

# Sensory encoding and memory retrieval are coordinated with propagating waves in the human brain

Received: 20 February 2025

Accepted: 19 January 2026

Cite this article as: Yang, Y., Leopold, D.A., Duyn, J.H. *et al.* Sensory encoding and memory retrieval are coordinated with propagating waves in the human brain. *Nat Commun* (2026). <https://doi.org/10.1038/s41467-026-69068-x>

Yifan Yang, David A. Leopold, Jeff H. Duyn & Xiao Liu

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

# Sensory encoding and memory retrieval are coordinated with propagating waves in the human brain

Yifan Yang<sup>1</sup>, David A. Leopold<sup>2,3</sup>, Jeff H. Duyn<sup>4</sup>, Xiao Liu<sup>1,5\*</sup>

## Affiliations:

<sup>1</sup>Department of Biomedical Engineering, The Pennsylvania State University, University Park, PA, 16802, USA.

<sup>2</sup>Section on Cognitive Neurophysiology and Imaging, National Institute of Mental Health, Bethesda, MD, 20892, USA.

<sup>3</sup>Neurophysiology Imaging Facility, National Institutes of Health, Bethesda, MD, 20892, USA.

<sup>4</sup>Advanced MRI Section, Laboratory of Functional and Molecular Imaging, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA

<sup>5</sup>Institute for Computational and Data Sciences, The Pennsylvania State University, University Park, PA, 16802, USA.

\* Corresponding Author: Xiao Liu, PhD  
431 Chemical and Biomedical Engineering Building  
The Pennsylvania State University  
University Park, PA 16802-4400  
Tel: +1 814 863 4419  
Fax: +1 814 863 0490  
E-mail: [xxl213@psu.edu](mailto:xxl213@psu.edu)

**Keywords:** infra-slow brain waves; semantic coding; memory encoding; memory recall; global signal; brain states; hippocampal ripples.

## This file includes:

Main Text  
Figures 1 to 4

## Abstract

Complex behavior entails a balance between taking in sensory information from the environment and utilizing previously learned internal information. Studies on mice show the brain continually alternates between outward and inward cognitive modes every few seconds, accompanied by stereotyped cascade of neuronal spiking. Our analysis of large fMRI datasets revealed a similar mechanism in human. Human brain activity was punctuated every several seconds by coherent, propagating waves emerging in the exteroceptive sensorimotor regions and terminating in the interoceptive default mode network. As in mouse, these waves in human fMRI are accompanied by phase-specific enhancements in sensory information encoding and memory retrieval. These findings suggest a conserved feature of mammalian brain physiology that bears directly on the integration of sensory and mnemonic information during everyday behavior.

ARTICLE IN PRESS

## Introduction

The human brain undergoes slow, spontaneous fMRI fluctuations during rest, in the absence of external stimulation and task engagement (1, 2). While this activity has been used widely to characterize the functional connectivity between brain regions (2–5), its contribution to the normal operation of the brain has remained elusive. Two curious features of this activity that have drawn attention in recent years are its manifestation as discrete, quasi-periodic events and its spatiotemporal propagation across the brain (6–9). Recent work describes such propagation as moving from low-order sensory-motor (SM) regions to high-order default mode network (DMN) (7, 10). This traversing of the cortical hierarchy has been compared to cross-layer error back-propagation required for optimizing artificial neural networks (11, 12), raising the prospect that these waves may play a physiological role in learning and memory consolidation.

Analogous global brain dynamics have been observed at the single neuron level in the mouse. These widespread fluctuations in neuronal firing are often coupled to spontaneous behaviors such as pupil dilation and locomotion, and show distinct delays between two major clusters of neurons (13, 14). More recently, these dynamics were shown to manifest as massive spiking cascades involving ~70% of recorded neurons across the forebrain and playing out over several seconds during stationary rest (15). During both spontaneous activity and periods of visual stimulation, spiking cascades were coordinated in time with hippocampal sharp-wave ripples (SPW-Rs), a neurophysiological event known to be involved in memory functions (16). In the case of visual stimulation, each cascade cycle involved a transition from a phase of high-efficiency sensory encoding to a phase of lower encoding during heightened SPW-Rs (17). These transitions appear to reflect switches between exteroceptive sensory sampling and internal mnemonic processes over the timescale of seconds. This type of continuous alternations between external and internal processing modes has been observed across various brain systems, particularly the hippocampus, at multiple temporal scales (18, 19), and is hypothesized to play a critical role in learning (20).

One conceptually appealing possibility is that the fMRI waves in the human brain reflect the same or homologous neurophysiological processes as the spiking cascades in the mouse brain. Indeed, they share common features. For example, both phenomena manifest as quasi-period events that occur over seconds time scales, affect global forebrain activity, and couple with measurable arousal fluctuations (15, 21), as evidenced by their correlation with pupil size (10, 22–24). In the absence of external stimulation, fMRI waves in the human brain propagate between two sets of brain networks showing opposite responses to cognitive tasks (25–27). Similarly, the spiking cascade sequences in mice involves the interplay between two groups of neurons with opposite activity modulations during locomotion (15). The hippocampus appears to be central to this large-scale coordination of activity. When hippocampal SPW-Rs were measured together with concurrent fMRI in the monkey, they were synchronized with fMRI changes across the brain (28, 29). Analysis of this synchronization revealed distinct delays between sensory/motor and higher-order regions that suggest cross-hierarchy propagation (28, 29). Nevertheless, it remains unknown whether propagating fMRI events are the macroscopic counterpart of neural firing cascades. More importantly, it is also unclear whether fMRI waves, like neural firing cascades, continue to occur during stimulation and relate to sensory and memory functions during wakefulness.



In the present study, we analyzed multiple human fMRI and mouse neuronal recording datasets to address this topic. We found that, like the spiking cascades in the mouse, the propagating fMRI waves in the human brain persisted during the performance of a visual memory task. The fMRI wave cycle was also similarly marked by alternating phases of stronger sensory information encoding and greater efficiency in memory retrieval. The cascades and fMRI waves were similarly synchronized to pupil dilations in humans and mice, suggesting a shared neuromodulatory basis. These findings thus demonstrate that similar, internally generated physiological cycles are coordinated with switches between exteroceptive sensory sampling and internal mnemonic processes in the human and mouse brain, suggesting an evolutionarily conserved principle of mammalian forebrain function.

## Results

### *Fluctuating arousal entrains brain-wide events across the mouse and human forebrain*

Pupil diameter is a surrogate signal for fluctuating arousal that is readily measured in both human and mouse subjects during rest (30, 31). In both species, we observed that the dynamic changes in pupil diameter were matched to the occurrence of brain-wide events, thus providing a means to compare spiking cascades and fMRI waves.

In the mouse, pupil size fluctuations, indicative of changes in arousal state, were prominent during periods of immobility, with or without visual stimulation, as evident in data from the Allen Institute Visual Coding project (32) (Fig. 1A and 1B). Across the brain, we found that pupil dilations coincide with moments of widespread spiking events, in which neurons fire sequentially in reproducible patterns (Fig. 1B, C, red arrows). The similar dynamics were derived previously without using pupil data and described as brain-wide spiking cascades (15). We repeated the same analysis on a two-photon calcium imaging dataset (33) and another large-scale Neuropixel dataset with broader coverage of mouse brain (34), revealing that these pupil-associated cascades span across widespread brain regions and involve multiple neuron subtypes (Fig. S1 and S2).

In the human, we similarly found that pupil size changes are associated with highly-structured resting-state fMRI changes across the brain (Fig. 1D and 1E). For this analysis, we used the Human Connectome Project (HCP) 7T dataset (35). Alignment of fMRI time courses across the cortex to pupil dilation onset revealed a spatiotemporal sequence progressing along a principal gradient (PG) direction (Fig. 1F and S3; Supplementary Movie 1), which approximates the cortical hierarchy gradient (36). These events were manifest as infra-slow (multi-second) waves moving gradually from SM to DMN regions. These SM-to-DMN waves have been derived previously without using pupil data (7, 10). The cortical changes were accompanied by corresponding thalamic progression from posterior/lateral sensorimotor nuclei to anterior/medial progression limbic nuclei (Fig. 1G and S3F). Therefore, spontaneous pupil dilation events during immobile rest are synchronized with sequential brain dynamics of global involvement, observed as spiking cascades in mice and propagating fMRI waves in humans. Simultaneous EEG-fMRI recordings also revealed associated changes in delta-band (1-4 Hz) activity across the fMRI wave cycle, similar to those observed across the cascade cycle (Fig. S4). The duration of the waves appears to be constant (~11s) and independent to the brain size (Fig. S5).

While the function of these brain-wide events is poorly understood, evidence in the mouse suggests that the spiking cascades mediate alternation between periods of external stimulus encoding and internal memory operation (17). Might the fMRI waves in humans likewise be linked to the switch between exteroceptive and interoceptive modes of brain function? To address this question, we investigated the occurrence of spontaneous propagating fMRI waves as human subjects performed a cognitive task involving memory. Specifically, we asked whether the sensory encoding of stimuli and successful memory retrieval performance varied as a function of these spontaneous events.

### ***Visual stimulus encoding predicts subsequent memory function.***

To systematically investigate the role of propagating SM-to-DMN waves in human cognition, we first needed to establish a reliable method for evaluating the encoding of visual stimuli from fMRI responses across the brain. We developed a method to do this using the Natural Scenes Dataset (NSD) (37), in which a series of 10,000 captioned natural images were shown, in the form of 4-s trials, to each of 8 subjects with each image being presented three times over 40 scan sessions on different days. For each trial, the subjects needed to indicate whether they had seen the stimulus before (Fig. 2E).

To quantify the level of sensory stimulus encoding, we developed a deep learning model to decode semantic information of each image stimulus based on fMRI responses it evoked (Fig. 2A and S6A). The model comprised an fMRI encoder, which extracted latent representation from the fMRI responses, i.e., the fMRI embeddings, and a caption decoder (38), which translated the fMRI embeddings into descriptive text captions. The fMRI encoder was trained to align the fMRI embeddings with the contrastive language-image pre-training (CLIP) embedding space (39) through contrastive learning (40). We then derived a composite semantic similarity (CSS) score to quantify the semantic similarity between the fMRI-decoded caption and the original caption. An image trial is considered correctly decoded if the generated caption shows significant ( $p < 0.05$ ) similarity compared to a null distribution created from randomly sampled captions (see Methods for details). The performance of the decoding model thus serves as a measure of the accuracy of brain encoding of semantic information.

Our deep learning model successfully decoded the semantic information associated with the visual stimuli based on the fMRI responses they evoked. The representation similarity analysis confirmed the alignment between fMRI and caption embeddings (Figs. 2B, S6B, and S6C), and the fMRI embeddings after training are organized as distinct categories in a low-dimensional space (Fig. 2C) (41). The trained model generated correctly decoded captions — those significantly similar to the ground truth captions — in  $33.0 \pm 4.2\%$  (mean  $\pm$  SD) of trials (see Fig. 2D and Fig. S7A for examples), compared to the 5% chance-level performance of the untrained model (Fig. 2F,  $p = 2.8 \times 10^{-7}$ ; and Fig. S7B). In the context of the cognitive task, the semantic encoding accuracy faithfully predicted subsequent memory performance: a higher CSS score at the first appearance of an image stimulus led to a higher rate of correctly recalling it at its second repeat (Fig. 2G and S8). This relationship held true across subjects, with subsequent memory performance significantly better for trials with correct initial encoding compared to those without (Fig. 2H). Given this trained model, it was next possible to evaluate whether the occurrence of spontaneous propagating fMRI waves might bear on the quality of stimulus encoding, subsequent memory recall, or both.

### *Alternating stimulus encoding versus memory recall during propagating fMRI waves*

To explore the potential relevance of propagating fMRI waves for encoding and memory performance, we first established their presence during the cognitive task. These waves were identified directly from task fMRI data without using pupil data (7) (Fig. 3A). Similar to the resting state, propagating waves remained closely tied to pupil diameter fluctuations, which were affected to a lesser extent by task events (Fig. S9A). The duration of the SM-to-DMN waves (~10-15 seconds) is much longer than the task trials (4 seconds), and their occurrence and relationship to pupil fluctuations are dissociated from the structure of the concurrent cognitive task (Fig. S9B).

To evaluate the potential influence of these propagating fMRI waves on sensory encoding efficiency, we used the fMRI deep learning-based decoding method described above to assess the quality of encoding for each stimulus presentation. We found that the accuracy of such encoding varied systematically across the SM-to-DMN propagation cycle (Fig. 3C, red trace). Accounting for hemodynamic delays (see Methods), the stimulus encoding was strongest at the SM-activated phase of the propagating wave, showing a significant improvement to the DMN-activated phase ( $p = 9.8 \times 10^{-5}$ ,  $N = 8$ , two-sided pair-wise t-test). Similar results were consistently reproducible across various semantic similarity metrics (Fig. S10A).

