Neuron

Article

Working memory readout varies with frontal theta rhythms

Highlights

- Theta phase in FEF predicts working memory readout
- Performance varies with theta phase and target location
- Theta "scans" visual space top to bottom
- Theta couples to spiking and beta on opposing phases

Authors

Hio-Been Han, Scott L. Brincat, Timothy J. Buschman, Earl K. Miller

Correspondence

ekmiller@mit.edu

In brief

Han et al. show that frontal theta oscillations rhythmically control access to working memory. The theta rhythm sweeps across the mental image, shaping behavior by coordinating spikes and beta oscillations. This reveals a wave-based mechanism for how the brain reads out stored information.

Han et al., 2026, Neuron *114*, 159–166
January 7, 2026
https://doi.org/10.1016/j.neuron.2025.09.031

CellPress

Neuron

CellPress

Article

Working memory readout varies with frontal theta rhythms

Hio-Been Han,[1,2] Scott L. Brincat,[1] Timothy J. Buschman,[1,3] and Earl K. Miller[1,4,*]
[1]The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]School of Convergence, Seoul National University of Science and Technology, Seoul 01811, Republic of Korea
[3]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540, USA
[4]Lead contact
*Correspondence: ekmiller@mit.edu
https://doi.org/10.1016/j.neuron.2025.09.031

SUMMARY

Increasing evidence suggests that attention varies rhythmically, phase locked to ongoing cortical oscillations. Here, we report that the phase of theta oscillations (3–6 Hz) in the frontal eye field (FEF) is associated with the spatiotemporal variation of information readout from working memory (WM). Non-human primates were briefly shown a sample array of colored squares. A short time later, they viewed a test array and were rewarded for identifying which square changed color (the target). Behavioral performance varied systematically with theta phase at the time of test array onset, as well as with the target's location. This is consistent with theta "scanning" across the FEF and thus visual space from top to bottom. Theta was coupled, on opposing phases, to both spiking and beta (12–20 Hz). These results could be explained by a wave of activity that moves across the FEF, modulating the readout of information from WM.

INTRODUCTION

Cortical activity fluctuates rhythmically, which has consequences for its function. This begins with sensory systems that sample the external world with periodicity.[1,2] It is also evident in visual attention. Even when trained to sustain steady visual attention on a single location, attention nonetheless fluctuates. The ability of non-human primates (NHPs) to detect a target at that location waxes and wanes on a 3–6 Hz (theta) cycle.[3] These alternations of better and worse performance align with the phase of theta local field potential (LFP) oscillations in the frontal cortex. This periodicity of perceptual and attentional processes raises the possibility that other cortical functions, not just those involved in selecting and processing external inputs, might synchronize to particular cortical rhythms.

We examined working memory (WM), which is linked with attention.[4–7] Neural spiking during WM retention shows burstiness that co-varies with LFP rhythms across a wide range of frequencies.[8–14] But, it is not known whether WM function per se cycles at a base frequency, like attention does in theta. Thus, we sought to test whether there was a rhythmic component to WM-dependent behavior.

We analyzed neural activity recorded from the frontal eye field (FEF) in NHPs performing a change identification WM task (Figure 1A).[15,16] NHPs were shown a sample array of colored squares (set size, 2–5), followed by an 800–1,000 ms memory delay. Then, a test array appeared in which one of the squares had changed color (the target). The NHPs were required to maintain central eye gaze until the test array appeared (see STAR Methods). Then, they made a direct saccade toward the target to receive a juice reward.

WM performance depended on both the FEF theta phase at test array onset and the target's position. The findings suggest a traveling wave of activity across the FEF, leading to a top-to-bottom spatial sampling that influenced WM readout.

RESULTS

Behavioral performance cycled with FEF theta phase

Each NHP completed 14 recording sessions (NHP 1, 16,940 trials in total; NHP 2, 15,762 trials in total). Behavioral performance declined as a function of set size, consistent with the limited capacity of WM (Figure 1B).[15,17] LFPs and spiking activity were recorded in FEF during task performance. Overall LFP power was characterized by prominent theta oscillations extending through the trial (Figures 1C and S1A).

To determine whether behavior varied with LFP, we determined the instantaneous phase for frequencies from 2 to 64 Hz at the time of the test array onset. Note that due to the unpredictable time of the test array (Figure 1D), LFP phase in FEF at test array onset was not time locked to external events. LFP phase was compared to the NHP's reaction time (RT) and accuracy (percent correct change identification) by measuring the Kullback-Leibler (KL) divergence from a circular uniform

