Article

Current Biology

Oscillatory control over representational geometry of sequence working memory in macaque frontal cortex

Highlights

- Multi-channel synchronous recording of single units and LFP in macaque frontal cortex
- Theta power transiently encodes each stimulus in a 2-item sequential WM task
- Theta phase-spike interaction is rank-selectively associated with memory subspaces
- Such interaction is disrupted in error trials and is generalizable to 3-item task

Authors

Wen Fang, Xi Jiang, Jingwen Chen, Cong Zhang, Liping Wang

Correspondence

liping.wang@ion.ac.cn

In brief

Fang et al. found that frontal cortical theta-range spike-field coherence is rank-selectively associated with unit spiking-based memory subspaces, supporting the binding of incoming sequential working memory items to their appropriate ordinal ranks. This control process is disrupted in error trials and is generalizable to longer sequences.

Fang et al., 2025, Current Biology *35*, 1495–1507
April 7, 2025
https://doi.org/10.1016/j.cub.2025.02.031

CellPress

Current Biology

CellPress

Article

Oscillatory control over representational geometry of sequence working memory in macaque frontal cortex

Wen Fang,[1,3] Xi Jiang,[1,3] Jingwen Chen,[1] Cong Zhang,[1] and Liping Wang[1,2,4,*]

[1]Institute of Neuroscience, Key Laboratory of Brain Cognition and Brain-Inspired Intelligence Technology, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

[2]Shanghai Academy of Natural Sciences (SANS), Fudan University, Shanghai 200031, China

[3]These authors contributed equally

[4]Lead contact

*Correspondence: liping.wang@ion.ac.cn

https://doi.org/10.1016/j.cub.2025.02.031

SUMMARY

To process sequential streams of information, e.g., language, the brain must encode multiple items in sequence working memory (SWM) according to their ordinal relationship. While the geometry of neural states could represent sequential events in the frontal cortex, the control mechanism over these neural states remains unclear. Using high-throughput electrophysiology recording in the macaque frontal cortex, we observed widespread theta responses after each stimulus entry. Crucially, by applying targeted dimensionality reduction to extract task-relevant neural subspaces from both local field potential (LFP) and spike data, we found that theta power transiently encoded each sequentially presented stimulus regardless of its order. At the same time, theta-spike interaction was rank-selectively associated with memory subspaces, thereby potentially supporting the binding of items to appropriate ranks. Furthermore, this putative theta control can generalize to length-variable and error sequences, predicting behavior. Thus, decomposed entry/rank-WM subspaces and theta-spike interactions may underlie the control of SWM.

INTRODUCTION

Daily events of our lives all occur within the river of time. Our perception and memory of an event are based not only on the event's constituent features but also on its temporal context. Writing down a friend's phone number, for instance, requires one to sequentially memorize the identity of each digit and its ordinal position, then bind all digits into sequence working memory (SWM) for later action. However, it remains unclear how exactly the brain coordinates the entry of multiple temporally ordered external stimuli into SWM. To explore the neural mechanisms of stimulus item-ordinal rank binding under SWM, we need to find out how the brain (1) represents sensory inputs (rank-independent) and their associated rank-dependent WMs and (2) selectively controls and transforms rank-independent stimulus entry into rank-dependent memory.

Regarding the first question (on item/rank representations), our colleagues recently discovered that sequential visuospatial items were first entered into a common entry subspace consisting of frontal cortical neurons and then sent into rank-selective SWM subspaces,[1] whose activity could be maintained throughout the whole delay period following stimulus presentation.[2] As for the second question (on item-rank binding), it was often proposed that such binding of multiple WM aspects could be achieved by oscillations synchronizing neurons that represent different features of the same perceptual object.[3–5] Experimental evidence suggests a more general role of oscillatory control over neural activity in multiple cognitive functions, including attention,[6,7] perception,[8,9] and memory.[10–14] It remains unclear whether compositional item-rank representations could be identified from oscillatory neural signals as in neuronal unit populations, nor whether oscillatory control plays a role in the contextual binding of incoming sequential stimuli to appropriate ordinal ranks. Advancements on these fronts could potentially lead to non-invasive monitoring applications that allow decoding of complex compositional neural representations, e.g., those involved in language production.

Among the various frequency bands where WM-relevant neural oscillations have been found, theta (typically 3/4–8/9 Hz) and higher-frequency gamma (typically greater than 50–60 Hz), in particular, are known to modulate population coding in feature-specific ways during cognitive processing.[15,16] Theta oscillation has been widely observed in the formation and storage of associative memories, which require the binding of different co-occurring elements.[17] Theta-coordinated neural activities are also often found among distant brain regions, and such global activation is associated with conscious perception of sensory stimuli.[14,18] Regarding SWM, certain human temporal lobe neurons appear to encode consecutive items by precise firing patterns at distinct theta oscillatory phases, resembling spatial sequence encoding by the rodent hippocampus.[19,20]

Computational studies have proposed that the control of information selection in visual signal processing and WM is embedded in the communication and transformation between

representative subspaces,[21,22] in line with our earlier hypothesized linkage between entry and memory subspaces.[2] Despite its clear involvement, whether theta oscillation reorganizes sequential items in an ordinal rank-dependent manner remains an open question. Since sequence representation can be regarded as a specific binding problem to link the physical stimuli to temporal order, the selective binding of sequential stimuli to proper ordinal ranks may rely on oscillatory control over different SWM subspaces. To address this hypothesis, we analyzed neural data from macaque monkeys performing visuospatial delayed sequence reproduction tasks. The data were collected using a 157-channel high-throughput electrophysiological recording system, including spiking activities and local field potential (LFP) across the macaque frontal cortex.

RESULTS

Behavioral paradigm and performance

Three macaque monkeys, O, L, and G, were trained to learn an SWM task (Figure 1A). Each trial was initiated by visual fixation on the center of a monitor/touchscreen facing the animals. During the sample period, two stimuli were presented sequentially on the monitor screen while the monkeys kept fixation on the center. Each stimulus only remained on the screen for 250 ms after onset, so the monkeys had to maintain the sequence item(s) in memory during both the post-S1 inter-stimulus interval and the post-S2 delay. After the delay period, the monkeys were prompted to reproduce memorized sequences by making consecutive touches to the appropriate locations on the screen. The behavioral performance of length-2 sequences in all monkeys was significantly better than chance and balanced across both stimuli identities (i.e., six dot positions on screen) and ordinal ranks (two ranks: R1 for stimulus S1, R2 for stimulus S2) (Figures 1B, S1A, and S1B). Parts of the behavioral data were published in our previous studies.[1,2,23,24] Each monkey received implantation of a 157-channel micro-drive[25] over the left frontal cortex (Figures 1C and S1C) upon training completion, allowing us to simultaneously record both single neuron spikes and LFP (Figures 1D and 1E). Due to relatively poor behavioral response (Figure 1B, right column) and lower recording quality (~20 units per recording session vs. >100 in monkeys O/G), we excluded monkey L from the rest of this work.

Task-relevant single-channel LFP activities

According to previous findings,[1,2] in SWM, initially rank-independent representations of external stimuli (item) are assigned to a given context (order), thereby turning entry signals into rank-dependent memories (Figure 1F). While neuronal unit spiking-based representations of the two primitives, i.e., rank-independent "entry" and rank-dependent "memory" in SWM, had already been found in the macaque frontal cortex,[2] no similar LFP features that contained task information had been identified. We first performed time-frequency analysis using channel-wise data aligned to stimulus onsets, thereby obtaining an overview of average LFP activity across correctly performed trials (Figure 1E, with evoked potential subtracted[26]; see Figure S1D for unadjusted examples; subsequent analyses using unadjusted data were comparable to present results [data not shown]). Only channels with spiking units were used for this and all subsequent analyses involving LFP. In each recording session, prominent post-stimulus-onset (0–500 ms) increases of event-related spectral power in classical theta and/or high gamma (HG) bands were observed across a substantial portion of channels (on average over sessions, 56.1% ± 4.10% and 49.0% ± 7.63% of all channels for monkeys O and G, respectively). We then examined the item tuning of theta (4–8 Hz) and HG (80–120 Hz; with spike removal preprocessing to reduce contamination; STAR Methods) Hilbert amplitudes (Figures 1G, 1H, and S1F). We found that, as was the case for spike tuning (Figure S1G; see also previous work[2]), theta and HG amplitude changes within trials also encoded target item identities, but the two frequency bands yielded distinct response patterns. LFP channels with item selectivity in theta amplitude specifically displayed rank-independent transient tuning (Figure 1G). By contrast, channels with selectivity in HG amplitude tended to have two separable dynamic components similar to those found in neurons: a rank-independent transient component (Figure S1F) and a sustained rank-dependent memory component (Figure 1H).

To quantitatively evaluate task-relevant LFP activities, we next extracted channel-wise theta/HG power to compute the ordinal rank-specific percentage of explained variance (PEV) across the encoding period. As shown in the example from one recording session (Figures 2A and S1E) and overall distributions of significant PEV timing (Figures 2B, 2C, S2A, and S2B), theta response preferentially occurred early and in a transient manner befitting stimulus entry and/or temporarily required WM control. By contrast, while some channels did show transient burst-like HG PEV, relatively prolonged/sustained memory-like activity occurred more often in HG, corresponding to the spatial tuning patterns seen previously (Figures 1G and 1H). This divergence in WM representation patterns was also observed via categorizing rank-independent vs. rank-dependent item tuning significance (Figure S2C), where theta channels predominantly fell within the former and unit/HG the latter.

While the temporal dynamics of theta and gamma activity at the single-channel level reflected characteristics of entry and memory in our data, it remained unclear at the group level whether they genuinely encode entry and memory information during the sequential WM processing—analogous to phenomena previously observed at the population neuronal level.[2] To remedy this, we performed targeted dimensionality reduction (TDR) over multi-unit/channel data (firing rates for units, instantaneous powers for theta/gamma) to reduce the neural signals into their most representative two dimensions. This decomposition yielded a rank-independent subspace and multiple orthogonal (as guaranteed via QR factorization) rank-dependent subspaces in a time period-targeted manner (Figures 2D–2H) for each of the three data types. To identify the rank-independent entry subspaces, we stacked post-S1-onset and post-S2-onset 0–250 ms signals together (Figure 2D) as independent trials when building the initial unit/channel-by-trial matrices for TDR. This allowed us to locate for each data type a subspace that could simultaneously describe spatial location encoding under both ordinal ranks R1 and R2 (Figure 2E). When projecting neural activity at each time point across the trial time to the identified entry subspaces found by this procedure, the resulting temporal dynamics—transient increase and rapid decay post-stimuli—indeed resembled the earlier tuning dynamics of rank-independent/putative entry neurons and theta/HG channels