We also used human memory performance to assess how fMRI waves affected both stimulus encoding and memory recall. For encoding, accurate memory of a given stimulus during its second appearance (Repeat #2 in Fig. 3B) was taken as a measure of strong encoding during the initial presentation (Repeat #1 in Fig. 3B), where the wave dynamics were examined. We found that the strength of encoding (i.e. subsequent memory) was closely related to the phase of the SM-to-DMN wave at the time of stimulus presentation. Stimuli presented during the SM-activated phase had a higher efficacy of encoding than those presented during the DMN-activated phase ( $p = 2.2 \times 10^{-3}$ ,  $N = 8$ , two-sided pair-wise t-test; Fig. 3C, purple trace), thus matching the fMRI deep learning-based measure of stimulus encoding described just above (Fig. 3C, red trace). By contrast, evaluation of recall performance, which was estimated at the 2<sup>nd</sup> and 3<sup>rd</sup> presentations of a stimulus along with fMRI wave dynamics, revealed a significant peak performance later in the wave cycle, when the subject recall coincided with the DMN-activated phase ( $p = 1.2 \times 10^{-3}$ ,  $N = 8$ , two-sided pair-wise t-test; Fig. 3C, green trace). These results were similar for both short-term (i.e., within the same day) and long-term (i.e., across days) memory types (Fig. S10B and S10C).

The modulation of stimulus encoding and memory recall was further supported by subcortical changes during propagating fMRI waves. The hippocampus was activated at the late wave phase but its activation level was only weakly associated with memory recall performance (Fig. S11C and S11D). Moreover, anterior and medial thalamic regions, which are associated with memory and limbic functions (42, 43), are activated more at the late phase, whereas posterior thalamic regions most involved in sensory processing are activated earlier in the wave cycle (Fig. S11E).

These cyclic modulations of stimulus encoding and memory recall in humans resembled analogous observations in mice during different phases of the spiking cascades (17) (Fig. 3D). Specifically, the SM-activated phase of the fMRI wave matched a period within the cascade cycle (0–0.5 sec) of improved stimulus encoding, whereas the DMN-activated phase aligned with a different period within the cascade cycle (0.5–2 seconds) of increasing hippocampal SWP-R rate, which was also associated with pupil dilation (Fig. 3C and 3D). While the hippocampal SWP-R rate and memory

performance are clearly different measures, they may point to similar processes that transpire during more introspective modes of brain activity, commonly associated with activity of the DMN (44, 45).

### ***Visual semantic information coding in multiple brain regions is similarly modulated by the SM-to-DMN wave cycle***

Repeating the semantic decoding using only regional fMRI data suggested that the semantic information was encoded across a wide range of brain regions, with the highest encoding accuracy observed in the visual cortex (Fig. 4A and 4B). Importantly, the fMRI encoding accuracy in all individual regions was systematically modulated over the SM-to-DMN wave cycle (Fig. 4C and Fig. S12) in a similar way as the whole-brain finding (Fig. 3C). Interestingly, the DMN, particularly its C division that encompasses the hippocampal complex and adjacent to visual association areas, exhibited its peak encoding accuracy at the SM-activated phase, while its activation was relatively low, suggesting a dissociation between sensory encoding and regional activation level. These region-specific results on visual semantic encoding are consistent with those on cascade-dependent visual encoding (17), further suggesting that the spiking cascades and cross-hierarchy waves represent the same underlying neurophysiological process present in mice and humans, and possibly conserved across mammals more broadly.

## **Discussion**

Here we showed that infra-slow ( $<0.1$  Hz) fMRI waves propagating over the human cortical surface are associated with a counter-acting modulation of encoding and retrieval of information conferred by visual stimuli, resembling those previously tied to spiking cascades in mice. First, we analyzed electrophysiological and fMRI measures of brain activity, demonstrating that spontaneous pupil dilations are similarly accompanied by spiking cascade dynamics in mice and SM-to-DMN propagating waves in humans, thereby unifying these two types of infra-slow ( $<0.1$ Hz) global brain activity across different spatial scales and species. We then assessed the semantic encoding of visual stimuli using a CLIP-based deep learning model, finding that the SM-to-DMN propagating waves continued to occur during task performance and were associated with counter-valent modulation in both encoding and retrieval of the stimulus content. The encoding of semantic information and memory peaked at the early phase of SM-activation, whereas memory retrieval accuracy reached the maximum at the DMN-activated phase. Together with previous findings from mice, these results suggested that the highly structured infra-slow global brain activity may serve as an evolutionarily conserved mechanism by which the brain orchestrates the execution of exteroceptive sensory sampling and internal mnemonic processes on the timescale of seconds.

The brain's response to identical sensory stimuli is known to vary over time even on the timescale of seconds. Previous studies have shown how pre-stimulus ongoing activity and arousal state may contribute to this variability (13, 31, 46–52). Our findings align with and extend these previous reports. Leveraging recent advances in deep learning techniques, our study goes beyond a simple quantification of response amplitude (2, 49, 50) and assesses the accuracy of the brain's encoding of semantic information. Importantly, most previous studies have presumed that ongoing brain activity and changes in arousal occur spontaneously and randomly. As a result, much focus has

been on the response modulation of ongoing activity that is temporally locked (prior to the stimulus) and spatially restricted (confined to the same local brain region). In contrast, we consider the effects of internal fluctuation in the context of highly structured brain dynamics (i.e., the spiking cascade or propagating wave) involving the large-scale coordination of activity. The initiation of these recurring global brain events is independent of visual stimulation and memory tasks, and they are associated with sensory processing quasi-periodically in a continuous and persistent way.

We further found that memory retrieval was modulated over the SM-to-DMN wave cycle in a manner opposite to that of stimulus encoding, matching our previously observed counter-modulation of hippocampal SPW-R rate and visual encoding (17). This previous study did not, however, identify specific memory functions or other cognitive operations associated with SPW-Rs during the task, since SPW-Rs are usually observed during rest and sleep and often linked to offline memory consolidation (29, 53, 54). By comparing our human study results with these prior findings in mice, we found a correspondence between the cascade phase of high SPW-R rate to the wave phase of fMRI DMN activation, which is associated with a better performance in memory retrieval. This observation largely agrees with a series of recent studies on different species that linked SPW-Rs during tasks to memory retrieval (55–57), as well as the marked fMRI DMN activations (44, 45, 58, 59).

Notably, we applied a global 6-sec shift to the fMRI BOLD signals to infer underlying network activity with accounting for hemodynamic delays (Fig. 3C and 4C). A 1-2 second deviation of this modeled delay from the true value would be minimal relative to the duration of fMRI waves and, therefore, is not expected to affect our interpretation of their temporal relationship with modulations in sensory and memory functions. Without this time shift, however, the periods of stronger semantic encoding and poorer memory retrieval would correspond to the baseline preceding the SM-to-DMN waves, rather than to the SM-activated phase.

The observed modulation of sensory and memory functions over the cascade/wave cycle may be associated with a change in the direction of information flow, particularly between the cortex and hippocampus. Memory retrieval during tasks and memory consolidation during rest and sleep likely require information flow from the hippocampus to the cortex, whereas the encoding of sensory information and memory would be facilitated by a reversed flow (29, 60, 61). Thus, the cascade/wave phases optimized for sensory encoding and memory retrieval may be dominated by opposite directions of information transmission, which may rely on distinct spatial gradients in activation level. In fMRI, such activation gradients are obvious for SM-to-DMN waves with dominant SM or DMN activation at different phases. This is less clear for cascades, since the negative- and positive-delay neurons were found in all recorded brain regions (15). However, the hippocampal regions, especially CA1 and the dentate gyrus (DG), contain a much higher number of negative-delay neurons compared to any other areas, including all visual areas, whereas the thalamus has the least. The fMRI signals may arise from the summation of activity of these two types of neurons, and the apparent wave propagation may relate to their distinct composition across different brain regions and across cortical hierarchies. Thus, the activation gradient between these two neuronal groups can be translated into spatial gradients between the hippocampus, cortex, and thalamus. We hypothesize that these gradients, alternating on the multi-second scale, determine the dominant direction of information flow, which itself occurs on much faster (millisecond)

timescales. This hypothesis remains to be tested by future studies. It is worth noting that artificial neural networks also feature alternating forward/backward information flows across hierarchical layers during training (11, 12), which may thus represent a mechanism essential to the learning of all connection-based intelligence systems.

The cascade and wave dynamics reported here might represent a fundamental mechanism by which the brain coordinates the opposing operations of exteroceptive sensory sampling and internal mnemonic processes. The switching between these processes over seconds may establish a dynamic balance essential for optimizing cognitive/mental processes of the same seconds timescale (62–65), likely achieved under states of intermediate arousal (66, 67). Highly aroused states could break this balance by terminating this infra-slow global dynamic. Locomotion, presumably associated with heightened arousal, has been found to replace cascade dynamics with sustained firing of the positive-delay neurons (17) that are expected to enhance sensory and memory encoding but impede memory retrieval (68, 69). Toward the other end of this spectrum, during drowsiness, the infra-slow global dynamic may prolong the memory consolidation phase whereas hinder encodings. The SM-to-DMN waves have been found to occur more frequently during various sleep stages and be associated with learning-related features (i.e., the rapid eye movements and possibly Ponto-Geniculo-Occipital (PGO) waves) during rapid eye movement (REM) sleep (70). Though not directly focused on the cascade and waves, recent studies convergently point out an essential role of infra-slow neural dynamics in learning and memory. In addition to hippocampal SPW-Rs, infra-slow dynamics have been found to simultaneously coordinate the density of sleep spindles, an electrophysiological feature that has relevance for learning and memory (71, 72). Importantly, the amplitude of infra-slow dynamics during sleep, defined through spindle density and cardiac rate, is not only correlated with memory performance on the subsequent day (73), but optogenetically enhancing it also leads to improved memory (74). Similar to the modulation of brain activity during cascades and waves described here, such spindle-based infra-slow dynamic alternates between an offline phase, characterized by higher spindle and hippocampal SPW-Rs rates with low arousal, and an online phase, marked by lower spindle and ripple rates with higher arousal and susceptibility to external stimulation (73).

The neural mechanisms underlying these global brain dynamics remain unclear. However, they are unlikely to be mediated purely through corticocortical axon conduction due to the highly mismatched speed. Instead, the neuromodulatory systems, particularly the cholinergic system, may play a crucial role in the general of these global brain dynamics. Consistent with the associated arousal modulation, the fMRI waves are accompanied by subcortical de-activation specifically at the basal forebrain nucleus basalis (NB) and brainstem arousal-relating nuclei, including the locus coeruleus (75). Pharmaceutical deactivating the NB on one side of monkey brains suppressed the global mean BOLD (gBOLD), whose peaks the SM-to-DMN waves are manifested as, on the ipsilateral side (76). The global brain dynamics might originate from interactions among neurons at this subcortical region, which are then broadcasted to the cortex and other brain regions through widespread cholinergic projections (77–79) to affect the dynamics of the two distinct groups of neurons, i.e., the negative- and positive-delay neurons. These neuron-level dynamics, i.e., the spiking cascades, are then translated to the apparent fMRI waves across hierarchies due to the systematic modulation of their composition along this same direction. However, the validation of this hypothesis would require future studies with neural recordings of distinct scales.

The SM-to-DMN propagating wave and its effect on sensory and memory functions may offer explanations for some previous task fMRI observations. Graph-theory metrics based on fMRI connectivity/correlations, such as cartography and network flexibility, have been used to quantify brain dynamics and found associations with various cognitive components, particularly learning (64, 80, 81). Most of these metrics focused on assessing the integration and segregation of the large-scale networks, which are expected to be profoundly affected by the presence of the global SM-to-DMN waves. Thus, the waves could be an important contributor to these metrics of network dynamics. Another related phenomenon is the so-called encoding/retrieval flip, in which the de-activation and activation of the posteromedial cortex, a key component of DMN, are preferentially associated with successful memory encoding and retrieval respectively (82–84). This phenomenon can be explained by our finding that memory encoding and recall were oppositely modulated over the wave cycle with distinct DMN activations. Importantly, the present study expands this early research by incorporating the previous findings into the framework of highly structured cross-hierarchy propagating waves, which persist under various brain conditions beyond tasks.

Finally, the SM-to-DMN waves may also relate to memory dysfunction in Alzheimer's disease (AD). The gBOLD signal has been repeatedly linked to various AD pathologies (85–87). The gBOLD peaks (also SM-to-DMN waves (21)) have been found to be coupled by strong cerebrospinal fluid (CSF) movements, known to be essential for peri-vascular waste clearance (88–90). The strength of this gBOLD-CSF coupling is indeed associated with the accumulation of amyloid-beta and tau (86, 87). Particularly, the failure of the SM-to-DMN waves to reach the DMN appeared to account for preferential amyloid-beta accumulation at these higher-order regions at the early stage of AD (86). The link between the fMRI waves and waste clearance may be at least partly attributed to associated non-neural physiological modulations, particularly infra-slow vasomotion and CSF dynamics (21, 89, 91, 92). Besides the toxic protein accumulation, AD also features dysfunctions in memory and subcortical neuromodulatory systems (93–96), both of which are linked to the cascades and global waves (10, 15). Thus, changes in this infra-slow global dynamic may also relate to the dysfunction of the memory and arousal systems in AD.