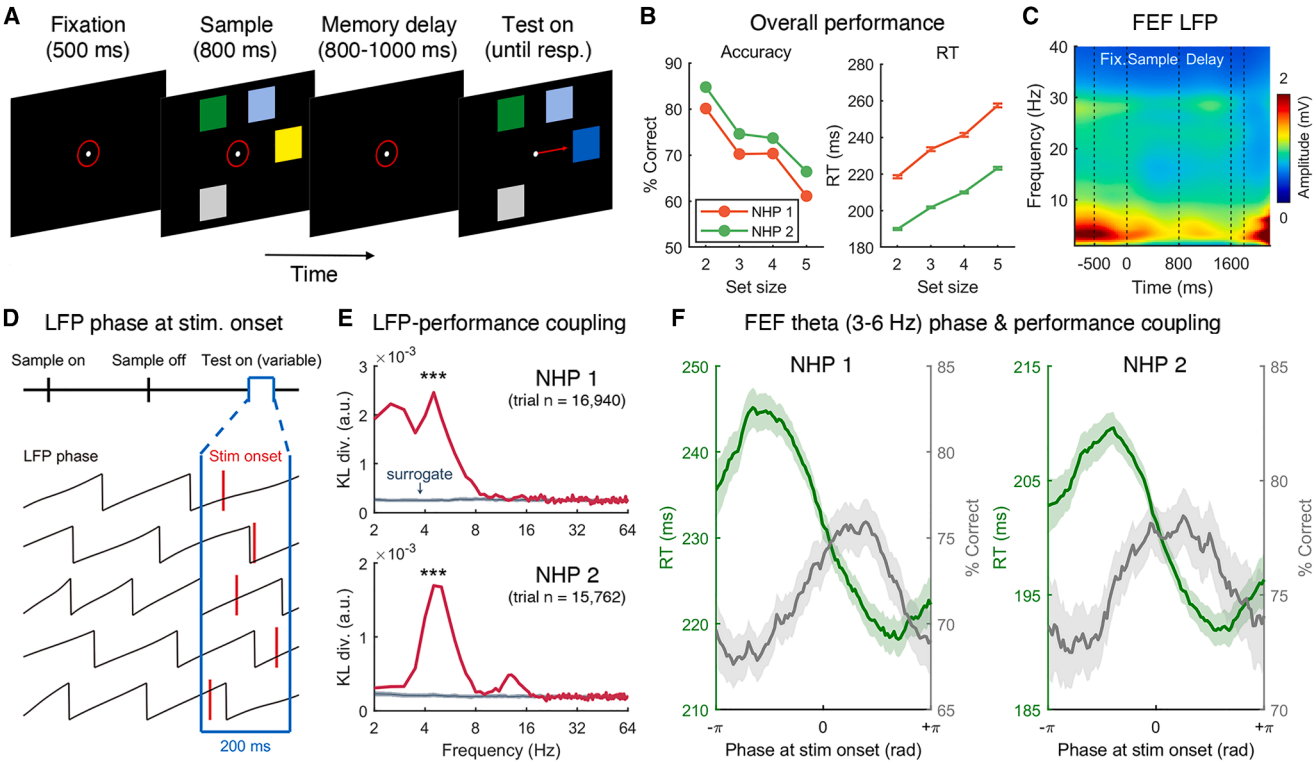

Figure 1. FEF theta modulates WM task performance

(A) WM task sequence. NHPs indicated which square changed color (the target) by making a saccade to its location. Timing of the test array was randomly determined because of the variable length of memory delay period.

(B) Performance of two NHPs as a function of the number of items held in WM. Error bars show ±1 standard error.

(C) Oscillatory amplitude (grand averaged) across the time course of the WM task.

(D) Schematic illustration of the LFP phase at the timing of test stimulus onset. Across all trials, LFP phase was roughly uniformly distributed at test array onset (Figures S1E–S1H).

(E) Non-uniformity of RT distribution over LFP phase measured by KL divergence in FEF (top, NHP 1; bottom, NHP 2). Non-uniformity of the distribution of RTs peaked at 4–5 Hz in FEF. Surrogate data were obtained by trial-shuffled bootstrapping (1,000 samplings).

(F) Behavioral performance (green, RT; gray, percent correct) as a function of FEF theta phase for NHP 1 (left) and NHP 2 (right). Shaded area shows 1 SD, calculated by bootstrapping with 33% subsampling, 10,000 iterations. ***$p < 0.001$ for the result of Wilcoxon's signed-rank test.

distribution across LFP phases, separately at each frequency. Large KL divergence indicate performance reliably differs between LFP phases at test array onset.

This revealed a relationship between FEF theta and task performance. There was a clear peak within the theta range (~5 Hz) for both NHPs, indicating a correspondence between behavior and FEF theta phase when the test array appeared (Figure 1E). Behavioral performance—both RT and accuracy—showed significant modulation by the phase of FEF theta (3–6 Hz; $p < 0.001$ for RT and accuracy, for both NHPs) (Figures S1B and S1C). RT was faster and accuracy was higher if the test array appeared during the falling ("good") phases (i.e., 0 to $+\pi$ rad) of theta relative to the rising ("poor") phases (i.e., $-\pi$ to 0 rad) of theta (Figure 1F). This theta modulation was stronger in trials with higher WM load (Figure S1D). We confirmed that LFP phase at test array onset was not phase locked to trial events and thus uniformly distributed across trials (see Figures S1E–S1H for event-related potentials and inter-trial phase coherence). Phase reset driven by the test array onset only became significant around 150–200 ms after its presentation (e.g., Figure S1G) and did not correlate with behavioral performance before this (see Figures S1I–S1P). This confirmed that behavior was influenced by the intrinsic FEF theta phase at test array onset.

WM errors were spatially organized

We found that behavior not only depended on the theta phase at test array onset, it also depended on the location of the target. Errors were not randomly distributed in space. Rather, incorrect choices tended to be near the correct target. Figure 2B shows a distribution of choices when the target was at the 3 o'clock position. When they chose the wrong array item, that item was more likely to be at 1 and 5 o'clock compared with further locations (e.g., 9 o'clock). Figure 2C shows the choice distribution after rotating and aligning saccadic landing positions relative to the target (for results of each location; see Figures S2A and S2B). The proximity of items to the target location and the frequency of choice errors showed a negative correlation (Figure 2D; NHP 1, $r = -0.70$, and $p < 0.001$; NHP 2, $r = -0.65$, and $p < 0.001$).

