Distributed fMRI dynamics predict distinct EEG rhythms across sleep and wakefulness

Leandro P. L. Jacob[1], Sydney M. Bailes[1,3], Stephanie D. Williams[1,3], Carsen Stringer[4], Laura D. Lewis[1,2,5]

[1] Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA

[2] Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, USA

[3] Boston University, Boston, MA, USA

[4] HHMI Janelia Research Campus, Ashburn, VA, USA

[5] Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston MA USA

Abstract:

The brain exhibits rich oscillatory dynamics that vary across tasks and states, such as the EEG oscillations that define sleep. These oscillations play critical roles in cognition and arousal, but the brainwide mechanisms underlying them are not yet described. Using simultaneous EEG and fast fMRI in subjects drifting between sleep and wakefulness, we developed a machine learning approach to investigate which brainwide fMRI dynamics predict alpha (8-12 Hz) and delta (1-4 Hz) rhythms. We predicted moment-by-moment EEG power from fMRI activity in held-out subjects, and found that information about alpha power was represented by a remarkably small set of regions, segregated in two distinct networks linked to arousal and visual systems. Conversely, delta rhythms were diffusely represented on a large spatial scale across the cortex. These results identify distributed networks that predict delta and alpha rhythms, and establish a computational framework for investigating fMRI brainwide dynamics underlying EEG oscillations.

INTRODUCTION

As the brain navigates tasks and states, neurons frequently fire in synchrony, generating rich oscillatory dynamics with various temporal and spatial distributions. These neural signals have been studied with scalp EEG for almost a century, and have formed a core aspect of modern neuroscience due to their links to specific cognitive computations. Oscillations are particularly fundamental to the study of sleep, which is defined by the appearance of distinct rhythmic EEG patterns with well-studied physiological outcomes. As the brain transitions between wakefulness and sleep stages, oscillations wax and wane, signaling distinct arousal states. Arousal-related oscillations are linked to diverse cognitive processes, including several forms of memory [1–9] and attention [10,11], along with basic physiological processes such as brain waste clearance [12]. EEG oscillations during sleep thus have strong relationships to cognition and physiology [13]; however, the large-scale brain network activity giving rise to these oscillations remains poorly understood.

Two key oscillations with broad relevance for the study of brain arousal states are delta (1-4 Hz) and alpha (8-12 Hz) rhythms. Alpha (8-12 Hz; Fig. 1a), which is maximally detected during eyes-closed wakeful rest, is implicated in diverse aspects of cognition [11,14,15]. Delta (1-4 Hz; Fig. 1a) is most prominent during slow wave sleep and strongly linked to its cognitive benefits [3,4,8,16]. Extensive work has sought to identify the focal generators of these oscillations [17–25], but the brainwide states that they represent—which need not oscillate at the same frequency—are not established.

Identifying the brainwide dynamics that underlie variations in EEG activity is technically difficult to investigate; scalp EEG suffers from low spatial resolution and is unable to resolve deep brain activity, while invasive methods such as electrocorticography have limited spatial coverage. A way to obtain greater spatial information is to simultaneously acquire EEG and functional magnetic resonance imaging (fMRI), which provides high spatial resolution and the ability to image deeper brain regions. Analyzing EEG-fMRI data, however, poses a significant challenge: the BOLD (blood-oxygen-level-dependent) signal measured by fMRI is an indirect measure of brain activity, and while its relationship to the

underlying neural signal can be modeled using a hemodynamic response function (HRF), the true relationship is complex. The coupling between BOLD and the underlying neural activity varies across subjects and brain areas [26], and the HRF fails to account for many temporal properties of the fMRI signals [27–29]. Alternative approaches that fit a more flexible HRF are susceptible to noise and may overfit the data. To avoid overfitting, cross-validation is essential, yet often neglected—even traditional correlation approaches using linear regression may overfit the data and not be generalizable to the greater population [30]. Finally, standard methods primarily conduct univariate analyses that consider each region or voxel separately, which can identify individual regions correlated with EEG rhythms, but leave open the question of how joint activity across networks is linked to fluctuating EEG dynamics. Predictive methods have become a powerful approach in fMRI [31,32] as they overcome these limitations: they can learn network-level information with flexible data-driven functions, and they inherently involve cross-validation to ensure results are generalizable. These approaches have identified relationships between neural activity and fMRI signals [33], and can predict a vigilance index from fMRI [34,35]. However, it is not yet known how brainwide dynamics are linked to specific EEG oscillations.

To identify the brainwide fMRI dynamics underlying alpha and delta rhythms, we designed a novel machine learning framework to predict EEG oscillations from fMRI data. We leveraged two independent datasets of simultaneous EEG and fast fMRI (TR<400 ms), exploiting the high temporal sampling of fast fMRI data to identify the dynamics underlying EEG oscillatory states as subjects spontaneously drifted in and out of sleep. We first show that we can successfully predict EEG dynamics from fMRI on out-of-sample subjects. We then find that the distinct fMRI patterns that predict each rhythm are replicated across datasets, and that predictive information is present even at the single-region level. Finally, we develop predictive approaches that probe the joint fMRI information across brain regions, and use them to identify distinct networks patterns underlying each EEG rhythm. Importantly, rather than attempting to localize the generators of these oscillations, we characterized the underlying network states associated with these rhythms. By probing relationships involving multiple brain regions, we uncovered how fMRI information predictive of EEG power is spatially distributed across the brain—not just identifying which regions are individually predictive (i.e. as in a mass univariate approach), but also understanding how the information contained in these regions is organized in networks. We found a large, diffuse network of primarily cortical regions that predict variations in delta power, whereas alpha power was predicted by two separable networks primarily representing the distinct dynamics of the visual system and the subcortical arousal control.

