered by the synaptic theory of WM (Mongillo et al., 2008). According to this model, WM information is coded to content-specific patterns of functional connectivity, which result from activity-based STSP in the synaptic terminals of neurons that are strongly activated at the encoding stage (Mongillo et al., 2008). Instead of persistent neuronal activity, this model predicts that only sparse bursts of neuronal oscillations and firing activity are necessary to maintain this otherwise activity-silent mnemonic brain state (Buschman and Miller, 2022; Lundqvist et al., 2016; Postle, 2015; Stokes, 2015). Notably, although synaptic states are not directly measurable by non-invasive recordings, computational modeling predicts that a circuit that maintains information by content-specific changes of synaptic weights responds slightly differently to other stimuli (until the STSP decays) (Mongillo et al., 2008). These content-specific responses to external impulses could not only provide a way for the maintained content to be read out at the recall stage, but also allow one to "ping" the maintained content. Experimental evidence for this theory has been obtained in recent studies analyzing content-specific changes in responses to unrelated "impulse stimuli" (Mamashli et al., 2021; Wolff et al., 2015; Wolff et al., 2017), including TMS pulses (Rose et al., 2016).

Whereas the original synaptic theory refers to local circuits in prefrontal cortices (Mongillo et al., 2008), it has also been proposed that synaptic model can also provide a way to explain how WM information might be represented in sensory areas , which presumably cannot support persistent neuronal firing in the absence of sensory stimulation (Comeaux et al., 2023; Mejias and Wang, 2022; Wang, 2022). However, previous studies probing activity-silent states of WM have been limited to TMS-based perturbation of association areas (Rose et al., 2016), or to using auditory or visual "impulse stimuli" (Mamashli et al., 2021; Wolff et al., 2020) that might activate a wide array of brain networks beyond sensory areas involved in orienting to task-irrelevant stimulus changes. TMS may offer a more direct way to perturb cortical neurons (Murphy et al., 2016; Silvanto and Cattaneo, 2010), to ensure that the presumed content-related change comes from the sensory areas rather than from a higher-order brain region. The present results suggest that WM content can be decoded at a high accuracy from EEG responses to TMS pulses directed to pSTC. They could thus offer new insights into the role of activity-silent WM processes at the sensory level.

Using measures such as EEG to probe synaptic processes is supported by the notion that EEG signals primarily result from post-synaptic processes in apical dendrites of cortical pyramidal neurons (Buzsaki et al., 2012; Hämäläinen et al., 1993; Murakami and Okada, 2006). These neurons constitute fundamental components of the canonical cortical circuit that presumably supports WM (Kilpatrick et al., 2013; Miller et al., 2018; Vogel et al., 2022; Wang, 1999). An inherent limitation of non-invasive measures, however, is that they do not conclusively rule out other alternative explanations. In addition to an activity-silent synaptic state, the present enhancement of WM decoding by TMS pulses could also result from perturbation of an activity-based maintenance process. Subcellular-level studies demonstrate that single-pulse TMS activates a broad population of cell bodies in the cortex (Murphy et al., 2016; Pashut et al., 2014). The rapid firing of these neurons after the TMS pulse, which is followed by a refractory period, disrupts the cortical network, resetting the stimulated region (Murphy et al., 2016; Pashut et al., 2014). Our results could therefore also be speculated to be explained by a subthreshold attractor model (Christophel et al., 2018; Jaffe and Constantinidis, 2021). The subthreshold attractor model suggests that WM-related persistent activity tends to be attracted to a bump state that emerges in varying locations across this network (Itskov et al., 2011; Wimmer et al., 2014). By interfering

with such activation patterns, TMS pulses might result in content-specific signal changes that are recordable by EEG. However, a challenge for such a model in the present context is that the TMS pulse, which tends to overwrite the neuronal activity in the stimulated area, would disrupt the content-specific population activity in the stimulated network. Neurophysiological studies provide experimental evidence indicating that distractor events disrupt content-specific firing activities rather than amplifying them to a discernible level in the mass action of neurons (Miller et al., 1993). Therefore, while it is not entirely incompatible with the synaptic theory of WM, the subthreshold attractor model does not adequately describe our results because we found no evidence of impaired WM performance in the active TMS vs. Control condition.

Another alternative for a hidden state (whether it is due to synaptic plasticity or to other means) of WM content in sensory areas is that the WM maintenance is carried through a recurrent neural network where the PFC shapes and transforms the WM representations according to task demands (Mejias and Wang, 2022; Panichello and Buschman, 2021). The recurrent model is another possible explanation for how the synaptic weights could be formed and maintained in the posterior and sensory brain areas. Indeed, it has been recently shown that the WM content can be effectively maintained by a neuronal behavior explained by a combination of activity-based and activity-silent models of WM (Barbosa et al., 2020).