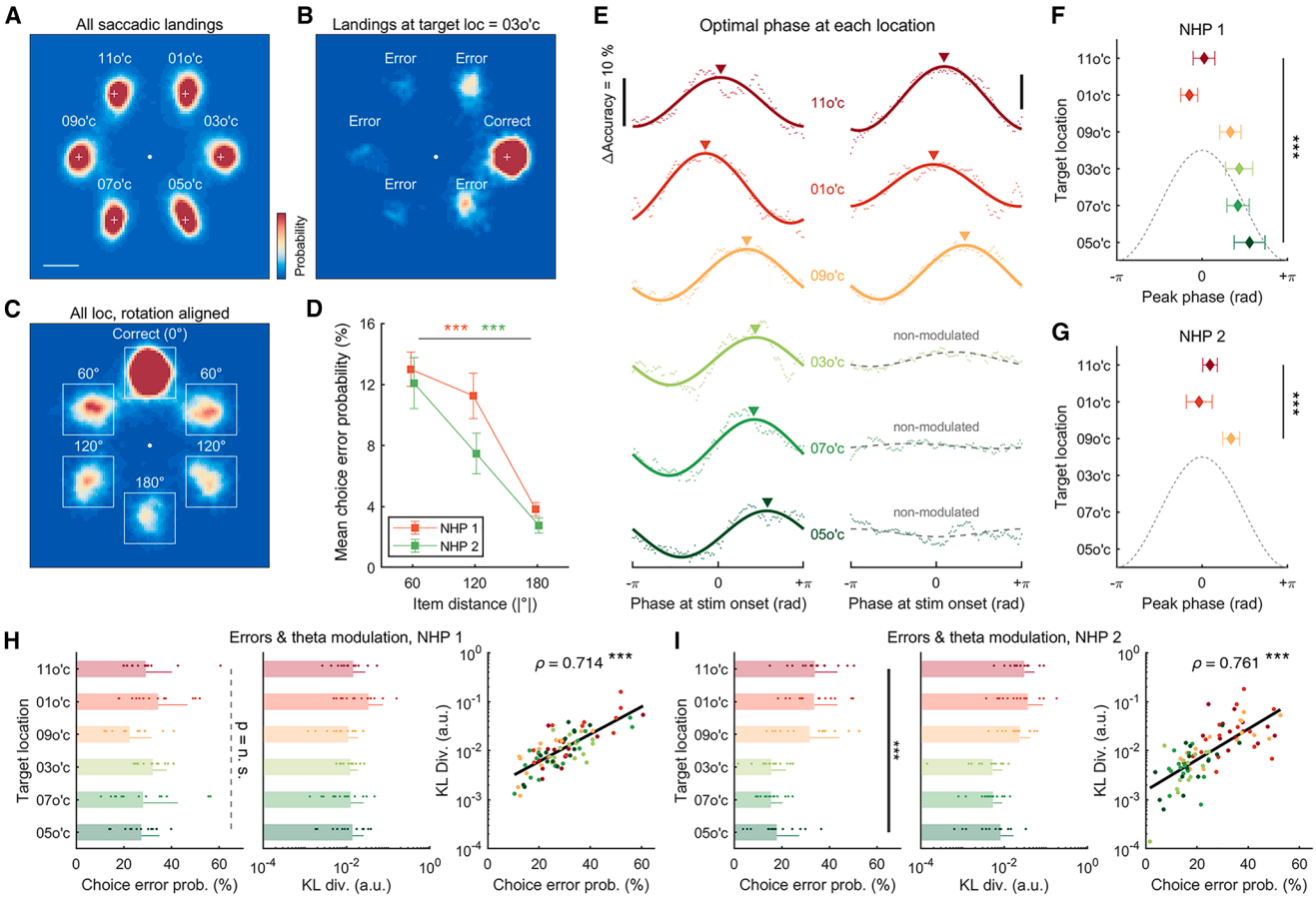

Figure 2. Choice errors and location-dependent theta modulation of WM task performance

(A) Two-dimensional histogram of saccadic landing position for all possible target locations (all trials from two NHPs aggregated, trial n = 32,702).

(B) Saccadic landing histogram for the trials with target location at 3 o'clock.

(C) Saccadic landing histogram for all possible target locations, rotated to align target location at 0° (top).

(D) Monotonic decrease of choice errors (areal mean of histogram values) as a function of item distance from target. ***$p < 0.001$ for Pearson correlation test. Error bars indicate ± 1 SEM.

(E) Fluctuation of WM task performance as a function of FEF theta phase at test display onset, for 6 possible target locations. Each line shows sinusoidal curve fits of accuracy distribution of one theta cycle. Colored triangles denote the optimal phase from the sinusoidal fits.

(F) Peak phase at each target location. The optimal phase was estimated using bootstrapping and defined as the circular mean of peak phases obtained from the fitted sine functions in (E). Error bars indicate ± 1 SD.

(G) Same as (F), but for NHP 2. **$p < 0.01$, ***$p < 0.001$ for the result of circular ANOVA test in (F) and (G).

(H) Negative correlation between daily performance for each target location and level of theta modulation at the location.

(I) Same as (H), but for NHP 2. NHP 2 showed a better performance for certain target locations than other locations; ANOVA ***$p < 0.001$. Error bars indicate 1 SEM.

See also Figure S2.

Theta sampled visual space sequentially

The good and poor phases of theta seemed to sample visual space sequentially from the top to bottom of the array. This was revealed by an analysis of performance as a function of theta phase and position of the target in the array. We fitted a sinusoidal function to the distribution of accuracy over FEF theta phase at test array onset to estimate the peak of accuracy (i.e., optimal phase), separately for each target location. For both NHPs, a circular ANOVA showed significant differences in the optimal phase across locations (bootstrapped with 33% subsampling, 10,000 iterations; $p < 0.001$ for both NHP 1 and NHP 2). For NHP 1, theta phase modulated performance at all six array locations in a systematic fashion. Performance was better for the top two array locations (11 and 1 o'clock) when the array appeared at or near the peak of FEF theta. At lower locations (9 and 3 o'clock, then 7 and 5 o'clock), performance peaked when the test array appeared at progressively later phases of theta (Figures 2E and 2F, NHP 1). NHP 2 showed significant theta modulation at 3 out of 6 locations (Figure 2G). The other locations (3, 5, and 7 o'clock) were non-modulated, and KL-divergence spectra did not show a clear peak in theta range (Figures S2C–S2F). This difference between NHPs may stem from their different task strategies. NHP 1 performed equally well across all target locations (Figure 2H; $F(5,78) = 2.210$, $p = 0.062$). In contrast,

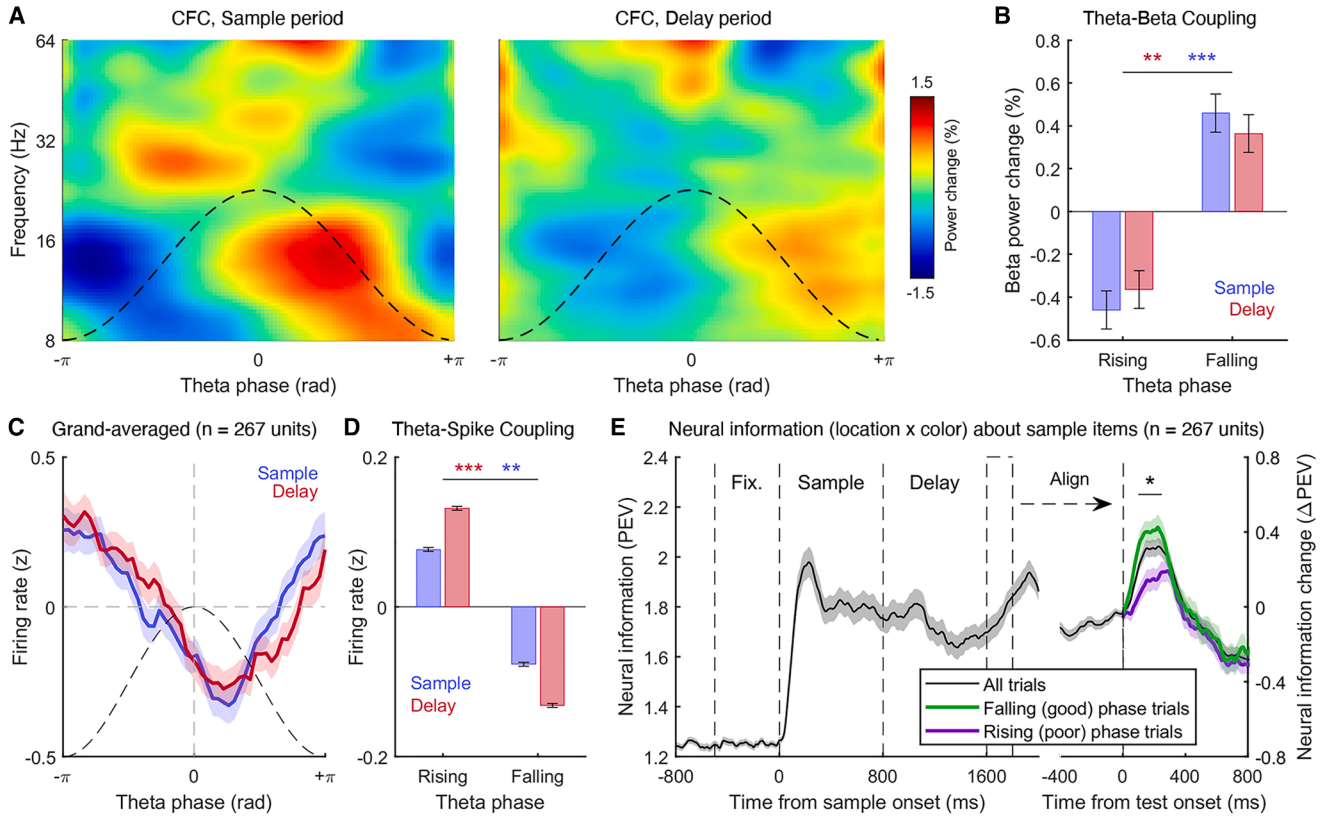

Figure 3. Theta rhythmically modulates beta power and spike rate in FEF

(A) Cross-frequency coupling between theta phase and high-frequency oscillation power during sample (0–800 ms, left) and delay period (800–1,600 ms, right). Dashed lines denote one cycle of theta.