RESULTS

EEG alpha and delta power were predicted from brainwide fMRI data in held-out subjects

We first investigated whether it is possible to predict EEG oscillatory dynamics over time from simultaneously acquired fast fMRI data. Prior work successfully predicted EEG-based vigilance metrics (a ratio between two frequency bands) from simultaneous fMRI data and achieved a mean Pearson's linear correlation value of r=0.31 between predictions and targets when training on the same subjects [35]. However, predicting distinct oscillations is a more challenging task, as fluctuations in individual frequency bands are noisier and more complex than global variations in vigilance metrics, which collapse multiple distinct EEG properties into a single simplified metric. Additionally, our goal of generalizing to out-of-sample subjects could make predicting individual frequency bands particularly difficult as EEG power magnitudes exhibit strong individual differences. Thus, based on that prior work and the variability and noise of individual EEG frequency bands, we expected a ceiling in the range of r=0.3 (Pearson's linear correlation between held-out subject data and ground truth) on predictive performance. We hypothesized that by designing a flexible predictive model and training with the rich high temporal resolution data provided by fast fMRI, we could predict individual EEG oscillations.

We tested our models on held-out subjects—subjects which the model had never seen during training—to ensure overfitting did not inflate our performance estimates and to assess the generalizability of the model across subjects. We employed two separate datasets collected at different sites, to further confirm our methods and inferences were reproducible (dataset A: n=12, TR=0.378s; dataset B: n=9, TR=0.367s). For alpha analyses, spectrograms were first inspected to identify whether subjects had clear alpha rhythms since some individuals do not exhibit spontaneous alpha

(dataset A: n=11; dataset B: n=5). Each dataset used fast fMRI acquisitions to provide rich temporal information within short (~20 s, 60 TRs) time windows. We trained a linear machine learning model to predict EEG alpha and delta power, calculated in sliding 5-second windows from the same occipital electrodes (which are less susceptible to MR-induced noise), from simultaneously collected fMRI data, in subjects drifting in and out of sleep (Fig. 1b). The fMRI data was parcellated in 84 regions, consisting of 31 bilateral cortical regions, 7 bilateral subcortical regions, and 8 non-gray matter regions (white matter and ventricles).

The model was trained on sliding windows (length of 60 TRs, ~20s) of the parcellated fMRI data, with each segment yielding a single-point prediction of normalized EEG power (at the temporal center of the window). We centered the window of fMRI predictors on the EEG point being predicted to allow the model to flexibly learn whichever fMRI temporal patterns were most relevant, rather than assuming a specific lag between EEG and fMRI. The model was thus trained on 5,040 predictors (84 regions with 60 time points per region). Training was carried out separately for each dataset, iteratively on all of its subjects but one, with testing then performed on the held-out subject. The performance was compared to a shuffled control, with training and testing conducted using fMRI data circularly shifted in time (by 2000 TRs) in relation to the EEG prior to training and testing, breaking the true relationship between the modalities.

This whole-brain model successfully predicted EEG power in out-of-sample subjects for both alpha ('all' vs 'control', Fig. 1c, $p<0.01$) and delta ('all' vs 'control', Fig. 1e, $p<0.001$). Notably, our delta predictions surpassed our expected performance ceiling, displaying a mean correlation between predictions and ground truth of $r=0.37$, while our alpha predictions reached the expected performance ceiling at $r=0.30$, demonstrating robust prediction of dynamic fluctuations in EEG oscillations. Representative predictions obtained from these regional subsets demonstrate that the model successfully captured short- and long-timescale fluctuations in alpha and delta power (Fig. 1d, 1f).

