

# Distinct functions for beta and alpha bursts in gating of human working memory

Received: 14 December 2023

Accepted: 7 October 2024

Published online: 17 October 2024

 Check for updates

Johan Liljefors <sup>1</sup>, Rita Almeida <sup>1,2</sup>, Gustaf Rane<sup>1</sup>, Johan N. Lundström <sup>1,3</sup>, Pawel Herman <sup>4</sup> & Mikael Lundqvist <sup>1</sup> ✉

Multiple neural mechanisms underlying gating to working memory have been proposed with divergent results obtained in human and animal studies. Previous findings from non-human primates suggest prefrontal beta frequency bursts as a correlate of transient inhibition during selective encoding. Human studies instead suggest a similar role for sensory alpha power fluctuations. To cast light on these discrepancies we employed a sequential working memory task with distractors for human participants. In particular, we examined their whole-brain electrophysiological activity in both alpha and beta bands with the same single-trial burst analysis earlier performed on non-human primates. Our results reconcile earlier findings by demonstrating that both alpha and beta bursts in humans correlate with the filtering and control of memory items, but with region and task-specific differences between the two rhythms. Occipital beta burst patterns were selectively modulated during the transition from sensory processing to memory retention whereas prefrontal and parietal beta bursts tracked sequence order and were proactively upregulated prior to upcoming target encoding. Occipital alpha bursts instead increased during the actual presentation of unwanted sensory stimuli. Source reconstruction additionally suggested the involvement of striatal and thalamic alpha and beta. Thus, specific whole-brain burst patterns correlate with different aspects of working memory control.

Working memory (WM) is a key cognitive component that allows us to hold and manipulate information online in our mind<sup>1–5</sup>. WM has a very limited capacity and thus we need to control that only relevant information enters our WM<sup>2,6,7</sup>. Work on non-human primates and humans has, however, suggested distinct mechanisms of such WM control. Here we resolve this discrepancy and demonstrate multiple novel task- and region-specific neural correlates of human WM control on the single-trial level.

We previously observed high power beta (14–35 Hz) bursts in prefrontal cortex of non-human primates as a single-trial correlate of executive WM control<sup>8–12</sup>. At the conceptual level, these studies implied that cognition is supported by brief events rather than slowly

changing dynamics and proposed that burst analysis may better capture moment-by-moment interactions between regions as cognitive commands unfold<sup>13–15</sup>. Specifically, beta bursting was reduced when information was encoded, particularly in cortical sites where neurons subsequently retained the WM information<sup>8,12</sup>. This suggests that spatial patterns of prefrontal beta bursts reflect the filtering of unwanted stimuli. Similarly, prefrontal beta bursts were suppressed when information in WM was accessed and subsequently elevated when it was cleared out.

In contrast, human neuroimaging studies focused on alpha (8–12 Hz) power in occipital and parietal areas as a correlate of filtering out visual information<sup>16–21</sup>. In particular, alpha power was shown to

<sup>1</sup>Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden. <sup>2</sup>Stockholm University Brain Imaging Centre, Stockholm University, Stockholm, Sweden. <sup>3</sup>Monell Chemical Senses Center, Philadelphia, PA, United States of America. <sup>4</sup>School of Electrical Engineering and Computer Science, and Digital Futures, KTH Royal Institute of Technology, 10044 Stockholm, Sweden. ✉ e-mail: [Mikael.Lundqvist@ki.se](mailto:Mikael.Lundqvist@ki.se)

increase during the processing of irrelevant versus relevant stimuli. This elevation occurred in specific cortical locations depending on the visuo-spatial location of distractors or their sensory modality<sup>19,20,22–24</sup>. In addition, alpha oscillations in human occipital and parietal regions were considered to have an inhibitory role in other WM processes such as removal and selective prioritization of information<sup>25,26</sup>.

Taken together, the control-related patterns observed in beta bursting in the frontal cortex of non-human primates<sup>8,10,12</sup> appear to be analogous to the ones observed in parieto-occipital alpha power fluctuations in humans<sup>19,23,24</sup>. This difference in frequency could be attributed to species differences, analysis methods (power fluctuations vs bursts) or the distinct areas studied. The observation that beta oscillations are gradually shifted towards slower frequencies lower in the cortical hierarchy, with the alpha dominance in visual area V4, would suggest the latter<sup>9,27–29</sup>.

To test the hypothesis about regional specificity of WM control-related alpha and beta activity in human participants, we deployed a sequential WM task that required gating of input to WM. To determine single-trial neural correlates of WM gating, we recorded whole-scalp magnetoencephalogram (MEG) and frequency tagged each WM item in the sequence<sup>21,30,31</sup>. Frequency tagging involved modulating the luminance of stimuli at a known frequency entraining cerebral activity. The resulting frequency specific entrainment allowed us to estimate how strongly individual items were processed depending on their status as a target or distractor, and how this related to alpha and beta bursting in different cortical regions. The goal was to reconcile the proposed roles of beta bursting in macaques with alpha power in humans.

In particular, we asked whether both alpha and beta reflect the gating of distractor items? Do they correlate with gating information into WM observable on single trials and in behaviour? If so, do they have distinct roles and cortical origins? We found evidence that high-power bursts of both alpha and beta gate information in and out of WM but with partially distinct roles between the two frequency bands and between cortical locations. Overall, alpha seemed consistent with the suppression of sensory processing, whereas beta bursts more selectively appeared to gate information from sensory processing into WM and proactively removed information already retained in WM.

## Results

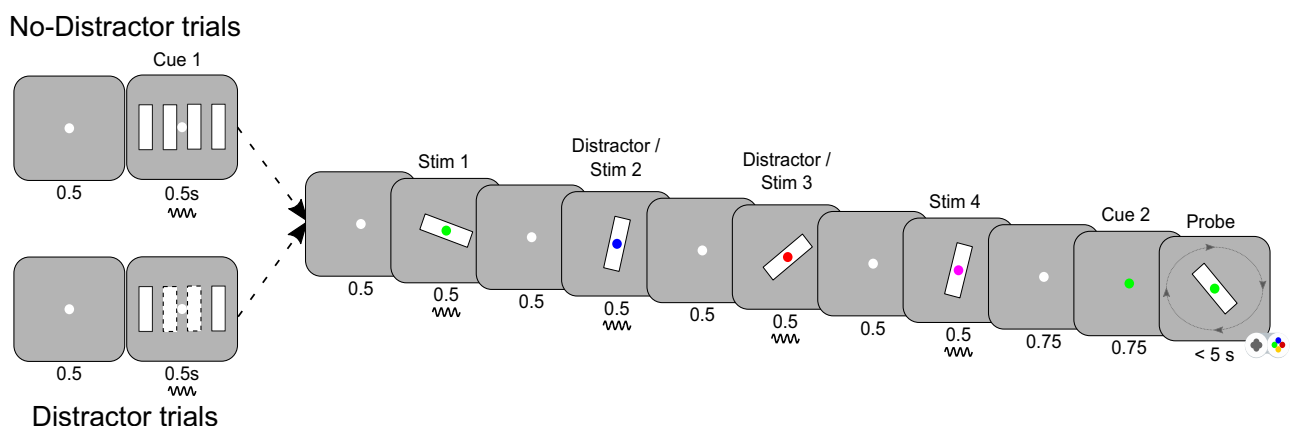
We recorded MEG data from 17 healthy volunteers while they performed a serial WM task (Fig. 1). The task required them to hold fixation on the centre of the screen throughout the whole trial. There were two sets of trials, with or without distractors. On No-Distractor trials,

four randomly oriented bars were sequentially presented foveally for 500 ms with inter-stimulus periods of 500 ms showing a fixation dot. Each bar was associated with a unique colour. Following a delay of 750 ms, the fixation dot turned into one of the four colours, acting as a retro-cue signalling which of the four bars would be tested. After an additional 750 ms, a test probe appeared on the screen and subjects were to rotate the test probe using a control pad to match the orientation of the memorized and cued bars within 5 s. On Distractor trials bar stimuli 2 and 3 acted as distractors, which should not be remembered and hence were never probed. Distractor and No-Distractor trials were randomly interleaved and were discriminated by a pre-cue just prior to each trial (Fig. 1).

## Behaviour suggested that distractors were successfully filtered out

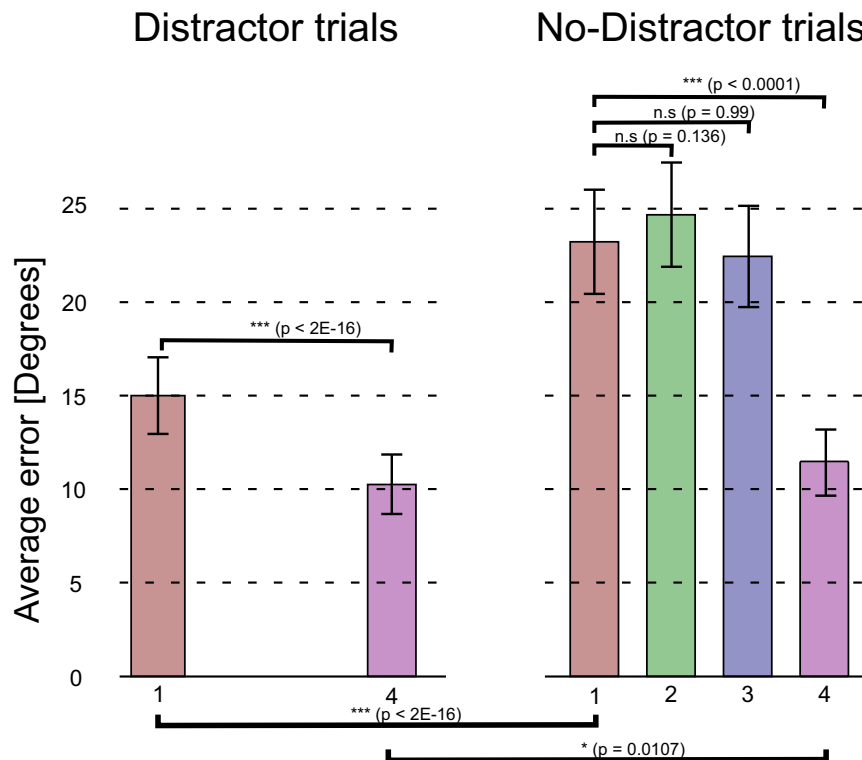
To establish that subjects correctly performed the task and selectively encoded target bars, we first analysed behaviour. We measured performance as the absolute value of the angle between the orientations of the probed and reported bars. The data was modelled using a linear mixed-effects model, with subject as random effect ( $n = 17$ ). To meet the general linear model assumptions the performance data was first transformed (see Methods). Condition (Distractor or No-Distractor), order of the bar probed and the interaction between these two variables were modelled as fixed effects. In the first model, trials probing bar 2 and 3 in the No-Distractor condition were removed to match the trials of the Distractor condition. Effect of condition was significant ( $\beta = 0.56539$ ,  $t(5080) = 9.89$ ,  $p < 2e-16$ ), as well as order ( $\beta = -0.46529$ ,  $t(5080) = -9.968$ ,  $p < 2e-16$ ). Further, the interaction between condition and order of the probed bar was significant ( $\beta = -0.42525$ ,  $t(5080) = -5.26$ ,  $p < 2e-7$ ), meaning that the effect of condition was not the same for bar 1 and bar 4. The performance in the Distractor condition was better than on the No-Distractor condition, although the task difference for bar 1 ( $\beta = 0.5654$ ,  $t(2532) = 9.585$ ,  $p < 2e-16$ ) was larger than for bar 4 ( $\beta = 0.1401$ ,  $t(2352) = 2.554$ ,  $p = 0.011$ ). This suggested that participants indeed treated bar 2 and bar 3 as distractors on these trials, freeing up resources to encode the two target bars on Distractor trials with higher precision.