(B) Beta power comparison between rising (behaviorally poor) and falling (good) phase of theta.

(C) Grand-averaged spike rate as a function of theta phase (delay period only).

(D) Firing rate comparison between rising and falling phase of theta.

(E) Time course of neural information contained in spikes calculated by PEV, time-aligned by sample array onset (left) and test array onset (right). PEV analysis was performed using correct trials only. Shaded area and error bars indicate 1 SEM. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ for the result of Wilcoxon's signed-rank test. See also Figure S3 and Table S1.

NHP 2 showed uneven performance across locations with higher accuracy at the lower location that did not show sequential sampling (Figure 2I; $F(5,78) = 14.312$, $p < 0.000$). Thus, NHP 1 seemed to split processing evenly across all locations (and thus sampled all locations). NHP 2 instead focused mainly on the lower locations and thus did not "scan" the full visual space equally like NHP 1. Thus, we observed sequential theta modulation at visual field locations where NHPs divided their attention. To further examine this apparent difference, we analyzed the relationship between daily error rates and theta modulation for each target location. We found a positive correlation between theta modulation strength and daily error rates (Figures 2H and 2I; Pearson $r > 0.714$, $p < 0.001$ in both NHPs; see also Figures S2G and S2H), suggesting stronger theta modulation emerged on sessions when cognitive resources were strained.

To test whether FEF theta exhibits traveling wave properties, we analyzed theta phase differences (3–6 Hz) from single-electrode data and their pairs (before averaging, ~8 electrodes per session) during the memory delay. We found theta phase difference is correlated to physical distance between electrode pairs (Figures S2I–S2K), propagating 0.5°/mm–0.7°/mm in both animals. A general linear model (GLM) analysis further revealed that phase differences were predicted by mediolateral and anteroposterior distances (Table S1), indicating a consistent direction of propagation across FEF.

Neural information cycled with theta

Theta was composed of alternating excitatory and inhibitory cortical states (Figure 3). Phase-amplitude coupling revealed that theta phase modulated beta (12–20 Hz) power. Beta power was lowest during the rising theta phases and higher during the falling theta phases (Figure 3A). This was significant for both the sample and memory delay (Figure 3B; $p < 0.01$). Single unit spike rate was also coupled to theta (Figures 3C and 3D; $p < 0.01$). Spike rate was highest at the troughs of theta (near $-\pi$ and $+\pi$) (Figure 3C). This was when beta was lowest (around $-\pi$) and late in the falling theta phase (around $+\pi$). Spike rate was lowest just after the peak of theta when beta power was highest.

We measured neural information about WM items (i.e., the sample array) using FEF spiking activity to examine its interaction

with theta phase. Using a GLM, we predicted item location and color, quantifying explanatory power with percent explained variance (PEV). Figure 3E shows that neural information increased after sample onset and remained stable during the delay period, as expected. After test array onset, it increased again, suggesting WM retrieval triggered by the test array. The effect of theta phase in the memory delay emerged at test array onset. More information was present when the test array appeared during the falling (good) phase of theta than during the rising (poor) phase (Figure 3E). The good vs. poor phases were defined by the instantaneous theta phase at test array onset (Figure 1F). Visual information arrived in the FEF with a latency of about 100 ms (see "sample" period in Figure 3E). Thus, information about the test array would arrive, not during the instantaneous (inhibitory) phase but in the opposite excitatory phase. A similar effect was observed in single unit spiking rates. A subpopulation of neurons showed greater increases in spiking following test onset on good-phase trials (Figure S3).

DISCUSSION

Our results indicate that the phase of frontal theta during a memory delay was associated with the readout of information from WM. The ability to detect a change (a target) in a visual scene (an array of items) from a similar scene held in WM fluctuated with theta oscillations in the FEF. Performance depended on both the theta phase when the comparison scene appeared as well as the location of the target. It appeared as if a theta wave was "scanning" the WM representation of the scene from top to bottom. This could be explained by a traveling wave moving across a retinotopic FEF. Performance improved when the theta wave happened to align with the location of the target when the comparison scene appeared. Cortical excitatory/inhibitory states and neural information also cycled with frontal theta, suggesting a possible mechanism for its effects on behavior.

Previous work has shown that selective attention waxes and wanes in theta in correspondence with frontal theta rhythms. During periods when attention is ostensibly sustained at a constant location in space, behavioral performance varies with a temporal periodicity of ~4–5 Hz.[18,19] Studies of human electroencephalogram (EEG)/electrocorticogram (ECoG)[20–22] and NHP LFPs[23–25] have shown this rhythmicity reflects ongoing theta oscillations in the frontoparietal cortex. Theta oscillations are often induced when attention is directed toward one of multiple competing stimuli.[26,27] Depending on the phase of these theta oscillations when a probe stimulus is shown, behavioral performance can vary dramatically.[3,20–23,25] These results show that ongoing cortical theta oscillations can modulate attention to external sensory inputs.

Our results indicate that *internal* WM representations also cycle with frontal theta. This is consistent with human behavior[28–30] and EEG[31,32] studies. In our case, the theta cycling modulated the behavioral readout of WM. Our results share many commonalities with the attention literature. Prior work on attention also found the strongest modulation within a similar 4–5 Hz band.[21–23,25] Attention studies have likewise found behavior is optimal when probed during the falling phase of frontal theta.[23] Prior work also showed similar theta modulation of spiking and higher-frequency activity.[22,23,25] Overall, similarities between our results and previous studies of attention strongly suggest a common mechanism may be at play. We propose that it reflects shared control mechanisms deployed for both attention and WM. This is consistent with many previous proposals suggesting shared control of attention and WM.[3–7]

Our results also build on previous work by demonstrating that theta modulation has an orderly structure across visual space. Previous studies have typically contrasted single locations inside and outside the focus of attention. Their results have been interpreted as good and poor theta phases alternating at the attended location, while poor and good theta phases oscillate in anti-phase at the unattended location.[3,33] This can be equivalently thought of as good and poor theta phases alternating back and forth between the attended and unattended locations. Our study generalizes this idea to a structured shift of theta phases when resources are not focused on a single location but divided across visual space. The optimal theta phase for behavior varied by retinotopic target location, progressing from the top to the bottom of the visual field.