## Methods

### *Datasets*

**Allen Institute Visual Coding Neuropixel Dataset:** The dataset comprises high-density extracellular neuron recordings of 58 mice (13 females) using Neuropixel probes (32). Each mouse was implanted up to six Neuropixel probes, which targeted the primary visual cortex and five high-order visual cortical areas. The silicon probes were inserted to a depth of up to 3.5mm into the brain, enabling the recording of spiking activity within two visual thalamic nuclei, i.e., the lateral posterior nucleus (LP) and the lateral geniculate nucleus (LGN), as well as other regions that the probes traversed, such as the hippocampus.

**Allen Institute Visual Coding Two-photon Calcium Imaging Dataset:** The dataset comprises single-neuron recordings from the mouse visual cortex, obtained through 2-photon fluorescence imaging. Utilizing transgenic tools, these recordings specifically targeted the activities of distinct populations of Cre-defined neurons. The dataset includes a total of 63,251 neurons from 14

different transgenic lines, covering 6 cortical areas and 4 cortical layers. For our study, we focused on two Cre lines (Cux2-CreERT2 and Emx1-IRES-Cre) that had the highest number of neurons recorded. Further details can be found in (33).

**UCL Mice Neuropixel Recording Dataset:** The UCL mice dataset includes recordings from approximately 30,000 neurons across 43 brain regions in mice, utilizing Neuropixels probes to cover the entire left hemisphere. In each mouse, two or three probes were inserted simultaneously, allowing for the concurrent recording of hundreds of neurons during each session. The study comprised 92 probe insertions across 39 sessions from 10 mice (6 females), with an average of approximately 747 neurons recorded per session. Additional details are available in (34).

**Human Connectome Project (HCP):** We utilized the WU-UMinn HCP 7T dataset, a subset of the HCP S1200 release (35), comprising 7T fMRI data from 184 subjects (112 females) within the age range of 22 to 35. Our analysis focused on two 15-minute, eyes-open resting-state fMRI sessions, with repetition time (TR) of 1s and 1.6 mm isotropic voxels. Simultaneous eye tracking was conducted using an EyeLink device with a sampling rate of 1000 Hz. The resting-state HCP fMRI volumetric data were preprocessed using the minimal preprocessing pipeline (97) and artifacts were further removed with ICA-FIX denoising (98). fMRI data were then spatially smoothed with 2mm Full Width at Half Maximum (FWHM) Gaussian kernel and temporal filtering within a bandpass range of 0.001-0.15Hz. Finally, signal from each voxel was standardized by subtracting the mean and dividing by the standard deviation.

**Natural Scenes Dataset (NSD):** The NSD includes whole-brain 7T fMRI scans with a repetition time (TR) of 1.6s and 1.8mm isotropic voxels, conducted during a visual memory task (37). This dataset involved eight human participants (6 females) who were shown between 9000 and 10000 natural scene images. Each image was presented three times, resulting in a total of 22000 to 30000 trials across a span of one year. For every trial, as participants viewed an image stimulus, they were required to report whether they perceived the image as novel. The NSD fMRI data underwent initial preprocessing to correct for head motion, EPI distortion, gradient nonlinearities, and alignment across scan sessions. Analyses of fMRI data for each subject were performed in the subject-native space. For the fMRI time series analysis, i.e. propagation, nuisance parameters such as motion, white matter, and cerebrospinal fluid (CSF) signals were regressed out from the volumetric fMRI data, which were then smoothed spatially with 2mm FWHM Gaussian kernel and temporally with bandpass filtering of 0.001-0.15Hz. Each voxel's signal was then standardized to have zero mean and unit standard deviation. For analyzing the hemodynamic response of single trials (inputs of the decoding models), GLMdenoise—a generalized linear model (GLM) approach was used to provide estimates of the BOLD amplitude while reducing noise by integrating nuisance regressor (37, 99).

**Simultaneous EEG-fMRI Resting-state Dataset:** Resting-state fMRI data were collected simultaneously with electroencephalography (EEG) for 27 subjects (14 females, average age:  $22.1 \pm 3.1$  years). For our analysis, we only utilized the 10-minute resting-state scan. The fMRI imaging data were acquired using a 3T scanner, with a repetition time (TR) of 2.1s and 3mm isotropic voxels. The EEG data were gathered using a 32-channel MR-compatible EEG system, with a recording sampling rate of 5000Hz. Additional details can be found in (21). The resting-state fMRI BOLD data was preprocessed using script from the 1000 Functional Connectomes



Project with slight modification (100). Nuisance parameters, including linear and quadratic trends, motion parameters, white matter, and CSF signals, were regressed from the fMRI data. The volumetric data then was smoothed spatially with 2mm FWHM Gaussian kernel, and temporally filtered with a bandpass range of 0.001-0.15Hz.

For every fMRI dataset analyzed, we utilized the Schaefer 400 Parcellations (Yeo-17 Network version)(101) to obtain cortical signals from the volumetric fMRI data. This was accomplished by averaging the standardized voxel signals located within each parcel. Similarly, signals from thalamic nuclei/regions were extracted utilizing the Morel Atlas (102), and signals from the brainstem nuclei, part of the ascending arousal network, using the Harvard AAN atlas (103).

Informed written consent was obtained from all participants in the human EEG-fMRI dataset. Data analysis adhered to usage agreements, human data collection was conducted in compliance with protocols approved by the Institutional Review Board at Pennsylvania State University (protocol numbers: STUDY00005969 and STUDY00015305).

### ***Pupil Size***

Pupil areas recorded for humans were obtained using the EyeLink device, with raw data sampled at 2000Hz for the NSD dataset and 1000Hz for the HCP dataset. We converted the pupil area into pupil diameters and then down sampled to 50Hz. Missing pupil data, resulting from false detections or eye blinks, were interpolated using data from the nearest time points. Subsequently, we synchronized the pupil data with each TR of the concurrent fMRI signal. The periods with significant missing data (more than 50%) were removed from subsequent analyses.

Pupil areas recorded for mice were captured using cameras, with a sampling rate of 100Hz for the UCL Neuropixel dataset and 30Hz for both the Allen Institute Neuropixel and two-photon datasets. We converted the pupil area into diameters and then resampled the data to a uniform rate of 30Hz. Missing pupil data were then interpolated using the nearest time points to ensure continuity in the dataset.

To identify dilation events within infra-slow pupil fluctuations, we utilized a low-pass filter on the pupil size time series, setting a cut-off frequency at 0.15Hz for human datasets and 0.3Hz for mouse datasets. Pupil dilation events were defined as the periods where dilation lasted for at least 1 second in the filtered pupil size data.

### ***EEG Analysis***

The EEG data were preprocessed to remove the gradient and ballistocardiogram artifacts from each channel, utilizing algorithms detailed in (104). Following this, the data were subjected to low-pass filtering with a cut-off frequency of 125Hz. Pulse artifacts were removed through independent component analysis (ICA), and signals were corrected for distortions to account for distortions attributable to head motion (105). More comprehensive description of the EEG signal preprocessing can be found in (21).

Delta-band power for each channel was computed by first applying a band-pass filter within the 1-4Hz range and then calculating the amplitude of the Hilbert-transformed signal. Then the power for each channel is individually normalized by subtracting its mean and dividing by its standard deviation. The averaged delta power is obtained by taking the mean across all the recording channels.

### ***Local Field Potentials (LFPs) Analysis***

**Delta-band Power:** Delta power was computed for LFPs across all recorded channels. To calculate delta power, a band-pass filter (1-4 Hz) was applied to the LFP signal of each channel, followed by rectification and lowpass filtering (<0.72 Hz, corresponding to  $\pi$  cycles of the mean band-pass frequencies).

**Hippocampal Sharp Wave Ripples (SWRs):** Hippocampal sharp wave ripples (SWRs) are brief, high-frequency oscillations (110-200Hz) that can be observed in the LFP recorded from hippocampal recording sites. For ripple detection in this study, we employed an offline method (15, 106) utilizing the LFP signal (1250 Hz) captured from the hippocampal CA1 region. Ripple events was detected individually for each CA1 recording site (channel), resulting in extensively overlapping ripple detection across the channels. Detected ripple events were considered valid only if they were identified in over 40% of the CA1 channels.

### ***Semantic Decoding Model***

The semantic decoding model is designed to evaluate the semantic information contained in the stimuli-evoked BOLD responses, generating text captions that describe the image stimuli presented to the subject. The decoding model consists of two main components: an fMRI encoder and a caption decoder (Fig. 2A).

The fMRI encoder is used to extract latent representation from the BOLD response. The detailed architecture of the fMRI encoder, as shown in Fig. S6A, utilizing convolutional layers and residual connections, aims to transform the high-dimensional BOLD response into 512-dimensional fMRI embeddings. To address the challenge of fMRI data scarcity, the fMRI encoder was trained to align the fMRI embeddings with the CLIP embedding space which has been extensively pre-trained using 400 million (image, text) pairs, thus offering a rich, 512-dimensional target. Therefore, we trained the fMRI encoder in contrastive learning paradigm (39), aiming to maximize the alignment between the fMRI embeddings and the corresponding CLIP text embeddings, while minimizing the alignment with mismatched pairing. To achieve this, we use the contrastive training loss with loss function defined for  $i^{\text{th}}$  fMRI embedding  $\mathbf{Z}_i$  and  $j^{\text{th}}$  CLIP text embedding  $\mathbf{T}_j$  within a batch  $B$  as:

$$\mathcal{L}_{\text{contrast}}(\mathbf{Z}_i, \mathbf{T}_j) = -\log \frac{\exp(\cos(\mathbf{Z}_i, \mathbf{T}_j)/\tau)}{\sum_{i,j \in B, i \neq j} \exp(\cos(\mathbf{Z}_i, \mathbf{T}_j)/\tau)} \quad (1)$$

Here,  $\tau$  represents the temperature, a hyperparameter hyperparameter, and  $\cos(\cdot, \cdot)$  computes vector similarity. In addition, we also maximize the alignment between fMRI embeddings and CLIP embeddings by incorporating the cosine loss defined as

$$\mathcal{L}_{align}(\mathbf{Z}_i, \mathbf{T}_i) = \lambda_1(1 - \cos(\mathbf{Z}_i, \mathbf{T}_i)). \quad (2)$$

Therefore, the total loss is defined as the summation of alignment loss and contrastive loss:

$$\mathcal{L}_{total} = \mathcal{L}_{align} + \lambda_2 \mathcal{L}_{contrast} \quad (3)$$

where  $\lambda_1$  and  $\lambda_2$  are tuning hyperparameters. In our training regimen, we set  $\lambda_1 = 0.35$ ,  $\lambda_2 = 0.65$  and  $\tau = 0.45$ . We also incorporate dropout before the final layers, with a dropout ratio of 0.3.

We utilized a pre-trained caption decoder, DeCap (38), to generate captions from fMRI embeddings. DeCap was initially trained to produce captions using CLIP text embeddings based on large text corpus. Since the fMRI embeddings were aligned with the CLIP space through the fMRI encoder, we directly employed DeCap to decode these fMRI embeddings, for the generation of captions that describe the content of image stimuli shown to the subjects.

In our study, unless otherwise specified, we use the decoding model to decode the response of voxels within the "nsdgeneral" ROI which includes occipital regions that are generally responsive in the NSD experiment (37). For region-wise decoding analysis as shown in Fig. 4 and S12, we adopted the regions of interest (ROIs) as delineated by the Yeo-17 network (101). For each specific ROI, only the voxels that are defined by the corresponding ROI mask are considered. The responses from these selected voxels are then utilized as inputs to the decoding model to generate captions.

### ***Representation Similarity Analysis (RSA)***

We analyzed the alignment between CLIP text embeddings and fMRI embeddings through RAS. For each validation fold, we constructed representation dissimilarity matrices (RDMs) for both the fMRI and CLIP text embeddings. To facilitate visualization and interpretation of the RDMs' structure, we employed t-SNE techniques (41) to project CLIP text embeddings into a 2-dimensional space. Subsequently, we applied k-means clustering to group stimuli with minimal Euclidean distance in this 2-dimensional representation into the same cluster. The RDMs for both fMRI and CLIP text embeddings, as shown in Fig. 2A, are organized and sorted according to this clustering scheme.

The similarity between the fMRI and CLIP text RDMs is quantified using Pearson's correlation coefficient. To evaluate the statistical significance of this similarity, we followed the permutation testing approach described by Kriegeskorte et al. (107). Specifically, we generated a null distribution by randomly permuting stimulus labels—thereby shuffling the rows and columns of one of the RDMs—and calculating the correlation between the permuted and original matrices. This process was repeated 500 times to approximate the null hypothesis that the two RDMs are unrelated. The p-value was then estimated by comparing the observed correlation to this null distribution.