Separating the data according to the condition, we found a clear serial order effect for both conditions, with much smaller errors on the last bar in the sequence relative to the first bar (Distractor trials ( $\beta = -0.46529$ ,  $t(3382) = -10.2$ ,  $p < 2e-16$ ), No-Distractor trials ( $\beta = -0.89054$ ,  $t(1682) = -12.96$ ,  $p < 2e-16$ ). Finally, we analysed performance on the No-Distractor condition in relation to the bar probed, considering all four bars. There was a significant effect of bar ( $p < 0.0001$ ).



**Fig. 1 | Experimental design.** A sequential array of four bars (stim 1–4) with random orientation was presented. Each bar was randomly associated with a colour (red, green, blue or magenta dot in the middle of the bar). Later, a retro-cue with one of the four colours was presented, identifying the bar orientation to recall. In half of

the trials (randomized), bar 2 and 3 in the sequence were never tested and acted as distractors. These trials were indicated by a pre-cue. The pre-cue and all four bar stimuli were frequency tagged (Methods). The subjects had 5 s to submit a response before the next trial. Each subject performed 400 trials.



**Fig. 2 | Performance by trial type and the sequential position of the bar stimulus.** Average absolute errors (angular distance between probed angle and response) are shown with their  $\pm 1$  standard error of mean (SEM,  $n = 17$  subjects), for each bar in the sequence and by trial type (left Distractor trials, right No-Distractor trials). Significance analysis was conducted using linear mixed-effects models,

which included fixed effects for trial type and sequential position, and random effect for subjects. Planned comparisons were conducted using Tukey's method. All hypothesis tests were two-sided. Significance levels are indicated as follows: \*\*\* denotes  $p < 0.001$ , \*  $p < 0.05$ . Other statistical details are provided in the main text. Source data are provided as a Source Data file.

Planned comparisons using Tukey's method showed that the effect was driven by better performance on the last bar when compared to the three first ones, for which there were no statistically significant differences in performance (Fig. 2).

Taken together, the results suggested that the subjects appeared to encode all target bars and selectively skipped to encode distractors. The encoding of additional bars in the sequence in No-Distractor trials degraded the memory representations of the bars already encoded, leading to a strong order effect. Thus, for the last bar in the sequence there was relatively good performance regardless of trial type.

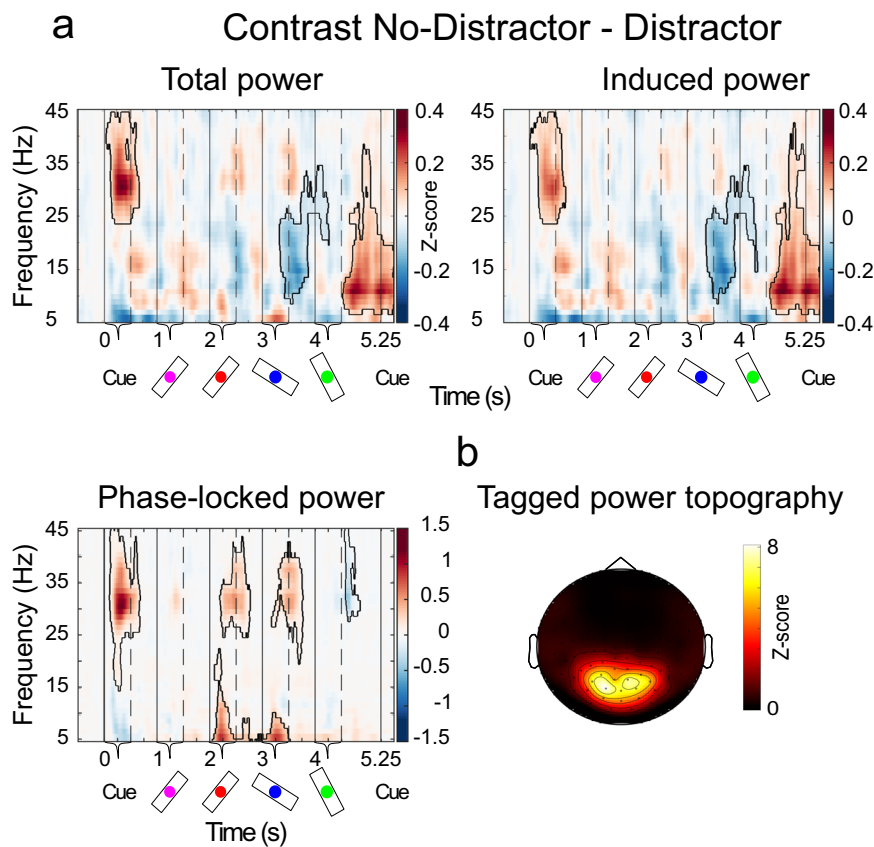
### Frequency tagging revealed correlates of selective encoding

We next analysed the MEG signals at the sensor level. Each oriented bar stimulus and the pre-cue were frequency tagged (alternating between 31.1 or 37.1 Hz, see Methods). Because the frequency tagging was phase-locked to the stimulus onset, we separately analysed total, phase-locked and induced power. Neural activity corresponding to the processing of the tagged stimuli would be most strongly observed in the phase-locked power as the tagging itself was phase-locked to the stimulus onset (Fig. 3a; see contours  $p < 0.05$ ). We assessed neural substrates of gating by contrasting power in the No-Distractor and Distractor trials. For phase-locked power this analysis demonstrated that distractors entrained cortical activity in the tagged frequencies to a significantly lower degree than targets around stimulus presentations (Fig. 3a). The topography of these effects suggested that the main difference in entrainment of tagged frequencies were around the occipital sensors (Fig. 3b). In the induced power, more strongly reflecting endogenous oscillations, we observed significant differences in the alpha and beta frequency ranges between these two conditions around the same time (Fig. 3a). Thus, it suggested the

involvement of alpha and beta oscillations as correlates of filtering information into WM.

### Alpha and beta bursts correlate with selective suppression of distractors

To directly connect our current results to prior work on beta bursts in non-human primates<sup>8,12</sup>, we extracted bursts in three frequency ranges (alpha: 8-12 Hz; low beta: 12-18 Hz; high beta: 18-26 Hz) for each sensor (Methods). The choice of distinct frequency bands were motivated by our recent findings in non-human primates that the "functional" beta frequency tend to be lower further down in the cortical hierarchy<sup>9</sup>. Thus, we suspected that there could be important differences between sub-bands in terms of cortical origin and function. Further, our analysis was bounded upwards to avoid overlap with the tagged frequencies. The average burst rates, for all sensors and all trials, for both alpha and beta frequencies showed strong similarities with prefrontal beta burst rates in non-human primates (Fig. 4a, b). Namely, the burst rates were elevated during fixation and delays, and suppressed during stimulus presentations. This is generally consistent with their proposed inhibitory role in filtering cortical bottom-up processing<sup>8,12</sup>. There were however two striking differences between the alpha and beta burst rates over time. First, alpha peaked just before stimulus onset, was suppressed during each stimulus presentation but then started to smoothly rebound and peak just before the next presentation. This rebound started even before the stimulus was removed. In the beta bands, the pattern was similar but whereas alpha had preserved levels of bursting between the different stimuli in the sequence, beta burst rates were gradually lower as the sequence progressed (Fig. 4b, c). One-sample t-tests confirmed a negative burst gradient for low beta ( $t(16) = -4.03$ ,  $p < 0.001$ , Cohen's  $d = 0.93$ ) and high beta ( $t(16) = -7.04$ ,  $p < 0.001$ , Cohen's  $d = 1.63$ ), while alpha was non-significant



**Fig. 3 | Time-frequency analysis of Distractor and No-Distractor trials. a** Power contrasts between the Distractor and No-Distractor trials. Spectral decompositions were done using adaptive superlets and z-scores were calculated on a trial level normalized to a pre-stimulus window of 250 ms before the onset of the cue (see Methods). Induced power was calculated by subtracting the event-related field from the raw MEG data before spectral decomposition and phase-locked power as

the difference between total and induced power. Vertical lines show the onset/offset of the cues and bars. Boundaries delineate areas as statistically significant using cluster-based non-parametric tests at the 95% confidence level. **b** The topographical sensor plot shows the average phase-locked power for the tagged frequencies (31.1 Hz and 37.1 Hz).

( $t(16) = -1.39$ ,  $p = 0.09$ , Cohen's  $d = 0.32$ ). We have previously observed the same for beta bursting in the prefrontal cortex of non-human primates<sup>8</sup>. Second, in contrast to alpha, beta burst rates were timed to both stimulus onset and offset, not just onset (Fig. 4d). Both these differences were clearer in the higher beta band compared to the lower one, gradually changing from alpha to higher beta.

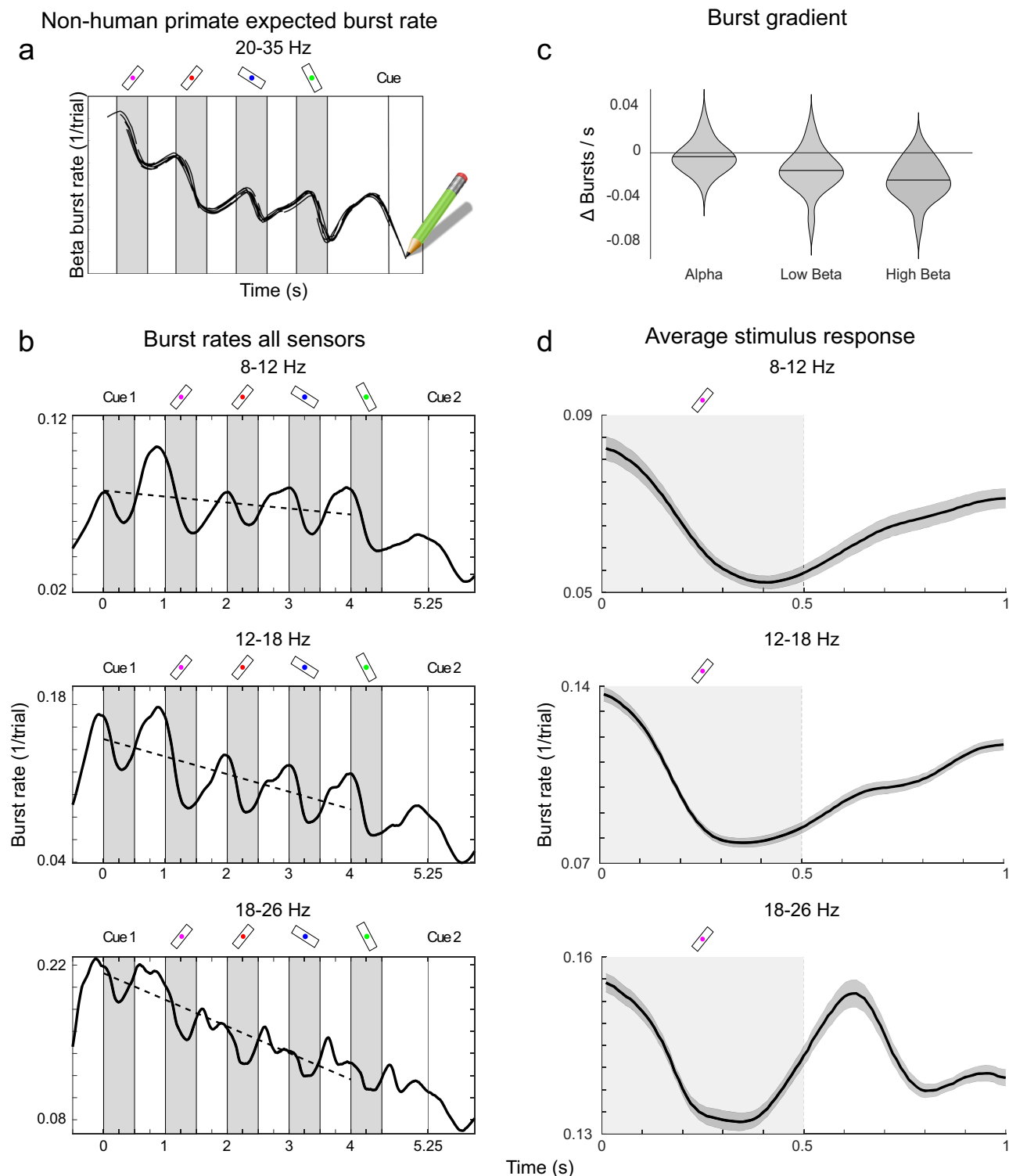
### Alpha and beta bursts exhibited distinct functions depending on cortical origin