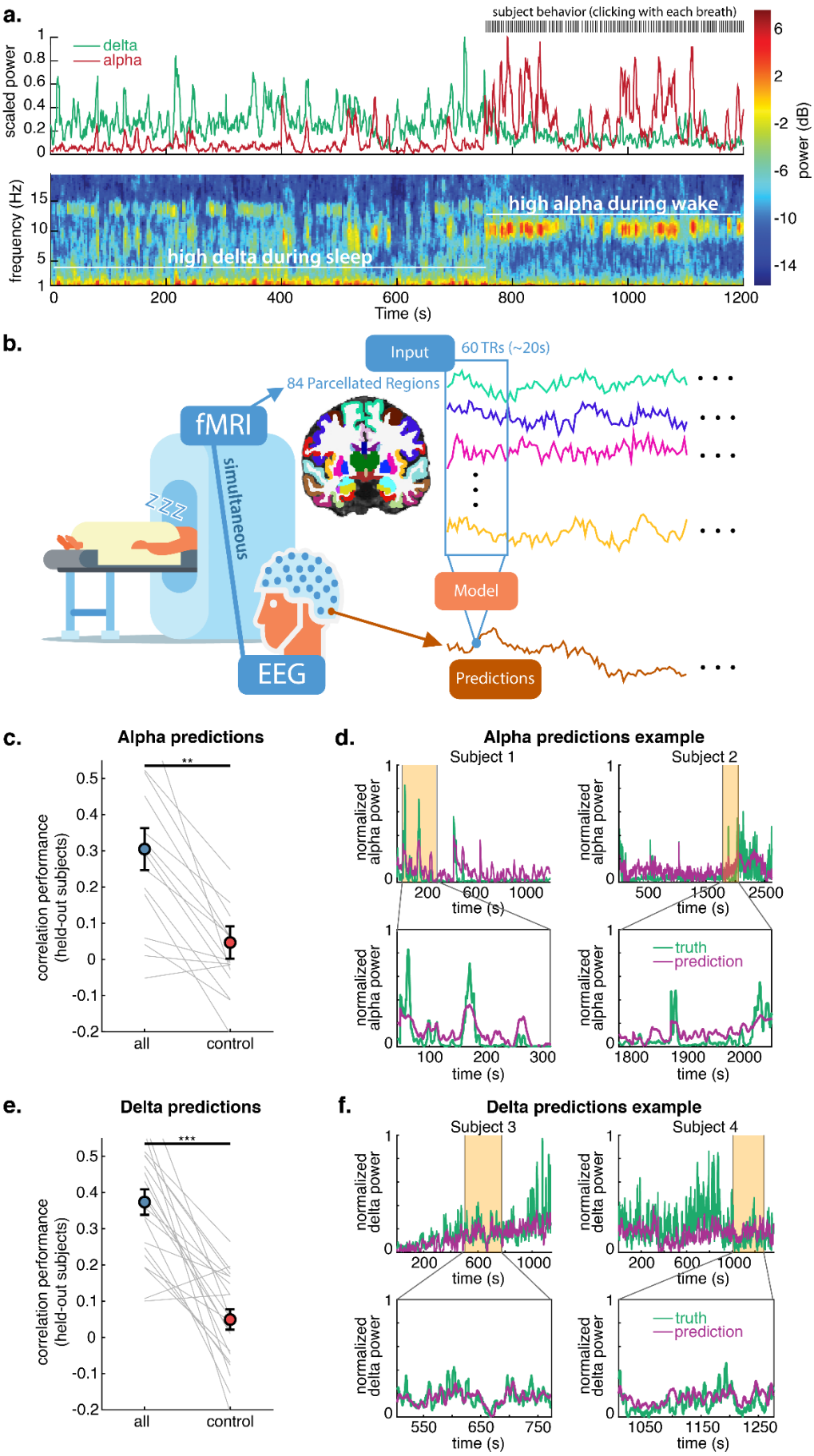


Fig. 1: Machine learning can predict dynamic fluctuations in EEG oscillations using brainwide fast fMRI timeseries. ***a.*** *Example of occipital alpha and delta EEG power fluctuating across sleep and wakefulness in one subject.* ***b.*** *EEG alpha and delta power were separately predicted from simultaneous fMRI data, collected while subjects rested inside the scanner with their eyes closed, drifting in and out of sleep. fMRI data was parcellated into 84 cortical, subcortical, and non-gray matter regions. Each point in the EEG power series, interpolated to match TR times, was predicted by 60 TRs of parcellated fMRI data.* ***c.*** *Correlation between alpha predictions (on held-out subjects) and ground truth, demonstrating that predictions using brainwide fMRI data ('all') are significantly better than the 'control' condition, in which the true relationship between EEG and fMRI was broken with a 2000 TR circular shift. ** p<0.01, t-test. Gray lines indicate held-out subject prediction performance. Error bars are SEM; N=16.* ***d.*** *Examples of alpha prediction in individual subjects, showing tracking of short- and long-timescale alpha fluctuations.* ***e.*** *Delta prediction performance, across the same conditions as panel c, shows significantly above-chance delta prediction. N=21. *** p<0.001, t-test.* ***f.*** *Examples of delta prediction in individual subjects.*

Cortex, subcortex, and non-gray matter regions differentially predict alpha and delta power, with patterns replicated across datasets

Given that we were able to predict EEG dynamics on out-of-sample subjects, we next explored how we could use this approach to identify which regions carry information about EEG power variations. Importantly, both neuronal and non-neuronal signals (such as the blood and cerebrospinal fluid) carry information about general fluctuations in arousal state [12]. We therefore first examined whether this predictive information was specific to gray matter regions in cortex or subcortex, or whether it was also carried in non-neuronal regions such as the white matter and the ventricles. We trained and tested new models using three conditions: cortex only, subcortex only, and non-gray matter regions only. These conditions were then compared with the 'all regions' and 'control' (fMRI data circularly shifted to break the true relationship with EEG) conditions described above. We tested performance in each dataset independently in order to assess whether predictive patterns would be replicated, and we also display statistical results when combining held-out subject performance from both datasets.

Alpha power (Fig. 2a) was most strongly predicted by the subcortex, with the subcortex-only model achieving significant prediction performance in each individual dataset ($p=0.006$, dataset A; $p=0.02$, dataset B). Combining the held-out subject values from the two datasets (which were trained and tested independently) allowed the cortical condition ($p=0.02$) and the condition utilizing all regions ($p=0.005$) to reach statistical significance due to the greater number of subjects. Alpha was not predictable by non-gray matter regions on their own. These results demonstrated that subcortical fMRI signals were a strong predictor of alpha oscillations, despite the lower signal-to-noise ratio of these brain regions, and that these predictions were not attributable to non-neuronal sources.

Delta power (Fig. 2b) was linked to a substantially different spatial pattern of fMRI signals. Delta was most strongly predicted by cortex, with significant predictive power in each dataset ($p=0.006$, dataset A; $p<0.001$, dataset B). Delta cortical predictions were on a similar level as predictions using the whole brain, suggesting that these cortical dynamics were sufficient to explain delta fluctuations on their own. Using subcortex-only, delta predictions reached statistical significance only when combined across datasets ($p=0.007$), demonstrating that information was also present in these regions. Interestingly, delta predictions using only the non-gray matter areas were also significantly better than control ($p=0.019$, dataset B; $p=0.024$, combined datasets), likely owing to the known coupling between delta power, cerebrospinal fluid flow (in the ventricles), and global BOLD [12].

This modeling approach therefore enabled identifying distinct spatial information linked to each EEG rhythm, which was replicated across datasets, with strong and statistically significant subcortical contributions to alpha power, and cortical dominance for delta power.