### ***Semantic Similarity Metrics***

To evaluate the fidelity of generated captions to their semantic content, we computed semantic similarity score between the predicted captions and ground truth captions labeled by human. This evaluation employs several established metrics widely used in computer vision and natural language processing research: Bleu\_1, Bleu\_2 (108), CIDEr (109), METEOR (110), and ROUGE\_L (111), with each of these metrics offering a different perspective on the semantic alignment between generated and ground truth text. Given the limitation inherent to each metric, we additionally defined a composite semantic metric (CSS) derived by averaging the scores from the aforementioned metrics, thereby providing a more holistic evaluation of caption fidelity.

To assess the statistical significance of the model-generated captions, for each trial/image, the semantic similarity score, for example, CSS score, is compared against a null distribution. This null distribution is constructed from the CSS scores between the ground truth caption and 1,000 randomly generated captions, which are generated by randomly sampling from the CLIP embedding space (39). Notably, this null distribution exhibits variability across different trials, reflecting the diverse complexity levels associated with the semantic content of each image stimulus. Such approach ensures the evaluations accurately reflecting the model's performance in generating semantically coherent captions and invariant to the varying degrees of semantic complexity present across trials. The same evaluative framework is applied across all aforementioned metrics to assess the accuracy of the generated captions. A trial is considered correctly decoded if the metric score is significant compared to the null distribution, using a significance level of 0.95; otherwise, it is considered incorrectly decoded. Encoding accuracy is quantified as the proportion of trials that are correctly decoded.

### ***Memory Encoding and Recall***

The NSD dataset includes visual memory tasks, where each image stimulus was presented to participants three times. During each presentation, participants are prompted to indicate whether they have previously seen the image. By dissecting the memory task performance, we derived two memory measurements for evaluating the memory function based on participant response: memory encoding and memory recall.

Memory encoding, as a proxy for measuring the sensory coding efficacy, tends to evaluate how effectively the participant encodes a novel image into memory such that the participant can correctly recognize the image as previously seen. We formally define the memory encoding accuracy for each novel image as the memory recognition accuracy at the second presentation of the image. Memory recall, on the other hand, tends to evaluate how effectively the participants is able to accurately retrieve and recognize an image previously seen. For each image, we can derive two recall accuracy, corresponding to the participant's recognition performance during the second and third presentations of the image, which is not novel to the participant.

### ***Neural Population Sensory Decoding Analysis***

To assess the sensory information encoded by neuronal populations, we analyzed spiking data obtained during natural scene image stimulation sessions from the Allen Mice Neuropixel dataset.

In these sessions, mice were subjected to passive viewing of a sequence of images, each displayed for a duration of 250 milliseconds. Strictly following the approach in (17), we defined the neural code as a population response to each displayed image and employed support vector machines (SVMs) based on neural code to decode which one of the 118 images was viewed by the mouse. The sensory efficacy is quantified by the decoding accuracy of the SVMs.

### *fMRI Infra-slow Propagating Waves*

To detect SM-to-DMN fMRI propagating waves, we adopt a template-matching approach (7). The propagations, which typically involve a majority of cortical regions, are assumed entrained within the fMRI global signal fluctuations. Accordingly, we defined set of candidate events based on the low-frequency ( $<0.15\text{Hz}$ ) components of the global signal, with the event boundaries defined by the adjacent troughs. We derived a delay profile for each candidate event and a candidate event was considered as SM-to-DMN propagation if its delay profile closely matched the principal gradient profile (PG)  $\mathbf{u}_{\text{PG}} \in \mathbb{R}^N$  (36). The degree of similarity was quantitatively assessed using Pearson's correlation coefficient, and we considered a candidate event as SM-to-DMN propagation if the correlation was significant ( $p < 0.001$ ). See Supplementary Methods for further details.

### *Neural Spiking Cascades*

In line with prior studies (15, 17), we first computed neural spike rates using 200ms time bins and identified candidate infra-slow neural events by segmenting the spiking rate data based on the troughs of the filtered global mean spike rate (low-pass, 0.25Hz). We then, using the delay-profile decomposition method, extract delay profile for each candidate event and the principal delay profile representing the predominant sequential pattern within infra-slow brain activity. A candidate event was considered as a valid cascade event when the principal delay (PD) profile matched with its delay profile with a significant correlation ( $p < 0.001$ ). Similar to (15, 17), we categorized neurons into two distinct groups based on their positioning within the PD profile: positive-delay neurons exhibiting significant PD values and negative-delay neurons characterized by significantly negative PD values ( $p < 0.001$ , one-sample t-test). The slow cascade features a sharp increase in the spiking activity of positive-delay neurons in the middle, following the definition in (15), we thus identified and defined these time points as the local peak of the first-order temporal derivative of the mean spiking time course of the positive-delay neurons. See Supplementary Methods for further details.

### *Modulation Across Propagating Wave Cycles*

To assess the fluctuation of the trial-based measurements, i.e., semantic encoding accuracy, memory encoding/decoding accuracy across the propagating waves cycle, as shown in Fig. 3C, we constructed time series  $f_{\text{acc}}(t)$  for each measurement during fMRI scan session defined as

$$f_{\text{acc}}(t) = \sum_{k=1}^n f_{\text{acc}}^{(k)}(t) \quad (4)$$

$$\text{with } f_{\text{acc}}^{(k)}(t) = \begin{cases} C_k & \text{if } T_k \leq t \leq T_k + \Delta T \\ \text{NaN} & \text{otherwise,} \end{cases}$$

where  $T_k$  is the onset time of stimulus and  $C_k$  is the accuracy for the  $k$ -th trial, and  $\Delta T$  is the time window following the stimulus onset, which was set to 2 seconds throughout our study. The NaN stands for "Not a Number" and is omitted in the summation in equation.

Therefore, by formulating these trial-based assessments as time series, they are analyzed exactly same to other continuous measurements, such as pupil diameter and delta power. For the analysis of modulation across the propagation wave cycle, we aligned and averaged these time-series metrics relative to the propagation center (defined as the global signal peak) and normalized to the change in percentage relative to baseline, defined from -21s to -11s prior to the global signal peak.

To present the temporal relationship between these trial-based measurements and the neural dynamics underlying fMRI waves, we shifted the fMRI BOLD wave backward in time by 6 seconds to account for the hemodynamic delay.

### **Predicting Subsequent Memory Performance with Semantic Encoding**

We conducted two analyses using the NSD dataset to study how initial image presentation encoding accuracy influences subsequent memory task performance. In the first analysis (Fig. 2G), stimuli from all eight participants were grouped into bins based on the percentile of the first encoding accuracy, incremented by 20%, and the memory task accuracy during the second presentation was averaged within each bin. In the second analysis (Fig. 2H), for each participant, trials were grouped based on whether they were correctly decoded. The subsequent memory accuracy for each group was then averaged and compared across groups.

### **Relationship Between Propagating Waves and Brain Size**

We examined the effect of brain size on SM-to-DMN propagating waves using the HCP-7T dataset. Brain size was estimated based on brain volume (BrainSegNotVent, generated in FreeSurfer (112)). The duration (segment length) of each propagation event was calculated and averaged for each subject. Propagation speed was determined by dividing brain volume by the average event duration for each subject. Based on brain size, the 184 subjects were divided into three groups: small (lower 1/3), medium (middle 1/3), and large (upper 1/3). Group analyses were then conducted to compare propagation speed and durations across these groups (Fig. S5).

### **Software and implementation**

All code was developed in Python. Deep learning models were implemented using PyTorch (v1.13.1) and dalle-pytorch (v1.6.4). Data processing and analysis were performed with SciPy (v1.10.0), NumPy (v1.24.2), scikit-learn (v1.2.0), and pandas (v1.5.3). Visualization was carried out using Matplotlib (v3.6.3) and Seaborn (v0.13.2). fMRI data were processed and visualized with Nilearn (v0.10.0) and NiBabel (v5.3.2). The CLIP model was implemented based on the open-source repository from OpenAI (<https://github.com/openai/CLIP>).

### **Data Availability**

For mice single neuron analysis, we used the Neuropixels Visual Coding Neuropixels and two-photon calcium imaging datasets from the Allen Institute (32, 33), accessible at <https://portal.brain-map.org/overview>. For resting-state human fMRI analysis, we used HCP-7T dataset from <https://www.humanconnectome.org>. We shared our EEG-fMRI dataset at <https://openneuro.org/datasets/ds003768>. For task human fMRI analysis, we used NSD dataset available at <https://naturalscenesdataset.org>. Source data are provided with this paper. The processed data in this study is available upon request.

## Code Availability

The Python code that produced the major results of this paper is available at <https://zenodo.org/records/17065604>. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

## Reference

1. B. Biswal, F. Zerrin Yetkin, V. M. Haughton, J. S. Hyde, Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magn Reson Med* 34, 537–541 (1995).
2. M. D. Fox, M. E. Raichle, Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8, 700–711 (2007).
3. M. Yu, O. Sporns, A. J. Saykin, The human connectome in Alzheimer disease — relationship to biomarkers and genetics. *Nat Rev Neurol* 17, 545–563 (2021).
4. S. M. Smith, T. E. Nichols, D. Vidaurre, A. M. Winkler, T. E. J. Behrens, M. F. Glasser, K. Ugurbil, D. M. Barch, D. C. Van Essen, K. L. Miller, A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat Neurosci* 18, 1565–1567 (2015).
5. D. Zhang, M. E. Raichle, Disease and the brain’s dark energy. *Nat Rev Neurol* 6, 15–28 (2010).
6. T. Matsui, T. Murakami, K. Ohki, Transient neuronal coactivations embedded in globally propagating waves underlie resting-state functional connectivity. *Proceedings of the National Academy of Sciences* 113, 6556–6561 (2016).
7. Y. Gu, L. E. Sainburg, S. Kuang, F. Han, J. W. Williams, Y. Liu, N. Zhang, X. Zhang, D. A. Leopold, X. Liu, Brain Activity Fluctuations Propagate as Waves Traversing the Cortical Hierarchy. *Cerebral Cortex* 31, 3986–4005 (2021).
8. B. Yousefi, J. Shin, E. H. Schumacher, S. D. Keilholz, Quasi-periodic patterns of intrinsic brain activity in individuals and their relationship to global signal. *Neuroimage* 167, 297–308 (2018).
9. B. Yousefi, S. Keilholz, Propagating patterns of intrinsic activity along macroscale gradients coordinate functional connections across the whole brain. *Neuroimage* 231, 117827 (2021).
10. R. V Raut, A. Z. Snyder, A. Mitra, D. Yellin, N. Fujii, R. Malach, M. E. Raichle, Global waves synchronize the brain’s functional systems with fluctuating arousal. *Sci Adv* 7, eabf2709 (2021).
11. T. P. Lillicrap, A. Santoro, L. Marris, C. J. Akerman, G. Hinton, Backpropagation and the brain. *Nat Rev Neurosci* 21, 335–346 (2020).

12. D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors. *Nature* 323, 533–536 (1986).
13. C. Stringer, M. Pachitariu, N. Steinmetz, C. B. Reddy, M. Carandini, K. D. Harris, Spontaneous behaviors drive multidimensional, brainwide activity. *Science* (1979) 364, eaav7893 (2019).
14. M. Okun, N. A. Steinmetz, A. Lak, M. Dervinis, K. D. Harris, Distinct Structure of Cortical Population Activity on Fast and Infralow Timescales. *Cerebral Cortex* 29, 2196–2210 (2019).
15. X. Liu, D. A. Leopold, Y. Yang, Single-neuron firing cascades underlie global spontaneous brain events. *Proceedings of the National Academy of Sciences* 118, e2105395118 (2021).
16. Y. Yang, D. A. Leopold, J. H. Duyn, X. Liu, Hippocampal replay sequence governed by spontaneous brain-wide dynamics. *PNAS Nexus*, pgae078 (2024).
17. Y. Yang, D. A. Leopold, J. H. Duyn, G. O. Sipe, X. Liu, Intrinsic forebrain arousal dynamics governs sensory stimulus encoding. *bioRxiv*, doi: 10.1101/2023.10.04.560900 (2023).
18. M. E. Hasselmo, Neuromodulation and cortical function: modeling the physiological basis of behavior. *Behavioural Brain Research* 67, 1–27 (1995).
19. G. Buzsáki, Two-stage model of memory trace formation: A role for “noisy” brain states. *Neuroscience* 31, 551–570 (1989).
20. C. J. Honey, E. L. Newman, A. C. Schapiro, Switching between internal and external modes: A multiscale learning principle. *Network Neuroscience* 1, 339–356 (2017).
21. Y. Gu, F. Han, L. E. Sainburg, M. M. Schade, O. M. Buxton, J. H. Duyn, X. Liu, An orderly sequence of autonomic and neural events at transient arousal changes. *Neuroimage* 264, 119720 (2022).
22. D. Yellin, A. Berkovich-Ohana, R. Malach, Coupling between pupil fluctuations and resting-state fMRI uncovers a slow build-up of antagonistic responses in the human cortex. *Neuroimage* 106, 414–427 (2015).
23. T. Bolt, S. Wang, J. S. Nomi, R. Setton, B. P. Gold, B. deB. Frederick, R. N. Spreng, S. D. Keilholz, L. Q. Uddin, C. Chang, A Unified Physiological Process Links Global Patterns of Functional MRI, Respiratory Activity, and Autonomic Signaling. *bioRxiv*, doi: 10.1101/2023.01.19.524818 (2023).
24. E. Podvalny, R. Sanchez-Romero, M. W. Cole, Functionality of arousal-regulating brain circuitry at rest predicts human cognitive abilities. *Cerebral Cortex* 34, bhae192 (2024).
25. M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, M. E. Raichle, The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences* 102, 9673–9678 (2005).
26. A. Mitra, A. Z. Snyder, T. Blazey, M. E. Raichle, Lag threads organize the brain’s intrinsic activity. *Proceedings of the National Academy of Sciences* 112, E2235–E2244 (2015).
27. T. Bolt, J. S. Nomi, D. Bzdok, J. A. Salas, C. Chang, B. T. Thomas Yeo, L. Q. Uddin, S. D. Keilholz, A parsimonious description of global functional brain organization in three spatiotemporal patterns. *Nat Neurosci* 25, 1093–1103 (2022).
28. N. K. Logothetis, O. Eschenko, Y. Murayama, M. Augath, T. Steudel, H. C. Evrard, M. Besserve, A. Oeltermann, Hippocampal–cortical interaction during periods of subcortical silence. *Nature* 491, 547–553 (2012).
29. G. Buzsáki, Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus* 25, 1073–1188 (2015).