To better understand the relevance of burst patterns observed in the alpha and beta bands for WM filtering, we next investigated how they differed in occipital, parietal and prefrontal sensors, and between Distractor and No-Distractor trials (Fig. 5; see Figure S1 for the same analysis based on power). There were several significant differences between the two conditions. First, there was increased beta and alpha bursting during and immediately following the presentations of distractors when compared to targets (Figs. 5, 6). This difference was only seen in occipital sensors (Fig. 6a; main significant clusters: alpha  $p$  value  $< 0.001$ , low beta  $p$  value  $< 0.001$ , high beta  $p$  value  $< 0.001$ ; see Figure S2 for parietal and prefrontal sensors) and coincided with the decreased processing of distractors as measured by the phase-locked power in the tagged frequencies over occipital sensors (Fig. 3a). The elevated bursting in occipital sensors during distractors had different characteristics depending on the frequency band (Fig. 6a). Alpha bursting was more suppressed during and following the presentation of targets relative distractors. Beta bursting, by contrast, had a distinct peak at around the time when distractor

presentations ended (Figs. 5, 6a). This was consistent with their distinct temporal dynamics, i.e. beta (high beta in particular) was timed to both the onset and the offset of stimuli and not just the onset as in the case of alpha. Thus, the difference in beta bursting between distractors and targets was more focused around the offset of the distractor, in the transition from sensory to WM processing.

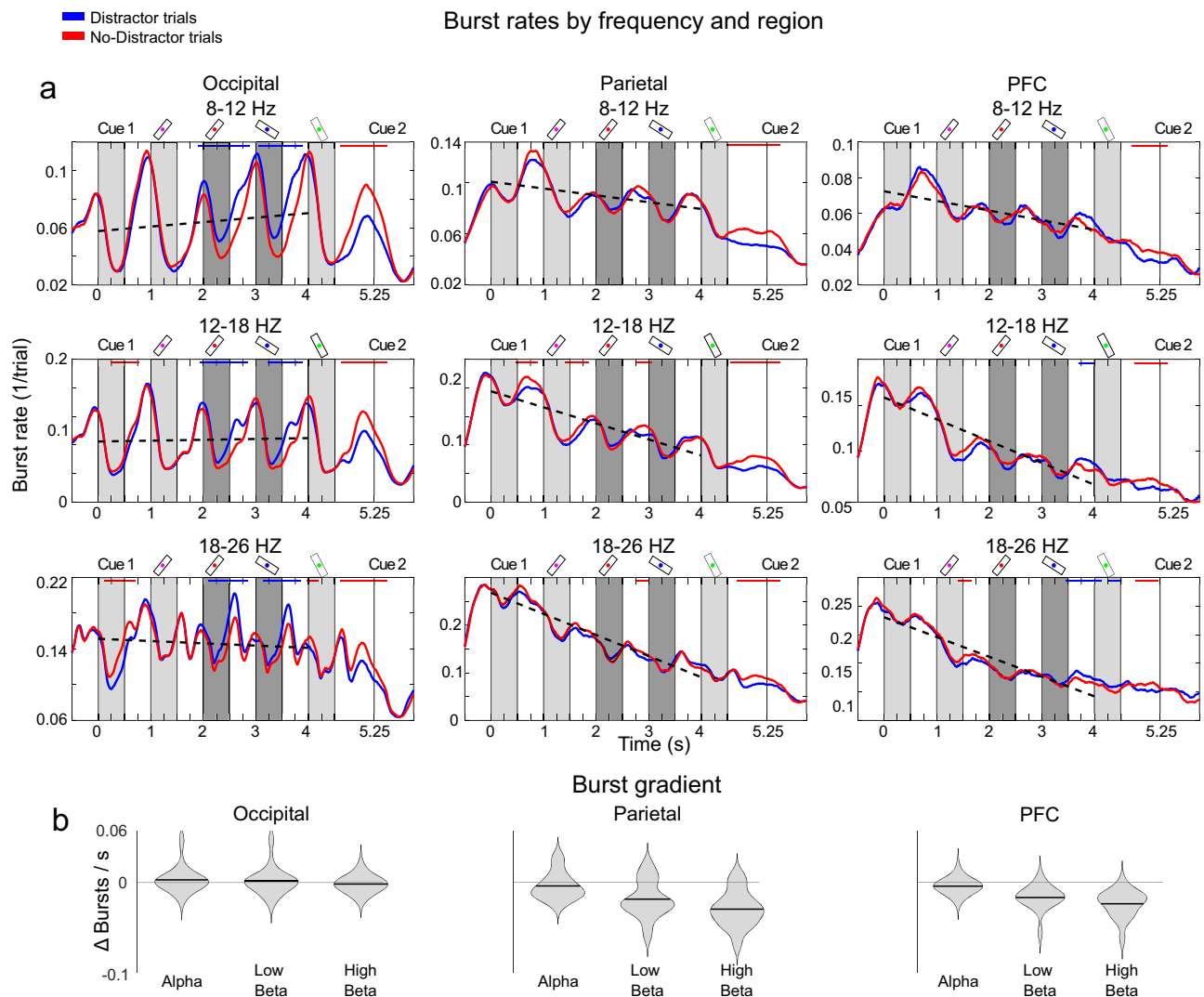
Second, in contrast to the above, there was less bursting in Distractor trials in the delay periods preceding the distractors compared to the delay period prior to the corresponding target stimuli (bar 2 and 3) in No-Distractor trials. This was primarily seen in parietal and prefrontal sensors and mostly in the beta bands (Fig. 6b; parietal main significant clusters: low beta  $p$  value  $< 1e-6$ , high beta  $p$  value  $< 1e-6$ , PFC main significant clusters: low beta  $p$  value = 0.001, high beta  $p$  value  $< 1e-6$ ; similar qualitative differences in alpha were not significant, see Fig. 5a and Figure S2). Thus, in the period before target items were about to be presented, there was elevated bursting in prefrontal and parietal sensors. This is again consistent with the suggested inhibitory role of beta bursting: it proactively freed up resources in higher-order regions for the upcoming WM item by downregulating information already held in WM. Prefrontal sensors had overall similar behaviour as the parietal, with a key difference that there was also a brief period of elevated bursting after the second distractor consistent with disregarding or removal of that task-irrelevant information.

Third, the gradual decrease in beta bursting over the course of the sequence was only observed over parietal and prefrontal, not occipital sensors (Fig. 5b). One-sample t-tests showed no significant gradients in



**Fig. 4 | Burst rates across all sensors.** **a** Expected beta burst rate in Macaque monkeys based on prior research where monkeys were tasked to memorize 3 stimuli for periods of 0.3 s<sup>8</sup>. The illustration has been adapted to be comparable to our experiment. **b** Distinct temporal dynamics of alpha and beta bursts across all sensors. Plots show the grand average burst rates (across all sensors, trial types and subjects) for the three frequency bands. Burst rate is calculated as the fraction of trials where a burst was detected at that given part of the trial (Methods). The greyed areas represent the bar stimulus presentation, while the dotted line

indicates the linear trend for burst rate decay calculated using least squares. **c** Burst gradient by frequency band for all trial types, represented as the slope of the dotted lines in Fig. 4b. Violin plots are based on subject level averages, horizontal line indicates mean. **d** Burst rate from c, averaged over 1 s periods including the presentation and the following delay period for all 4 bar stimuli. Grey regions indicate stimulus presentation periods, error bands represent  $\pm 1$  SEM. Source data are provided as a Source Data file.



**Fig. 5 | Bursting by frequency and cortical region.** **a** Burst rates for trial types and regions. Burst rates per frequency band (rows) and region (columns) are shown. Burst rates for Distractor (blue) and No-Distractor (red) trials are plotted independently. Blue bars denote when Distractor trial burst rates were significantly above No-Distractor trial burst rates, and red bars the opposite, using two-sided cluster-based permutation test at the  $p < 0.001$  level. The dotted line indicates the

burst gradient, calculated through a linear fit using least squares to all trials. Distractor presentation periods in Distractor trials (bar 2 and bar 3) are indicated by dark (as opposed to light) grey areas. **b** Burst gradient by frequency band and region for all trial types, represented as the slope of the dotted lines in **a**. Violin plots are based on subject level averages, horizontal line indicates mean. Source data are provided as a Source Data file.