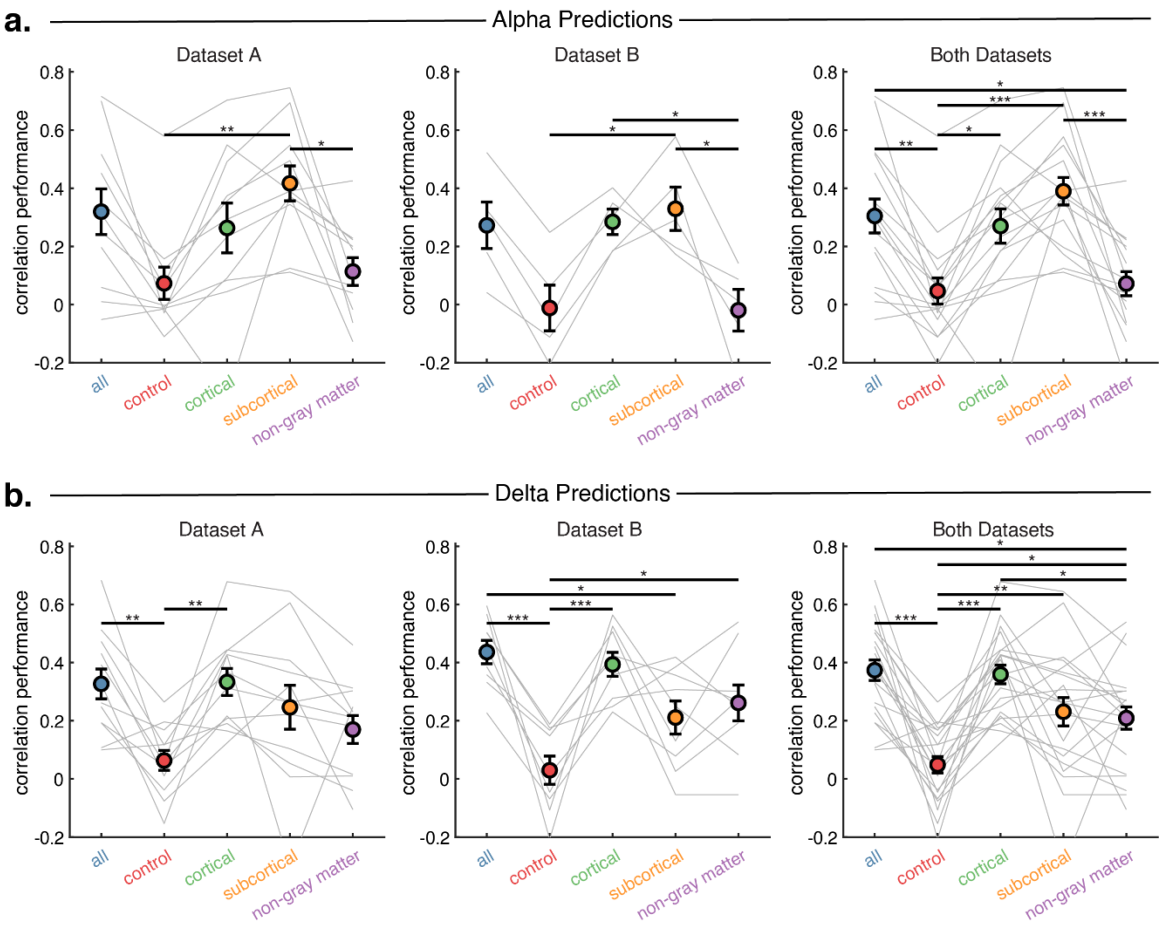


Fig. 2: Alpha and delta oscillations show distinct relationships to cortical, subcortical, and non-gray-matter fMRI signals. a. *Subcortical regions alone provide significant prediction of alpha rhythms. Circles show mean correlation between alpha power ground truth and model predictions (on held-out subjects) in the two independent datasets, and when combining held-out subject performance. Model was trained and tested under five conditions: cortical regions only, subcortical regions only, non-gray matter regions (ventricles and global white matter) only, all regions, and a control condition using all regions but with a 2000 TR circular shift of the fMRI data, breaking the true relationship between EEG and fMRI. * p<0.05; ** p<0.01; *** p<0.001; Tukey's HSD. Error bars are SEM. Gray lines indicate individual held-out subject prediction performance.* ***b.*** *Delta power is strongly predicted by cortical fMRI signals, but also by subcortex and non-gray matter regions, as shown by mean correlation between delta power ground truth and model predictions (on held-out subjects).*

Predictive fMRI information is detected in individual brain regions, with distinct spatial patterns for alpha and delta EEG power

To next identify finer-scale information beyond the coarse anatomical divisions of cortex and subcortex, we tested the information contained within each region in isolation by training and testing the model on a single bilateral gray matter region at a time. This prediction approach allowed us to learn flexible temporal functions shared across subjects linking the EEG to each fMRI area, rather than assuming a fixed pre-specified relationship. We generated spatial maps of individual brain regions' predictive performances for each EEG frequency band, to visualize the large-scale pattern of information content across the brain (Fig. 3). For visualization, the delta results combined the performance on held-out subjects for both datasets (n=21; model training was carried out independently in each dataset due to different imaging parameters). Alpha spatial maps included only Dataset A (n=11) as only 5 subjects in Dataset B displayed alpha rhythms, too few to assess prediction using only a single region. These maps demonstrated that single brain regions could be sufficient to achieve successful prediction of EEG oscillations (Fig. 3), but with substantially poorer performance than considering a larger network. A control analysis (Extended Data 1) confirmed the significance of the global predictive maps by comparing its patterns to shuffled data.

This individual-region analysis re-confirmed the predictive power of subcortical regions for alpha rhythms (Fig. 3a), revealing that alpha could be significantly predicted using only the thalamus (r=0.29, p<0.01), the dorsal striatum (separately as the putamen, r=0.28, p<0.01, and caudate, r=0.24, p<0.01), or the pallidum (r=0.27, p<0.01). This result highlighted that predictions can be successful even in individual deep brain regions with lower signal-to-noise ratio. The thalamus in particular, which plays key roles in both arousal control and sensory information flow, has consistently been associated with alpha power variations [17–21]. We also found that the cuneus (V1) could predict alpha variations (Fig. 3a; r=0.20, p=0.018), consistent with alpha's established role in the visual system [17–22]. However, our method also identified novel subcortical regions linked to alpha power—the dorsal striatum and pallidum—likely reflecting the model's ability to learn more flexible functions relating EEG and fMRI signals.