30. R. E. Yoss, N. J. Moyer, R. W. Hollenhorst, Pupil size and spontaneous pupillary waves associated with alertness, drowsiness, and sleep. *Neurology* 20, 545 (1970).
31. J. Reimer, E. Froudarakis, C. R. Cadwell, D. Yatsenko, G. H. Denfield, A. S. Tolias, Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. *Neuron* 84, 355–362 (2014).
32. J. H. Siegle, X. Jia, S. Durand, S. Gale, C. Bennett, N. Graddis, G. Heller, T. K. Ramirez, H. Choi, J. A. Luviano, P. A. Groblewski, R. Ahmed, A. Arkhipov, A. Bernard, Y. N. Billeh, D. Brown, M. A. Buice, N. Cain, S. Caldejon, L. Casal, A. Cho, M. Chvilicek, T. C. Cox, K. Dai, D. J. Denman, S. E. J. de Vries, R. Dietzman, L. Esposito, C. Farrell, D. Feng, J. Galbraith, M. Garrett, E. C. Gelfand, N. Hancock, J. A. Harris, R. Howard, B. Hu, R. Hytnen, R. Iyer, E. Jessett, K. Johnson, I. Kato, J. Kiggins, S. Lambert, J. Lecoq, P. Ledochowitsch, J. H. Lee, A. Leon, Y. Li, E. Liang, F. Long, K. Mace, J. Melchior, D. Millman, T. Mollenkopf, C. Nayan, L. Ng, K. Ngo, T. Nguyen, P. R. Nicovich, K. North, G. K. Ocker, D. Ollerenshaw, M. Oliver, M. Pachitariu, J. Perkins, M. Reding, D. Reid, M. Robertson, K. Ronellenfitch, S. Seid, C. Slaughterbeck, M. Stoecklin, D. Sullivan, B. Sutton, J. Swapp, C. Thompson, K. Turner, W. Wakeman, J. D. Whitesell, D. Williams, A. Williford, R. Young, H. Zeng, S. Naylor, J. W. Phillips, R. C. Reid, S. Mihalas, S. R. Olsen, C. Koch, Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* 592, 86–92 (2021).
33. S. E. J. de Vries, J. A. Lecoq, M. A. Buice, P. A. Groblewski, G. K. Ocker, M. Oliver, D. Feng, N. Cain, P. Ledochowitsch, D. Millman, K. Roll, M. Garrett, T. Keenan, L. Kuan, S. Mihalas, S. Olsen, C. Thompson, W. Wakeman, J. Waters, D. Williams, C. Barber, N. Berbesque, B. Blanchard, N. Bowles, S. D. Caldejon, L. Casal, A. Cho, S. Cross, C. Dang, T. Dolbeare, M. Edwards, J. Galbraith, N. Gaudreault, T. L. Gilbert, F. Griffin, P. Hargrave, R. Howard, L. Huang, S. Jewell, N. Keller, U. Knoblich, J. D. Larkin, R. Larsen, C. Lau, E. Lee, F. Lee, A. Leon, L. Li, F. Long, J. Luviano, K. Mace, T. Nguyen, J. Perkins, M. Robertson, S. Seid, E. Shea-Brown, J. Shi, N. Sjoquist, C. Slaughterbeck, D. Sullivan, R. Valenza, C. White, A. Williford, D. M. Witten, J. Zhuang, H. Zeng, C. Farrell, L. Ng, A. Bernard, J. W. Phillips, R. C. Reid, C. Koch, A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nat Neurosci* 23, 138–151 (2020).
34. N. A. Steinmetz, P. Zatka-Haas, M. Carandini, K. D. Harris, Distributed coding of choice, action and engagement across the mouse brain. *Nature* 576, 266–273 (2019).
35. D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. E. J. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. W. Curtiss, S. Della Penna, D. Feinberg, M. F. Glasser, N. Harel, A. C. Heath, L. Larson-Prior, D. Marcus, G. Michalareas, S. Moeller, R. Oostenveld, S. E. Petersen, F. Prior, B. L. Schlaggar, S. M. Smith, A. Z. Snyder, J. Xu, E. Yacoub, The Human Connectome Project: A data acquisition perspective. *Neuroimage* 62, 2222–2231 (2012).
36. D. S. Margulies, S. S. Ghosh, A. Goulas, M. Falkiewicz, J. M. Huntenburg, G. Langs, G. Bezgin, S. B. Eickhoff, F. X. Castellanos, M. Petrides, E. Jefferies, J. Smallwood, Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences* 113, 12574–12579 (2016).
37. E. J. Allen, G. St-Yves, Y. Wu, J. L. Breedlove, J. S. Prince, L. T. Dowdle, M. Nau, B. Caron, F. Pestilli, I. Charest, J. B. Hutchinson, T. Naselaris, K. Kay, A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nat Neurosci* 25, 116–126 (2022).

38. W. Li, L. Zhu, L. Wen, Y. Yang, “DeCap: Decoding CLIP Latents for Zero-Shot Captioning via Text-Only Training” in *The Eleventh International Conference on Learning Representations* (2023; <https://openreview.net/forum?id=Lt8bMlhiwx2>).
39. A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, I. Sutskever, “Learning Transferable Visual Models From Natural Language Supervision” in *Proceedings of the 38th International Conference on Machine Learning*, M. Meila, T. Zhang, Eds. (PMLR, 2021; <https://proceedings.mlr.press/v139/radford21a.html>) vol. 139 of *Proceedings of Machine Learning Research*, pp. 8748–8763.
40. T. Chen, S. Kornblith, M. Norouzi, G. Hinton, “A Simple Framework for Contrastive Learning of Visual Representations” in *Proceedings of the 37th International Conference on Machine Learning*, H. D. III, A. Singh, Eds. (PMLR, 2020; <https://proceedings.mlr.press/v119/chen20j.html>) vol. 119 of *Proceedings of Machine Learning Research*, pp. 1597–1607.
41. L. van der Maaten, G. Hinton, Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9, 2579–2605 (2008).
42. J. P. Aggleton, S. M. O’Mara, S. D. Vann, N. F. Wright, M. Tsanov, J. T. Erichsen, Hippocampal–anterior thalamic pathways for memory: uncovering a network of direct and indirect actions. *European Journal of Neuroscience* 31, 2292–2307 (2010).
43. K. Hwang, J. M. Shine, M. W. Cole, E. Sorenson, Thalamocortical contributions to cognitive task activity. *Elife* 11, e81282 (2022).
44. C. Higgins, Y. Liu, D. Vidaurre, Z. Kurth-Nelson, R. Dolan, T. Behrens, M. Woolrich, Replay bursts in humans coincide with activation of the default mode and parietal alpha networks. *Neuron* 109, 882–893.e7 (2021).
45. Y. Norman, O. Raccach, S. Liu, J. Parvizi, R. Malach, Hippocampal ripples and their coordinated dialogue with the default mode network during recent and remote recollection. *Neuron* 109, 2767–2780.e5 (2021).
46. A. Arieli, A. Sterkin, A. Grinvald, A. Aertsen, Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science (1979)* 273, 1868–1871 (1996).
47. M. S. Livingstone, D. H. Hubel, Effects of sleep and arousal on the processing of visual information in the cat. *Nature* 291, 554–561 (1981).
48. A. Hasenstaub, R. N. S. Sachdev, D. A. McCormick, State Changes Rapidly Modulate Cortical Neuronal Responsiveness. *Journal of Neuroscience* 27, 9607–9622 (2007).
49. M. D. Fox, A. Z. Snyder, J. M. Zacks, M. E. Raichle, Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nat Neurosci* 9, 23–25 (2006).
50. B. J. He, Spontaneous and Task-Evoked Brain Activity Negatively Interact. *Journal of Neuroscience* 33, 4672–4682 (2013).
51. M. J. McGinley, M. Vinck, J. Reimer, R. Batista-Brito, E. Zagha, C. R. Cadwell, A. S. Tolias, J. A. Cardin, D. A. McCormick, Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron* 87, 1143–1161 (2015).
52. W. Chen, K. Park, Y. Pan, A. P. Koretsky, C. Du, Interactions between stimuli-evoked cortical activity and spontaneous low frequency oscillations measured with neuronal calcium. *Neuroimage* 210, 116554 (2020).

53. M. A. Wilson, B. L. McNaughton, Reactivation of Hippocampal Ensemble Memories During Sleep. *Science (1979)* 265, 676–679 (1994).
54. D. R. Euston, M. Tatsuno, B. L. McNaughton, Fast-Forward Playback of Recent Memory Sequences in Prefrontal Cortex During Sleep. *Science (1979)* 318, 1147–1150 (2007).
55. Y. Norman, E. M. Yeagle, S. Khuvis, M. Harel, A. D. Mehta, R. Malach, Hippocampal sharp-wave ripples linked to visual episodic recollection in humans. *Science (1979)* 365, eaax1030 (2019).
56. A. P. Vaz, S. K. Inati, N. Brunel, K. A. Zaghloul, Coupled ripple oscillations between the medial temporal lobe and neocortex retrieve human memory. *Science (1979)* 363, 975–978 (2019).
57. J. J. Sakon, M. J. Kahana, Hippocampal ripples signal contextually mediated episodic recall. *Proceedings of the National Academy of Sciences* 119, e2201657119 (2022).
58. J. Karimi Abadchi, M. Nazari-Ahangarkolaee, S. Gattas, E. Bermudez-Contreras, A. Luczak, B. L. McNaughton, M. H. Mohajerani, Spatiotemporal patterns of neocortical activity around hippocampal sharp-wave ripples. *Elife* 9, e51972 (2020).
59. R. Kaplan, M. H. Adhikari, R. Hindriks, D. Mantini, Y. Murayama, N. K. Logothetis, G. Deco, Hippocampal Sharp-Wave Ripples Influence Selective Activation of the Default Mode Network. *Current Biology* 26, 686–691 (2016).
60. J. G. Klinzing, N. Niethard, J. Born, Mechanisms of systems memory consolidation during sleep. *Nat Neurosci* 22, 1598–1610 (2019).
61. Y. Dudai, A. Karni, J. Born, The Consolidation and Transformation of Memory. *Neuron* 88, 20–32 (2015).
62. J. M. Palva, S. Palva, Infra-slow fluctuations in electrophysiological recordings, blood-oxygenation-level-dependent signals, and psychophysical time series. *Neuroimage* 62, 2201–2211 (2012).
63. A. Kucyi, M. J. Hove, M. Esterman, R. M. Hutchison, E. M. Valera, Dynamic Brain Network Correlates of Spontaneous Fluctuations in Attention. *Cerebral Cortex* 27, 1831–1840 (2017).
64. J. M. Shine, P. G. Bissett, P. T. Bell, O. Koyejo, J. H. Balsters, K. J. Gorgolewski, C. A. Moodie, R. A. Poldrack, The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. *Neuron* 92, 544–554 (2016).
65. M. J. McGinley, S. V. David, D. A. McCormick, Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron* 87, 179–192 (2015).
66. J. Bayer, J. Gläscher, J. Finsterbusch, L. H. Schulte, T. Sommer, Linear and inverted U-shaped dose-response functions describe estrogen effects on hippocampal activity in young women. *Nat Commun* 9, 1220 (2018).
67. E. Baldi, C. Bucherelli, The Inverted “U-Shaped” Dose-Effect Relationships in Learning and Memory: Modulation of Arousal and Consolidation. *Nonlinearity Biol Toxicol Med* 3, nonlin.003.01.002 (2005).
68. G. Zerbès, F. M. Kausche, J. C. Müller, K. Wiedemann, L. Schwabe, Glucocorticoids, Noradrenergic Arousal, and the Control of Memory Retrieval. *J Cogn Neurosci* 31, 288–298 (2019).
69. D. J.-F. de Quervain, B. Roozendaal, J. L. McGaugh, Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature* 394, 787–790 (1998).