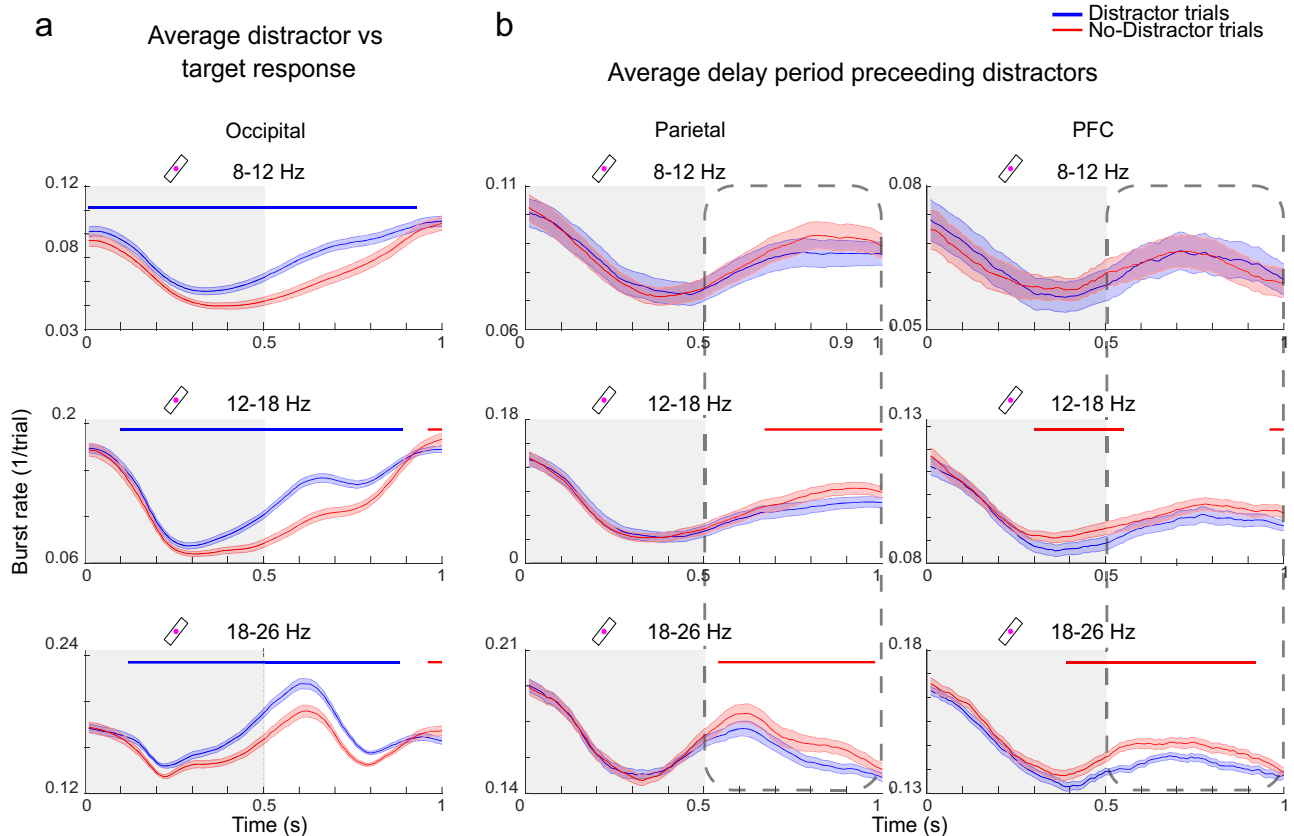
occipital cortex: alpha ( $t(16) = 0.71$ ,  $p = 0.76$ , Cohen's  $d = 0.16$ ), low beta ( $t(16) = 0.31$ ,  $p = 0.62$ , Cohen's  $d = 0.07$ ), high beta ( $t(16) = -1.21$ ,  $p = 0.12$ , Cohen's  $d = 0.28$ ). In parietal cortex there were negative gradients for low beta ( $t(16) = -4.22$ ,  $p < 0.001$ , Cohen's  $d = 0.98$ ) and high beta ( $t(16) = -7.36$ ,  $p < 0.001$ , Cohen's  $d = 1.70$ ), but not for alpha ( $t(16) = -1.40$ ,  $p = 0.09$ ). In PFC there was a significant gradient in all frequency bands: alpha ( $t(16) = -2.34$ ,  $p < 0.05$ , Cohen's  $d = 0.54$ ), low beta ( $t(16) = -5.81$ ,  $p < 0.001$ , Cohen's  $d = 1.34$ ) and high beta ( $t(16) = -7.55$ ,  $p < 0.001$ , Cohen's  $d = 1.74$ ). Finally, in the delay period following the sequence, just prior to the retro-cue there was a larger elevation of bursts for No-Distractor trials in all three frequency bands and all sensors (see Fig. 5a for statistical differences at  $p < 0.001$ ). This difference likely reflected WM load (2 vs 4 items), and disappeared after the retro-cue, when in both conditions only a single item remained relevant for the task.

#### Source reconstruction suggested the involvement of sub-cortical structures

We performed the main analysis on sensor space data since it yielded robust differences between regions while requiring minimal

processing and choices in the analytical pipeline. However, source reconstruction may further highlight important differences between alpha and beta bursts. Using subject T1 scans, we applied minimum norm estimate (MNE) to project sensor level activity onto an MNI normalized source space with  $1 \text{ cm}^3$  resolution (Methods). The MNI source space was parcellated into volumes of interest using the AAL atlas, and for each volume we performed burst analysis. First, we compared our source reconstructed data to our sensor space findings by taking the average burst rates per condition across all volumes within our three regions of interest. This yielded overall highly similar results to the analysis performed in sensor space (Figure S3). However, a striking difference was in the occipital regions, where high beta bursts rates almost exclusively seemed to respond to stimulus offset, accentuating the differences with alpha burst rates in the same region.

Next, focusing on the differential bursting between distractor and target presentations we tested for significant clusters in each brain volume (from the AAL atlas) for each frequency band independently (using  $p < 0.01$  for cluster threshold and significance alpha during



**Fig. 6 | Bursting before and after distractor presentations.** **a** Burst rates as shown in Fig. 5a but averaged over 1 s periods starting from the onset of bars 2 and 3 (serving as distractors in Distractor trials) for each frequency band and the occipital area. The grey shaded area indicates the presentation period of bars 2 and 3. **b** Burst rates as shown in Fig. 5a but averaged over 1 s periods starting from the onset of bars 1 and 2 for each frequency band and the parietal and PFC areas. The dotted

boxes represent the delay periods preceding the presentation of distractors (blue, in Distractor trials) or targets (red, in No-Distractor trials) for bar 2 and 3 in the sequence. Blue bars denote when Distractor trial burst rates were significantly above No-Distractor trial burst rates, and red bars the opposite, using two-sided cluster based permutation test at the  $p < 0.001$  level. Error bands reflect  $\pm 1$  SEM. Source data are provided as a Source Data file.

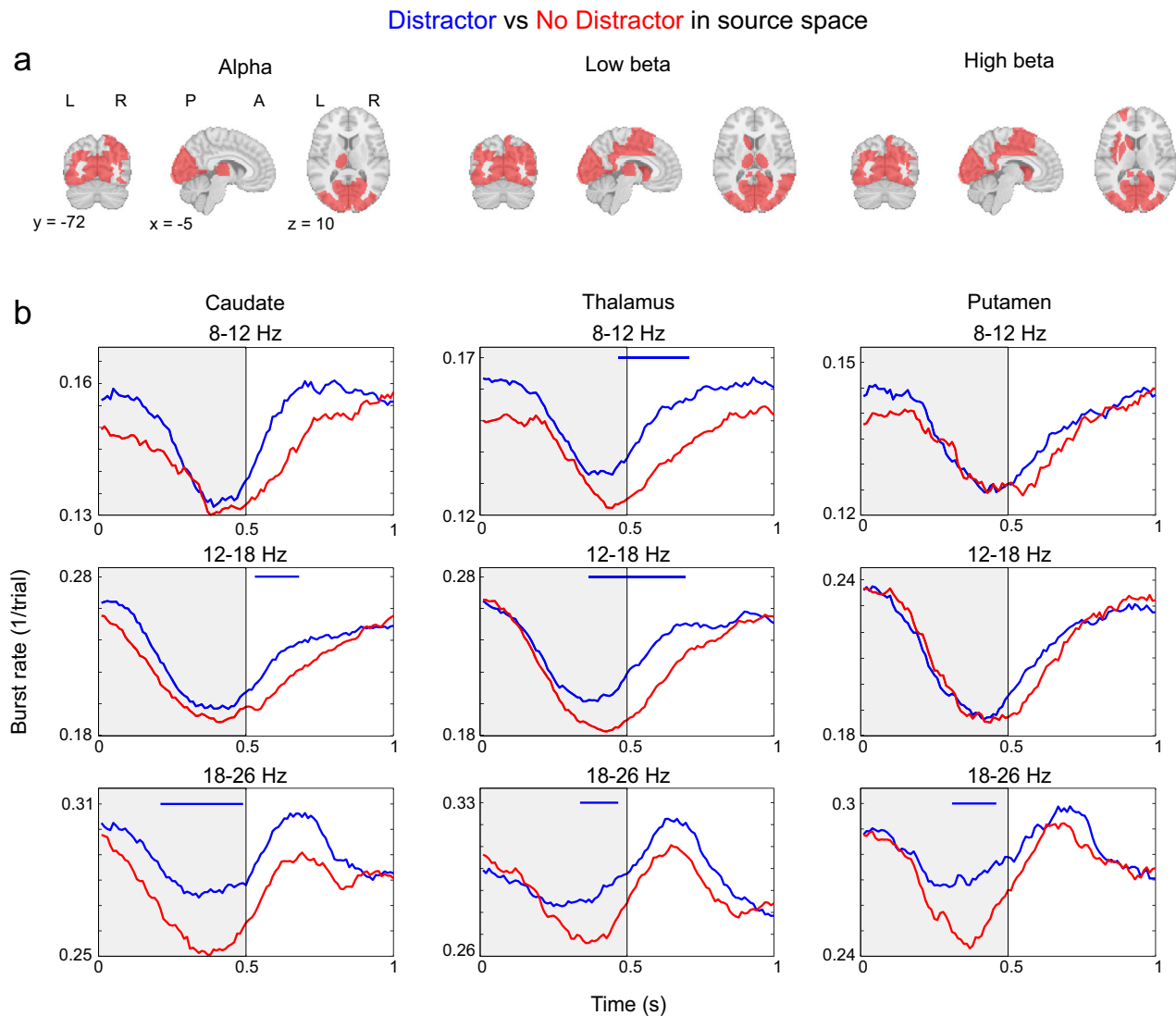
encoding, see Methods). This revealed that effects were focused on the occipital volumes, particularly for alpha (Fig. 7a). For all three frequencies we also observed effects in thalamus and fusiform gyrus within temporal cortex (Fig. 7a). For beta frequency bursting we additionally observed differences in parietal and dorso-striatal volumes (Fig. 7a, b), with elevated bursting for distractors. The involvement of thalamus and striatum is of particular interest as we have speculated that beta bursts might arise from cortico-subcortical interactions<sup>5,15</sup>. To investigate temporal relationships between bursts in different brain regions we performed burst-triggered analysis between pairs of brain regions. It allowed us to discern if the onset of a burst in a certain region tended to precede or follow bursts in another. Overall, our findings suggested that bursts in prefrontal and occipital regions led parietal and sub-cortical bursts (Figure S4). These patterns were mostly stable over different phases of the task. However, prefrontal cortex was the highest in the hierarchy during encoding, and occipital cortex during retention.

### Beta and alpha bursts correlated with encoding on single trials

The burst analysis offered an opportunity to study beta and alpha bursts as single-trial correlates of sensory filtering. We first analysed how the amplitude of phase-locked power in the tagged frequencies during stimulus presentations was modulated by the presence of bursts (burst-triggered phase-locked power). For this we used all bursts from occipital sensors, i.e. the sensors in which tagged frequencies were modulated by distractors, occurring during stimulus processing (between 200 to 700 ms from stimulus onset, the period in

which tagged frequencies were elevated). Since both burst rates and phase-locked power showed strong modulation over the course of a stimulus presentation, we compared burst-triggered power to surrogate data to reveal a single-trial relationship (Methods). The surrogate data was produced by shuffling trial indices (for each sensor, subject and condition independently) for which the associated phase-locked power was taken from. Thus, importantly, the temporal structure was preserved in surrogate data. This revealed that bursts in all three frequency bands had a small but significant inhibitory effect on phase-locked power (see Fig. 8 for significant clusters for 31.1 Hz tagging frequency; see Figure S5 for 37.1 Hz tagging frequency). Again, there were some differences between bands. For alpha, power was reduced following the onset of bursts. For beta, there was in addition significantly elevated phase-locked power prior to the onset of bursts. We interpreted this as beta bursts timely responding to the increases of power in tagged frequencies occurring around stimulus presentations, consistent with beta's role in active regulation of processing. As a control analysis we performed the same analysis but based on the frequency that was not tagged at each presentation (Figure S6). It revealed similar effects (likely resulting from spectral leakage from the tagged frequency) but an order of magnitude smaller, thus ruling out a possibility that the effects were based on spectral leakage from bursts themselves.