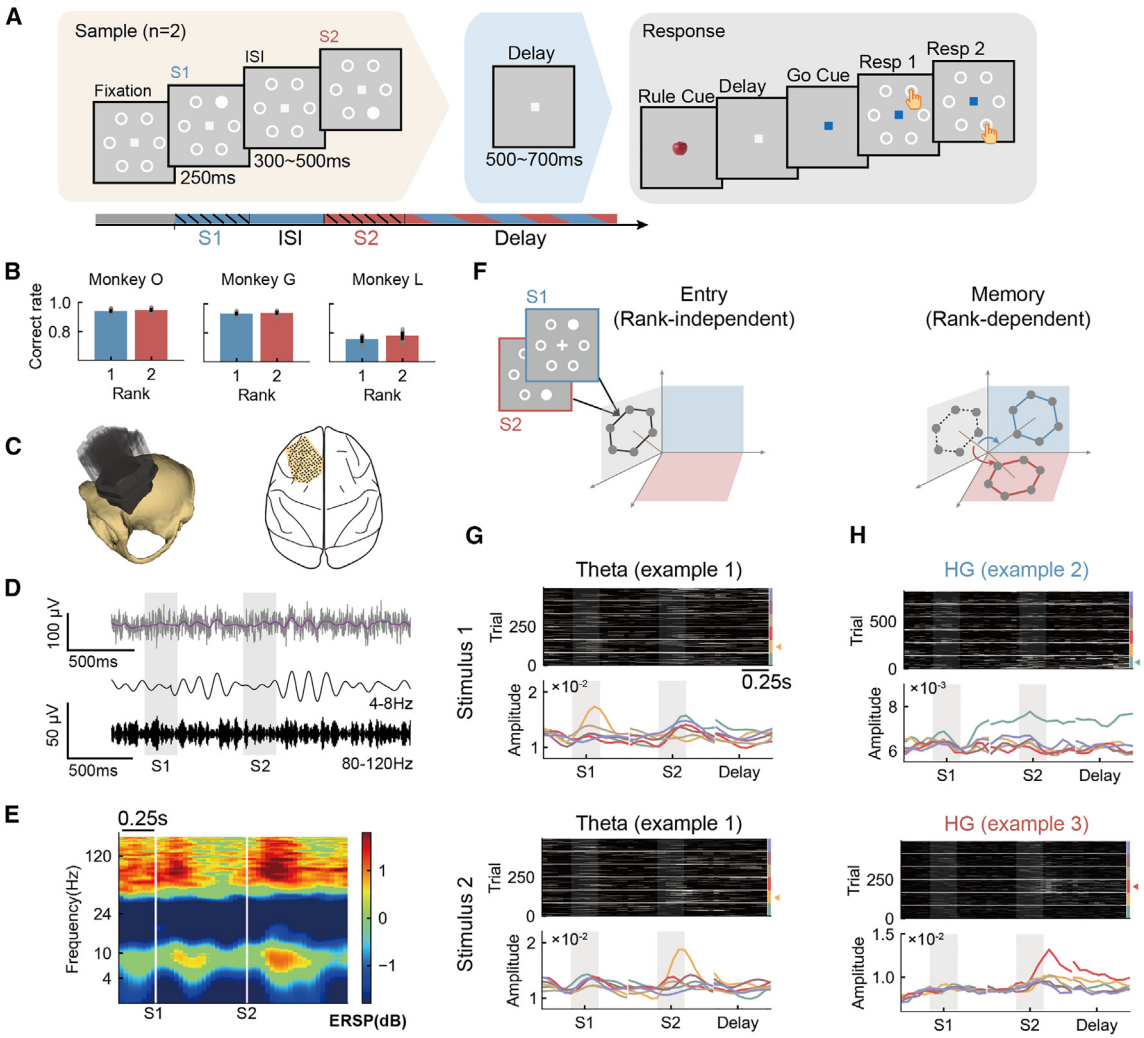

Figure 1. Macaque SWM task structure and sample electrophysiological recording data

(A) Task design. After the delay period, the sequence was reproduced from memory by touch. Each stimulus's corresponding ordinal rank is marked by a set color across all figures (blue, S1; red, S2). Repeat, report stimuli in received order; mirror, report stimuli in reversed order. ISI, inter-stimuli-interval.

(B) Behavioral performance: correct rates of responses at each ordinal rank (values for individual sessions marked by dots).

(C) Illustrations of monkey O's implanted microdrive with inserted electrodes (left) and of the electrode tip positions flattened against the lateral cortical surface of the left hemisphere (right).

(D) Example single-trial local field potential (LFP) data recorded from one channel in the frontal cortex (ventral area 46). Top: raw trace with Gaussian-smoothed (100 ms window) overlay. Middle/bottom: theta-/high gamma (HG)-bandpassed (4–8 Hz/80–120 Hz) LFP, respectively. Gray shading: stimulus on-screen times.

(E) Example time-frequency plot of LFP (with evoked potential subtraction; see Figure S1D for plot with unadjusted data). Data here are from (D)'s channel in correctly performed trials. Both theta and HG bands show significant responses (0 = non-significance mask) to S1/S2 (onsets aligned to vertical white lines). ERSP, event-related spectral perturbation.

(F) Diagram depicting distinct representations of ordinal rank-independent stimulus entry (left) and of rank-dependent memories (right) by orthogonal neural subspaces.

(G and H) Sample bandpassed LFP recordings showing rank-independent theta (G) and rank-dependent HG (H) tuning patterns. For each panel, top: trial-wise instantaneous amplitudes grouped by item identity (vertical color blocks, with colored triangles for preferred item identity); bottom: mean amplitudes over item-grouped trials.

See also Figure S1.

(Figures 2G and S2D). Furthermore, as expected from rank-independent subspaces, we found no particular difference in theta decoding accuracy between ranks (Figure S2F).

To compute rank-dependent memory subspaces, we extracted neural activity within the last 250 ms of the late delay period (Figure 2D), applying a generalized linear model to obtain subspaces that encoded S1-memory and S2-memory in a sustained manner (Figure 2F). We then projected data from the entire trial onto these late delay-derived subspaces. This approach enabled us to assess whether memory-related signals remained stably maintained from post-stimulus onset all the way through to the delay period. Previous studies have suggested

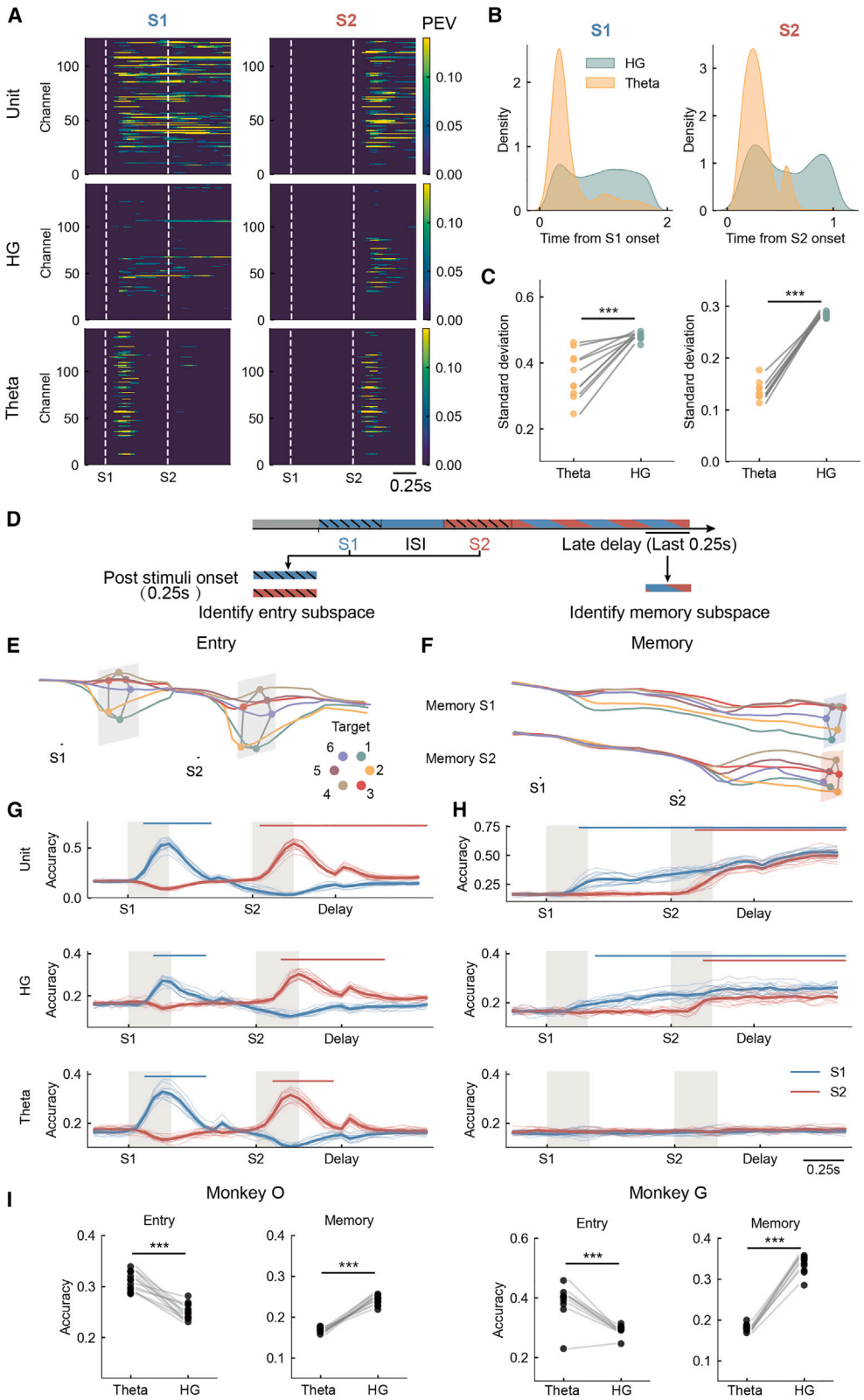

Figure 2. SWM representation by oscillatory neural activities

(A) Rank-specific significant percent explained variance (PEV) computed from unit spiking and oscillatory neural activity (evoked potential subtracted) across monkey O's frontal channels in one recording session. For theta PEV using ERP-unadjusted data, see Figure S1E.

(legend continued on next page)

that gamma activity encodes working memory information in a burst-like manner at the single-channel level.[27] By contrast, our subspace analysis revealed that at the population level, gamma signals, akin to neuronal ensembles,[2] can maintain sequential working memory information across multiple orthogonal and stable subspaces (Figures 2H and S2E).

Our results above on the representational geometry of SWM demonstrated that the entry and rank WMs could be held in separate low-dimensional subspaces based on different oscillatory components of neural activity, even though entry and memory information appears intensively mixed at the single neuronal level (Figure S2C). In addition, putative functional segregation between theta and unit/HG was further supported by selectively higher decoding accuracies using subspace projection magnitudes (Figure 2I), with theta's rank-independent entry subspace being more informative than that of HG, and HG's memory subspaces being more informative than those of theta. To verify that our neural subspaces identified above faithfully captured stimulus item information from each neural data type, we also performed classical linear support vector machine (LSVM)-based cross-time and cross-rank decoding using unit firing rates and channel-wise theta/HG power. We found the decoding results to be similar to those of the TDR analysis (Figures S3A–S3L). Besides theta and HG, beta oscillations were also found to be involved in tasks requiring control over WM[27,28]; therefore, we performed cross-time and cross-rank decoding using beta power also but did not find sufficient evidence of beta's involvement in our task (Figures S3M and S3N). Overall, our previous analyses indicated that, while HG appeared generally similar to units in terms of WM representations, i.e., supporting both entry and memory subspaces, but preferentially the latter, theta predominantly carried item information in a transient manner following stimuli onsets (Figures 1E, 2A–2C, 2G, 2I, S2A, S2B, and S2D).