70. X. Liu, D. Picchioni, Y. Yang, H. Mandelkow, J. A. de Zwart, J. H. Duyn, X. Liu, Sleep-stage dependent patterning of slowly propagating brain activity. *npj Biological Timing and Sleep* 2, 1 (2025).
71. A. G. Siapas, M. A. Wilson, Coordinated Interactions between Hippocampal Ripples and Cortical Spindles during Slow-Wave Sleep. *Neuron* 21, 1123–1128 (1998).
72. S. Diekelmann, J. Born, The memory function of sleep. *Nat Rev Neurosci* 11, 114–126 (2010).
73. S. Lecci, L. M. J. Fernandez, F. D. Weber, R. Cardis, J.-Y. Chatton, J. Born, A. Lüthi, Coordinated infraslow neural and cardiac oscillations mark fragility and offline periods in mammalian sleep. *Sci Adv* 3, e1602026 (2017).
74. C. Kjaerby, M. Andersen, N. Hauglund, V. Untiet, C. Dall, B. Sigurdsson, F. Ding, J. Feng, Y. Li, P. Weikop, H. Hirase, M. Nedergaard, Memory-enhancing properties of sleep depend on the oscillatory amplitude of norepinephrine. *Nat Neurosci* 25, 1059–1070 (2022).
75. X. Liu, J. A. de Zwart, M. L. Schölvinck, C. Chang, F. Q. Ye, D. A. Leopold, J. H. Duyn, Subcortical evidence for a contribution of arousal to fMRI studies of brain activity. *Nat Commun* 9, 395 (2018).
76. J. Turchi, C. Chang, F. Q. Ye, B. E. Russ, D. K. Yu, C. R. Cortes, I. E. Monosov, J. H. Duyn, D. A. Leopold, The Basal Forebrain Regulates Global Resting-State fMRI Fluctuations. *Neuron* 97, 940-952.e4 (2018).
77. M.-M. Mesulam, E. J. Mufson, A. I. Levey, B. H. Wainer, Cholinergic innervation of cortex by the basal forebrain: Cytochemistry and cortical connections of the septal area, diagonal band nuclei, nucleus basalis (Substantia innominata), and hypothalamus in the rhesus monkey. *Journal of Comparative Neurology* 214, 170–197 (1983).
78. L. Pinto, M. J. Goard, D. Estandian, M. Xu, A. C. Kwan, S.-H. Lee, T. C. Harrison, G. Feng, Y. Dan, Fast modulation of visual perception by basal forebrain cholinergic neurons. *Nat Neurosci* 16, 1857–1863 (2013).
79. E. C. Ballinger, M. Ananth, D. A. Talmage, L. W. Role, Basal Forebrain Cholinergic Circuits and Signaling in Cognition and Cognitive Decline. *Neuron* 91, 1199–1218 (2016).
80. D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, S. T. Grafton, Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences* 108, 7641–7646 (2011).
81. U. Braun, A. Schäfer, H. Walter, S. Erk, N. Romanczuk-Seiferth, L. Haddad, J. I. Schweiger, O. Grimm, A. Heinz, H. Tost, A. Meyer-Lindenberg, D. S. Bassett, Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proceedings of the National Academy of Sciences* 112, 11678–11683 (2015).
82. W. Huijbers, P. Vannini, R. A. Sperling, P. C.M., R. Cabeza, S. M. Daselaar, Explaining the encoding/retrieval flip: Memory-related deactivations and activations in the posteromedial cortex. *Neuropsychologia* 50, 3764–3774 (2012).
83. S. Daselaar, S. Prince, N. Dennis, S. Hayes, H. Kim, R. Cabeza, Posterior midline and ventral parietal activity is associated with retrieval success and encoding failure. *Front Hum Neurosci* 3 (2009).
84. P. Vannini, J. O'Brien, K. O'Keefe, M. Pihlajamäki, P. LaViolette, R. A. Sperling, What Goes Down Must Come Up: Role of the Posteromedial Cortices in Encoding and Retrieval. *Cerebral Cortex* 21, 22–34 (2011).

85. F. Han, J. Chen, A. Belkin-Rosen, Y. Gu, L. Luo, O. M. Buxton, X. Liu, the A. D. N. Initiative, Reduced coupling between cerebrospinal fluid flow and global brain activity is linked to Alzheimer disease–related pathology. *PLoS Biol* 19, e3001233- (2021).
86. F. Han, X. Liu, R. B. Mailman, X. Huang, X. Liu, Resting-state global brain activity affects early  $\beta$ -amyloid accumulation in default mode network. *Nat Commun* 14, 7788 (2023).
87. F. Han, J. Lee, X. Chen, J. Zientz, S. M. Landau, S. L. Baker, T. M. Harrison, W. J. Jagust, Reduced coupling between cerebrospinal fluid flow and global brain activity is linked to tau pathology. *Alzheimer's & Dementia* 19, e075860 (2023).
88. J. J. Iliff, M. Wang, D. M. Zeppenfeld, A. Venkataraman, B. A. Plog, Y. Liao, R. Deane, M. Nedergaard, Cerebral Arterial Pulsation Drives Paravascular CSF Interstitial Fluid Exchange in the Murine Brain. *Journal of Neuroscience* 33, 18190–18199 (2013).
89. S. J. van Veluw, S. S. Hou, M. Calvo-Rodriguez, M. Arbel-Ornath, A. C. Snyder, M. P. Frosch, S. M. Greenberg, B. J. Bacskaï, Vasomotion as a Driving Force for Paravascular Clearance in the Awake Mouse Brain. *Neuron* 105, 549-561.e5 (2020).
90. L. Xie, H. Kang, Q. Xu, M. J. Chen, Y. Liao, M. Thiagarajan, J. O'Donnell, D. J. Christensen, C. Nicholson, J. J. Iliff, T. Takano, R. Deane, M. Nedergaard, Sleep Drives Metabolite Clearance from the Adult Brain. *Science* (1979) 342, 373–377 (2013).
91. T. Broggini, J. Duckworth, X. Ji, R. Liu, X. Xia, P. Mächler, I. Shaked, L. P. Munting, S. Iyengar, M. Kotlikoff, S. J. van Veluw, M. Vergassola, G. Mishne, D. Kleinfeld, Long-wavelength traveling waves of vasomotion modulate the perfusion of cortex. *Neuron* 112, 2349-2367.e8 (2024).
92. X. Liu, Decoupling Between Brain Activity and Cerebrospinal Fluid Movement in Neurological Disorders. *Journal of Magnetic Resonance Imaging* 60, 1743–1752 (2024).
93. P. J. Whitehouse, D. L. Price, R. G. Struble, A. W. Clark, J. T. Coyle, D. M. R, Alzheimer's Disease and Senile Dementia: Loss of Neurons in the Basal Forebrain. *Science* (1979) 215, 1237–1239 (1982).
94. R. T. Bartus, On Neurodegenerative Diseases, Models, and Treatment Strategies: Lessons Learned and Lessons Forgotten a Generation Following the Cholinergic Hypothesis. *Exp Neurol* 163, 495–529 (2000).
95. D. A. Drachman, J. Leavitt, Human Memory and the Cholinergic System: A Relationship to Aging? *Arch Neurol* 30, 113–121 (1974).
96. M. E. Hasselmo, The role of acetylcholine in learning and memory. *Curr Opin Neurobiol* 16, 710–715 (2006).
97. M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, M. Jenkinson, The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80, 105–124 (2013).
98. G. Salimi-Khorshidi, G. Douaud, C. F. Beckmann, M. F. Glasser, L. Griffanti, S. M. Smith, Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* 90, 449–468 (2014).
99. K. Kay, A. Rokem, J. Winawer, R. Dougherty, B. Wandell, GLMdenoise: a fast, automated technique for denoising task-based fMRI data. *Front Neurosci* 7 (2013).
100. B. B. Biswal, M. Mennes, X.-N. Zuo, S. Gohel, C. Kelly, S. M. Smith, C. F. Beckmann, J. S. Adelstein, R. L. Buckner, S. Colcombe, A.-M. Dogonowski, M. Ernst, D. Fair, M. Hampson, M. J. Hoptman, J. S. Hyde, V. J. Kiviniemi, R. Kötter, S.-J. Li, C.-P. Lin, M. J. Lowe, C. Mackay, D. J. Madden, K. H. Madsen, D. S. Margulies, H. S. Mayberg, K.

- McMahon, C. S. Monk, S. H. Mostofsky, B. J. Nagel, J. J. Pekar, S. J. Peltier, S. E. Petersen, V. Riedl, S. A. R. B. Rombouts, B. Rypma, B. L. Schlaggar, S. Schmidt, R. D. Seidler, G. J. Siegle, C. Sorg, G.-J. Teng, J. Veijola, A. Villringer, M. Walter, L. Wang, X.-C. Weng, S. Whitfield-Gabrieli, P. Williamson, C. Windischberger, Y.-F. Zang, H.-Y. Zhang, F. X. Castellanos, M. P. Milham, Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences* 107, 4734–4739 (2010).
101. A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, B. T. T. Yeo, Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex* 28, 3095–3114 (2018).
  102. A. Krauth, R. Blanc, A. Poveda, D. Jeanmonod, A. Morel, G. Székely, A mean three-dimensional atlas of the human thalamus: Generation from multiple histological data. *Neuroimage* 49, 2053–2062 (2010).
  103. B. L. Edlow, E. Takahashi, O. Wu, T. Benner, G. Dai, L. Bu, P. E. Grant, D. M. Greer, S. M. Greenberg, H. C. Kinney, R. D. Folkerth, Neuroanatomic Connectivity of the Human Ascending Arousal System Critical to Consciousness and Its Disorders. *J Neuropathol Exp Neurol* 71, 531–546 (2012).
  104. Z. Liu, J. A. de Zwart, P. van Gelderen, L.-W. Kuo, J. H. Duyn, Statistical feature extraction for artifact removal from concurrent fMRI-EEG recordings. *Neuroimage* 59, 2073–2087 (2012).
  105. M. Falahpour, C. Chang, C. W. Wong, T. T. Liu, Template-based prediction of vigilance fluctuations in resting-state fMRI. *Neuroimage* 174, 317–327 (2018).
  106. E. Stark, L. Roux, R. Eichler, Y. Senzai, S. Royer, G. Buzsáki, Pyramidal Cell-Interneuron Interactions Underlie Hippocampal Ripple Oscillations. *Neuron* 83, 467–480 (2014).
  107. N. Kriegeskorte, M. Mur, D. A. Ruff, R. Kiani, J. Bodurka, H. Esteky, K. Tanaka, P. A. Bandettini, Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron* 60, 1126–1141 (2008).
  108. K. Papineni, S. Roukos, T. Ward, W.-J. Zhu, “BLEU: a method for automatic evaluation of machine translation” in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (Association for Computational Linguistics, USA, 2002; <https://doi.org/10.3115/1073083.1073135>) *ACL '02*, pp. 311–318.
  109. R. Vedantam, C. Zitnick, D. Parikh, “CIDEr: Consensus-based image description evaluation” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE Computer Society, Los Alamitos, CA, USA, 2015; <https://doi.ieeecomputersociety.org/10.1109/CVPR.2015.7299087>), pp. 4566–4575.
  110. M. Denkowski, A. Lavie, “Meteor Universal: Language Specific Translation Evaluation for Any Target Language” in *Proceedings of the Ninth Workshop on Statistical Machine Translation*, O. Bojar, C. Buck, C. Federmann, B. Haddow, P. Koehn, C. Monz, M. Post, L. Specia, Eds. (Association for Computational Linguistics, Baltimore, Maryland, USA, 2014; <https://aclanthology.org/W14-3348>), pp. 376–380.
  111. C.-Y. Lin, “ROUGE: A Package for Automatic Evaluation of Summaries” in *Text Summarization Branches Out* (Association for Computational Linguistics, Barcelona, Spain, 2004; <https://aclanthology.org/W04-1013>), pp. 74–81.
  112. B. Fischl, FreeSurfer. *Neuroimage* 62, 774–781 (2012).
  113. B. T. Thomas Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fischl, H. Liu, R. L. Buckner, The

organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106, 1125–1165 (2011).

## Acknowledgments

This work was supported by the Brain Initiative award (1RF1MH123247-01 to XL), the NIH R01 award (1R01NS113889-01A1 to XL), the Intramural Research Program of the National Institute of Mental Health (ZIA-MH002838 to DAL), and the Intramural Research Program of the National Institute of Neurological Disorders and Stroke (ZIA-NS003027 to JHD).

## Author Contributions Statement

Conceptualization: Y.Y., X.L. Methodology: Y.Y., X.L. Investigation: Y.Y., X.L. Supervision: X.L. Writing – original draft: Y.Y., D.A.L., J.H.D. and X.L. Writing – review & editing: Y.Y., J.H.D. and X.L.