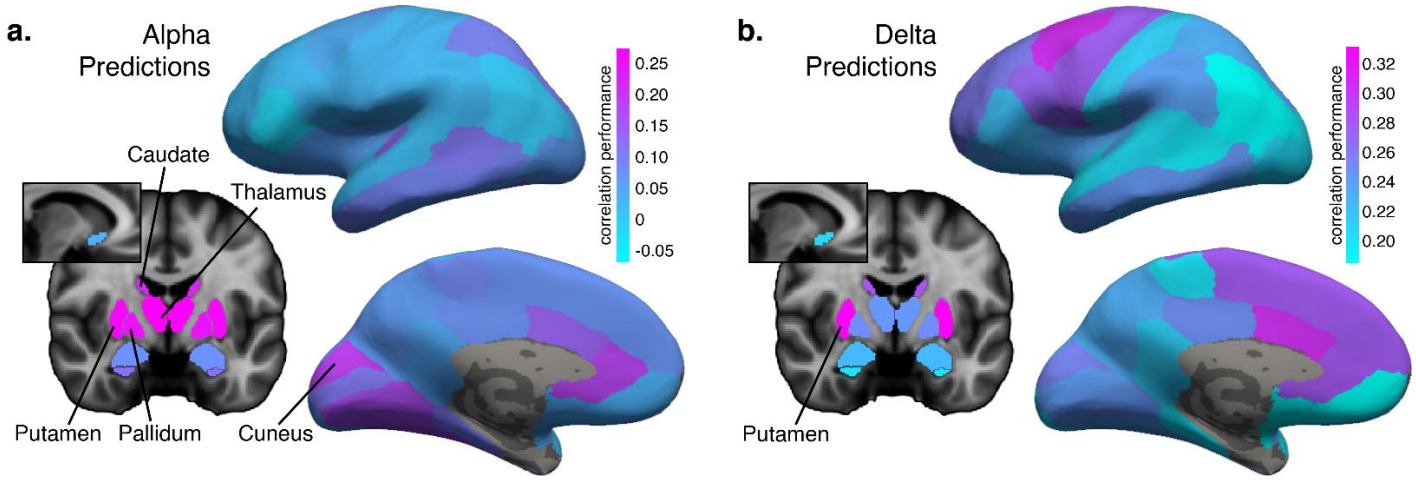


Fig. 3: Single-region predictions performance maps show which regions individually carry significant information about alpha and delta rhythms. a*. Prediction performance (mean correlation between alpha predictions on held-out subjects and ground truth) when model was trained on single bilateral regions to predict alpha. Text-labeled regions yielded alpha predictions that were significantly better than control.* ***b.*** *Prediction performance for delta power. Model was trained on single bilateral regions combined with non-gray matter areas to predict delta. Text-labeled regions combined with non-gray matter regions yielded predictions that were significantly better than using non-gray matter alone.*

When predicting delta, given that the model was able to predict EEG delta power from non-gray matter regions alone (Fig. 2), we included these regions in the delta prediction model in addition to each bilateral gray matter region. We then compared the prediction performance of this model to the model fit with only non-gray matter regions to identify gray matter regions which contained information about delta which was distinct from the non-neuronal components of the BOLD signal. In other words, this approach allowed us to identify predictive information that is likely to be specifically neural, to overcome a common challenge with analyses of the BOLD signal (which is heavily influenced by non-neuronal sources), thus providing an advantage over traditional correlation analyses. Only the putamen generated predictions significantly better than the non-gray matter regions by themselves (r=0.33, compared to non-gray matter r=0.21; p=0.017), indicating that, in isolation, it was the only region capable of offering information beyond what is contained in the ventricles and global white matter. We highlight that this finding does not mean that the putamen is the only region coupled to delta power, but rather that its activity possessed delta-predictive components that were substantially distinct from the non-neuronal signals. However, the poor predictions from individual cortical regions were in marked contrast to the coarse-grained analysis, which showed that the complete array of cortical regions generated superior delta predictions (Fig. 2). This suggests that a large integrated network of fMRI regions may be critical to accurately predict delta power variations—a question we sought to answer next.

Distinct fMRI spatial network sizes are needed to predict alpha and delta rhythms

Our individual-region approach demonstrated that unique predictive information capable of generalizing across subjects is present even on a single-region level, with distinct predictive maps for alpha and delta rhythms. However, oscillatory states likely reflect brainwide network dynamics spanning both the cortex and subcortex; thus, single-regions

predictive maps could miss larger-scale integrated networks that carry joint information about EEG oscillations. We therefore turned to network-level predictions, considering multiple brain regions in tandem to identify brainwide patterns that predict alpha and delta.

We first sought to assess the size of fMRI networks required to predict each EEG rhythm: in other words, how many distinct fMRI regions are needed for optimal predictions? If many regions are required for optimal predictions, then that would imply that predictive information is widely distributed across the brain, with low redundancy between regions. On the other hand, if only a small subset of regions is required for optimal predictions, this would suggest that predictive information is less widespread, and/or many regions contain redundant information about EEG rhythms. Simply imposing sparsity on the weights of our model via L1 regularization in order to reduce the number of regions used for prediction did not yield good performance. Therefore, to identify the number of optimal regions, we designed an iterative procedure to remove redundant information across regions. We iteratively removed low-weight regions from the model (starting with all 84 regions) while tracking performance changes. Low-weight regions imply a lack of relevant *unique* information; while such regions may still be predictive in isolation, their pattern of activity does not contain information that is not already substantially present in other regions.

We found that removing low-impact regions enabled us to achieve improved predictions of both alpha and delta power (r=0.49; Fig. 4a-d), substantially better than using all regions, demonstrating that we could achieve surprisingly robust prediction of EEG oscillations using a focal network of fMRI regions. A control analysis (Extended Data 2), in which randomly selected regions were dropped instead of uninformative regions, revealed no systematic improvements. This demonstrates that reducing the number of brain regions presented to the model is only beneficial if these brain regions had low weights associated with them. This performance increase eventually plateaued once the model reached a low number of regions, and subsequent reductions in region numbers caused performance to drop, reverting the benefits of feature selection. This turning point happened only at very low numbers for alpha predictions (Fig. 4a), indicating that unique alpha-predictive information is found in a remarkably small set of brain regions. Meanwhile, for delta predictions, the model plateaued at a much larger number of regions (Fig. 4d) with a more pronounced performance loss following the plateau, indicating that delta predictions benefit from a much larger spatial scale.