Some inherent limitations of EEG interpretations during a combined TMS-EEG study include several types of artifacts such as direct muscle stimulation, somatosensory sensation related to the vibration of the coil, and acoustical clicks. There is a possibility that the improved decoding following a TMS pulse is attributable not only to its neurophysiological effects on the target area but also to non-specific effects associated with unrelated physiological events. Here, we attempted to control these biases by using a single-blind control session, with a hard plastic block that was shaped exactly like the figure-of-eight coil directly attached to the coil. The idea was to provide as identical of a somatosensory sensation and acoustic effects in the Control condition as possible to the active condition without exceeding the threshold of direct brain stimulation. To mitigate the effect of acoustical clicks we also used a continuous noise masker stimulus and TMS

(and fMRI) compatible insert earphones accompanied with earplugs that attenuate the background noise.

To conclude, with the help of TMS stimulation and EEG recordings as well as multivariate pattern analysis, the present study found that WM content is carried in an "activity-silent" state in auditory brain areas. Here we have provided evidence for the theory that WM maintenance involves brain states that are effectively hidden relative to other intrinsic population processes visible in EEG. However, future studies with more specific neurophysiological methods are necessary to confirm the exact neuronal mechanism.

References

Ahveninen, J., Uluc, I., Raij, T., Nummenmaa, A., Mamashli, F., 2023. Spectrotemporal content of human auditory working memory represented in functional connectivity patterns. Commun Biol 6, 294.

Al, E., Stephani, T., Engelhardt, M., Haegens, S., Villringer, A., Nikulin, V.V., 2023. Cardiac activity impacts cortical motor excitability. PLoS Biol 21, e3002393.

Barbosa, J., Stein, H., Martinez, R.L., Galan-Gadea, A., Li, S., Dalmau, J., Adam, K.C.S., Valls-Sole, J., Constantinidis, C., Compte, A., 2020. Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. Nat Neurosci 23, 1016-1024.

Bigelow, J., Poremba, A., 2014. Achilles' ear? Inferior human short-term and recognition memory in the auditory modality. PLoS One 9, e89914.

Buonomano, D.V., Maass, W., 2009. State-dependent computations: spatiotemporal processing in cortical networks. Nat Rev Neurosci 10, 113-125.

Buschman, T.J., Miller, E.K., 2022. Working Memory Is Complex and Dynamic, Like Your Thoughts. J Cogn Neurosci 35, 17-23.

Buzsaki, G., Anastassiou, C.A., Koch, C., 2012. The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13, 407-420.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2, 1-27.

Christophel, T.B., Haynes, J.D., 2014. Decoding complex flow-field patterns in visual working memory. Neuroimage 91, 43-51.

Christophel, T.B., Iamshchinina, P., Yan, C., Allefeld, C., Haynes, J.D., 2018. Cortical specialization for attended versus unattended working memory. Nat Neurosci 21, 494-496.

Christophel, T.B., Klink, P.C., Spitzer, B., Roelfsema, P.R., Haynes, J.D., 2017. The Distributed Nature of Working Memory. Trends Cogn Sci 21, 111-124.

Comeaux, P., Clark, K., Noudoost, B., 2023. A recruitment through coherence theory of working memory. Progress in Neurobiology 228, 102491.

Constantinidis, C., Klingberg, T., 2016. The neuroscience of working memory capacity and training. Nat Rev Neurosci 17, 438-449.

Constantinidis, C., Williams, G.V., Goldman-Rakic, P.S., 2002. A role for inhibition in shaping the temporal flow of information in prefrontal cortex. Nature Neuroscience 5, 175-180.

Czoschke, S., Fischer, C., Bahador, T., Bledowski, C., Kaiser, J., 2021. Decoding Concurrent Representations of Pitch and Location in Auditory Working Memory. The Journal of Neuroscience 41, 4658-4666.

D'Esposito, M., Postle, B.R., 2015. The Cognitive Neuroscience of Working Memory. Annual Review of Psychology 66, 115-142.

Fritz, J., Elhilali, M., Shamma, S., 2005. Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. Hear Res 206, 159-176.

Fritz, J.B., David, S.V., Radtke-Schuller, S., Yin, P., Shamma, S.A., 2010. Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. Nat Neurosci 13, 1011-1019.

Froemke, R.C., Merzenich, M.M., Schreiner, C.E., 2007. A synaptic memory trace for cortical receptive field plasticity. Nature 450, 425-429.