## Competing Interests Statement

The authors declare no competing interests.

## Figure Captions

**Figure 1. Association of spontaneous pupil dilations and brain-wide sequential activity in mice and humans.** (A) Locations of neuropixel probes from all mice with major recording sites color-coded: visual cortex (blue), hippocampus (green), and thalamus (pink). Top: 3D illustration. Bottom: 2D projection on a middle brain slice. (B) Example pupil and spiking data from a representative mouse during stationary visual stimulation. Top: spontaneous fluctuation of pupil diameter with alternating dilation (red) and constriction (blue) phases, with dilation onsets marked by red dashed lines and triangles. Bottom: normalized spiking activity of recorded neurons sorted by the principal delay profile, revealing spiking cascades of sequential activations from negative-delay neurons (blue-symbolic-neurons) to positive-delay neurons (red-symbolic-neurons). (C) The normalized pupil diameter (top, mean  $\pm$  SEM) and spiking activity (bottom) from this representative mouse were averaged based on pupil dilation onsets over an 8-s time window. (D) Schematic of a resting-state fMRI scan from HCP 7-Tesla dataset. (E) Example pupil and fMRI data from a representative human HCP subject at rest. Top: pupil diameter fluctuations. Bottom: concurrent fMRI signals of various brain regions sorted by principal gradient values (36). (F) The normalized pupil diameter (top, mean  $\pm$  SEM) and fMRI signals (bottom) were averaged at pupil dilation onsets over an 18-s time window, summarized from all 184 subjects. (G) The pupil-dilation-associated fMRI changes mapped onto a brain surface, shown at 7 time lags (0–12s post-dilation onset). Rows display cortical surface maps (first two rows) and thalamic volume maps (third row). Directional arrows denote dorsal (D), anterior (A), and anatomical left (L) directions.

**Figure 2. Stimulus information encoding was assessed through fMRI responses using deep learning-based model.** (A) Framework for training and evaluating a CLIP-based semantic decoder. An fMRI encoder was trained to map stimulus-evoked fMRI responses into CLIP embedding space via contrastive learning. For evaluation, the trained fMRI encoder encodes fMRI responses to fMRI embeddings in CLIP space, which are decoded by pre-trained caption decoder (38) to generate text descriptions, which were then compared to ground truth captions for evaluation. (B) Representational dissimilarity matrices based on cosine similarity for semantic CLIP embeddings (left) and fMRI embeddings (right) showed highly similar ( $p = 0$ ; the permutation test) structures. (C) Visualization of trained fMRI embeddings using t-SNE reveals categorical distinctions, with categories color-coded. (D) Examples of correctly and incorrectly decoded samples, from top to bottom: image stimuli IDs in the NSD dataset, ground truth captions, and captions decoded from fMRI response. (E) Schematic of task design in the NSD dataset. Eight participants viewed 10,000 unique images, each presented three times over 30–40 scans over a year, using a 4-second event-related design. (F) Box plot comparing the encoding accuracy, i.e. the proportion of correctly decoded samples, between the trained and untrained fMRI encoders ( $p=2.8 \times 10^{-7}$ ,  $N=8$ , two-sided paired t-test). Dots represent participants, and the dotted line indicates 5% chance level. (G) Influence of encoding accuracy on subsequent memory task performance. Stimuli are binned by initial encoding accuracy (20% increments). Memory performance at the second presentation showed significant differences across bins ( $p=2.8 \times 10^{-7}$ , one-way ANOVA,  $N = 195,000$  trials, 5 groups). Error bars represent the standard error of the mean (SEM). (H) Box plot showing significant difference ( $p=0.004$ ,  $N=8$ , two-sided one-sample t-test) in subsequent memory performance between trials with correct and incorrect initial encoding. Dots represent participants. Box plots show the interquartile range (boxes), minima and maxima (whiskers), and the median (center line). Source data are provided as a Source Data file.

**Figure 3. Semantic encoding and memory retrieval are oppositely modulated over the fMRI SM-to-DMN wave cycle.** (A) Detection of SM-to-DMN propagating waves during the task. The waves were detected using template-matching methods (7), with the principal gradient (PG) map as the template. Six examples of detected waves, with similarity values showing correlations between their delay profile and the PG template. (B) Schematics of memory encoding and recall measurements. (C) Opposite modulations of semantic/memory encoding and memory recall over the cycle of the SM-to-DMN waves. The averaged pattern of the detected SM-to-DMN waves (first row) was shifted 6 seconds backward in time to account for the hemodynamic response delay. Time zero marks the onset of the global mean signal increase (dashed line), which appears to correspond to the cascade center (D), as indicated by the timing of the upswing in semantic and visual encoding accuracies for human and mouse data. Both pupil size (second row) and memory recall (green, bottom row) change significantly across the wave cycle and peak at the DMN-activated phase, whereas the semantic encoding (third row) and memory encoding (purple, bottom row) are modulated in an opposite manner. Shadows represent areas within one SEM ( $N=8$  subjects). (D) Opposite modulations of visual encoding accuracy (third row) and hippocampal SPW-R rate (bottom row) across the spiking cascade cycle (first row) during stationary periods of continuous natural image stimulation in mice, adapted from (17). Time zero (dashed line) marks the onset of positive-delay neuron firing. Shadows represent areas within one SEM ( $N = 20$  mice). Source data are provided as a Source Data file.



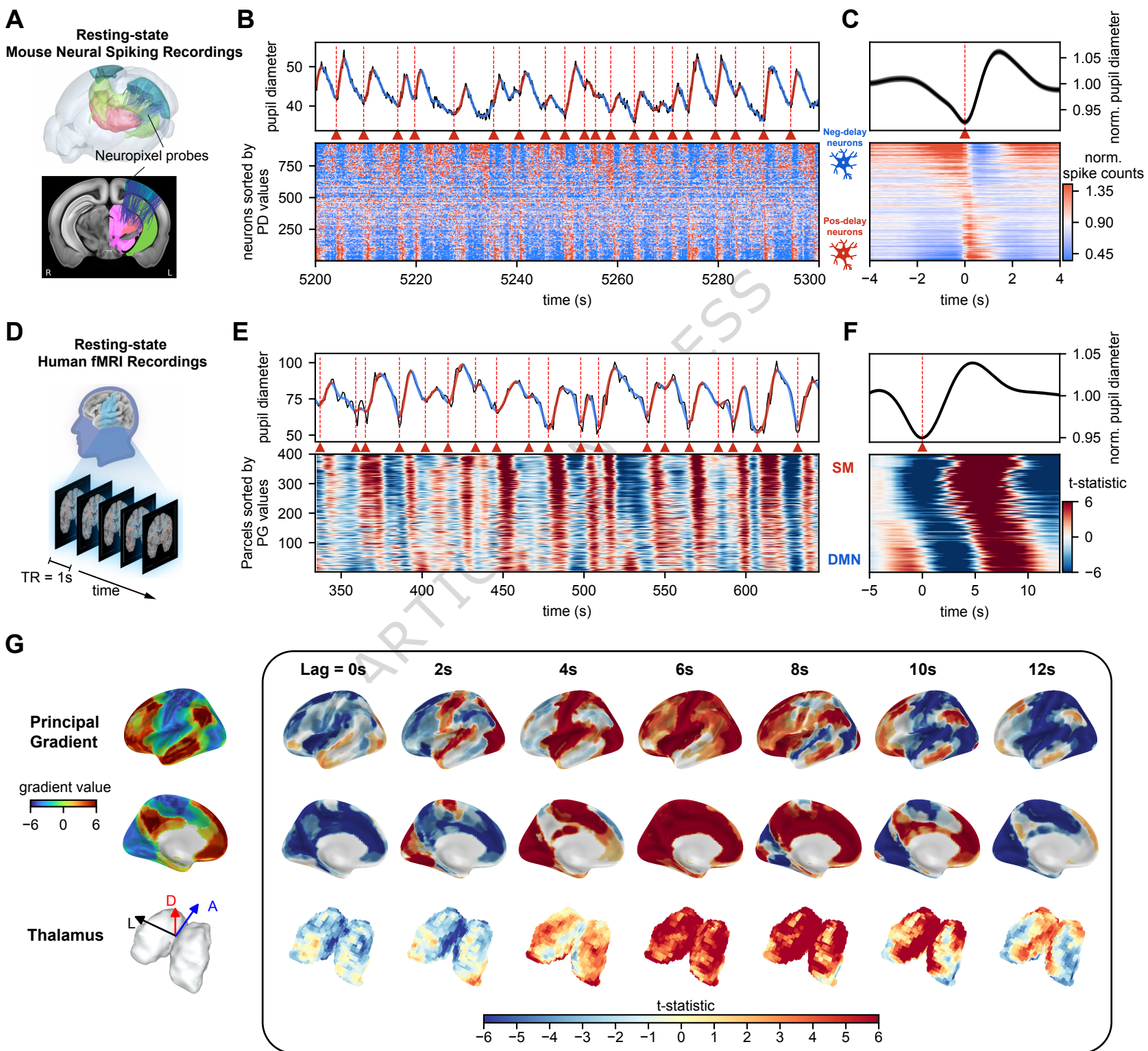
**Figure 4. Stimulus encoding and its modulation across the SM-to-DMN wave cycle in different brain regions.** (A) Cortical surface map showing the regional significance (paired t-tests: trained vs. untrained) of semantic encoding accuracy. (B) A box plot showing the semantic encoding accuracy estimated for different brain regions defined in the Yeo-17 networks atlas (101, 113). Each dot represents an individual participant. The significance of encoding accuracy was assessed by comparing the trained model to untrained model, similar to the analysis in Figure 2F. Asterisks denote the levels of statistical significance (two-sided paired t-test, N=8): \*,  $0.01 < p < 0.05$ ; \*\*,  $0.001 < p < 0.01$ ; \*\*\*,  $p < 0.001$ . Exact p-values are listed in Table S3. Box plots show the interquartile range (boxes), minima and maxima (whiskers), and the median (center line). (C) Semantic encoding accuracy (solid line) is consistently modulated over the SM-to-DMN wave cycle in six brain regions exhibiting the most significant encoding accuracy. There is no significant variation in the phase of peak encoding accuracy among these regions ( $p = 0.64$ , one way ANOVA,  $N = 48$ , 6 groups). For comparison, the average activation for each region is marked with a dashed line, which is shifted ahead of time by 6 seconds to account for the hemodynamic response delay. These regions are distinctly color-coded and their locations are indicated on a flattened cortical surface. Time series data are provided as the mean + SEM for eight participants (N=8). Source data are provided as a Source Data file.