Spatiotemporal characteristics of encoding theta befit putative role in WM control

WM control and maintenance are two distinct cognitive needs, likely served by different neural mechanisms.[28,29] This mechanistic segregation allows for flexible cognitive control that is not dependent on the specific WM content being controlled (i.e., rank-specific memory subspaces in our task). Given that theta does not appear to support rank-specific memory subspaces (Figure 2H) like unit/HG, and that theta frequency activities are nonetheless important in cognitive control,[30,31] it is reasonable to hypothesize that theta oscillations may be involved in regulating the stimulus entry-to-memory pipeline. Therefore, our subsequent analyses primarily focused on theta oscillations and theta-unit interactions. Building on established theories regarding WM control (e.g., spatial computing),[28,32] control-related activity (i.e., assigning items to corresponding ranks) may arise from a distinct source separate from the initial entry activity in neurons. We thus hypothesized that LFP oscillation-based top-down control signals would capture spatially organized information about the order of items. For theta to be such a candidate control signal, it would ideally possess the following spatiotemporal characteristics: first, it would preferentially occur over the time period where entry-to-memory transition happens (in our case, over the limited time windows after stimulus onsets), and second, the source of task-relevant theta activity would be spatially segregated from the memory representations to avoid interference.

To evaluate temporal relationships among entry and memory subspaces, we extracted the initial significance timings of each recording session's subspace geometry-based decoding profiles (as in Figures 2G, 2H, S2D, and S2E). We found as expected among spiking unit-based subspaces general temporal precedence of entry over memory (Figure 3A). Theta entry subspaces also became significantly informative at earlier times than unit memory subspaces (Figure 3B) and stayed informative over short post-stimulus-onset time windows where entry and memory subspace decoding profiles overlapped (Figure 3A).

To determine whether recording sites contributing to different subspaces exhibit distinct spatial organizations, we identified each neuron/channel's contribution toward entry and memory subspaces by its feature weight (FW; see STAR Methods subsection "subspace decomposition of unit and LFP activity") and then projected the FW onto the anatomical space of the recording electrode arrays (Figures 3C and S4A–S4C). We found that the segregation among neural subspaces in terms of

(B) Distributions of monkey O's post-S1-onset (left) and post-S2-onset (right) significant theta/HG power-based PEV times over the encoding period. For (B) and (C)'s monkey G counterparts, see Figures S2A and S2B.
(C) Average-across-channel standard deviations in significant post-stimuli PEV times over all recording sessions of monkey O. HG appears significantly more variable than theta (***$p < 0.001$, paired two-sample t tests).
(D) Illustration of trial-wise data segments used to compute two types of target dimensionality reduction (TDR)-based subspaces (rank-independent entry and rank-dependent memory).
(E and F) Item identity-specific unit activity across the encoding period generates a ring-like geometrical structure in each TDR subspace (data from one session per panel).
(E) Geometry of example subspace supporting rank-independent stimulus entry (same data as top panel of G).
(F) Geometry of example subspace supporting rank-dependent item memories (same data as top panel of H).
(G and H) Accuracy-over-time trajectories based on stimulus item (S1/S2) identity decoding with unit and LFP activity subspaces' projection magnitudes, and data here are from all recording sessions of monkey O ($n = 13$). For monkey G ($n = 10$), see Figures S2D and S2E. Bold lines: means across sessions; lighter lines: individual sessions' traces. Gray shading: 0–250 ms post-S1/S2 onsets. Blue/red horizontal bars atop each panel denote times of significant mean decoding accuracy for S1/S2, respectively ($\alpha = 0.05$, permutation tests).
(G) Theta/HG/unit activity-based rank-independent TDR subspaces showing transient (entry-like) decoding profiles.
(H) HG/unit activity-based rank-dependent TDR subspaces showing prolonged/sustained (memory-like) decoding profiles.
(I) Theta vs. HG subspace-based item-decoding accuracies. Except for one outlier recording session of monkey G, theta entry subspaces had better decoding performance than those of HG, and the opposite was true for memory subspaces. Accuracy values are averaged over time windows of subspace identification as in (D). Each pair of dots represents one recording session ($n = 10$ total). Paired two-sample t test, ***$p < 0.001$.
See also Figures S1–S3.

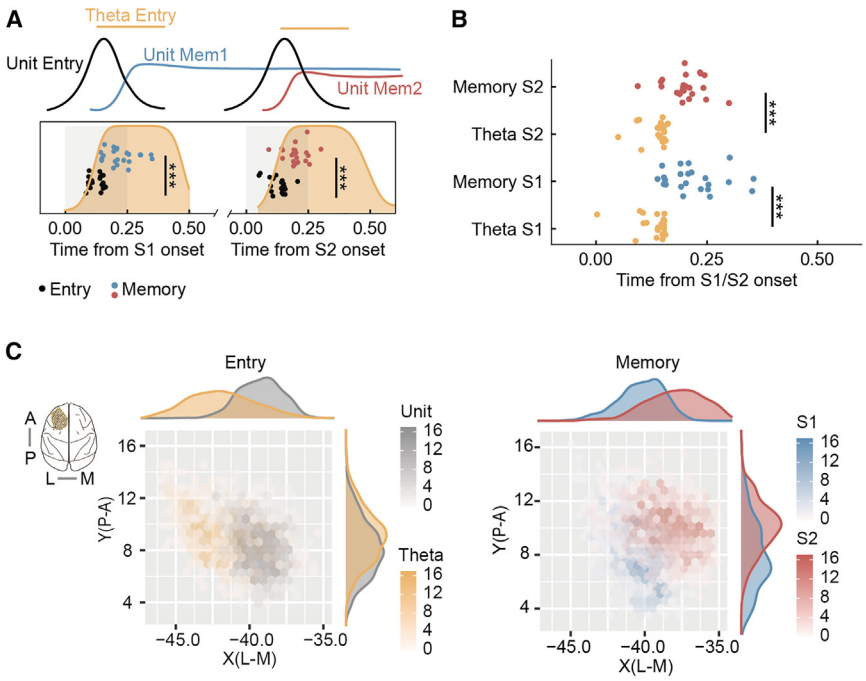

Figure 3. Spatiotemporal characteristics of theta-based rank-independent subspace

(A) Theta entry subspaces are significantly informative over likely entry-to-memory transition periods. Top: illustrations of unit activity-based rank-independent entry (black) and rank-dependent memory (blue, S1; red, S2) subspaces' target decoding accuracy-over-time profiles (cf. Figures 2G and 2H, averaged across sessions), alongside times of theta entry subspace decoding significance (yellow, significant times shared across sessions). Bottom: individual recording sessions' first significant times for unit entry and memory subspaces, overlaid with theta entry subspace's decoding accuracy-over-time profiles. Unit entry subspaces become informative at generally earlier times than memory subspaces (paired two-sample t tests, sessions from both monkeys combined). Gray shading: stimulus-on times. ***$p < 0.001$ (paired two-sample t tests).

(B) Theta entry subspaces become informative at significantly earlier times than unit memory subspaces (paired two-sample t tests as in A, ***$p < 0.001$).

(C) Estimated regions (in monkey O's microdrive array space) of highest contribution to theta/spike entry subspaces (left) and to spike memory subspaces with marginal histograms (right). For monkey G, see Figure S4B. Color bars indicate counts of permutations where peak FW locations fell within each spatial bin (see Figure S4C). See also Figure S4.

representational geometry could be extended to anatomical space as predicted. Specifically, channels with high FW of theta entry subspace often occupied areas different from those of spike entry subspace, hinting at potentially different roles during encoding (Figure 3C, left; Wilcoxon rank-sum tests, $p < 0.001$ for x/y-marginals, effect sizes = 0.812/0.310; for monkey G, see Figure S4B, top). Channels with high FW of spike memory S1 subspace and those with high FW of S2 subspace also tended to differ in spatial locations (Figure 3C, right; rank-sum tests between x/y-marginal histograms: both $p < 0.001$, effect sizes = 0.537/0.406; for monkey G, see Figure S4B, bottom). In summary, theta-based rank-independent activity was spatially distinct from unit-based WM representations and temporally situated at entry-to-memory transition windows, thereby appearing to be a suitable candidate WM control signal.

Theta-mediated selective control over neural states with rank-dependent memory information

Previously, it had been established that neurons carrying WM task-relevant information were more likely to be modulated by theta-range LFP.[26,33] We therefore investigated whether, as befitting our previous hypothesis on SWM control through segregated theta control signal, post-stimulus-onset differences in theta modulation patterns were associated with the rank-dependent unit memory subspaces. To evaluate if cross-channel theta-spike coordination actually reflected information transfer from stimulus entry to rank-dependent memories (Figure 4A), we verified the presence of LFP-organized spiking activity by computing pairwise cross-channel spike-field coherence (SFC). Each channel pair consisted of one LFP phase source and one unit source, and channels with no units or multiple units recorded were not included for analysis. We first used spikes/LFP from the same channels as those used in previous subspace-based analyses to evaluate SFC over a wide frequency range (4–140 Hz), with data from 0 to 500 ms post-stimulus onsets (Figure 4B). This allowed us to confirm that theta-range SFC appeared uniquely informative among the examined frequencies (Figures 4B and S4D). This result could not be explained by firing rate differences between target-informative/non-informative neuronal groups alone (Figure S4E). To establish further links between theta SFC and WM, we went on to tally unit-wise "SFC strength" as a metric for how much each unit's activity may be influenced by theta elsewhere (Figure 4C). Since we have previously identified rank-specific memory subspaces based on spiking units, it seemed straightforward to examine whether SFC strength could be associated with unit-wise contributions to spike memory subspaces (Figure 4D) in a rank-selective manner. More specifically, if SFC-based cross-channel connectivity did reflect rank-selective control, then SFC strengths (sums of SFC matrices across theta source dimension) computed at different ranks should selectively correspond to spike memory subspace FW at those ranks (Figures 4C and 4E).

To further quantify SFC × FW relationships and to test the prediction above, we then built multiple linear regression models with both ranks' SFC strength together as predictors and spike S1- or S2-memory subspace FW as the response variable, combining data from all sessions within each animal. To reduce confounds from irrelevant sites of the recording array, only channels with significant memory information under ANOVA analysis (Figure S2C) were included in these models. We found that, as expected, S1 SFC strength was significantly associated only with spike S1 memory subspace, and S2 SFC strength only with spike S2 memory subspace, indicating the presence of rank-selective theta modulation (Figure 4F). This rank selectivity