Finally, to relate bursts to behaviour in single trials, we constructed a linear mixed-effects model based on trials where probed items were the first or last bar. The performance (absolute magnitude of error) data was transformed to meet the general linear model



**Fig. 7 | Burst rates in source space starting from onset of bars 2 and 3 (serving as distractors in distractor trials).** **a** AAL regions with significant burst rate differences between trial types ( $p < 0.01$ ) during bar 2 and 3 presentation periods. **b** Burst rates averaged over 1 s periods starting from the onset of bar 2 and 3. The grey shaded area indicates the presentation period of bars 2 and 3. Blue bars denote

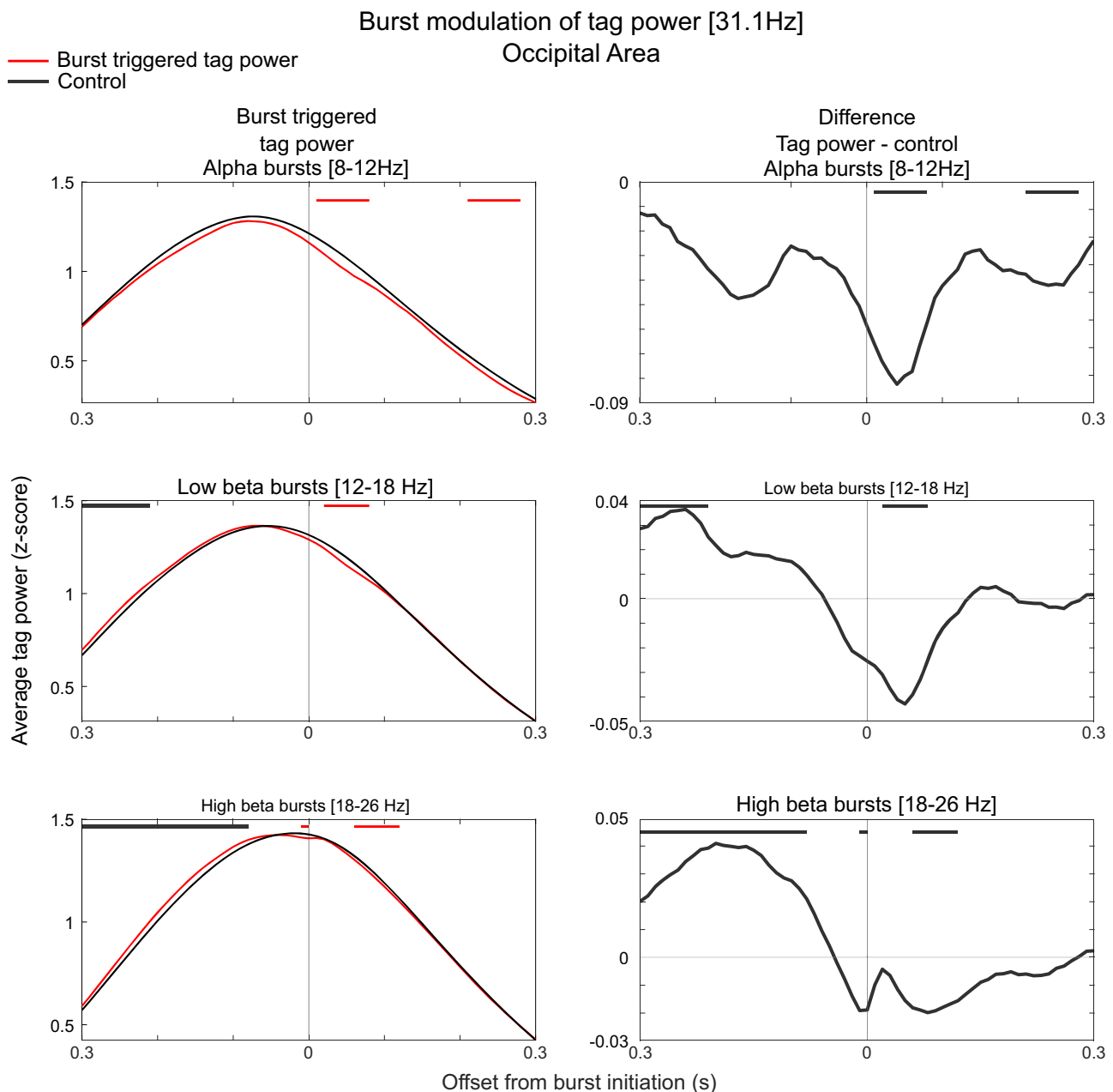
when Distractor trial burst rates were significantly above No-Distractor trial burst rates. AAL regions for Caudate (Caudate\_L, Caudate\_R), Thalamus (Thalamus\_L, Thalamus\_R) and Putamen (Putamen\_L, Putamen\_R). Test were done using two-sided cluster-based permutation test at the  $p < 0.01$  level. Source data are provided as a Source Data file.

assumptions (see Methods). The model had subject as random effect, condition, serial order of probed item, and occipital burst count during encoding (onset to 500 ms) as fixed effects. We performed this analysis for each frequency band separately. It suggested significant relationships in all frequency bands (alpha burst count (coefficient = 0.844,  $t = 4.18$ ,  $p = 3e-5$ ,  $df = 4908$ ), probed item order (coefficient = -0.435,  $t = -15.87$ ,  $p < 2e-16$ ,  $df = 4896$ ), trial type (coefficient = 0.343,  $t = 8.35$ ,  $p < 2e-16$ ,  $df = 4896$ ); low beta burst count (coefficient = 0.667,  $t = 3.42$ ,  $p = 0.0006$ ,  $df = 4907$ ), probed item order (coefficient = -0.430,  $t = -15.67$ ,  $p < 2e-16$ ,  $df = 4896$ ), trial type (coefficient = 0.344,  $t = 8.37$ ,  $p < 2e-16$ ,  $df = 4896$ ); high beta burst count (coefficient = 0.587,  $t = 2.71$ ,  $p = 0.0068$ ,  $df = 4904$ ), probed item order (coefficient = -0.428,  $t = -15.06$ ,  $p < 2e-16$ ,  $df = 4896$ ), trial type (coefficient = 0.345,  $t = 8.39$ ,  $p < 2e-16$ ,  $df = 4896$ )), such that higher burst counts during encoding subsequently led to lower WM accuracy. In addition, we performed the same analysis but using burst counts in prefrontal and parietal regions preceding the second and third stimuli. It suggested weak evidence in the high beta band, which was not significant (alpha burst count: coefficient = 0.0030,  $t = 1.16$ ,  $p = 0.25$ ,  $df = 4910$ ; low

beta burst count: coefficient = 0.148,  $t = 1.37$ ,  $p = 0.17$ ,  $df = 4906$ ; high beta burst count: coefficient = 0.247,  $t = 1.89$ ,  $p = 0.06$ ,  $df = 4901$ ).

## Discussion

We set out to resolve the distinct neural correlates of WM control proposed by earlier studies. Prior work has either suggested prefrontal beta bursts (non-human primates;<sup>5,8,10-12</sup>) or occipital power fluctuations in alpha (humans;<sup>17,19-21,23,25,26,32,33</sup>) as correlates of visual attention and WM control. We applied the single-trial burst analysis we had originally developed for the non-human primate data<sup>8</sup> on human data recorded with MEG. Overall, the quantification of cortical rhythmic activity with narrow-band bursts rather than with more traditional power has several advantages<sup>15</sup>. First and foremost, it is conceptually important whether discrete or slowly changing neural dynamics support cognitive processes. In addition, burst analysis provides enhanced signal-to-noise ratio by reducing the effect of aperiodic activity in the given band between the burst events. It also offers meaningful information about the timing of neural activity thereby helping to relate alpha/beta activity to behaviour or other metrics on a



**Fig. 8 | Burst modulation of tag power for each frequency band.** Left column shows the burst triggered phase-locked power ( $31.1 \pm 2$  Hz, z-scored, see Methods) in red. Time 0 denotes the time of the detected burst onset. Displayed is also surrogate data, calculated by shuffling trial labels (within each condition and sensor independently) to estimate the power in the tagged frequencies drawn from the same distribution of times as the observed bursts (grey). Red bars denote periods in

which power in the original data is lower than the surrogate data and grey bars indicate when power is higher in the original data than the surrogate data using two-sided permutation test at  $p < 0.001$ . The right column zooms in the difference between modulated tag power and surrogate data. The right column shows the difference between the actual data and the surrogate data. Source data are provided as a Source Data file.

moment-by-moment basis, more directly than smoother estimates of spectral power modulations.

Our single-trial burst analysis of MEG demonstrated that WM-related beta bursts had similar patterns in humans as in non-human primates. In addition, alpha bursts shared a similar behaviour to beta bursts but with some key distinctions. In particular, the reported contrasts between target and distractor processing were observed at different times and in different cortical regions for the two bands. Alpha bursts, primarily in occipital regions, suppressed distractor stimulus processing. Occipital beta bursts were instead elevated just as distractor stimuli were removed from the screen in the transition from sensory processing to WM retention. In addition, beta bursts in the

parietal and prefrontal regions tracked sequence order and were elevated *before* target items. The latter pattern is consistent with inhibitory beta bursts in higher-order cortex freeing up space in WM prior to additional encoding. Earlier studies have indeed suggested that while alpha appears to be important in the filtering of irrelevant sensory information there may also be several complementary mechanisms beyond gating at play<sup>34–36</sup>.

Our existing non-human primate model suggests that the level of beta bursting reflects the level of inhibitory cognitive control<sup>5,8,10,12</sup>. This model is based on simultaneous analysis of spiking and intracranial LFPs on the single trial level, where beta bursts suppress gamma bursts and spiking. It also accounts for overall patterns of beta bursting

during various cognitive operations, and how they differ on recording sites in which information is or is not encoded into the patterns of spiking. Thus, in this model beta bursting is suppressed during encoding of information into WM, especially on recording sites where information is encoded, and during read-out of information from WM. Beta bursting has intermediate levels when information is retained but not used, and then strongly elevated following each trial when information has to be cleared out, in particular on those sites where the information was encoded<sup>5,8,12</sup>. Here we observed analogous beta and alpha burst rate patterns in occipital, parietal and prefrontal regions. While we reported task-specific modulation of alpha and beta bursting in most of the brain, the patterns were relatively weak in prefrontal regions, at odds with the strong focus on the prefrontal cortex in non-human primate WM<sup>3,5</sup>. This likely stems from the nature of MEG recordings with relatively low signal-to-noise ratio over frontal regions. However, the timing of alpha and beta bursts in prefrontal regions tended to precede bursts in other areas, including subcortical structures, suggesting a key role for prefrontal bursting. Additionally, a recent human study suggested that persistent training over time, as is typically the case in non-human primates, makes prefrontal WM representations more pronounced, providing a potential explanation for the differential focus on brain areas between species<sup>37</sup>.