This may reflect a progressive attentional shift across the visual field during the memory delay. At the neural level, this could be implemented by a traveling wave of activity across the cortical surface. Traveling waves have been observed in a number of cortical areas, suggesting they may be a ubiquitous motif of cortical processing.[34–39] Our results would suggest a wave sweeping across the polar angle dimension of the FEF topographic map, arrayed along the anterior-posterior axis.[40] In fact, waves of theta oscillations propagating in the posterior-to-anterior direction have been observed in human frontal cortex.[41] Our results may also suggest traveling theta interacts with receptive field (RF) structure, as FEF has a retinotopic organization.[42] Although, traveling waves can also influence function independent of spiking via subthreshold modulation of excitability and synaptic weights.[35,43] This could not be tested directly with our methodology. It can be tested in future work employing RF mapping and denser sampling of FEF.

A simple explanation of the results is that a traveling wave of excitation enhances processing when it aligns with the target. The behaviorally good phase was the inhibitory phase (when spiking and gamma is falling and alpha/beta is higher). However, this was the *instantaneous* phase when the test array appeared. Spiking activity indicates that information reaches the FEF around 100–200 ms after onset (see sample period in Figure 3E), matching the latency of effects following test array onset (Figure 3E). This suggests that test array information in spiking arrives in the FEF not during the instantaneous inhibitory phase but during the following excitatory phase of theta (i.e., when spiking and gamma is rising and alpha/beta is lower). Thus, the excitatory phase was the good phase in terms of brain mechanisms. On the other hand, it is unclear exactly when this input becomes functionally relevant—it may instead arrive during an inhibitory phase. If so, theta could play a role in top-down WM processes such as stabilizing internal representations to reduce interference through beta. Beta, which was higher during the inhibitory theta phase, has been associated with stabilizing cortical representations[44] and top-down control of sensory processing.[20,22,45] In either case, our results suggest the theta cycle modulates how information is read out from WM.

Our findings build on current WM models[9,45] by suggesting that theta traveling waves could coordinate top-down and bottom-up processes by coordinating other frequencies. Beta during the inhibitory theta phase may facilitate feedback signaling, while the excitatory phase facilitates feedforward processing, as predicted by spatial computing models.[46]

Many theories have emphasized the role of theta as a temporal framework for structuring cognitive processes. The *rhythmic theory of attention*[3] proposes visual attention sequentially samples perceptual inputs within a theta cycle. The *theta-gamma neural code*[11] suggests that within a single theta cycle, distinct neural ensembles encoding different information are activated in succession across several nested gamma cycles, enabling multiplexed representation. More recently, the *rhythmic attentional scanning* model[47] suggests each theta cycle acts as a selection window, determining which of multiple competing representations is propagated downstream. These frameworks collectively suggest that theta actively segments cognitive processing into periodic sampling windows, which our results support.

Our results also suggest that theta modulation plays a key role under high cognitive demands, such as when memory load increases and resources are divided across multiple items or locations—a characteristic trait of theta oscillations.[26,27,48–51] Frontal theta, involved in active resource control,[52] is known to increase with cognitive load—much like a car engine straining uphill.[53] In visual attention studies, theta has been linked to the intermittent sampling of unattended locations.[25] This may explain why theta modulation is often stronger outside the primary focus of attention.[25] Sustained attention to a single location likely involves continuous resource allocation, reducing the need for theta-driven sampling. Our task required distribution of limited-capacity WM resources across multiple locations, but the NHPs appeared to vary in how they achieved this. Behavioral evidence suggests NHP 1 spreads resources evenly across the visual field. In contrast, NHP 2 appeared to focus resources more on lower-field locations, while deprioritizing locations in the upper field. Correspondingly, we observed theta modulation throughout the visual field for NHP 1 but only in the upper field for NHP 2. The less-prioritized locations may have required sequential sampling and thus exhibited theta modulation. In contrast, the prioritized locations may have allowed for continuous, non-rhythmic processing. These findings suggest theta may be involved in distributing resources across multiple locations. This theta mechanism may generalize to any context where the brain must manage multiple simultaneous representations, whether external or internal.

Our findings demonstrate that WM readout varies with frontal theta oscillations, with behavioral performance depending on FEF theta phase. Theta appeared to structure the spatial organization of WM, with retrieval performance varying systematically across retinotopic space. Our findings provide further evidence that cognition is intrinsically linked to cortical oscillatory dynamics.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Earl K. Miller (ekmiller@mit.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Electrophysiological data presented in this article can be obtained from the lead contact upon reasonable request. Additional details necessary to reproduce the analyses are also available from the lead contact.

ACKNOWLEDGMENTS

This work was supported by the Picower Institute for Learning and Memory; the Office of Naval Research MURI N00014-23-1-2768; NEI 1R01EY033430-01A1; Office of Naval Research N00014-22-1-2453; the Freedom Together Foundation; National Research Foundation of Korea grant (RS-2024-00460958, RS-2025-22802990); the Technology Innovation Program (RS-2025-04812973; Neuro-Semi-AI Fusion Superhuman Project: A Platform for Augmenting Physical Abilities Based on Central Nervous System Stimulation) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea); and Research Program funded by the SeoulTech (Seoul National University of Science and Technology).

AUTHOR CONTRIBUTIONS

H.-B.H. and E.K.M. designed the study; T.J.B. conducted the research; H.-B.H., S.L.B., and T.J.B. analyzed the data; H.-B.H., S.L.B., and E.K.M. wrote the manuscript; and E.K.M. supervised the study.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Behavioral protocol and data acquisition
  - LFP data acquisition and preprocessing
  - Spike data analysis
  - Theta modulation of behavioral performance
  - Neural information analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.neuron.2025.09.031.

Received: April 1, 2025
Revised: August 27, 2025
Accepted: September 22, 2025
Published: October 20, 2025

REFERENCES

1. Colgin, L.L. (2013). Mechanisms and Functions of Theta Rhythms. Annu. Rev. Neurosci. *36*, 295–312. https://doi.org/10.1146/annurev-neuro-062012-170330.
2. VanRullen, R. (2016). Perceptual Cycles. Trends Cogn. Sci. *20*, 723–735. https://doi.org/10.1016/j.tics.2016.07.006.
3. Fiebelkorn, I.C., and Kastner, S. (2019). A Rhythmic Theory of Attention. Trends Cogn. Sci. *23*, 87–101. https://doi.org/10.1016/j.tics.2018.11.009.