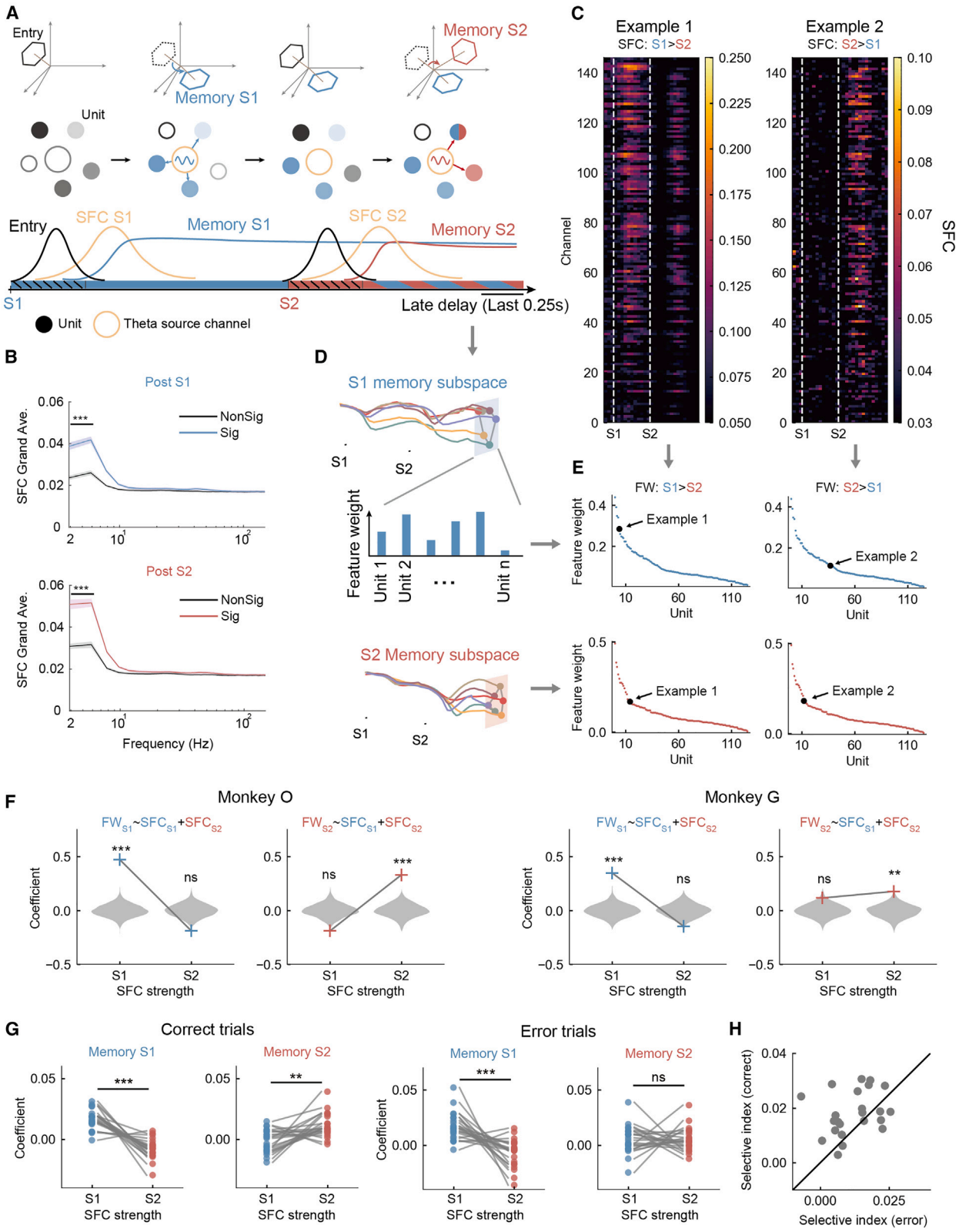

Figure 4. Theta modulation of spiking activity reveals rank-selective gating mechanism

(A) Upon stimulus entry, information held by the currently active neural population (left end of timeline) may be guided via oscillatory control into the ordinal rank-appropriate memory subspaces to achieve faithful SWM representation (right end of timeline). Black dashed lines on bottom timeline: stimulus presentation

(legend continued on next page)

was retained when the models were rebuilt using only SFC strength of units that did not have significantly different post-S1/S2 firing rates, such that rank-dependent SFC × FW associations could not be explained by differences in firing rates across ranks alone (Figure S4F). By contrast, we did not find consistently significant rank selectivity in control models where theta SFC strength was replaced with spike firing rates or with theta power/beta SFC (Figures S5A–S5C). We also did not find significant differences between ranks in terms of memory-encoding neurons' preferred theta phase distributions (Figure S5D). Notably, while the same neuronal units were used to compute S1/S2 SFC and S1/S2 memory subspaces, the neural activities used for memory subspace computations did not come from the same time period as SFC computations (late delay for TDR, post-S1/S2 onsets for SFC). Thus, the observed correlations between S1/S2-SFC and S1/S2-FW were unlikely to be due to overlapping data. Additional analyses with session-wise built multiple linear regression models provided further evidence of SFC strength-specific rank selectivity, allowing us to examine the difference between models built with data from correct trials (Figures 4G, left, and S5E) and error trials (Figures 4G, right, and S5F). While S1 SFC strength was preferentially associated with higher spike S1 FW even using error trial data, the same did not hold for S2 SFC strength × spike S2 FW (Figures 4G, right, and S5F). Quantifying the degree of rank selectivity of SFC × FW using correct vs. error trials also yielded comparatively greater selectivity with correct trials (Figure 4H). Overall, these results suggested that proper spike-theta coupling may be crucial to transforming rank-independent entry into appropriately rank-dependent memories.

Information binding generalization to 3 items

Thus far, using data from the 2-item SWM task across all correctly performed trials, we were able to identify orthogonal representations of rank-independent stimulus entry and rank-dependent memories from both unit and LFP signals. We had focused on spike-theta coupling as the main conduit through which rank-independent information was selectively gated into SWM. If our hypothetical theta control mechanism indeed operated as expected, we should be able to gather further supporting evidence through the maintenance of rank selectivity in tasks with longer sequence length. We therefore trained both monkeys to learn a 3-item sequence reproduction task (Figure 5A) to test whether such theta control can generalize to longer sequences. Both monkeys performed the task well (Figure 5B). We computed S1/S2/S3 SFC matrices using similar post-stimulus periods as before and extracted TDR entry/memory subspace FWs in the same manner as for the 2-item task. We found that the distinct anatomical segregation of spike memory subspaces remained significant as in the 2-item task (Figure 5C; $p < 0.001$ for all x/y-marginal difference tests except S2 vs. S3 x-marginal of monkey O; effect sizes: monkey O, S1 vs. S3: 0.613/0.623, S1 vs. S2: 0.617/0.140, S2 vs. S3: 0.0469/0.601; monkey G, S1 vs. S3: 1.06/1.02, S1 vs. S2: 0.245/1.03, S2 vs. S3: 1.12/0.238). We next evaluated whether the extended rank-dependent spike memory subspaces would also preferentially associate with SFC strength of corresponding ranks, in line with our proposed stimulus-to-memory pipeline via theta control. We constructed multiple linear regression models, each with the 3 rank-specific SFC strengths as predictor variables, and with one rank's spike memory subspace FW as response variable. For both monkeys, we found significant rank selectivity via stepwise regression, whereby S1/S2/S3 memory FW appeared most associated with S1/S2/S3 SFC, respectively (Figures 5D and S5G; Table S1). Stepwise models were chosen over simple 3-predictor linear models to better describe FW × SFC relationships in light of greater collinearity and with less prior assumption on predictor membership. As was the case under the 2-item task, this rank selectivity was in turn impaired in error trials (Figure S5H). Our key results based on 2-item sequences were therefore generalizable to 3-item sequences.

DISCUSSION

By decomposing unit and LFP data from macaque electrophysiological recordings, besides previously identified unit-based rank-independent entry and rank-dependent memory subspaces,[2] we discovered that theta band LFP could support entry

periods (0–250 ms post-onset). Solid-colored circles represent neurons involved in the entry or memory subspaces, while hollow circles represent neurons not responding to sequence stimuli at a given stage. Orange circles represent the source channels of theta oscillations in the SFC.

(B) Post-stimulus-onset theta-range (4–8 Hz) spike-field coherence (SFC) appeared consistently greater for task-relevant units than for other units. Solid traces represent grand averages over session-wise SFC (monkey O: $n = 13$; for monkey G, see Figure S4D) of either informative or non-informative units (Sig/NonSig, based on ANOVA analysis as in Figure S2C), with shading for SEM across sessions, and each session's SFC over frequencies (computed in ±2Hz frequency bins) was in turn averaged from unit-wise SFC means over all possible theta sources. Elevation of Sig SFC over NonSig could not be explained by firing rate differences alone (Figure S4E). ***$p < 0.001$, two-sample t tests (Sig vs. NonSig for each frequency bin).

(C) Example SFC from all possible theta source channel pairings of one unit source with generally higher SFC post-S1 (left) and another unit source with higher SFC post-S2 (right).

(D) Illustration of unit activity-based TDR subspace feature weights (FWs) as unit-wise contributions toward subspace geometry.

(E) Illustration of putative SFC × subspace FW relationship. If SFC strength sufficiently describes the degree of rank-selective theta control received by each unit, then higher SFC strength should in turn predict larger FW of the corresponding spike memory subspace. Example 1/2 units are the same ones in (B).

(F) SFC strength (here measured unit-wise as sums of pairwise SFC across all possible theta sources) displays rank selectivity in their correlations with spike memory FW, i.e., R1 SFC x S1 memory FW association is stronger than R2 SFC × S1 FW, and vice versa for S2 FW. SFC/TDR-FW was computed separately for each session and then concatenated. Gray violins indicate distributions of permutation-based linear model coefficients, and plus-sign markers show actual coefficient values. ***$p < 0.001$; **$p < 0.01$; ns, non-significant.

(G) Single-session results of correct (left) vs. error (right) trial SFC strength × FW linear models. Note that rank selectivity seen in (F) is not preserved in error trials. For monkey-wise results, see Figures S5E and S5F. Paired two-sample t tests: ***$p < 0.001$; **$p < 0.01$; ns, non-significant.

(H) SFC selective indices (sum of [β(SFC-S1) − β(SFC-S2)] magnitudes) of correct trials are generally higher than those of the same sessions' error trials (paired two-sample t test, $p = 0.00596$). Sessions from monkeys O and G are pooled here.

See also Figures S4 and S5.

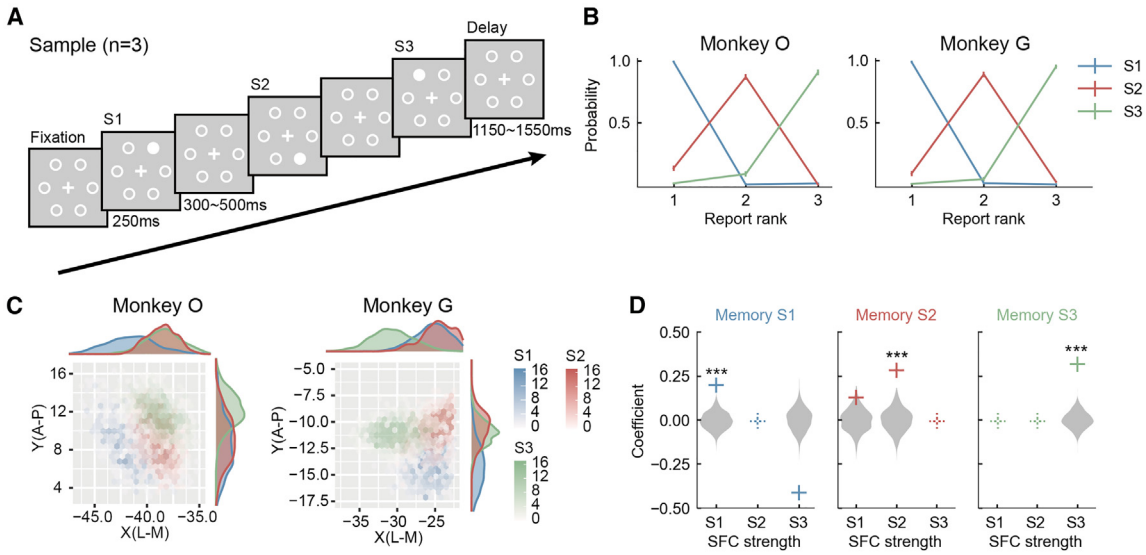

Figure 5. Generalization of SFC-based selective gating in 3-item trials

(A) Diagram for the encoding period of 3-item trials in additional recording sessions.

(B) 3-item task performance (x axis for the animal's actual reports, colored lines for the task demands). Note the relatively higher error chance when the animals should report ranks 2/3. Error bars: standard deviation across sessions (n = 10/11 for monkey O/G).

(C) Estimated regions of microdrive array space of highest contribution to 3-item task spike memory subspaces. Color bars indicate counts of permutations as in Figures 3C and S3B.