To establish a single-trial correlate of filtering of sensory information using only non-invasive measures we used frequency tagging<sup>21,30,31</sup> and also correlated bursts with behaviour. The latter revealed a single-trial link between the amount of bursting during presentation of target stimuli and subsequent precision in the WM task. Unlike in earlier studies, which analysed power fluctuations, we observed that occipital beta and alpha bursting suppressed sensory processing as estimated by tag power on the single-trial level<sup>21,38</sup>. Thus, our findings suggest a direct link between occipital alpha/beta bursts and cortical excitability. This is important as the lack of such a relationship calls for alternative interpretations of alpha and beta<sup>36</sup>. Here the burst analysis (as opposed to power analysis) bears particular relevance as there was only a transient window of reduced power at tagged frequencies following the onset of a burst. Thus, on the whole-trial level such relationship may be masked if, for instance, bursting is upregulated on trials in which there happens to be a strong cortical response to the stimulus. There was an interesting difference between beta and alpha bursts in this regard. Processing tended to be elevated just prior to each beta burst, and not just suppressed during their presence. We interpret this as beta being upregulated by a feedback mechanism when cortical levels of activity are too high. Alternatively, it was recently observed that beta bursts are preceded by a brief window of excitation before the longer-lasting inhibition<sup>39</sup>. This could also explain the elevated levels of processing preceding beta bursts.

There were several other differences between alpha and beta bursting, and between the cortical regions. We argue that these differences, taken together, suggest that occipital alpha reflects sensory filtering, while beta, particularly in higher-order areas, reflects instances when information is already being processed and cortical activity has to be subdued: occipital alpha was suppressed more during the presentation of target items compared to distractors. However, the overall rates were not modulated by other task factors such as load or the passage of time within trials. Occipital beta bursts rates were also modulated by the presence of distractors but more focused around the time of the removal of the distractor from the screen (thus when it would have to be encoded into WM). Activity over parietal and prefrontal sensors had quite a different pattern with significantly elevated beta bursting *before* the presentations of target items (as opposed to upcoming distractors). We interpret this as downregulation of existing WM-related activity to reallocate resources for the upcoming target items in a proactive way. We draw these conclusions primarily based on timing of the alpha and beta bursting. Further experiments, where the order of distractors and target items are varied, and in which

distractors are not always predictable, may shed further light on this potential distinction between alpha and beta bursting. If, for instance, distractors can appear both before and after target items are presented the proposed role of beta in regulating already encoded information may be directly tested.

In addition, beta (but not alpha) burst rates in parietal and prefrontal regions were strongly modulated by the passage of time within trials, with gradually lower burst rates as the trial advanced. Since this trend was equally strong in Distractor and No-Distractor trials we found it likely to be tracking the task structure over time, rather than being load dependent (which also changed with time but differently for the two types of trials). Thus, beta bursting in higher-order areas may be more directly linked to executive control functions, where keeping track of the various parts of a trial is essential. We have recently proposed that the spatio-temporal evolution of beta bursting patterns help implement cognitive operations by directing information flow to distinct patches of cortex during distinct parts of a task<sup>10,15</sup>. This shapes low-dimensional and task-related aspects of single neuronal spiking that are needed to solve the cognitive task at hand<sup>40,41</sup>.

We also observed a somewhat puzzling finding from the perspective of beta and alpha bursting providing transient inhibition. While at the end of the delay, the retro-cue induced more alpha and beta bursting in no-distractor trials, consistent with more inhibition when more items need to be suppressed around the retro-cue, or protected from sensory interference, there were no load effects prior to the final delay (i.e. between Distractor and No-Distractor trials). It could potentially be that alpha and beta bursting during the encoding sequence mainly reflected executive control for allocation of the WM resources, which in total were similar in the two sets of trials (with more resources per item and better performance in load 2 trials). The consistent reduction in beta bursting observed in the parietal and prefrontal sensors (and sources) with the presentation of each item regardless of cognitive status as distractor or target, might reflect a mechanism of tracking the sequence of events during the trial, as discussed above. Prior studies have reported load dependent alpha and beta power during retention but without any analysis relating to the gradual build-up of WM load during the encoding phase<sup>42,43</sup>.

The role of beta in WM executive function was also recently supported in clinical studies<sup>44,45</sup>. In patients with Parkinson's disease, insufficient cortico-striatal beta power suppression during encoding of information into WM was linked to diminished WM performance and correlated with symptom severity<sup>44</sup>. In obsessive-compulsive disorder, the diminished prefrontal beta power rebound following trials was linked to the impairment in the removal of information from WM<sup>45</sup>.

Changes in alpha and beta power during stimulus processing have been extensively studied<sup>46–49</sup>. These dynamics have been attributed among others to encoding of information into episodic and long-term memory<sup>50,51</sup>, and several theories have suggested a link between alpha and the dynamic functional architecture of cortex<sup>47,48,52</sup>. Specifically related to our findings, numerous human studies have suggested that modulation of alpha power reflects filtering of sensory processing in attention and WM tasks<sup>16–18,20,23,24,31,49</sup>, although not necessarily under strict top-down control<sup>21,36</sup>. Intracranial recordings have demonstrated that alpha activity was finely tuned spatially, consistent with a role in the selective suppression of unwanted information in a visual scene<sup>19,22</sup>. Furthermore, alpha power was reported to reflect the modality-specific suppression of distractors in a WM task involving visual and auditory inputs<sup>23</sup>. A load-dependent upregulation of alpha during WM retention has also been observed, consistent with a role in protecting WM information from distractors<sup>43,53,54</sup>. Thus the current and earlier studies align with the notion that whole-cortex burst (or power) patterns in alpha (primarily in sensory and parietal regions) and beta (primarily in parietal and prefrontal regions) dynamically evolve to orchestrate the flow of sensory information to be stored in or deleted from WM according to behavioural (task) demands<sup>10,48</sup>.

These cortical burst patterns are in turn likely coordinated in interactions with thalamus and basal ganglia<sup>5,15,39,44,55–60</sup>. This was supported by the observation that the differential beta bursting between distractor and no-distractor presentations observed in occipital cortex was mirrored in the thalamus and dorsal striatum in source reconstructed data. Interestingly however, prefrontal cortex tended to lead bursting in the sub-cortical structures. We also observed differential alpha and beta bursting for distractors in fusiform gyrus. Earlier studies implicated temporal cortex in WM, and fusiform gyrus specifically for face WM<sup>61–63</sup>. Recently the level of beta frequency coupling between prefrontal and inferotemporal cortex predicted performance on a WM task in non-human primates<sup>64</sup>.

In summary, our results help unite the body of findings regarding the role of beta and alpha in attention and WM in different cortical areas, across human and non-human primates. They appear to serve distinct roles, which can be further teased apart by future experiments.

## Methods

### Participants

We recruited 19 participants, 12 males and 7 females, aged 21–41 years (mean 27.1), with no known cognitive impairments and tested with Ishihara's test for colour deficiency without remarks. The participants were primarily recruited students and received 500 SEK in compensation vouchers. One male did not perform the experiment due to metal interfering with the MEG scanner, and one male was excluded due to poor data quality resulting from excessive movement during the experiment resulting in a total of 17 participants. The Swedish Ethical Review Authority (Dnr 2021-00336) approved the study. All participants were thoroughly briefed and provided written consent to their participation.

### Experimental paradigm

The purpose of the task was to study burst dynamics across regions and frequencies during the presentation of stimuli and the effect of distractors. The task was a sequential working memory task with 4 foveally presented, randomly oriented bars (Fig. 1). Before the onset of each sequence there was a pre-cue indicating if the current trial was a Distractor (50%) or No-Distractor trial (50%). Distractor and No-Distractor trials were randomly interleaved with the criterion that a maximum of 3 trials of a certain type could be presented in a row. On No-Distractor trials the participants were tasked with remembering the orientation of all four bars. On Distractor trials only the first and the fourth bars were to be remembered, and the second and third bars were distractors. The No-Distractor trials were indicated by a pre-cue with four parallel vertical solid bars, while the Distractor trials were indicated by the same bars except that the second and third bars were illustrated with dashed edges informing the participant that these were distractors (not to be remembered). Except for the pre-cue, the two types of trials were visually identical, which allowed us to study the neural mechanisms of encoding, filtering and removing information by comparing a to-be memorized item with a distractor. Each trial commenced on a black background, with a white fixation dot during 500 ms, followed by a pre-cue shown for 500 ms. Next, four bars with random orientation, each covering a visual angle of 7.1° were shown sequentially for 500 ms, separated by a brief delay period of 500 ms with a centrally placed white fixation dot. Each individual bar in the sequence was marked with a uniquely coloured fixation dot in the centre (the colour was randomly selected on each trial from red, green, blue and magenta). Following the presentation of the fourth bar, a white fixation dot was shown for 750 ms. The fixation thereafter changed colour (red, green, blue or magenta) acting as a retro-cue. The retro-cue identified the upcoming bar which would be tested for 750 ms after which the subject was presented with a randomly rotated probe. Using a 4-button control pad, the subjects had 5 seconds to rotate the probe to match the memorized bar. The subjects had to

perform 400 trials in blocks of 40 with pauses in between. In each pause the subject was asked to decide when to continue, and in total the task took between 60–70 minutes for the subjects to complete.

### Frequency Tagging

Frequency tagging involves manipulating the spectral power. This is achieved by oscillatory modulation of the intensity of the stimuli, which entrains neural activity to the frequency of the modulation. This provides insights into when and where the sensory stimuli are processed in the brain (as measured by MEG). Frequency tagging of the stimuli has been applied in both auditory<sup>65,66</sup> and visual perception<sup>30,33</sup>. We modulated the luminance of the cue and the four bars by a sine function phase-locked to the onset of the stimuli with tagging frequencies of either 31.1 or 37.1 Hz. We chose prime frequencies to avoid sub-harmonics, and the decimal was added from the result of a limited pilot study. Within each trial, the objects in the sequence were tagged at alternating frequencies (31.1 Hz and 37.1 Hz), with 50% of the trials tagging the first object with 31.1 Hz and the rest at 37.1 Hz. The stimuli with a frequency tag appeared as faint flickering and no visual difference between the two frequencies was seen or reported.

### Procedure, Materials and Data acquisition

Participants were sent information about the experiment in advance and upon arrival they were greeted and fully informed about the task. They signed a consent form, a MEG screening form and provided a suitable set of clothing to change into. All jewellery, hairpins and any other objects were removed at this stage. The participants were then prepared for the procedure by fitting electrodes and head digitalization. Next, they were led into a magnetically shielded room where the experiment was conducted. Before the MEG recordings were made, each subject was given time to practice the task for about 25–40 trials.

The MEG scanner was an Elekta Neuromag TRIUX 306-channel, located inside a 2-layer magnetically shielded room ([www.natmeg.com](http://www.natmeg.com)). The procedure was presented on a FL35 LED DLP Projector from Projection Design running at 32 bit colour in 1920\*1080 at 120 fps, and the experiment was performed using Presentation® software (Version 23.0, Neurobehavioural Systems, Inc., Berkeley, CA, [www.neurobs.com](http://www.neurobs.com)). Eye movements and eye-blinks were recorded with an Eyelink 1000 binocular tracker from SR Research.

Data were recorded at the sampling rate of 1000 Hz for all 306 MEG channels, eye tracker channels, 16 event code channels and 2 channels dedicated to electrooculography and electrocardiography. Subjects submitted their responses using a 4-button inline pad from Current Design.