4. Awh, E., and Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. Trends Cogn. Sci. *5*, 119–126. https://doi.org/10.1016/S1364-6613(00)01593-X.
5. Gazzaley, A., and Nobre, A.C. (2012). Top-down modulation: bridging selective attention and working memory. Trends Cogn. Sci. *16*, 129–135. https://doi.org/10.1016/j.tics.2011.11.014.
6. Panichello, M.F., and Buschman, T.J. (2021). Shared mechanisms underlie the control of working memory and attention. Nature *592*, 601–605. https://doi.org/10.1038/s41586-021-03390-w.
7. Han, S.-H., and Kim, M.-S. (2004). Visual search does not remain efficient when executive working memory is working. Psychol. Sci. *15*, 623–628. https://doi.org/10.1111/j.0956-7976.2004.00730.x.
8. Lundqvist, M., Herman, P., and Miller, E.K. (2018). Working Memory: Delay Activity, Yes! Persistent Activity? Maybe Not. J. Neurosci. *38*, 7013–7019. https://doi.org/10.1523/JNEUROSCI.2485-17.2018.
9. Miller, E.K., Lundqvist, M., and Bastos, A.M. (2018). Working Memory 2.0. Neuron *100*, 463–475. https://doi.org/10.1016/j.neuron.2018.09.023.
10. Stokes, M.G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. Trends Cogn. Sci. *19*, 394–405. https://doi.org/10.1016/j.tics.2015.05.004.
11. Lisman, J.E., and Jensen, O. (2013). The Theta-Gamma Neural Code. Neuron *77*, 1002–1016. https://doi.org/10.1016/j.neuron.2013.03.007.
12. Raghavachari, S., Kahana, M.J., Rizzuto, D.S., Caplan, J.B., Kirschen, M.P., Bourgeois, B., Madsen, J.R., and Lisman, J.E. (2001). Gating of Human Theta Oscillations by a Working Memory Task. J. Neurosci. *21*, 3175–3183. https://doi.org/10.1523/JNEUROSCI.21-09-03175.2001.
13. Sauseng, P., Klimesch, W., Schabus, M., and Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. Int. J. Psychophysiol. *57*, 97–103. https://doi.org/10.1016/j.ijpsycho.2005.03.018.
14. Sauseng, P., Griesmayr, B., Freunberger, R., and Klimesch, W. (2010). Control mechanisms in working memory: a possible function of EEG theta oscillations. Neurosci. Biobehav. Rev. *34*, 1015–1022. https://doi.org/10.1016/j.neubiorev.2009.12.006.
15. Buschman, T.J., Siegel, M., Roy, J.E., and Miller, E.K. (2011). Neural substrates of cognitive capacity limitations. Proc. Natl. Acad. Sci. USA *108*, 11252–11255. https://doi.org/10.1073/pnas.1104666108.
16. Kornblith, S., Buschman, T.J., and Miller, E.K. (2016). Stimulus Load and Oscillatory Activity in Higher Cortex. Cereb. Cortex *26*, 3772–3784. https://doi.org/10.1093/cercor/bhv182.
17. Luck, S.J., and Vogel, E.K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. Trends Cogn. Sci. *17*, 391–400. https://doi.org/10.1016/j.tics.2013.06.006.
18. Fiebelkorn, I.C., Saalmann, Y.B., and Kastner, S. (2013). Rhythmic Sampling within and between Objects despite Sustained Attention at a Cued Location. Curr. Biol. *23*, 2553–2558. https://doi.org/10.1016/j.cub.2013.10.063.
19. Landau, A.N., and Fries, P. (2012). Attention Samples Stimuli Rhythmically. Curr. Biol. *22*, 1000–1004. https://doi.org/10.1016/j.cub.2012.03.054.
20. Busch, N.A., and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. Proc. Natl. Acad. Sci. USA *107*, 16048–16053. https://doi.org/10.1073/pnas.1004801107.
21. Harris, A.M., Dux, P.E., and Mattingley, J.B. (2018). Detecting Unattended Stimuli Depends on the Phase of Prestimulus Neural Oscillations. J. Neurosci. *38*, 3092–3101. https://doi.org/10.1523/JNEUROSCI.3006-17.2018.
22. Helfrich, R.F., Fiebelkorn, I.C., Szczepanski, S.M., Lin, J.J., Parvizi, J., Knight, R.T., and Kastner, S. (2018). Neural mechanisms of sustained attention are rhythmic. Neuron *99*, 854–865.e5. https://doi.org/10.1016/j.neuron.2018.07.032.
23. Fiebelkorn, I.C., Pinsk, M.A., and Kastner, S. (2018). A Dynamic Interplay within the Frontoparietal Network Underlies Rhythmic Spatial Attention. Neuron *99*, 842–853.e8. https://doi.org/10.1016/j.neuron.2018.07.038.
24. Fiebelkorn, I.C., Pinsk, M.A., and Kastner, S. (2019). The mediodorsal pulvinar coordinates the macaque fronto-parietal network during rhythmic spatial attention. Nat. Commun. *10*, 215. https://doi.org/10.1038/s41467-018-08151-4.
25. Spyropoulos, G., Bosman, C.A., and Fries, P. (2018). A theta rhythm in macaque visual cortex and its attentional modulation. Proc. Natl. Acad. Sci. USA *115*, E5614–E5623. https://doi.org/10.1073/pnas.1719433115.
26. Kienitz, R., Schmiedt, J.T., Shapcott, K.A., Kouroupaki, K., Saunders, R.C., and Schmid, M.C. (2018). Theta Rhythmic Neuronal Activity and Reaction Times Arising from Cortical Receptive Field Interactions during Distributed Attention. Curr. Biol. *28*, 2377–2387.e5. https://doi.org/10.1016/j.cub.2018.05.086.
27. Rollenhagen, J.E., and Olson, C.R. (2005). Low-Frequency Oscillations Arising From Competitive Interactions Between Visual Stimuli in Macaque Inferotemporal Cortex. J. Neurophysiol. *94*, 3368–3387. https://doi.org/10.1152/jn.00158.2005.
28. Chota, S., Leto, C., Van Zantwijk, L., and Van Der Stigchel, S. (2022). Attention rhythmically samples multi-feature objects in working memory. Sci. Rep. *12*, 14703. https://doi.org/10.1038/s41598-022-18819-z.
29. Peters, B., Kaiser, J., Rahm, B., and Bledowski, C. (2021). Object-based attention prioritizes working memory contents at a theta rhythm. JEP G *150*, 1250–1256. https://doi.org/10.1037/xge0000994.
30. Pomper, U., and Ansorge, U. (2021). Theta-Rhythmic Oscillation of Working Memory Performance. Psychol. Sci. *32*, 1801–1810. https://doi.org/10.1177/09567976211013045.
31. Wöstmann, M., Lui, T.K.-Y., Friese, K.-H., Kreitewolf, J., Naujokat, M., and Obleser, J. (2020). The vulnerability of working memory to distraction is rhythmic. Neuropsychologia *146*, 107505. https://doi.org/10.1016/j.neuropsychologia.2020.107505.
32. Abdalaziz, M., Redding, Z.V., and Fiebelkorn, I.C. (2023). Rhythmic temporal coordination of neural activity prevents representational conflict during working memory. Curr. Biol. *33*, 1855–1863.e3. https://doi.org/10.1016/j.cub.2023.03.088.
33. VanRullen, R. (2018). Attention Cycles. Neuron *99*, 632–634. https://doi.org/10.1016/j.neuron.2018.08.006.
34. Bhattacharya, S., Brincat, S.L., Lundqvist, M., and Miller, E.K. (2022). Traveling waves in the prefrontal cortex during working memory. PLoS Comput. Biol. *18*, e1009827. https://doi.org/10.1371/journal.pcbi.1009827.
35. Muller, L., Chavane, F., Reynolds, J., and Sejnowski, T.J. (2018). Cortical travelling waves: mechanisms and computational principles. Nat. Rev. Neurosci. *19*, 255–268. https://doi.org/10.1038/nrn.2018.20.
36. Patel, J., Fujisawa, S., Berényi, A., Royer, S., and Buzsáki, G. (2012). Traveling Theta Waves along the Entire Septotemporal Axis of the Hippocampus. Neuron *75*, 410–417. https://doi.org/10.1016/j.neuron.2012.07.015.
37. Rubino, D., Robbins, K.A., and Hatsopoulos, N.G. (2006). Propagating waves mediate information transfer in the motor cortex. Nat. Neurosci. *9*, 1549–1557. https://doi.org/10.1038/nn1802.
38. Sato, T.K., Nauhaus, I., and Carandini, M. (2012). Traveling Waves in Visual Cortex. Neuron *75*, 218–229. https://doi.org/10.1016/j.neuron.2012.06.029.
39. Zanos, T.P., Mineault, P.J., Nasiotis, K.T., Guitton, D., and Pack, C.C. (2015). A Sensorimotor Role for Traveling Waves in Primate Visual Cortex. Neuron *85*, 615–627. https://doi.org/10.1016/j.neuron.2014.12.043.
40. Tehovnik, E.J., Sommer, M.A., Chou, I.H., Slocum, W.M., and Schiller, P.H. (2000). Eye fields in the frontal lobes of primates. Brain Res. Brain Res. Rev. *32*, 413–448. https://doi.org/10.1016/S0165-0173(99)00092-2.
41. Zhang, H., Watrous, A.J., Patel, A., and Jacobs, J. (2018). Theta and Alpha Oscillations Are Traveling Waves in the Human Neocortex. Neuron *98*, 1269–1281.e4. https://doi.org/10.1016/j.neuron.2018.05.019.