Funahashi, S., Bruce, C.J., Goldman-Rakic, P.S., 1989. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. Journal of Neurophysiology 61, 331-349.

Goldman-Rakic, P.S., 1995. Cellular basis of working memory. Neuron 14, 477-485.

Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., Hämäläinen, M., 2013. MEG and EEG data analysis with MNE-Python. Front Neurosci 7, 267.

Hämäläinen, M., Hari, R., Ilmoniemi, R., Knuutila, J., Lounasmaa, O., 1993. Magnetoencephalography -theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev Mod Phys 65, 413-497.

Harrison, S.A., Tong, F., 2009. Decoding reveals the contents of visual working memory in early visual areas. Nature 458, 632-635.

Itskov, V., Hansel, D., Tsodyks, M., 2011. Short-Term Facilitation may Stabilize Parametric Working Memory Trace. Front Comput Neurosci 5, 40.

Jääskeläinen, I.P., Ahveninen, J., Belliveau, J.W., Raij, T., Sams, M., 2007. Short-term plasticity in auditory cognition. Trends Neurosci 30, 653-661.

Jaffe, R.J., Constantinidis, C., 2021. Working Memory: From Neural Activity to the Sentient Mind. Compr Physiol 11, 2547-2587.

Kilpatrick, Z.P., Ermentrout, B., Doiron, B., 2013. Optimizing working memory with heterogeneity of recurrent cortical excitation. J Neurosci 33, 18999-19011.

Kozyrev, V., Staadt, R., Eysel, U.T., Jancke, D., 2018. TMS-induced neuronal plasticity enables targeted remodeling of visual cortical maps. Proc Natl Acad Sci U S A 115, 6476-6481.

Kumar, S., Joseph, S., Gander, P.E., Barascud, N., Halpern, A.R., Griffiths, T.D., 2016. A Brain System for Auditory Working Memory. J Neurosci 36, 4492-4505.

Lemus, L., Hernandez, A., Romo, R., 2009. Neural codes for perceptual discrimination of acoustic flutter in the primate auditory cortex. Proceedings of the National Academy of Sciences 106, 9471-9476.

Lundqvist, M., Rose, J., Herman, P., Brincat, S.L., Buschman, T.J., Miller, E.K., 2016. Gamma and Beta Bursts Underlie Working Memory. Neuron 90, 152-164.

Makarov, S.N., Noetscher, G.M., Raij, T., Nummenmaa, A., 2018. A Quasi-Static Boundary Element Approach With Fast Multipole Acceleration for High-Resolution Bioelectromagnetic Models. IEEE Trans Biomed Eng 65, 2675-2683.

Makarov, S.N., Wartman, W.A., Daneshzand, M., Fujimoto, K., Raij, T., Nummenmaa, A., 2020. A software toolkit for TMS electric-field modeling with boundary element fast multipole method: an efficient MATLAB implementation. J Neural Eng 17, 046023.

Mamashli, F., Khan, S., Hamalainen, M., Jas, M., Raij, T., Stufflebeam, S.M., Nummenmaa, A., Ahveninen, J., 2021. Synchronization patterns reveal neuronal coding of working memory content. Cell Rep 36, 109566.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164, 177-190.

Mejias, J.F., Wang, X.J., 2022. Mechanisms of distributed working memory in a large-scale network of macaque neocortex. Elife 11.

Miller, E.K., Li, L., Desimone, R., 1993. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. J Neurosci 13, 1460-1478.

Miller, E.K., Lundqvist, M., Bastos, A.M., 2018. Working Memory 2.0. Neuron 100, 463-475.

Mongillo, G., Barak, O., Tsodyks, M., 2008. Synaptic theory of working memory. Science 319, 1543-1546.

Murakami, S., Okada, Y., 2006. Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. J Physiol 575, 925-936.

Muri, R.M., Nyffeler, T., 2008. Using transcranial magnetic stimulation to probe decision-making and memory. Prog Brain Res 171, 413-418.

Murphy, S.C., Palmer, L.M., Nyffeler, T., Muri, R.M., Larkum, M.E., 2016. Transcranial magnetic stimulation (TMS) inhibits cortical dendrites. Elife 5.

Ng, C.W., Plakke, B., Poremba, A., 2014. Neural correlates of auditory recognition memory in the primate dorsal temporal pole. J Neurophysiol 111, 455-469.

Oosterhof, N.N., Connolly, A.C., Haxby, J.V., 2016. CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. Front Neuroinform 10, 27.