T1 MRI scans for all subjects were collected in a Siemens Prisma 3 Tesla whole-body MRI for source data analysis.

### Data pre-processing and frequency domain analysis

The recorded data files were run through Maxfilter software by Elekta Neuromag, with temporal signal space separation. Further analysis was done using Fieldtrip. Data files were segmented into trials starting one second before the trial start and one second after the submission of the response. Trials were demeaned, line noise removed and trials with jump artefacts, identified as a z-score greater than 80, were removed. This resulted in 15 trials removed on average per subject. Across the MEG channels 60 ICA components were calculated using FastICA and then used to automatically identify and remove EOG and muscle artefacts using procedures provided by Fieldtrip. ECG artefacts were identified and removed using a semi-automatic procedure as described by Fieldtrip. The data was down sampled to 250 Hz. Planar gradiometers were combined into synthetic sensors and grouped into regions of interests representing occipital, parietal and PFC regions (Figure S7).

Time frequency calculations were done for single trials in the range 5–45 Hz using adaptive, multiplicative superlets<sup>67</sup> with order

ranging from 1 to 10, and a base wavelet length of three cycles. Only data from planar gradiometers were used. Except for burst calculations, a trial-by-trial z-score baseline was calculated based on the 250 ms epoch before the pre-cue onset. Total power was calculated from the pre-processed activity data, and induced power was calculated by removing the ERF from the activity data and applying the same method. Phase-locked power was calculated by subtracting induced power from total power<sup>68</sup>.

### Burst extraction

To identify bursts of high power, we first specified three frequency bands of interest: 8-12 Hz (alpha), 12-18 Hz (low beta), or 18-26 Hz (high beta). The cut-off at 26 Hz was chosen to avoid spectral leakage from the tagging frequency at 31.1 Hz. Within each band and for each sensor we calculated the temporal profile of the induced power using adaptive superlets. Bursts were defined as intervals within each trial where instantaneous power was above 1.5 standard deviations above the mean (running average of the last 10 full trials), and with the duration of at least two cycles of the average frequency of the band<sup>8</sup>. For each subject, the burst events were averaged across the stimuli into a burst rate per subject, from which a grand average burst rate (per frequency band and cortical region) was calculated. Individual burst times were also kept for further single trial analysis.

### Burst triggered frequency tagged power

Burst events for each frequency band, subject, channel and trial were extracted for a window of 200-700 ms after the onset of a frequency tagged stimuli. Each event was associated with its descriptive information (subject, channel, trial type), and the spectral power within the tagged frequency (narrow band of 4 Hz) was extracted from phase-locked power, during a window of 30 ms before and after the onset of the burst event. The burst triggered frequency tagged phase-locked power was calculated by averaging events within subject, and then across subjects.

### Statistical analyses

To analyse behaviour, we collected all responses and compared their angles to the angles of the probed bars. Analysis was done in R using packages `tidyr`, `dplyr`, `tidyverse` and `circstats`. The circular error, defined as the angle of the response subtracted from the angle of the probed bar in a circular reference frame (from  $-90^\circ$  to  $+90^\circ$ ), was calculated for all trials. We then took the absolute value of the error and applied a transformation, to assure that the assumptions of the general linear model were met. Zero errors were set to 0.5 degrees and we applied a Box-Cox transformation with parameter  $L = 0.101$ . The parameter  $L$  was estimated to meet as close as possible the normality assumptions. Posterior inspection of the residuals showed a good fit with the assumptions of normality (Q-Q plots) and homogeneity of the variances. Mixed effects models were fitted in R<sup>69</sup>. The following packages were used: `lme4`<sup>70</sup>, `lmerTest`<sup>71</sup>, `Mass`<sup>72</sup>, `performance`<sup>73</sup>.

To relate bursts rates between the tasks (Figs. 5, 6 and 7), cluster based non-parametric tests<sup>74</sup> were performed for the burst activity across subjects. Trial types were shuffled within each channel and subject in order to preserve the structure of the data and a grand average was calculated across subjects and channels and t-statistics were calculated and filtered at the 99.9% percentile in sensor space, and 99% percentile in source space. Clusters were defined as the sum of temporally adjacent t values and for each sample the largest cluster was extracted. The procedure was repeated 10,000 times generating a histogram of clusters at which a limit of 99.9% in sensor space, and 99% in source space, was used to identify significant clusters in the original data.

The test for burst frequency tagged power required a different approach (Fig. 8). A matrix containing all burst events with their tagging power and descriptive information was constructed, and trial

numbers were shuffled 10,000 times within their descriptive groups. Averages were calculated and compared to actual data in order to calculate p-values, and tests were performed at the 99.9% percentile.

To compare burst events between source regions (Figure S4), cluster-based non-parametric tests were conducted to contrast burst rates in the target region before (0.3 s to onset) and after (onset to 0.3 s) the burst in the source region. Burst data in the target region, centred around the onset of the trigger in the source region, was divided into pre-burst and post-burst periods at the onset. The pre-burst data was mirrored along the time axis and shuffled with the post-burst data and 10,000 pairs of data were drawn and t-values calculated for each subject id, preserving the structure of the data. Two-sided cluster-based permutation test were performed on the t-values with a cluster threshold of 99%, and significance level of 99% contrasting pre-trigger bursts rates with post-trigger burst rates.

### Source Space Data Analysis

Source-level analysis was conducted using `Fieldtrip`<sup>75</sup>. Individual T1 MRI scans were co-registered with the MEG scanner, re-sliced, and segmented into brain, skull, and scalp volumes. The brain segment was used to construct a volume conduction head model. A spatial filter was calculated on a 10 mm grid, and individual anatomies were warped onto a standard MNI template. The spatial filter was employed to calculate trial-level activity data in source space using Minimum Norm Estimation<sup>76</sup>. Source-level activity was subsequently parcellated using the Anatomical Automatic Labelling (AAL) atlas<sup>77</sup> into 116 volumes. Time-frequency calculations and burst extraction of the parcellated volumes was performed as described above.

Interregional burst correlations were calculated for each frequency band, subject, trial, and source region of interest. This involved extracting burst onset times from the originating region of interest. For each event, a symmetric time window of  $\pm 0.3$  seconds was applied to detect burst onsets within the corresponding target region of interest during the same trial. The detected onsets were then aggregated across all subjects and trials. For visual representation in Figure S4, data preceding each burst event were mirrored along the y-axis to facilitate comparison and enhance interpretability. The volumes of interest used for parcellation using the AAL atlas were PFC (Frontal\_Sup\_L, Frontal\_Sup\_R, Frontal\_Sup\_Orb\_L, Frontal\_Sup\_Orb\_R, Frontal\_Mid\_L, Frontal\_Mid\_R, Frontal\_Mid\_Orb\_L, Frontal\_Mid\_Orb\_R, Frontal\_Inf\_Oper\_L, Frontal\_Inf\_Oper\_R, Frontal\_Inf\_Tri\_L, Frontal\_Inf\_Tri\_R, Frontal\_Inf\_Orb\_L, Frontal\_Inf\_Orb\_R, Frontal\_Sup\_Medial\_L, Frontal\_Sup\_Medial\_R, Frontal\_Med\_Orb\_L, Frontal\_Med\_Orb\_R), Parietal cortex (Parietal\_Sup\_L, Parietal\_Sup\_R, Parietal\_Inf\_L, Parietal\_Inf\_R), Occipital cortex (Calcarine\_L, Calcarine\_R, Cuneus\_L, Cuneus\_R, Occipital\_Sup\_L, Occipital\_Sup\_R, Occipital\_Mid\_L, Occipital\_Mid\_R, Occipital\_Inf\_L, Occipital\_Inf\_R, Precuneus\_L, Precuneus\_R, Basal Ganglia (Caudate\_L, Caudate\_R, Putamen\_L, Putamen\_R, Pallidum\_L, Pallidum\_R) and Thalamus (Thalamus\_L, Thalamus\_R).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The processed data, downsampled MEG data, MRI data and behavioural data used in this study are available in the OSF database under accession code CC BY ([https://osf.io/gu25f/?view\\_only=6412a8bd665e4ef082385dbaa3d33026](https://osf.io/gu25f/?view_only=6412a8bd665e4ef082385dbaa3d33026)). Source data are provided with this paper.

### Code availability

The code for this study is available for download from the Open Science Framework (OSF) homepage ([https://osf.io/gu25f/?view\\_only=6412a8bd665e4ef082385dbaa3d33026](https://osf.io/gu25f/?view_only=6412a8bd665e4ef082385dbaa3d33026)).