(D) Rank-selective gating is reflected among correlations between SFC strength values and rank-specific spike memory subspace FW for monkey O. Gray violins depict distributions of permutation-based linear model (based on final models in Table S1) coefficients, and plus-sign markers show the actual model coefficients (dash line markers indicate predictors removed during stepwise regression). For monkey G, see Figure S5G. ***$p < 0.001$.

See also Figure S5 and Table S1.

subspaces. Theta oscillations also displayed rank selectivity via unit-LFP coupling, with units contributing rank-dependently to different memories of constituent items within a sequence, providing a possible neural mechanism for the information transfer between entry and memory subspaces. Such a control process was apparently disrupted in error trials and could be generalized to longer sequences. Thus, our research combines the representational geometry of SWM in macaque frontal neuronal populations and neural oscillations to arrive at a potential mechanism of selective gating, i.e., unit-LFP coupling controls the ordinal rank-appropriate transformation of stimulus-entry representation into SWM.

Our sequence task with multiple WM items possessing visuospatial and ordinal aspects also gives us the opportunity to probe the more general question of contextual variable binding as a form of cognitive control. In our task, during rank-selective gating of incoming visuospatial stimuli, the unit-based entry and memory subspaces were the objects being controlled, and the theta SFC-based rank-to-item binding controls the transformation from entry-to-memory subspaces. The rank-dependent memory signal is the result of the binding process and is responsible for the maintenance of working memory after binding. Selective gating could thereby be seen as a special case of variable binding, which has been proposed to rely on the prefrontal cortex and is crucial to human cognition.[34,35] Our results regarding frontal theta control in SWM may therefore have broader implications on how the brain achieves flexible symbolic processing, namely that lower frequency oscillatory activity is capable of carrying rules of operation while higher-frequency oscillatory/spiking activity simultaneously carries the operand(s). Besides our own work, recent experimental evidence pointing in the same direction also exists in the form of cross-structural (frontal to hippocampal) theta-gamma/unit coupling to integrate WM storage and control.[36] On the more theoretical front, there have been ongoing attempts to generalize cognitive control along the stimulus-to-memory pipeline by integrating attention, WM, and long-term memory,[22,37,38] such that the putative selective gating mechanism we have identified in this work could, in principle, be extrapolated by follow-up studies onto more voluntary perceptive/attentional selection[22] and/or memory (re)-activation with flexible weights into fixed activation slots.[38]

While cognitive control over WM items via neural oscillations is conceptually well-established, particularly with regard to theta-modulated unit activity[18,26,36] and the theta-gamma code,[15,30,39,40] much of the currently accumulated evidence remains at the level of behavioral correlations with presence/absence of theta-unit modulation[18,26] or theta-gamma coupling,[15,16] and seldom venture into providing direct evidence for mechanistic descriptions of theta control. By exploiting the inherently composite characteristics of sequences in WM, we were able to extract from neural data (both unit and LFP) orthogonal subspaces representing stimulus entry and maintained memories of stimulus identity. Our group's earlier works with unit-based subspace analysis have shown that the formation of sequential memory involves a transition from a shared rank-independent entry subspace to rank-dependent memory subspaces that are

separate and orthogonal.[1,2] Therefore, the key to this perception-to-cognition transition lies in controlling the selective binding of rank-independent target representation to a given ordinal rank. As suggested by another recent study, the control signals (e.g., the rank of each item) modulating the item representations encoded by different locations in a network may be reflected in neural oscillations that act on separate neuronal populations.[28] Our results have explicitly demonstrated how the item information might be controlled to selectively bind to the corresponding rank through theta-unit coupling, and such coupling can flexibly allocate the weights of neurons corresponding to different rank-dependent memory subspaces to achieve the transformation of entry to specific memory subspaces. This SWM control mechanism could, in principle, be extended to studies involving (particularly frontal theta) oscillatory control beyond selective gating and SWM, such as conscious awareness,[12] decision-making,[41] or voluntary attention against distraction,[33] to improve our understanding of neural dynamics underlying key cognitive functions.

In our work, both memory maintenance and control could be implemented locally, i.e., within the frontal cortex, despite non-local theta-spike locking occurring at a relatively smaller scale as represented by spike-based subspaces and oscillation-based connectivity patterns. Although frontal theta activity is frequently investigated and has been our focus throughout this study, theta oscillations are by no means so limited in scope or origin.[42–44] Classical cross-regional theta-WM associations have been found over the neocortical frontoparietal network,[33,45,46] as well as cortico-subcortical coupling that involves the hippocampus and/or basal ganglia.[34,36,47,48] The binding of items to corresponding ranks might also occur in the hippocampus, which could organize the sequence received via cortical inputs and relay it back to the frontal cortex as needed.[20] Another possibility is that, rather than directly modulating frontal neurons, frontal theta oscillations might primarily modulate hippocampal neurons.[36] The frontal neurons modulated by theta oscillations may represent those that synchronize with hippocampal neuronal activity, likely facilitating communication between the two regions. While our recordings covered a larger extent of the frontal neocortex than would be possible with more traditional microelectrode arrays,[25] they leave room for further exploration of brain-wide large-scale theta modulation as a key aspect of cognitive control. The present results do not exclude the possibility that multiple cognitive control mechanisms may coexist with our proposed form of theta control, potentially sharing oscillatory components and/or participant brain regions.

Crucially, cross-channel frontal theta modulation over spikes in our results appeared to disambiguate between competing memories at a given ordinal rank in time, implying that control-related information may be carried by changing spatial coupling distribution. Similar conceptualization of "spatial computing," where changing spatial distributions of active neural components could encode aspects of WM task control, has only recently been proposed regarding beta-gamma activity.[28,49] Our work is among the first to provide direct evidence of such spatial computing and likely the first to extend this formulation to meso-scale cross-channel coupling. In addition, the anatomical segregation of theta entry channels from spike entry/memory indicates that theta oscillation may play a role as a driver for such spatial computing-based subspace transformation, which could theoretically allow generalization of spatially distributed oscillatory modulation to other cognitive control demands beyond rank-item associations. While our results differed from earlier demonstrations of beta-based spatial computing in that our putative control signal lies in the theta range (with no clear evidence for beta having a similar role), these findings are not in conflict. Besides potential frequency band-wise multiplexing of information processing as already observed in the gamma band,[50] beta-based WM control was potentially regulating selective inhibition of information,[49] such as attention selection in salience-goal conflict tasks[51] and sensory input selection in working memory tasks with distractors.[33] By contrast, our theta-based WM control mechanism is more closely associated with stimuli gating. Similarly, our results on similar preferred phase distributions across ranks were also obtained from the encoding period instead of the delay period during which earlier findings of different preferred phases for multiple WM items[26] were established, and present data are insufficient to claim that WM control via phasic modulation would preclude spatial computing, or vice versa. Regardless, since low-frequency oscillatory activity seems generally more suited to large-scale control due to relative ease of propagation,[52] future studies may be able to generalize our findings to other oscillatory couplings of the form "low-frequency controller/modulator and high-frequency (bottom-up) information-aspect carrier."

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to the lead contact, Liping Wang (liping.wang@ion.ac.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data used for analyses in this study are available from ScienceDB via the DOI listed in the key resources table as of the date of publication.
- All original code has been uploaded to ScienceDB and is publicly available via the DOI listed in the key resources table as of the date of publication.
- Any additional information required to reanalyze the data is available from the lead contact upon reasonable request.

ACKNOWLEDGMENTS

We thank Dahui Wang and Yang Xie for their comments on the manuscript. This work was supported by the National Science and Technology Innovation 2030 Major Project 2021ZD0204100 (2021ZD0204102), the National Science Fund for Distinguished Young Scholars (32225022), the CAS Project for Young Scientists in Basic Research (YSBR-071), and the CAS Strategic Priority Research Program XDB1010202 to L.W.; by the National Science and Technology Innovation 2030 Major Project 2021ZD0204200 (2021ZD0204204) and the CAS Youth Innovation Promotion Association (2023000024) to W.F.; and by the National Natural Science Foundation of China grants to W.F. (32100830) and to X.J. (32200834). L.W. is a SANS (Shanghai Academy of Natural Sciences) Exploration Scholar.

AUTHOR CONTRIBUTIONS

Conceptualization, L.W.; methodology, L.W., F.W., and X.J.; formal analysis, F.W. and X.J.; investigation, J.C. and C.Z.; writing – original draft, X.J. and

F.W.; writing – review & editing, L.W., F.W., and X.J.; funding acquisition, L.W., F.W., and X.J.; resources, J.C. and C.Z.; supervision, L.W.

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 task
  - Surgical procedures and implants
  - Electrophysiological recording and data preprocessing
  - Time-frequency analysis
  - Single neuron and band-limited LFP tuning properties
  - Cross-time/rank generalization decoding of stimulus location
  - Subspace decomposition of unit and LFP activity
  - Comparing distributions of subspace feature weights
  - Percent explained variance (PEV) analysis
  - Spike-field coherence (SFC)
  - Linear models for decoding weights and SFC

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2025.02.031.

Received: November 28, 2024
Revised: January 27, 2025
Accepted: February 17, 2025
Published: March 13, 2025