Panichello, M.F., Buschman, T.J., 2021. Shared mechanisms underlie the control of working memory and attention. Nature 592, 601-605.

Pashut, T., Magidov, D., Ben-Porat, H., Wolfus, S., Friedman, A., Perel, E., Lavidor, M., Bar-Gad, I., Yeshurun, Y., Korngreen, A., 2014. Patch-clamp recordings of rat neurons from acute brain slices of the somatosensory cortex during magnetic stimulation. Front Cell Neurosci 8, 145.

Postle, B.R., 2015. The cognitive neuroscience of visual short-term memory. Current Opinion in Behavioral Sciences, pp. 40-46.

Postle, B.R., Druzgal, T.J., D'Esposito, M., 2003. Seeking the Neural Substrates of Visual Working Memory Storage. Cortex 39, 927-946.

Romero, M.C., Davare, M., Armendariz, M., Janssen, P., 2019. Neural effects of transcranial magnetic stimulation at the single-cell level. Nat Commun 10, 2642.

Rose, N.S., 2020. The Dynamic-Processing Model of Working Memory. Current Directions in Psychological Science 29, 378-387.

Rose, N.S., LaRocque, J.J., Riggall, A.C., Gosseries, O., Starrett, M.J., Meyering, E.E., Postle, B.R., 2016. Reactivation of latent working memories with transcranial magnetic stimulation. Science 354, 1136-1139.

Saturnino, G.B., Puonti, O., Nielsen, J.D., Antonenko, D., Madsen, K.H., Thielscher, A., 2019. SimNIBS 2.1: A Comprehensive Pipeline for Individualized Electric Field Modelling for Transcranial Brain Stimulation. In: Makarov, S., Horner, M., Noetscher, G. (Eds.), Brain and Human Body Modeling: Computational Human Modeling at EMBC 2018, Cham (CH), pp. 3-25.

Scott, B.H., Mishkin, M., Yin, P., 2014. Neural correlates of auditory short-term memory in rostral superior temporal cortex. Curr Biol 24, 2767-2775.

Silvanto, J., Cattaneo, Z., 2010. Transcranial magnetic stimulation reveals the content of visual short-term memory in the visual cortex. Neuroimage 50, 1683-1689.

Sreenivasan, K.K., Curtis, C.E., D'Esposito, M., 2014. Revisiting the role of persistent neural activity during working memory. Trends in Cognitive Sciences 18, 82-89.

Stokes, M.G., 2015. 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. Trends Cogn Sci 19, 394-405.

Uluç, I., Peled, N., Paulk, A.C., Bush, A., Gumenyuk, V., Kotlarz, P., Lankinen, K., Mamashli, F., Matsuda, N., Richardson, M.R., Stufflebeam, S.M., Cash, S.S., Ahveninen, J., 2023. Decoding auditory working memory content from intracranial high frequency activity in humans. bioRxiv, 2023.2008.2004.552073.

Uluç, I., Schmidt, T.T., Wu, Y.H., Blankenburg, F., 2018. Content-specific codes of parametric auditory working memory in humans. Neuroimage 183.

van der Kouwe, A.J., Benner, T., Salat, D.H., Fischl, B., 2008. Brain morphometry with multiecho MPRAGE. Neuroimage 40, 559-569.

Vogel, P., Hahn, J., Duvarci, S., Sigurdsson, T., 2022. Prefrontal pyramidal neurons are critical for all phases of working memory. Cell Reports 39, 110659.

Wang, X.J., 1999. Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. J Neurosci 19, 9587-9603.

Wang, X.J., 2022. Theory of the Multiregional Neocortex: Large-Scale Neural Dynamics and Distributed Cognition. Annual Review of Neuroscience 45, 533-560.

Wimmer, K., Nykamp, D.Q., Constantinidis, C., Compte, A., 2014. Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. Nat Neurosci 17, 431-439.

Wolff, M.J., Ding, J., Myers, N.E., Stokes, M.G., 2015. Revealing hidden states in visual working memory using electroencephalography. Front Syst Neurosci 9, 123.

Wolff, M.J., Jochim, J., Akyurek, E.G., Stokes, M.G., 2017. Dynamic hidden states underlying working-memory-guided behavior. Nat Neurosci.

Wolff, M.J., Kandemir, G., Stokes, M.G., Akyurek, E.G., 2020. Unimodal and Bimodal Access to Sensory Working Memories by Auditory and Visual Impulses. J Neurosci 40, 671-681.

Zucker, R.S., Regehr, W.G., 2002. Short-term synaptic plasticity. Annu Rev Physiol 64, 355-405.