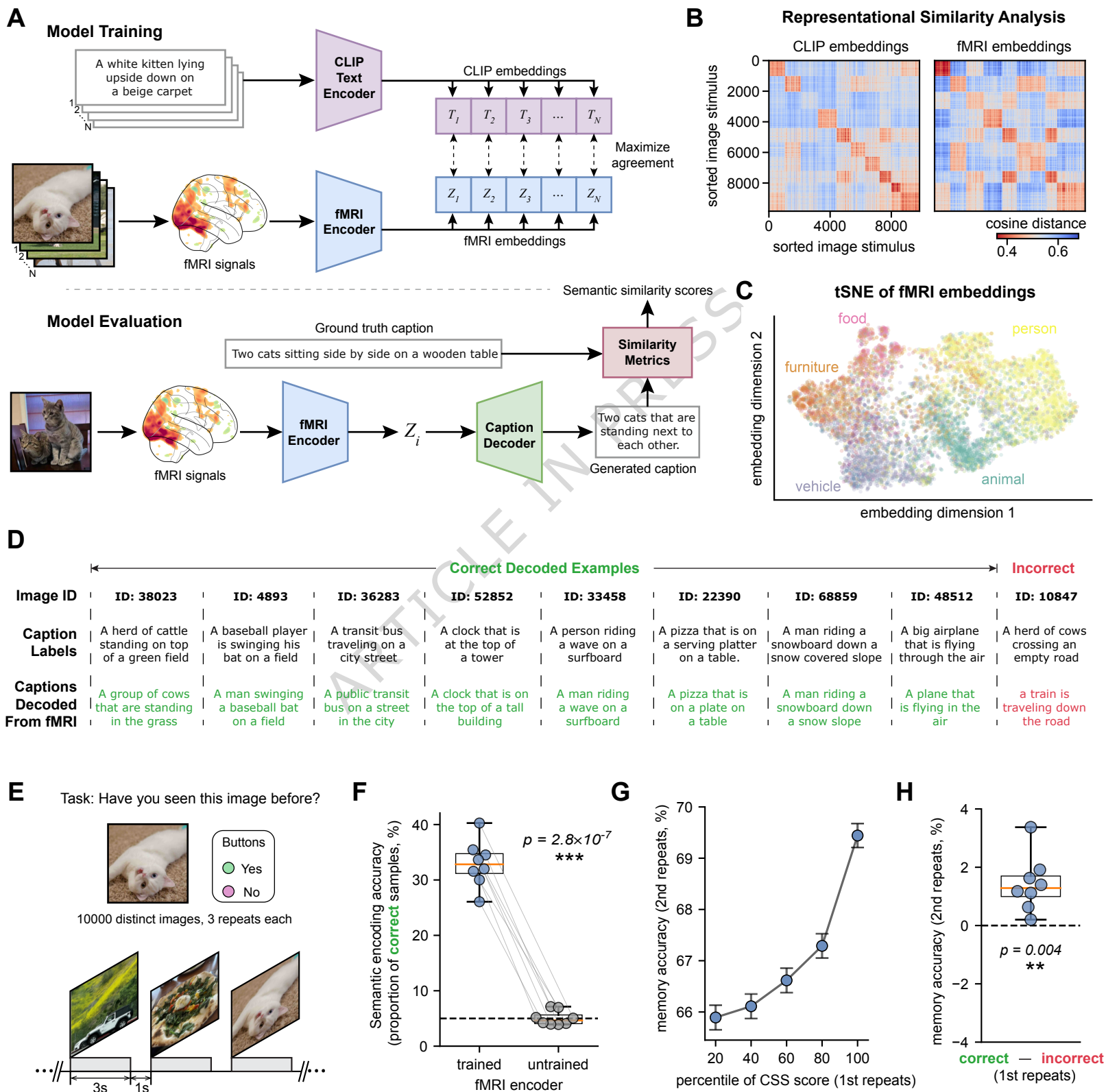
**Editor's Summary**

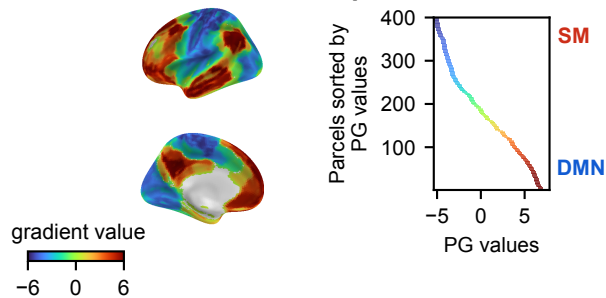
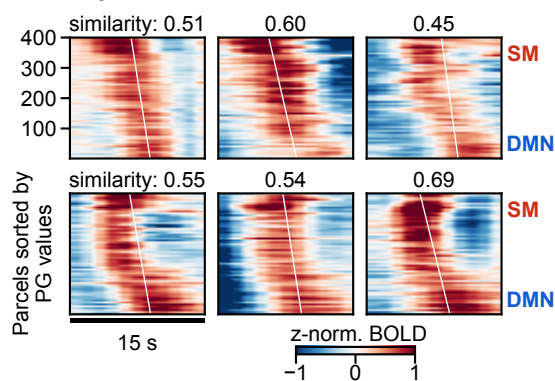
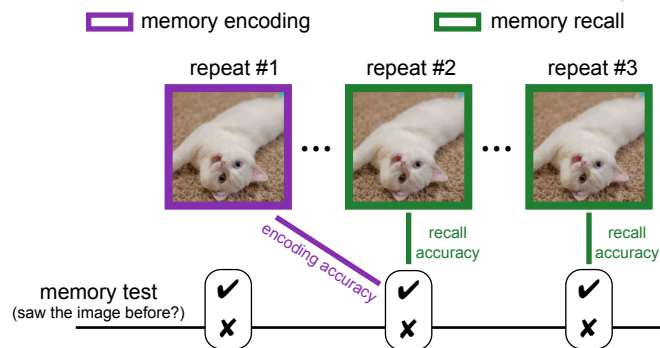
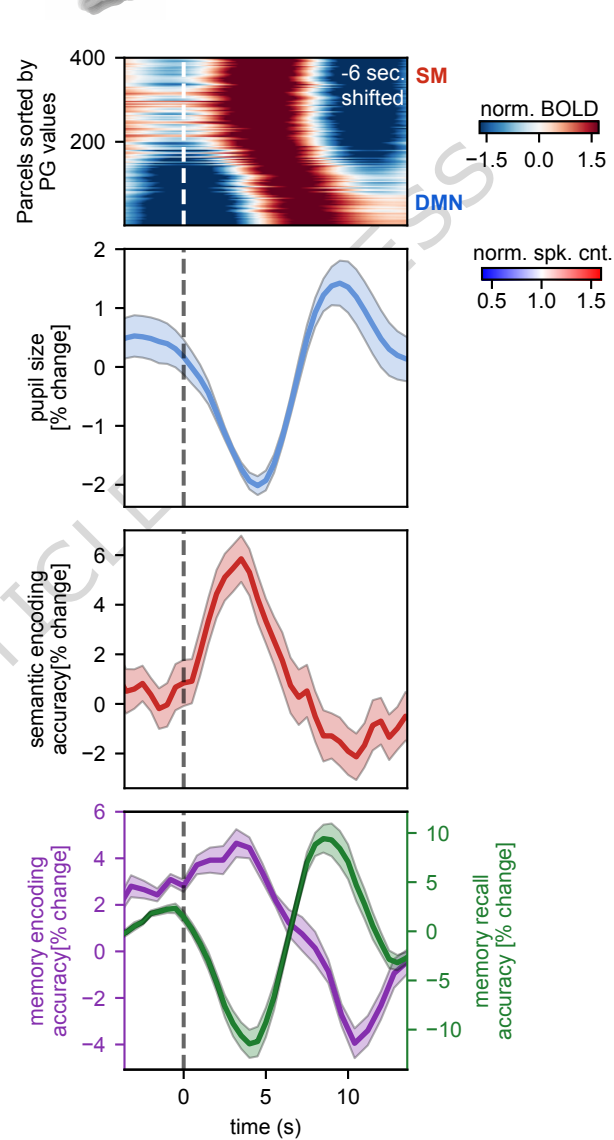
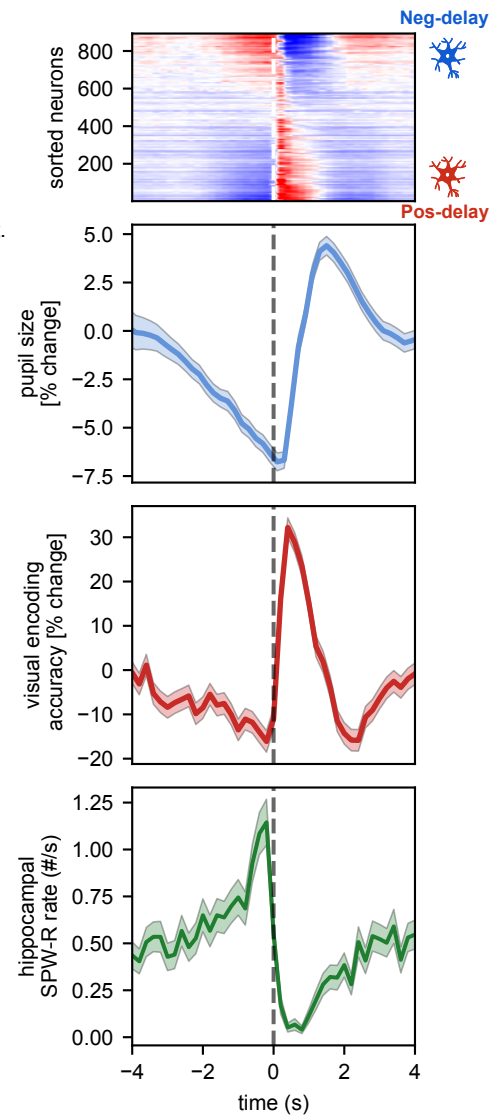
Sensory encoding and memory retrieval are oppositely modulated across the cycle of infra-slow brain dynamics in both humans and mice, revealing a conserved brain mechanism coordinating seconds-scale sensory and mnemonic processing.

**Peer Review Information:** *Nature Communications* thanks Ryan Raut and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

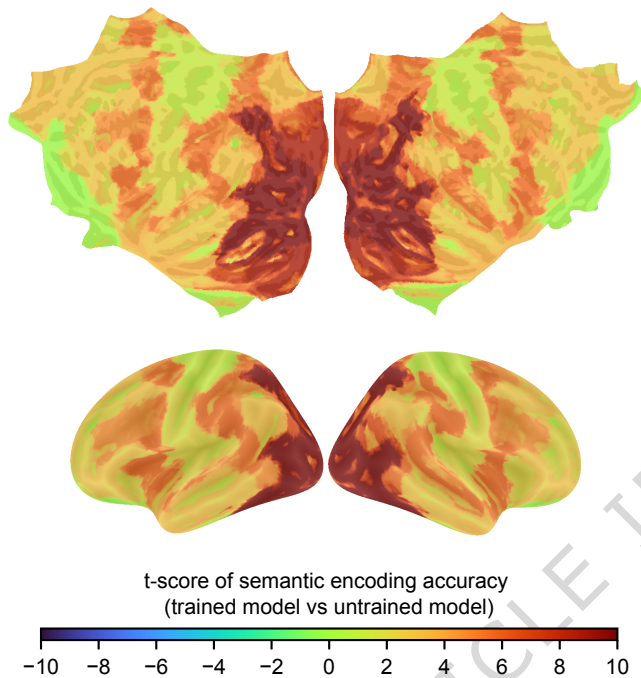
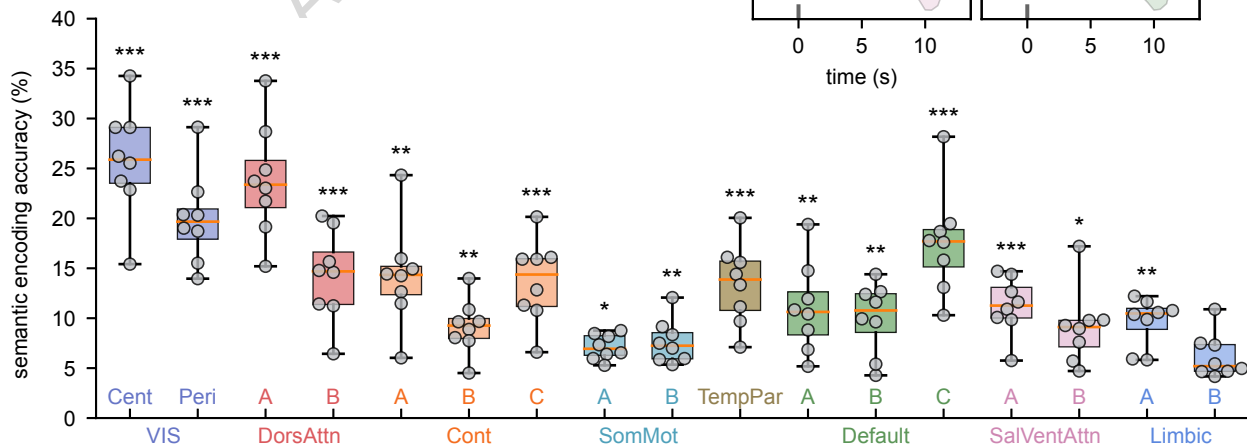
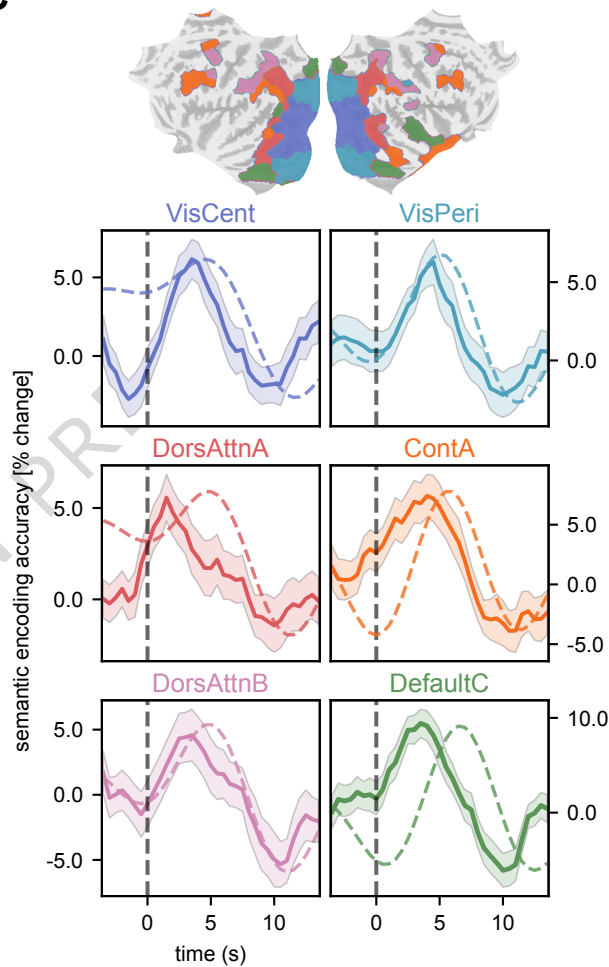
ARTICLE IN PRESS





**A****PG template****Examples of detected SM-to-DMN wave****B****Visual recognition memory metrics****C****Human SM-to-DMN Wave****D****Mouse Spiking Cascade**



**A****B****C****Yeo-17 Networks**