REFERENCES

1. Xie, Y., Hu, P., Li, J., Chen, J., Song, W., Wang, X.-J., Yang, T., Dehaene, S., Tang, S., Min, B., et al. (2022). Geometry of sequence working memory in macaque prefrontal cortex. Science *375*, 632–639. https://doi.org/10.1126/science.abm0204.
2. Chen, J., Zhang, C., Hu, P., Min, B., and Wang, L. (2024). Flexible control of sequence working memory in the macaque frontal cortex. Neuron *112*, 3502–3514.e6. https://doi.org/10.1016/j.neuron.2024.07.024.
3. Singer, W., Engel, A.K., Kreiter, A.K., Munk, M.H., Neuenschwander, S., and Roelfsema, P.R. (1997). Neuronal assemblies: necessity, signature and detectability. Trends Cogn. Sci. *1*, 252–261. https://doi.org/10.1016/S1364-6613(97)01079-6.
4. Fries, P. (2015). Rhythms for cognition: communication through coherence. Neuron *88*, 220–235. https://doi.org/10.1016/j.neuron.2015.09.034.
5. Gray, C.M., König, P., Engel, A.K., and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature *338*, 334–337. https://doi.org/10.1038/338334a0.
6. Steinmetz, P.N., Roy, A., Fitzgerald, P.J., Hsiao, S.S., Johnson, K.O., and Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. Nature *404*, 187–190. https://doi.org/10.1038/35004588.
7. Fries, P., Schröder, J.-H., Roelfsema, P.R., Singer, W., and Engel, A.K. (2002). Oscillatory neuronal synchronization in primary visual cortex as a correlate of stimulus selection. J. Neurosci. *22*, 3739–3754. https://doi.org/10.1523/JNEUROSCI.22-09-03739.2002.
8. Nakatani, H., and van Leeuwen, C. (2006). Transient synchrony of distant brain areas and perceptual switching in ambiguous figures. Biol. Cybern. *94*, 445–457. https://doi.org/10.1007/s00422-006-0057-9.
9. Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., and Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. J. Neurosci. *27*, 2858–2865. https://doi.org/10.1523/JNEUROSCI.4623-06.2007.
10. Klimesch, W., Freunberger, R., Sauseng, P., and Gruber, W. (2008). A short review of slow phase synchronization and memory: evidence for control processes in different memory systems? Brain Res. *1235*, 31–44. https://doi.org/10.1016/j.brainres.2008.06.049.
11. Ratcliffe, O., Shapiro, K., and Staresina, B.P. (2022). Fronto-medial theta coordinates posterior maintenance of working memory content. Curr. Biol. *32*, 2121–2129.e3. https://doi.org/10.1016/j.cub.2022.03.045.
12. Dwarakanath, A., Kapoor, V., Werner, J., Safavi, S., Fedorov, L.A., Logothetis, N.K., and Panagiotaropoulos, T.I. (2023). Bistability of prefrontal states gates access to consciousness. Neuron *111*, 1666–1683.e4. https://doi.org/10.1016/j.neuron.2023.02.027.
13. Liebe, S., Hoerzer, G.M., Logothetis, N.K., and Rainer, G. (2012). Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. Nat. Neurosci. *15*, 456–462. https://doi.org/10.1038/nn.3038.
14. Rey, H.G., Fried, I., and Quian Quiroga, R. (2014). Timing of single-neuron and local field potential responses in the human medial temporal lobe. Curr. Biol. *24*, 299–304. https://doi.org/10.1016/j.cub.2013.12.004.
15. Alekseichuk, I., Turi, Z., Amador de Lara, G., Antal, A., and Paulus, W. (2016). Spatial working memory in humans depends on theta and high gamma synchronization in the prefrontal cortex. Curr. Biol. *26*, 1513–1521. https://doi.org/10.1016/j.cub.2016.04.035.
16. Rajji, T.K., Zomorrodi, R., Barr, M.S., Blumberger, D.M., Mulsant, B.H., and Daskalakis, Z.J. (2017). Ordering information in working memory and modulation of gamma by theta oscillations in humans. Cereb. Cortex *27*, 1482–1490. https://doi.org/10.1093/cercor/bhv326.
17. Clouter, A., Shapiro, K.L., and Hanslmayr, S. (2017). Theta phase synchronization is the glue that binds human associative memory. Curr. Biol. *27*, 3143–3148.e6. https://doi.org/10.1016/j.cub.2017.09.001.
18. Quirk, C.R., Zutshi, I., Srikanth, S., Fu, M.L., Devico Marciano, N., Wright, M.K., Parsey, D.F., Liu, S., Siretskiy, R.E., Huynh, T.L., et al. (2021). Precisely timed theta oscillations are selectively required during the encoding phase of memory. Nat. Neurosci. *24*, 1614–1627. https://doi.org/10.1038/s41593-021-00919-0.
19. Reddy, L., Self, M.W., Zoefel, B., Poncet, M., Possel, J.K., Peters, J.C., Baayen, J.C., Idema, S., VanRullen, R., and Roelfsema, P.R. (2021). Theta-phase dependent neuronal coding during sequence learning in human single neurons. Nat. Commun. *12*, 4839. https://doi.org/10.1038/s41467-021-25150-0.
20. Buzsáki, G., and Tingley, D. (2018). Space and time: the hippocampus as a sequence generator. Trends Cogn. Sci. *22*, 853–869. https://doi.org/10.1016/j.tics.2018.07.006.
21. Li, H.-H., and Curtis, C.E. (2023). Neural population dynamics of human working memory. Curr. Biol. *33*, 3775–3784.e4. https://doi.org/10.1016/j.cub.2023.07.067.
22. 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.
23. Zhang, H., Zhen, Y., Yu, S., Long, T., Zhang, B., Jiang, X., Li, J., Fang, W., Sigman, M., Dehaene, S., et al. (2022). Working memory for spatial sequences: developmental and evolutionary factors in encoding ordinal and relational structures. J. Neurosci. *42*, 850–864. https://doi.org/10.1523/JNEUROSCI.0603-21.2021.
24. Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., and Wang, L. (2018). Production of supra-regular spatial sequences by macaque monkeys. Curr. Biol. *28*, 1851–1859.e4. https://doi.org/10.1016/j.cub.2018.04.047.

25. Dotson, N.M., Hoffman, S.J., Goodell, B., and Gray, C.M. (2017). A large-scale semi-chronic microdrive recording system for non-human primates. Neuron *96*, 769–782.e2. https://doi.org/10.1016/j.neuron.2017.09.050.

26. Siegel, M., Warden, M.R., and Miller, E.K. (2009). Phase-dependent neuronal coding of objects in short-term memory. Proc. Natl. Acad. Sci. USA *106*, 21341–21346. https://doi.org/10.1073/pnas.0908193106.

27. Lundqvist, M., Rose, J., Herman, P., Brincat, S.L., Buschman, T.J., and Miller, E.K. (2016). Gamma and beta bursts underlie working memory. Neuron *90*, 152–164. https://doi.org/10.1016/j.neuron.2016.02.028.

28. 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.

29. Tian, Z., Chen, J., Zhang, C., Min, B., Xu, B., and Wang, L. (2024). Mental programming of spatial sequences in working memory in the macaque frontal cortex. Science *385*, eadp6091. https://doi.org/10.1126/science.adp6091.

30. 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.

31. Senoussi, M., Verbeke, P., Desender, K., De Loof, E., Talsma, D., and Verguts, T. (2022). Theta oscillations shift towards optimal frequency for cognitive control. Nat. Hum. Behav. *6*, 1000–1013. https://doi.org/10.1038/s41562-022-01335-5.

32. Yu, D., Li, T., Ding, Q., Wu, Y., Fu, Z., Zhan, X., Yang, L., and Jia, Y. (2024). Maintenance of delay-period activity in working memory task is modulated by local network structure. PLoS Comput. Biol. *20*, e1012415. https://doi.org/10.1371/journal.pcbi.1012415.

33. Jacob, S.N., Hähnke, D., and Nieder, A. (2018). Structuring of abstract working memory content by fronto-parietal synchrony in primate cortex. Neuron *99*, 588–597.e5. https://doi.org/10.1016/j.neuron.2018.07.025.

34. Kriete, T., Noelle, D.C., Cohen, J.D., and O'Reilly, R.C. (2013). Indirection and symbol-like processing in the prefrontal cortex and basal ganglia. Proc. Natl. Acad. Sci. USA *110*, 16390–16395. https://doi.org/10.1073/pnas.1303547110.

35. Calmus, R., Wilson, B., Kikuchi, Y., and Petkov, C.I. (2020). Structured sequence processing and combinatorial binding: neurobiologically and computationally informed hypotheses. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190304. https://doi.org/10.1098/rstb.2019.0304.

36. Daume, J., Kamiński, J., Schjetnan, A.G.P., Salimpour, Y., Khan, U., Kyzar, M., Reed, C.M., Anderson, W.S., Valiante, T.A., Mamelak, A.N., et al. (2024). Control of working memory by phase–amplitude coupling of human hippocampal neurons. Nature *629*, 393–401. https://doi.org/10.1038/s41586-024-07309-z.

37. Latapie, H., Kilic, O., Thórisson, K.R., Wang, P., and Hammer, P. (2022). Neurosymbolic systems of perception and cognition: the role of attention. Front. Psychol. *13*, 806397. https://doi.org/10.3389/fpsyg.2022.806397.

38. Whittington, J.C.R., Dorrell, W., Behrens, T., Ganguli, S., and El-Gaby, M. (2023). On prefrontal working memory and hippocampal episodic memory: Unifying memories stored in weights and activation slots. Preprint at bioRxiv. https://doi.org/10.1101/2023.11.05.565662.

39. Ursino, M., Cesaretti, N., and Pirazzini, G. (2023). A model of working memory for encoding multiple items and ordered sequences exploiting the theta-gamma code. Cogn. Neurodyn. *17*, 489–521. https://doi.org/10.1007/s11571-022-09836-9.

40. Heusser, A.C., Poeppel, D., Ezzyat, Y., and Davachi, L. (2016). Episodic sequence memory is supported by a theta-gamma phase code. Nat. Neurosci. *19*, 1374–1380. https://doi.org/10.1038/nn.4374.

41. Tang, W., Shin, J.D., and Jadhav, S.P. (2021). Multiple time-scales of decision-making in the hippocampus and prefrontal cortex. eLife *10*, e66227. https://doi.org/10.7554/eLife.66227.

42. 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.

43. Solomon, E.A., Kragel, J.E., Sperling, M.R., Sharan, A., Worrell, G., Kucewicz, M., Inman, C.S., Lega, B., Davis, K.A., Stein, J.M., et al. (2017). Widespread theta synchrony and high-frequency desynchronization underlies enhanced cognition. Nat. Commun. *8*, 1704. https://doi.org/10.1038/s41467-017-01763-2.

44. Herweg, N.A., Solomon, E.A., and Kahana, M.J. (2020). Theta oscillations in human memory. Trends Cogn. Sci. *24*, 208–227. https://doi.org/10.1016/j.tics.2019.12.006.

45. Spoormaker, V.I., Gleiser, P.M., and Czisch, M. (2012). Frontoparietal connectivity and hierarchical structure of the brain's functional network during sleep. Front. Neurol. *3*, 80. https://doi.org/10.3389/fneur.2012.00080.

46. Rutishauser, U., Ross, I.B., Mamelak, A.N., and Schuman, E.M. (2010). Human memory strength is predicted by theta-frequency phase-locking of single neurons. Nature *464*, 903–907. https://doi.org/10.1038/nature08860.

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

48. Domanski, A.P.F., Kucewicz, M.T., Russo, E., Tricklebank, M.D., Robinson, E.S.J., Durstewitz, D., and Jones, M.W. (2023). Distinct hippocampal-prefrontal neural assemblies coordinate memory encoding, maintenance, and recall. Curr. Biol. *33*, 1220–1236.e4. https://doi.org/10.1016/j.cub.2023.02.029.

49. Lundqvist, M., Miller, E.K., Nordmark, J., Liljefors, J., and Herman, P. (2024). Beta: bursts of cognition. Trends Cogn. Sci. *28*, 662–676. https://doi.org/10.1016/j.tics.2024.03.010.

50. Khamechian, M.B., Kozyrev, V., Treue, S., Esghaei, M., and Daliri, M.R. (2019). Routing information flow by separate neural synchrony frequencies allows for "functionally labeled lines" in higher primate cortex. Proc. Natl. Acad. Sci. USA *116*, 12506–12515. https://doi.org/10.1073/pnas.1819827116.

51. Dubey, A., Markowitz, D.A., and Pesaran, B. (2023). Top-down control of exogenous attentional selection is mediated by beta coherence in prefrontal cortex. Neuron *111*, 3321–3334.e5. https://doi.org/10.1016/j.neuron.2023.06.025.

52. Myers, J.C., Smith, E.H., Leszczynski, M., O'Sullivan, J., Yates, M.J., McKhann, G., Mesgarani, N., Schroeder, C., Schevon, C., and Sheth, S.A. (2022). The spatial reach of neuronal coherence and spike-field coupling across the human neocortex. J. Neurosci. *42*, 6285–6294. https://doi.org/10.1523/JNEUROSCI.0050-22.2022.

53. Bokil, H., Andrews, P., Kulkarni, J.E., Mehta, S., and Mitra, P.P. (2010). Chronux: a platform for analyzing neural signals. J. Neurosci. Methods *192*, 146–151. https://doi.org/10.1016/j.jneumeth.2010.06.020.

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

55. Jun, J.J., Steinmetz, N.A., Siegle, J.H., Denman, D.J., Bauza, M., Barbarits, B., Lee, A.K., Anastassiou, C.A., Andrei, A., Aydın, Ç., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. Nature *551*, 232–236. https://doi.org/10.1038/nature24636.

56. Curzer, H.J., Perry, G., Wallace, M.C., and Perry, D. (2016). The three Rs of animal research: what they mean for the institutional animal care and use committee and why. Sci. Eng. Ethics *22*, 549–565. https://doi.org/10.1007/s11948-015-9659-8.

57. Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al. (2012). 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn. Reson. Imaging *30*, 1323–1341. https://doi.org/10.1016/j.mri.2012.05.001.

58. MacArthur Clark, J.A.M., and Sun, D. (2020). Guidelines for the ethical review of laboratory animal welfare People's Republic of China National Standard GB/T 35892–2018 [Issued 6 February 2018 Effective from 1 September 2018. Animal Model Exp. Med. *3*, 103–113. https://doi.org/10.1002/ame2.12111.

59. Zanos, T.P., Mineault, P.J., and Pack, C.C. (2011). Removal of spurious correlations between spikes and local field potentials. J. Neurophysiol. *105*, 474–486. https://doi.org/10.1152/jn.00642.2010.

60. Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

61. Wagner, J., Makeig, S., Gola, M., Neuper, C., and Müller-Putz, G. (2016). Distinct β band oscillatory networks subserving motor and cognitive control during gait adaptation. J. Neurosci. *36*, 2212–2226. https://doi.org/10.1523/JNEUROSCI.3543-15.2016.

62. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84. https://doi.org/10.1038/nature12742.

63. Rosenthal, R. (1991). Meta-Analytic Procedures for Social Research, Revised Edition (SAGE Publications).

64. Olejnik, S., and Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. Psychol. Methods *8*, 434–447. https://doi.org/10.1037/1082-989X.8.4.434.

65. 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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Data	This study	ScienceDB: https://doi.org/10.57760/sciencedb.20866
Experimental models: Organisms/strains		
Macaca mulatta	Xexbio, Topgene Biotechnology	N/A
Software and algorithms		
Matlab	Mathworks	RRID: SCR_001622
Python	Python Software Foundation	RRID: SCR_008394
R	The R Project for Statistical Computing	RRID: SCR_001905
EEGLAB	Swartz Center for Computational Neuroscience	RRID: SCR_007292
Chronux	Bokil et al.[53]	RRID: SCR_005547
Fieldtrip	Oostenveld et al.[54]	RRID: SCR_004849
JRCLUST	Jun et al.[55]	Github: https://github.com/JaneliaSciComp/JRCLUST
Analysis code	This study	Github: https://github.com/biodipus/Oscillation-SWM

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Three healthy male adult rhesus monkeys (Macaca mulatta, 7-8 years old, 9-10 kg, group housing with individually separate enclosures), with no prior history of research participation, were used in our experiments. All research procedures (surgical and experimental) were conducted following approval by the ethical committee of the Center for Excellence in Brain Science and Intelligence Technology (formerly the Institute of Neuroscience), Chinese Academy of Sciences, and conformed to the regulatory standards of the Chinese National Guidelines (GB/T 35892-2018) on animal welfare.

METHOD DETAILS

Behavioral task

After acclimation to custom-built primate chairs over 2-3 weeks, for each task session, monkeys were seated facing a 21.5'' LED monitor, with head position fixed and a liquid reward guide tube set in front of their lips. A horizontal slot on the chair allowed reaching and pulling of a rod/lever that could send pull/release time markers to the electrophysiology recording computer. Each trial was initiated when the performing monkey voluntarily pulled the rod and maintained visual fixation for 1 second over a yellow cross appearing at the center of the monitor. Eye positions were recorded by an infrared eye-tracking system (Eyelink, SR Research Ltd.) at >250 Hz sampling rate. Once the monkey achieved fixation and held the rod in place (both must then be maintained until the cue to respond appeared, or the current trial would terminate early), the center dot would become white ("cue-frame"), and the sample period for the current trial would begin after 500 ms.

In each trial, during the sample period, 2 stimuli identical in size (2° in diameter, 11° away from the screen center) and in color (white) were presented sequentially. The stimuli were randomly drawn from six candidate spatial locations—marked by empty circles 100 ms appearing after fixation—of a hexagon surrounding the center dot without replacement. The probability for all possible sequences (30 in total, since stimuli 1 and 2 cannot appear at the same location) to appear for each trial was balanced so that each unique sequence appeared in ~1/30 of trials of a given task session. Each stimulus was set to disappear after 250 ms, so that after the second stimulus was presented, the monkeys had to maintain the ordinal ranks and spatial locations of the stimuli in working memory in order to receive the reward at trial's end. The interval between stimulus 1's disappearance/offset and stimulus 2's appearance/onset was randomly set to between 300 and 500 ms for each trial.

After the delay period following stimulus 2's offset (500-700 ms), a "rule" cue (pictures of various fruits) was presented adjacent to the center dot to signal the correct way to reproduce the sequence, i.e. in either the sample order ("repeat" trial) or the reverse order ("mirror" trial), by making sequential touches to the appropriate locations on the screen. Each trial's reproduction rule was determined pseudo-randomly prior to trial start. Following the 3R principles of animal use in research,[56] the monkeys used in this study were taught to learn behavioral tasks most relevant to the largest number of research projects; therefore, while this work only utilized

data and behavioral records for the encoding period (prior to rule cue) of this task, the full task structure is reported here for completeness.

After the rule cue offset (250 ms) and another delay (500-800 ms, set randomly each trial), the center white dot turned blue ("go" cue), initiating the response period. From this point on, the monkeys were freed from fixation and could release the rod to touch the screen. Touching correct locations would turn them white, while touching erroneous locations would turn them blue. The trial termination criterion was set to be "error-stop": whenever an erroneous location was reported, the trial was terminated. When a sequence was reproduced correctly in full, a drop of water was given as reward. All monkeys (O, L, and G) were given length-2 sequences with reproduction order rules during electrophysiological recording, as well as additional sessions with length-3 sequences where the reproduction rule was set to forward repetition only. For monkeys O, L, and G, daily electrophysiological recording sessions (with task whenever animal condition permits) took place over 6, 8, and 5 months, respectively. Only recording sessions with >60% overall sequence reproduction accuracy and with more than 100 units were included for analysis. Due to poor recording quality (averaging 20 units per session), monkey L had no recording session that could fit both criteria.

Surgical procedures and implants

Each monkey received implantation of a 157-channel semi-chronic icrodrive recording system (Gray Matter Research USA)[25] with independently adjustable tungsten electrodes (AlphaOmega, ~1MΩ, 1.5 mm spacing) in the left hemisphere over the frontal cortex. Custom-built (i.e. form-fitted to the skull surface) recording chamber and head holder were designed for each monkey and aligned to pre-surgical MRI/CT in 3D Slicer,[57] so that the chamber and the enclosed microdrive could cover as large an extent of the frontal cortex as possible, and that the head holder over the right hemisphere could be securely attached to the skull alongside the chamber on the left side.

Four surgeries were performed per animal over 2-3 weeks: 1, implant the head holder; 2, implant the recording chamber; 3, perform craniotomy; 4, install the microdrive. To fix both the chamber and the head holder to the skull, screws (ceramic for head holder, titanium for chamber) and bone cement were applied during surgery. Immediately after chamber implantation, Single Bond Universal Adhesive resin cement (3M Inc., USA) was applied over the chamber-skull junction for sealing. One week after chamber implantation, biological fluid found in the recording chamber was removed by swab and cultured to verify sterility. The craniotomy surgery was then performed to remove only the skull within the chamber, trimming the boundary of the craniotomy to align with the chamber's internal walls while leaving the dura intact. A dummy plug was then applied to seal the chamber, and a sterility test was performed after another week. During the next (and last) surgery, the leftover dura within the craniotomy was removed; the microdrive was implanted afterwards, then mounted to the chamber with screws. The gap between the microdrive's bottom end and the cortical surface was filled with dura gel for protection. All pre-surgical and surgical procedures conformed to the regulatory standards of the Chinese National Guidelines (GB/T 35892-2018) on animal welfare.[58]

Electrophysiological recording and data preprocessing

Individual microelectrodes were gradually advanced into the cortical surface from two days post-microdrive implantation. All electrodes were moved across the electrically insulative dura gel and into the cortical surface (as evaluated by nanoZ impedance tester) within one week to avoid blockage due to dura hyperplasia. Electrode advancement was then slowed down to allow daily check of spike waveforms online, which were recorded alongside broadband electrophysiological signals at 40 kHz through the data acquisition system (OmniPlex, Plexon Inc.), where eye position and task event markers were also simultaneously registered. For offline spike sorting, the broadband signal was bandpassed via Butterworth filter between 300 and 8000 Hz, then clustered using Ironclust (variant of JRCLUST[55]) and manually curated with the Plexon Offline Sorter.

To prepare local field potential (LFP) data for subsequent analyses, offline preprocessing was performed using the Fieldtrip toolbox[54] and customs scripts under MATLAB 2020b. Broadband data were downsampled to 1000 Hz and notch-filtered (zero-phase) with 50 Hz harmonics up to 150 Hz. Electrode positions were identified by alignment of post-surgical CT with pre-surgical MRI, and corresponding channels where electrode tips fell within white matter or ambiguously identified regions were excluded from this study. Channels with frequent sharp transient artifacts due to animal movement and/or faulty cable contact were also excluded after visual inspection with Fieldtrip. A common median reference was then applied to the remaining channels. For 2-item task sessions, two additional preprocessing steps were performed: 1, due to common low-frequency noise from contact between headstage cable and port, a 32-channel-wise common average re-referencing was applied; 2, due to the prevalent occupancy of spikes in the majority of channels potentially introducing spurious spike-LFP coupling, unit spike waveforms were removed from the original 40 kHz signal using a Bayesian algorithm[59] before other LFP preprocessing steps. Prior to further analyses, stimulus-locked low-frequency "evoked potential" responses were subtracted from trial-wise trimmed LFP data to minimize interference with task-relevant oscillatory signals.[26]

A total of 13/10 2-item task recording sessions for monkeys O and G with the best overall behavioral performance and recording quality were included in this study. For monkey O, a total of 1,390 frontal neurons (originally 1,508; 1,390 is the number on channels with usable LFP and with one sorted neuron per electrode; rare occurrences (<<1% of data) with the same channel recording multiple units were excluded from this study) were recorded across days (one data set per day); for monkey G, 824 suitable neurons were recorded. 10/11 sessions of 3-item task data from monkeys O and G were also included in this study, containing a total of 937/939 neurons. After LFP preprocessing, 146 and 141 out of 157 total channels in monkeys O and G, respectively, were retained for further analyses.

Time-frequency analysis

Time-frequency matrices of sequence working memory task event-related spectral power (ERSP, as shown in Figure 1E) were generated from the preprocessed LFP using EEGLAB,[60] with ERP removal via subtraction of channel-wise mean LFP over trials.[26] Each channel's ERSP was calculated from 2 to 200 Hz in 1 Hz steps with Morlet wavelets (<=1-second window; linearly increasing cycles from 1 cycle onward up to 50 cycles, Hanning-tapered window applied) over LFP segments covering [-1500, 5500] ms around the post-fixation cue-frames (as detailed in the "Behavioral task" section above) of all correctly performed trials, then time-warped[61] to match subsequent visual stimuli occurrence times across trials. The resulting time-frequency matrix was then normalized with respect to mean power at -1000 to -500 ms before the cue-frame time, and masked with false discovery rate (FDR)-corrected two-tailed bootstrapping significance (n=200, α=0.05) using the same pre-cue-frame time range as baseline. Channels were deemed to contain significant event-related LFP signal if they covered more than +-2.5x ERSP versus baseline across all significant bins of the time-frequency matrix, or had >5 significant bins at >=80 Hz during encoding.