## References

1. Ma, W. J., Husain, M. & Bays, P. M. Changing concepts of working memory. *Nat. Neurosci.* **17**, 347–356 (2014).
2. Vogel, E. K., McCollough, A. W. & Machizawa, M. G. Neural measures reveal individual differences in controlling access to working memory. *Nature* **438**, 500–503 (2005).
3. Goldman-Rakic, P. S. Cellular basis of working memory. *Neuron* **14**, 477–485 (1995).
4. D'Esposito, M. & Postle, B. R. The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* **66**, 115–142 (2015).
5. Miller, E. K., Lundqvist, M. & Bastos, A. M. Working Memory 2.0. *Neuron* **100**, 463–475 (2018).
6. Chatham, C. H. & Badre, D. Multiple gates on working memory. *Curr. Opin. Behav. Sci.* **1**, 23–31 (2015).
7. Liesefeld, A. M., Liesefeld, H. R. & Zimmer, H. D. Intercommunication between prefrontal and posterior brain regions for protecting visual working memory from distractor interference. *Psychol. Sci.* **25**, 325–333 (2014).
8. Lundqvist, M. et al. Gamma and beta bursts underlie working memory. *Neuron* **90**, 152–164 (2016).
9. Lundqvist, M., Bastos, A. M. & Miller, E. K. Preservation and Changes in Oscillatory Dynamics across the Cortical Hierarchy. *J. Cogn. Neurosci.* **32**, 2024–2035 (2020).
10. Lundqvist, M. et al. Working memory control dynamics follow principles of spatial computing. *Nat. Commun.* **14**, 1429 (2023).
11. Bastos, A. M., Loonis, R., Kornblith, S., Lundqvist, M. & Miller, E. K. Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proc. Natl Acad. Sci. USA* **115**, 1117–1122 (2018).
12. Lundqvist, M., Herman, P., Warden, M. R., Brincat, S. L. & Miller, E. K. Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nat. Commun.* **9**, 394 (2018).
13. Hindriks, R. & Tewarie, P. K. B. Dissociation between phase and power correlation networks in the human brain is driven by co-occurrent bursts. *Commun. Biol.* **6**, 286 (2023).
14. Zich, C. et al. Spatiotemporal organisation of human sensorimotor beta burst activity. *eLife* **12**, e80160 (2023).
15. Lundqvist, M., Miller, E. K., Nordmark, J., Liljefors, J. & Herman, P. Beta: bursts of cognition. *Trends Cogn. Sci. (Regul. Ed.)* **28**, 662–676 (2024).
16. Sauseng, P. et al. A shift of visual spatial attention is selectively associated with human EEG alpha activity. *Eur. J. Neurosci.* **22**, 2917–2926 (2005).
17. Roux, F. & Uhlhaas, P. J. Working memory and neural oscillations:  $\alpha$ - $\gamma$  versus  $\theta$ - $\gamma$  codes for distinct WM information? *Trends Cogn. Sci. (Regul. Ed.)* **18**, 16–25 (2014).
18. Popov, T., Kastner, S. & Jensen, O. FEF-Controlled Alpha Delay Activity Precedes Stimulus-Induced Gamma-Band Activity in Visual Cortex. *J. Neurosci.* **37**, 4117–4127 (2017).
19. Popov, T., Gips, B., Kastner, S. & Jensen, O. Spatial specificity of alpha oscillations in the human visual system. *Hum. Brain Mapp.* **40**, 4432–4440 (2019).
20. Turner, W., Blom, T. & Hogendoorn, H. Visual Information Is Predictively Encoded in Occipital Alpha/Low-Beta Oscillations. *J. Neurosci.* **43**, 5537–5545 (2023).
21. Gutteling, T. P., Sillekens, L., Lavie, N. & Jensen, O. Alpha oscillations reflect suppression of distractors with increased perceptual load. *Prog. Neurobiol.* **214**, 102285 (2022).
22. Yuasa, K. et al. Precise spatial tuning of visually driven alpha oscillations in human visual cortex. *BioRxiv*. <https://doi.org/10.1101/2023.02.11.528137> (2023).
23. Zhou, Y. J., Ramchandran, A. & Haegens, S. Alpha oscillations protect working memory against distracters in a modality-specific way. *Neuroimage* **278**, 120290 (2023).
24. Haegens, S., Luther, L. & Jensen, O. Somatosensory anticipatory alpha activity increases to suppress distracting input. *J. Cogn. Neurosci.* **24**, 677–685 (2012).
25. Riddle, J., Scimeca, J. M., Cellier, D., Dhanani, S. & D'Esposito, M. Causal evidence for a role of theta and alpha oscillations in the control of working memory. *Curr. Biol.* **30**, 1748–1754.e4 (2020).
26. Wolff, M. J., Jochim, J., Akyürek, E. G. & Stokes, M. G. Dynamic hidden states underlying working-memory-guided behavior. *Nat. Neurosci.* **20**, 864–871 (2017).
27. Rosanova, M., Sarasso, S., Massimini, M. & Casarotto, S. Cortical Excitability, Plasticity and Oscillations in Major Psychiatric Disorders: A Neuronavigated TMS-EEG Based Approach. in *Non invasive brain stimulation in psychiatry and clinical neurosciences* (eds. Dell'Osso, B. & Di Lorenzo, G.) 209–222. [https://doi.org/10.1007/978-3-030-43356-7\\_15](https://doi.org/10.1007/978-3-030-43356-7_15) (Springer International Publishing, 2020).
28. Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M. & Engel, A. K. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nat. Neurosci.* **15**, 884–890 (2012).
29. Samaha, J., Gossesies, O. & Postle, B. R. Distinct oscillatory frequencies underlie excitability of human occipital and parietal cortex. *J. Neurosci.* **37**, 2824–2833 (2017).
30. Parkkonen, L., Andersson, J., Hämäläinen, M. & Hari, R. Early visual brain areas reflect the percept of an ambiguous scene. *Proc. Natl Acad. Sci. USA* **105**, 20500–20504 (2008).
31. Zhigalov, A. & Jensen, O. Alpha oscillations do not implement gain control in early visual cortex but rather gating in parieto-occipital regions. *Hum. Brain Mapp.* **41**, 5176–5186 (2020).
32. Ferrante, O., Zhigalov, A., Hickey, C. & Jensen, O. Statistical learning of distractor suppression downregulates prestimulus neural excitability in early visual cortex. *J. Neurosci.* **43**, 2190–2198 (2023).
33. Zhigalov, A., Herring, J. D., Herpers, J., Bergmann, T. O. & Jensen, O. Probing cortical excitability using rapid frequency tagging. *Neuroimage* **195**, 59–66 (2019).
34. Noonan, M. P. et al. Distinct mechanisms for distractor suppression and target facilitation. *J. Neurosci.* **36**, 1797–1807 (2016).
35. Kim, H., Smolker, H. R., Smith, L. L., Banich, M. T. & Lewis-Peacock, J. A. Changes to information in working memory depend on distinct removal operations. *Nat. Commun.* **11**, 6239 (2020).
36. Jensen, O. Distractor inhibition by alpha oscillations is controlled by an indirect mechanism governed by goal-relevant information. *Commun. Psychol.* **2**, 36 (2024).
37. Miller, J. A., Tambini, A., Kiyonaga, A. & D'Esposito, M. Long-term learning transforms prefrontal cortex representations during working memory. *Neuron* **110**, 3805–3819.e6 (2022).
38. Gundlach, C., Moratti, S., Forschack, N. & Müller, M. M. Spatial attentional selection modulates early visual stimulus processing independently of visual alpha modulations. *Cereb. Cortex* **30**, 3686–3703 (2020).
39. Law, R. G. et al. Thalamocortical Mechanisms Regulating the Relationship between Transient Beta Events and Human Tactile Perception. *Cereb. Cortex* **32**, 668–688 (2022).
40. Badre, D., Bhandari, A., Keglovits, H. & Kikumoto, A. The dimensionality of neural representations for control. *Curr. Opin. Behav. Sci.* **38**, 20–28 (2021).
41. Xie, Y. et al. Geometry of sequence working memory in macaque prefrontal cortex. *Science* **375**, 632–639 (2022).
42. Kornblith, S., Buschman, T. J. & Miller, E. K. Stimulus load and oscillatory activity in higher cortex. *Cereb. Cortex* **26**, 3772–3784 (2016).
43. Jensen, O., Gelfand, J., Kounios, J. & Lisman, J. E. Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb. Cortex* **12**, 877–882 (2002).

44. Paulo, D. L. et al. Corticostriatal beta oscillation changes associated with cognitive function in Parkinson's disease. *Brain* **146**, 3662–3675 (2023).
45. Boo, Y. J. et al. Altered prefrontal beta oscillatory activity during removal of information from working memory in obsessive-compulsive disorder. *BMC Psychiatry* **23**, 645 (2023).
46. Pfurtscheller, G. & Aranibar, A. Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroencephalogr. Clin. Neurophysiol.* **42**, 817–826 (1977).
47. Klimesch, W., Sauseng, P. & Hanslmayr, S. EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* **53**, 63–88 (2007).
48. Jensen, O. & Mazaheri, A. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* **4**, 186 (2010).
49. Worden, M. S., Foxe, J. J., Wang, N. & Simpson, G. V. Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J. Neurosci.* **20**, RC63 (2000).
50. Hanslmayr, S., Matuschek, J. & Fellner, M.-C. Entrainment of prefrontal beta oscillations induces an endogenous echo and impairs memory formation. *Curr. Biol.* **24**, 904–909 (2014).
51. Griffiths, B. J., Martín-Buro, M. C., Staresina, B. P. & Hanslmayr, S. Disentangling neocortical alpha/beta and hippocampal theta/gamma oscillations in human episodic memory formation. *Neuroimage* **242**, 118454 (2021).
52. Bastos, A. M., Lundqvist, M., Waite, A. S., Kopell, N. & Miller, E. K. Layer and rhythm specificity for predictive routing. *Proc. Natl Acad. Sci. USA* **117**, 31459–31469 (2020).
53. Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P. & Winkler, T. "Paradoxical" alpha synchronization in a memory task. *Brain Res. Cogn. Brain Res.* **7**, 493–501 (1999).
54. Bonnefond, M. & Jensen, O. Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Curr. Biol.* **22**, 1969–1974 (2012).
55. Jayachandran, M. et al. Nucleus reuniens transiently synchronizes memory networks at beta frequencies. *Nat. Commun.* **14**, 4326 (2023).
56. Sherman, M. A. et al. Neural mechanisms of transient neocortical beta rhythms: Converging evidence from humans, computational modeling, monkeys, and mice. *Proc. Natl Acad. Sci. USA* **113**, E4885–E4894 (2016).
57. Ketz, N. A., Jensen, O. & O'Reilly, R. C. Thalamic pathways underlying prefrontal cortex-medial temporal lobe oscillatory interactions. *Trends Neurosci.* **38**, 3–12 (2015).
58. de Mooij-van Malsen, J. G. et al. Task-specific oscillatory synchronization of prefrontal cortex, nucleus reuniens, and hippocampus during working memory. *iScience* **26**, 107532 (2023).
59. Saalman, Y. B., Pinsk, M. A., Wang, L., Li, X. & Kastner, S. The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* **337**, 753–756 (2012).
60. Hughes, S. W. & Crunelli, V. Thalamic mechanisms of EEG alpha rhythms and their pathological implications. *Neuroscientist* **11**, 357–372 (2005).
61. Miller, E. K., Li, L. & Desimone, R. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* **254**, 1377–1379 (1991).
62. Kornblith, S., Quiñero, R., Koch, C., Fried, I. & Mormann, F. Persistent Single-Neuron Activity during Working Memory in the Human Medial Temporal Lobe. *Curr. Biol.* **27**, 1026–1032 (2017).
63. Ranganath, C., Cohen, M. X., Dam, C. & D'Esposito, M. Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J. Neurosci.* **24**, 3917–3925 (2004).
64. Rezayat, E. et al. Frontotemporal coordination predicts working memory performance and its local neural signatures. *Nat. Commun.* **12**, 1103 (2021).
65. Bharadwaj, H. M., Lee, A. K. C. & Shinn-Cunningham, B. G. Measuring auditory selective attention using frequency tagging. *Front. Integr. Neurosci.* **8**, 6 (2014).
66. Manting, C. L., Gulyas, B., Ullén, F. & Lundqvist, D. Steady-state responses to concurrent melodies: source distribution, top-down, and bottom-up attention. *Cereb. Cortex* **33**, 3053–3066 (2023).
67. Moca, V. V., Bărgan, H., Nagy-Dăbăcan, A. & Mureşan, R. C. Time-frequency super-resolution with superlets. *Nat. Commun.* **12**, 337 (2021).
68. Cohen, M. X. *Analyzing neural time series data: theory and practice*. <https://doi.org/10.7551/mitpress/9609.001.0001>. (The MIT Press, 2014).
69. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org> (2023).
70. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
71. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).
72. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S*. <https://doi.org/10.1007/978-0-387-21706-2>. (Springer New York, 2002).
73. Lüdtke, D., Ben-Shachar, M., Patil, I., Waggoner, P. & Makowski, D. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *JOSS* **6**, 3139 (2021).
74. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
75. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).
76. Ou, W., Hämäläinen, M. S. & Golland, P. A distributed spatio-temporal EEG/MEG inverse solver. *Neuroimage* **44**, 932–946 (2009).
77. Tzourio-Mazoyer, N. et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289 (2002).

## Acknowledgements

This work was funded by ERC starting grant 949131, and Swedish Research Council (Vetenskapsrådet) grants 2018-04197, 2022-02328 (M.L.).

## Author contributions

Concept: M.L. Experimental design: M.L. and R.A. Data collection: J.Li and G.R. Methodology: M.L. and P.H. Software: P.H. Formal analysis: J.Li and R.A. Supervision: J.Lu., M.L., R.A., and P.H. Visualization: J.L. Wrote first draft: M.L. Text edits: P.H., R.A., J.Lu., J.Li., and M.L. Resources: M.L.

## Funding

Open access funding provided by Karolinska Institute.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-53257-7>.

**Correspondence** and requests for materials should be addressed to Mikael Lundqvist.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024