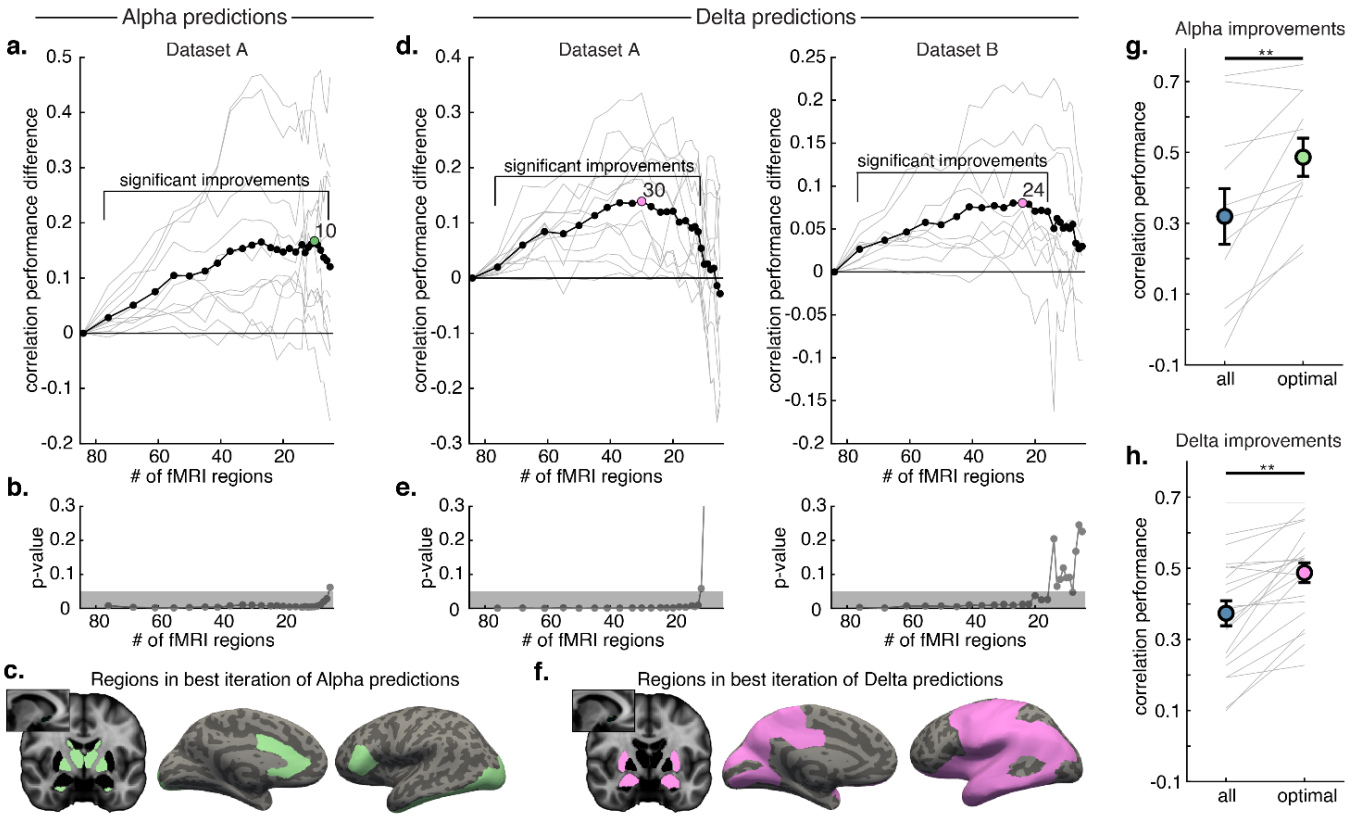


Fig. 4: Alpha rhythms can be predicted from a small spatial network, whereas delta requires information from a broader set of regions. *Iteratively removing fMRI regions associated with low weights increased model performance up until a plateau, with different spatial scales associated with the best alpha and delta predictions. As dataset B only contained 5 subjects with clear alpha rhythms, it was not included in this analysis.* ***a.*** *Alpha prediction performance peaked at 10 regions. The lowest-weight fMRI regions at each iteration were removed from the model, and then training and testing were performed again for the next iteration. Gray lines represent individual subjects. The iteration with best performance is marked in green. Successive iterations with significant improvements (as compared to all regions) are indicated.* ***b.*** *P-values (from t-tests) of performance improvements relative to using all regions, showing that improvements cease to be significant when model contains 5 or fewer fMRI regions.* ***c.*** *Representative regions present in the best iteration of alpha predictions are displayed in green (collapsed across hemispheres for visualization), demonstrating alpha's small spatial scale.* ***d.*** *Delta prediction performance, calculated as in panel a, showing performance peaking between 24-30 regions. The iteration with best performance is marked in pink.* ***e.*** *Delta p-values, calculated as in panel b, showing that performance ceases to be better than all regions when model contains 11 fMRI region for dataset A and 14 fMRI regions for dataset B.* ***f.*** *Regions shared by the two datasets in the mean of their best iterations of delta predictions are displayed in pink (collapsed across hemispheres for visualization), demonstrating delta's large spatial scale. See Extended Data 3 for details.* ***g.*** *Mean alpha prediction performance when using 'all' regions vs. when using the regions display in panel c ('optimal') demonstrate significant improvements with feature selection. ** p<0.01. Error bars are SEM; gray lines are individual subjects.* ***h.*** *Delta prediction performance also demonstrates significant improvements with feature selection. ** p<0.01.*

This analysis demonstrated a distinct spatial scale for each EEG oscillation, with delta predictions relying on a large number of brain regions, whereas alpha rhythms could be optimally predicted by a much smaller set of regions. To further illustrate the variable tipping points at which reducing the number of regions causes a loss of relevant information, we conducted paired t-tests between the held-out correlation performance of feature selection in each iteration, compared to when all regions were present. The resulting pattern of p-values again demonstrated that alpha predictions only degraded at a very low number of regions (Fig. 4b; increase when the model contained 5 regions), while delta predictions degraded at a higher number (Fig. 4e; increase at 11 regions for dataset A and at 14 regions for dataset B). Visualizing the regions in the model with highest performance showed distinct networks of cortical and subcortical regions that optimally predicted alpha (Fig. 4c) and delta (Fig. 4f). Performance improvements in this optimal iteration in relation to using all

regions were significant for both alpha (Fig. 4g, r=0.49 vs. r=0.32, p<0.01) and delta (Fig. 4h, r=0.49 vs. r=0.37, p<0.01), with more dramatic improvements for alpha, likely owing to its smaller spatial scale more strongly benefitting from feature selection.