42. Mayo, J.P., DiTomasso, A.R., Sommer, M.A., and Smith, M.A. (2015). Dynamics of visual receptive fields in the macaque frontal eye field. J. Neurophysiol. *114*, 3201–3210. https://doi.org/10.1152/jn.00746.2015.

43. Butler, K., and Cruz, L. (2025). Neuronal traveling waves form preferred pathways using synaptic plasticity. J. Comp. Neurosci. *53*, 181–198. https://doi.org/10.1007/s10827-024-00890-2.

44. Engel, A.K., and Fries, P. (2010). Beta-band oscillations — signalling the status quo? Curr. Opin. Neurobiol. *20*, 156–165. https://doi.org/10.1016/j.conb.2010.02.015.

45. Bastos, A.M., Lundqvist, M., Waite, A.S., Kopell, N., and Miller, E.K. (2020). Layer and rhythm specificity for predictive routing. Proc. Natl. Acad. Sci. USA *117*, 31459–31469. https://doi.org/10.1073/pnas.2014868117.

46. Lundqvist, M., Brincat, S.L., Rose, J., Warden, M.R., Buschman, T.J., Miller, E.K., and Herman, P. (2023). Working memory control dynamics follow principles of spatial computing. Nat. Commun. *14*, 1429. https://doi.org/10.1038/s41467-023-36555-4.

47. Fries, P. (2023). Rhythmic attentional scanning. Neuron *111*, 954–970. https://doi.org/10.1016/j.neuron.2023.02.015.

48. Missonnier, P., Deiber, M.P., Gold, G., Millet, P., Gex-Fabry Pun, M., Fazio-Costa, L., Giannakopoulos, P., and Ibáñez, V. (2006). Frontal theta event-related synchronization: comparison of directed attention and working memory load effects. J. Neural Transm. *113*, 1477–1486. https://doi.org/10.1007/s00702-005-0443-9.

49. Jensen, O., and Tesche, C.D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. Eur. J. Neurosci. *15*, 1395–1399. https://doi.org/10.1046/j.1460-9568.2002.01975.x.

50. Zakrzewska, M.Z., and Brzezicka, A. (2014). Working memory capacity as a moderator of load-related frontal midline theta variability in Sternberg task. Front. Hum. Neurosci. *8*, 399. https://doi.org/10.3389/fnhum.2014.00399.

51. Han, H.-B., Lee, K.E., and Choi, J.H. (2019). Functional Dissociation of θ Oscillations in the Frontal and Visual Cortices and Their Long-Range Network during Sustained Attention. eNeuro *6*. ENEURO.0248-19.2019. https://doi.org/10.1523/ENEURO.0248-19.2019.

52. Cavanagh, J.F., and Frank, M.J. (2014). Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. *18*, 414–421. https://doi.org/10.1016/j.tics.2014.04.012.

53. Clayton, M.S., Yeung, N., and Kadosh, R.C. (2015). The roles of cortical oscillations in sustained attention. Trends Cogn. Sci. *19*, 188–195. https://doi.org/10.1016/j.tics.2015.02.004.

54. Asaad, W.F., and Eskandar, E.N. (2008). Achieving behavioral control with millisecond resolution in a high-level programming environment. J. Neurosci. Methods *173*, 235–240. https://doi.org/10.1016/j.jneumeth.2008.06.003.

55. Asaad, W.F., and Eskandar, E.N. (2008). A flexible software tool for temporally-precise behavioral control in Matlab. J. Neurosci. Methods *174*, 245–258. https://doi.org/10.1016/j.jneumeth.2008.07.014.

56. Hwang, J., Mitz, A.R., and Murray, E.A. (2019). NIMH MonkeyLogic: Behavioral control and data acquisition in MATLAB. J. Neurosci. Methods *323*, 13–21. https://doi.org/10.1016/j.jneumeth.2019.05.002.

57. Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. J. Stat. Software *31*, 1–21.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Macaca mulatta (NHP 1)	Covance	RRID: NCBITaxon_9544
Macaca fascicularis (NHP 2)	Merck	RRID: NCBITaxon_9541
Software and algorithms		
MATLAB	MathWorks Inc	https://www.mathworks.com/products/matlab; RRID: SCR_001622
MonkeyLogic software	NIMH; Asaad and Eskandar[54,55]	https://monkeylogic.nimh.nih.gov/
CircStat toolbox	Berens et al.[56]	https://github.com/circstat/circstat-matlab; RRID:SCR_016651
Other		
Microelectrodes	FHC	Cat#UEWLFDSMNN1E

All procedures followed the guidelines of the Massachusetts Institute of Technology Committee on Animal Care and the National Institutes of Health. Detailed methods for experimental procedures and data processing pipelines are available in Methods S1.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

One adult rhesus macaque (*Macaca mulatta*, NHP 1: male, 13 kg) and one adult male cynomolgus monkey (NHP 2: male, 6 kg) were trained to perform the task. For neural recording, multiple epoxy-insulated tungsten electrodes (FHC, Bowdoin, ME, USA 04287) were inserted using custom-built manual screw microdrives. FEF was targeted via co-registration of structural MRI scans with standard atlases, and confirmed by microstimulation-driven saccades. The electrodes were acutely lowered at the beginning of every recording session (n = 14 sessions for each NHP) and settled for at least 2 hours before recording, then retracted after it. For further details on surgical procedures and animal handling, please see our previous publication with the same dataset.

METHOD DETAILS

Behavioral protocol and data acquisition

The behavioral paradigm was controlled with the MonkeyLogic program.[54–56] Each trial began with a 500 ms fixation period, followed by an 800 ms sample period where an array of colored squares was displayed. After a variable memory delay period (800–1000 ms), a test array appeared, identical to the sample array except for a color change in one randomly selected target. NHPs were required to make a single saccade to the changed item. Trials with eye movements during the sample or delay period, reflecting a failure to maintain fixation, were not included in the analysis. Eye movements were tracked using an ISCAN infrared system (240 Hz) throughout all sessions. Stimuli were 1° colored squares, with two possible colors (color A, and color B) at each location, randomized daily to prevent long-term memorization. Six item locations (roughly 1, 3, 5, 7, 9, and 11 o'clock) were used each session that were within ±75 angular degrees from the horizontal meridian and 4° to 6° from the fixation point. For the analysis, gaze coordinates at saccadic landings on six target locations were translated to ensure symmetry and then rotated so that the target location was positioned at 0 degrees (top of the screen). NHPs completed at least 720 correct trials per session. Invalid trials (e.g., failure to fixate before test array onset) were excluded (trial survival rate: 76.29% for NHP 1, 71.09% for NHP 2). As a result, total 16,940 trials for NHP 1 (load 2: 29.56%, load 3: 28.56%, load 4: 21.46%, load 5: 20.43%) and total 15,762 trials for NHP 2 (load 2: 27.44%, load 3: 26.95%, load 4: 20.21%, load 5: 18.44%) were included. For more details on the behavioral paradigm, please see our previous publication that used the same dataset[14] and Methods S1.

LFP data acquisition and preprocessing

Continuous LFP data was amplified, band-passed filtered (3.3–88 Hz), and digitized at 1 kHz sampling rate (Plexon Multi-channel Acquisition Processor). Although this hardware filter partially overlapped our primary band-of-interest (3–6 Hz), the filter was not sharp. Manual inspection confirmed that substantial power remained in the theta band. All signals were referenced to ground. Any 60 Hz line noise, 85 Hz noise related to the monitor refresh rate, and their harmonics were estimated and removed offline using an adaptive sinusoid fit method. For most of LFP analyses, LFPs were averaged across all simultaneously recorded FEF electrodes (n = 7.93 ± 2.00 per session) to estimate a representative LFP signal for FEF. Analyses using single-electrode data (e.g., Figure S2; Table S1) are noted explicitly.

To obtain amplitude (power) spectrograms, a fast Fourier transform was applied with a sliding Hanning window (window size = 1024, 100 ms step). For cross-frequency coupling analysis, the LFP signal was narrow-band-filtered (0.5 Hz step with 2 Hz bandwidth for 8–16 Hz; 1 Hz step with 4 Hz bandwidth for 16–32 Hz; 2 Hz step with 8 Hz bandwidth for 32–64 Hz) using Butterworth 5^{th} order filter.

Spike data analysis

Spikes were sorted into isolated single units manually using waveform features (Plexon offline Sorter), as previously described.[14] Then, two criteria were applied for excluding units from the analysis: (1) units with less than 30 trials sampled for each condition were excluded; (2) units with a firing rate less than 1 Hz during the task period (–0.5 to +1.6 s relative to sample stimulus onset) were excluded. As a result of applying these two criteria, 76.19% of units (leaving n = 267 units) from FEF were included in the analysis.

Theta modulation of behavioral performance

To estimate theta phase at the timing of test array onset, LFP was band-pass filtered at 3–6 Hz with zero-phase FIR filters (MATLAB's *filtfilt.m* function). After band-pass filtering, the Hilbert transform was used to obtain instantaneous angle (MATLAB's *hilbert.m* and *angle.m* functions). The theta phase was extracted at the time point minimally preceding the test display command sent by the computer (0–1 ms prior in 1 kHz recordings), to capture ongoing activity just before stimulus onset. Theta phase was divided into discrete phase bins (100 bins) for further analyses. The strength of coupling between LFP phase and behavioral performance was measured by the KL divergence between the observed histogram of the behavioral measure (RT or accuracy) and a circular uniform distribution. Detailed procedures are provided in Methods S1.

Neural information analysis

To investigate how spiking data encodes WM information (location x color), we analyzed individual single unit spiking using GLM and calculated its effect size (i.e., PEV measured by η^2, which quantifies how much of a unit's firing rate variability is attributed to stimulus location and identity). Since the color of each object was unknown to the NHP before sample array onset, we assessed how well spiking represented this information at each time point. PEV was computed using a single ANOVA model, where all locations and item identities were included in a unified model with dummy-coded variables representing location $\times$ item identity: $y = \Sigma\, bX$ where X = 0, 1, 2, represented absent, color A, and color B, respectively, for each location. PEV was calculated using the η^2 formula and summed across all model variables: PEV = (1 / SS_Total) * Σ SS_X * 100. Only neural data from correct trials were included in the PEV analysis. This approach provides a comprehensive assessment of how spiking activity encodes WM information across all spatial locations, resulting in a measure of the overall strength of WM representations. Then, PEV (Figure 3) and single unit spiking rate (Figure S3) were divided into either 'rising' or 'falling' phase trials based on the theta phase immediately preceding test array onset (rising-phase trials: -π to 0 rad; falling-phase trials: 0 to +π rad).

QUANTIFICATION AND STATISTICAL ANALYSIS

For testing sample means (or medians) against a null value (appropriate for a one-sample t-test), non-parametric Wilcoxon's signed rank test was used. Circular ANOVA was used for angular phase data, implemented by CircStat toolbox[57] in MATLAB (function *circ_wwtest.m*). For correlation analyses, Pearson's method was used except for logarithmic scale data (Spearman's method was used for KL divergence in Figure 2).