Single neuron and band-limited LFP tuning properties

For each unit defined by spike sorting (or each channel's bandpassed LFP), its tuning property was visualized (Figures 1G, 1H, S1F, and S1G) by computing mean firing rates (mean Hilbert amplitude for LFP) of all correct trials in 100 ms consecutive time bins with 50 ms steps, then computing across-trial averages grouped by stimulus identity (dot positions 1-6 on the screen). To further evaluate neuronal/LFP tuning with respect to participation in stimulus entry and/or in memory maintenance during the two-item sequence task, three distinct time periods were defined: 0-250 ms after first stimulus presentation (stimulus period S1), 0-250 ms after second stimulus presentation (stimulus period S2), and -250-0 ms before the rule cue (delay period). Each neuron's mean firing rate (FR) (or each channel's mean Hilbert amplitude) for each period across each trial group (first or second stimulus/target, dot positions 1-6) was then computed, and an ANOVA test was conducted over all periods and firing rates/amplitudes to identify significant selectivity (Figure S2C). Neurons/channels with selective spatial tuning in both stimulus periods and no significant interaction of stimulation position and rank were designated as stimulus entry-related, while those with tuning in the delay period were designated as memory-related; these designations are not mutually exclusive.

Cross-time/rank generalization decoding of stimulus location

Time generalization decoding (Figures S3A, S3C, S3E, S3G, S3I, S3K, and S3M) was performed with support vector machine classifiers implemented using the scikit-learn toolbox. Classifiers were first trained to identify the stimulus location (1 out of 6 possible) of rank-1 or rank-2 from each time window, then applied to predict the target location at all time windows across the whole trial. Three-fold cross-validation was implemented by splitting the neural data into 3 subsamples, each randomly drawn from the entire dataset. Decoders were then trained on 2 of the subsamples and tested on the remaining subsample, with this process repeated using all 3 subsamples as test data once. This cross-validation process was repeated 1000 times and the overall decoding accuracy was taken as the mean across the 1000 repetitions. Statistical significance for decoding accuracy was determined by comparing mean decoding accuracy from the original data of each time window against the accuracy from the shuffled data of the corresponding time. For the rank generalization decoding (Figures S3B, S3D, S3F, S3H, S3J, S3L, and S3N), we trained classifiers of rank-1 (rank-2) stimulus location (i.e., target item identity) from each time window to predict the stimulus location of rank-2 (rank-1) across the whole trial.

Subspace decomposition of unit and LFP activity

Targeted dimensionality reduction (TDR)[62] was performed to identify low-dimensional entry and memory subspaces capturing variance related to the stimulus location of each rank (Figures 2E–2I and S2D–S2F). To construct the TDR space, multivariate linear regression was applied to build generalized linear models (GLM), thereby estimating how the stimulus location in different ranks affected the responses of each unit (power of each channel). The neural data were first averaged across the given time windows, and the activity of a given neuron/LFP channel was centered by subtracting its mean response from all trials. To determine the memory subspace, the neural activity was averaged from the last 250 ms of the post-S2 delay period, and the time windows 250 ms post stimuli onset were used for entry subspace. Since the target location *l* is a categorical variable, dummy coding (using x-1 dimensions to describe x categories) was used to represent the variables in the linear model. The z-scored responses of unit (channel) *n* as a linear combination of these two ranks could be described as follows:

$$\boldsymbol{R_n}(i) = \sum_{1}^{r}\sum_{1}^{5}\beta_{\boldsymbol{n}}(r,l)\boldsymbol{x_r}(l) + \beta_{\boldsymbol{n}}(0)$$

Where $R_n(i)$ is the z-scored response of unit/channel *n* on trial *i*, and $\boldsymbol{x_r}(l)$ is the stimulation location in rank-*r*. The regression coefficient $\beta_{\boldsymbol{n}}(r,l)$ estimates how much trial-by-trial firing rate of unit/channel *n* depends on the corresponding rank.

To identify the subspace of each rank, the regression coefficients $\beta(r,l)$ (dimension: neuron/channel number $n \times 5$) were used to represent the low dimensional subspaces containing the most response variance at different ranks. For each rank (e.g. $\beta(r,l), r = 1, l = 1,\ldots,5$), principal component analysis was performed to identify the first two axes that captured the most response variance. The rank-r subspace is:

$$\boldsymbol{W_r} = \beta_{\boldsymbol{n}}(r,l)\,\boldsymbol{U_2}$$

Where U_2 holds the first two dimensions of the PCA projection matrix.

The orthogonal subspaces of different ranks were obtained by orthogonalizing the two-dimension subspaces of all ranks with the QR-decomposition:

$$\boldsymbol{W} = \boldsymbol{QR}$$

Where $\boldsymbol{W}$ is a matrix whose columns correspond to the dimensions of all rank subspaces, $\boldsymbol{Q}$ is an orthogonal matrix, and $\boldsymbol{R}$ is an upper triangular matrix. The first two columns of Q ($\boldsymbol{W_1^{\perp}}$) correspond to the orthogonalized axes of rank-1 subspace, while the third and forth columns of Q ($\boldsymbol{W_2^{\perp}}$) correspond to the orthogonalized axes of rank-2 subspace.

The TDR analysis is similar for determining entry subspaces, with several changes to the GLM design. To capture the variance of target stimulus identity while ignoring the rank information, neural activity of both ranks (0-250 ms after rank-1 and rank-2 stimuli onsets) were grouped. Thus, the linear model is:

$$\boldsymbol{R_n}(i) = \sum_{1}^{I} \boldsymbol{\beta_n}(I)\boldsymbol{x}(I) + \boldsymbol{\beta_n}(0)$$

Where $\boldsymbol{x}(I)$ is the stimulus location of rank-1 and rank-2 in all trials with dimension of (trial number * 2) × 5, and $\boldsymbol{R_n}(i)$ is the neural activity of each neuron/channel.

To evaluate the variance explained by each subspace, we performed SVM decoding on the data projected onto the memory and entry subspaces. The overall decoding accuracy was taken as the mean across all recording sessions/days. The chance level was estimated from the shuffled data of the corresponding time. 99% confidence interval of the shuffled data was determined by averaging the confidence intervals from all days. Then, the statistical significance for decoding accuracy was determined by comparing mean decoding accuracy from the original data of each time window against the accuracy from the shuffled data of the corresponding time.

To determine the contribution of each neuron/channel to a given subspace, the corresponding vector of the original neuron/channel axis was projected onto a given rank subspace. The norm of the projection vector on the rank subspace measured the degree of neuron/channel $\boldsymbol{n}$'s contribution to the rank subspace:

$$\boldsymbol{fw_n} = \boldsymbol{norm}(\boldsymbol{W}^{\perp}(\boldsymbol{n}))$$

Where $\boldsymbol{fw_n}$ is the feature weight of neuron/channel $\boldsymbol{n}$. $\boldsymbol{W}^{\perp}(\boldsymbol{n})$ is the $\boldsymbol{n}$-th projection vector corresponding to the neuron/channel $\boldsymbol{n}$. TDR subspace feature weights could then serve as summaries of how much individual theta channels and single neurons contributed to higher-dimensional neural population dynamics; the degree to which theta and unit feature weights for subspaces of similar nature could share channel origins could then describe also their functional/anatomical overlap or the lack thereof (e.g., Figure 3C).

Comparing distributions of subspace feature weights

Mean (over sessions) feature weights of theta entry, spike entry, and spike memory subspaces were plotted onto 2D projection maps of the microelectrode array grid in anatomical space (defined by lateral-medial and anterior-posterior axes) to obtain spatial distributions (Figures S4A and S4C). For better evaluation of the true anatomical distributions of peak mean feature weight locations, 1,000 random resamples were taken over the original mean feature weights (40 locations per resample, computing mean location of top 4 within) to generate 2D spatial distributions of estimated peak locations and the corresponding 1D marginal histograms (Figures 3C and S4B). Statistical evaluations of marginal histogram center location differences were performed via two-sample Wilcoxon rank sum tests, with effect sizes computed accordingly (test metric divided by square root of sample size).[63]

Percent explained variance (PEV) analysis

The percentage of explained variance (PEV) was used to evaluate the degree to which the variance of a single neuron/LFP channel can be used to explain visuospatial stimulus locations. Generally, PEV can be expressed as a statistical value of η^2, i.e., the variance ratio between groups to the total variance. As the statistical value of η^2 has a strong positive bias for a small sample, the unbiased ω^2 statistical value (ωPEV) was used.[64,65] The ωPEV was calculated in each time bin to characterize the temporal dynamics of neural information of stimulation location. The baseline for each ordinal rank was defined as the period of 0-250 ms before that rank's stimulus onset. Permutation tests using 1,000 random resamplings of trials were then performed to determine whether ωPEV of each time bin significantly differs from the baseline ωPEV. Only time periods that remained significant for three consecutive time windows were marked as having significant ωPEV, i.e. encoding information regarding stimulus location (Figures 2A and S1E).

Spike-field coherence (SFC)

To evaluate the coupling between theta activity and neuronal spiking during the encoding period of sequence working memory tasks, cross-channel SFC (both magnitude and overall preferred phase) was computed using the Chronux toolbox.[53] For each ordinal rank, 0-500 ms (except for Figure 4C where 500 ms size, 50 ms step sliding windows were taken) post-stimulus-onset was chosen to build a time-by-trial LFP matrix and a trial-wise spike train set for every channel pair: one channel as the source of LFP from 2-140 Hz with shifting 4 Hz bins (2 Hz step, for Figures 4B and S4D) or from the theta band (4-8 Hz, for other analyses), the other as the source of spikes. Channel x unit pairings are unique; rare cases where multiple units were recorded by the same channel had been excluded

from this study (see ''Electrophysiological recording and data preprocessing'' subsection above). To obtain an estimate of the theta-range spectrum over a given spike train, for instance, the time-bandwidth product was set to 3, using 5 tapers spanning 2-10 Hz. The multi-taper estimate of spike spectrum $S(f)$ at frequencies f across all T taper functions is usually computed as:

$$S(f) = \frac{1}{T}\sum_{t=0}^{T-1} E_T^D(f)$$

with E_T^D representing eigenspectral estimates obtained from individual taper functions over D samples. Given the limited post-stimulus time window, a single taper equivalent to a Hamming window was chosen from the multi-taper estimate that could best cover the range of 4-8 Hz. SFC could then be computed for each channel pair as coherence magnitude between the chosen spectral estimate of the current spike train set and the corresponding LFP, yielding a square matrix of trial-average SFC at each ordinal rank, with rows for theta-origin and columns for spike-origin. Sums across columns, denoted as SFC strength (i.e., higher value indicating local unit being influenced by more widespread theta activity), would thereby allow evaluation of selective association between SFC and different neural subspaces on a channel-wise basis.

Linear models for decoding weights and SFC

Linear regression models were built to examine the relationships among unit/LFP TDR subspace feature weights (i.e. mean β values over all trained linear support vector machine classifiers and across the time period used for subspace decomposition) and SFC strength. Modeling SFC strength with spike memory subspace feature weights to establish theta modulation as a putative selective gating mechanism required multiple linear regression models (data concatenated across days; spike feature weights normalized by z-score before concatenation) where all ordinal ranks' SFC strength were set as predictor variables, with each rank's memory feature weights as response variable. To better illustrate rank-selective associations between spike subspace and SFC strength, permutation tests with 1,000 sets of channel identity-shuffled feature weights were performed to compare each model coefficient against a random distribution (Figures 4, 5, S4F, S5A–S5C, S5G, and S5H). Models using three-item task data were produced via stepwise regression (Table S1) with linear terms and with SSE (p-value for an F-test of the change in the sum of squared error) as criterion to add (p <= 0.05) or remove (p >= 0.1) terms.