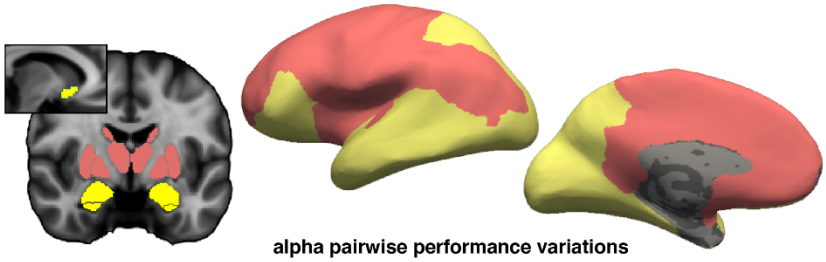
Since the model is stochastic and many regions are highly correlated with each other, the regions present in the best iterations (Fig. 4c, 4f) are representative samples of specific predictive networks, rather than the only uniquely predictive regions; they are shown only to illustrate the distinct spatial scales of predictive information. As an example, the putamen, which was shown to be highly predictive of alpha (Fig. 3a), was not present in the best iteration of alpha predictions (Fig. 4c); this is likely because its alpha-predictive information is largely redundant in relation to other regions that were preserved. This observation leads to the question of which larger networks are being represented by the subset of regions present in the best iteration; to answer this, we next developed a novel analysis to uncover networks of shared information that predict each EEG rhythm.

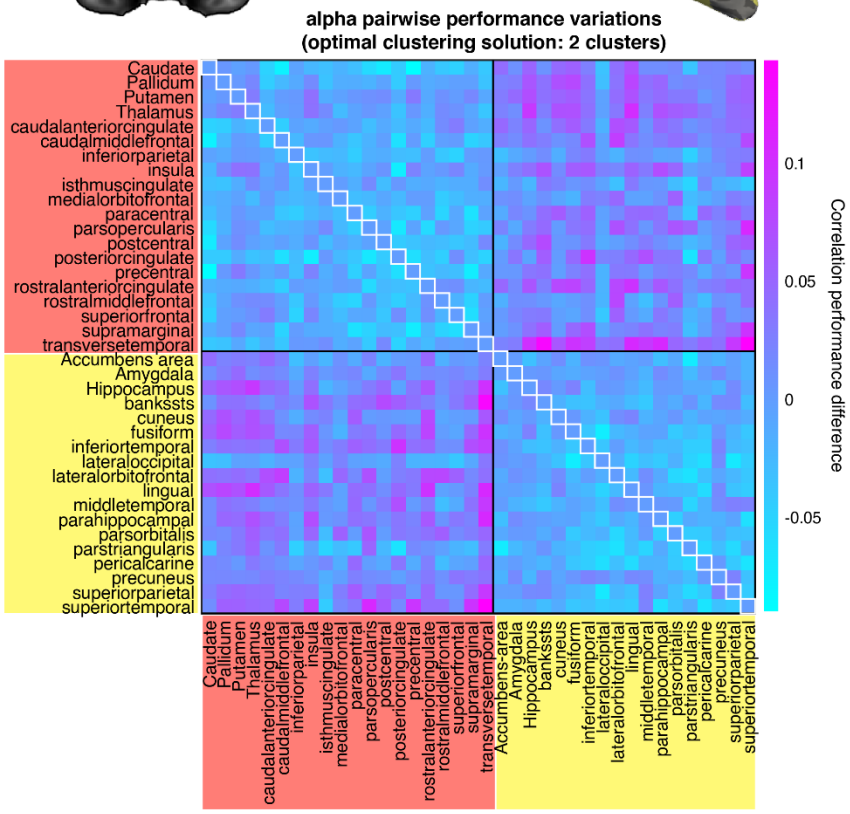
Two distinct networks predict alpha, while delta information is diffuse

To identify the specific network structure predictive of each oscillation, we developed a novel analysis to identify shared information across brain regions. We systematically assessed the predictive performance of each pairwise combination of bilateral brain regions in relation to the performance of its individual components, in order to evaluate whether regions contained distinct and complementary information (expressed as greater performance for the pair) or similar and redundant information (expressed as worse performance for the pair). In other words, we trained and tested the model on each pair of regions, and investigated the performance gains (or losses) of each pair of regions (each pair is shown in Extended Data 4), as compared to the individual regions within the pair (Fig. 3). As before, non-gray matter regions were also used when generating delta predictions, to enable us to identify fMRI information beyond that present in the white matter or ventricles. Next, we performed k-means clustering on these results, using the gap statistic to evaluate the optimal clustering solution (including the null model of no suitable clustering). By clustering the performance variations, we could identify the presence of different networks of information—groups of regions that share redundant information among themselves (resulting in performance losses when paired together), but which generate predictions better than the maximum of their pairs when paired with regions from other networks.

We found that alpha pairwise performance variations were optimally clustered in two groups (Extended Data 5), identifying two separate networks with distinct information content for predicting alpha power (Fig. 5a). This clustering behavior implies that regions were separable based on commonality of shared information: as a result, pairs of within-cluster regions yielded no performance benefits in relation to the maximum of their individual components (Fig. 5b; average of r=-0.0163 across all regions), while pairs of between-cluster regions yielded performance benefits (Fig. 5b; average of r=0.0265 across all regions; p<.001). Every region saw an increase in alpha-predictive performance when comparing between the mean of its within-cluster pairings and the mean of its between-cluster pairing (Fig. 5b).

a.





b.

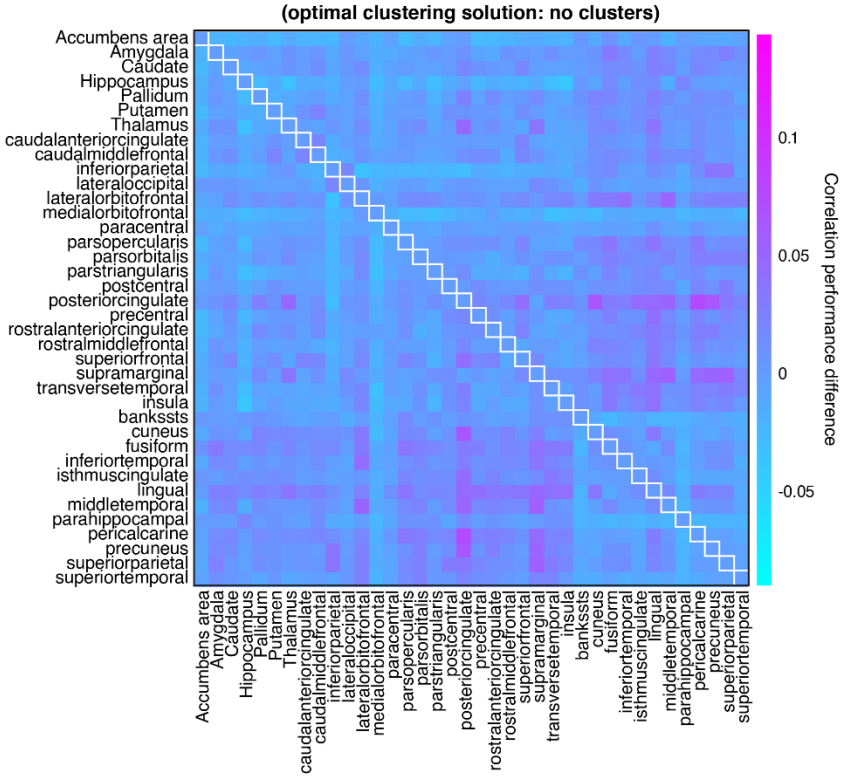
alpha mean performance variations for each region

	Within-cluster	Between-cluster
Caudate	-0.033	0.022
Pallidum	-0.012	0.038
Putamen	0.005	0.039
Thalamus	-0.001	0.045
caudalanteriorcingulate	-0.015	0.034
caudalmiddlefrontal	-0.022	0.032
inferiorparietal	-0.018	0.001
insula	-0.003	0.035
isthmuscingulate	-0.009	-0.006
medialorbitofrontal	-0.016	0.009
paracentral	-0.024	0.026
parsopercularis	-0.021	0.031
postcentral	-0.017	0.016
posteriorcingulate	-0.035	0.027
precentral	-0.021	0.020
rostralanteriorcingulate	-0.017	0.045
rostralmiddlefrontal	-0.016	0.007
superiorfrontal	-0.011	0.009
supramarginal	-0.024	0.027
transversetemporal	-0.015	0.073
Accumbens area	-0.012	0.010
Amygdala	-0.003	0.016
Hippocampus	-0.005	0.037
bankssts	-0.011	0.040
cuneus	-0.006	0.025
fusiform	-0.020	0.038
inferiortemporal	-0.021	0.044
lateraloccipital	-0.029	-0.007
lateralorbitofrontal	-0.016	0.037
lingual	-0.019	0.045
middletemporal	-0.008	0.022
parahippocampal	-0.016	0.035
parsorbitalis	-0.022	0.029
parstriangularis	-0.034	-0.005
pericalcarine	-0.029	0.009
precuneus	-0.001	0.017
superiorparietal	-0.015	0.033
superiortemporal	-0.029	0.052

0.07, 0.06, 0.05, 0.04, 0.03, 0.02, 0.01, 0, -0.01, -0.02, -0.03

Correlation performance difference

c.



d.

delta mean performance variations for each region

	Within-cluster	Between-cluster
Accumbens area	-0.015	-0.009
Amygdala	-0.002	0.010
Caudate	-0.001	0.008
Hippocampus	-0.015	-0.005
Pallidum	-0.005	0.009
Putamen	-0.002	0.006
Thalamus	0.000	0.007
caudalanteriorcingulate	-0.002	0.003
caudalmiddlefrontal	0.000	0.007
inferiorparietal	-0.015	0.001
lateraloccipital	-0.001	-0.001
lateralorbitofrontal	0.004	0.020
medialorbitofrontal	-0.017	-0.017
paracentral	-0.007	-0.003
parsopercularis	-0.003	0.014
parsorbitalis	-0.003	0.013
parstriangularis	-0.008	0.002
postcentral	-0.001	0.009
posteriorcingulate	0.007	0.033
precentral	-0.001	0.011
rostralanteriorcingulate	0.011	-0.001
rostralmiddlefrontal	0.008	-0.004
superiorfrontal	0.006	0.004
supramarginal	0.025	0.002
transversetemporal	0.014	-0.004
insula	0.010	-0.007
bankssts	-0.007	0.001
cuneus	0.003	0.010
fusiform	0.008	0.016
inferiortemporal	0.006	0.008
isthmuscingulate	0.003	0.008
lingual	0.017	0.020
middletemporal	0.007	0.010
parahippocampal	-0.011	0.001
pericalcarine	0.006	0.016
precuneus	0.007	0.011
superiorparietal	0.008	0.011
superiortemporal	-0.003	0.003

0.07, 0.06, 0.05, 0.04, 0.03, 0.02, 0.01, 0, -0.01, -0.02, -0.03

Correlation performance difference

Fig. 5: Clustering analysis on alpha pairwise performance variations reveal two distinct networks of alpha-predictive fMRI information, while delta-predictive information is not segregated in clusters. *Pairwise performance (correlation between predictions on held-out subjects and ground truth) variations were calculated as the performance gain or loss of a pair in relation to the maximum performance of its individual components; see Fig. 3 for individual performance values.* ***a.*** *Identified clusters (represented in yellow and red) and all pairwise performance variations for alpha.* ***b.*** *Mean alpha-predictive performance of each region based on its pair category. When regions were paired with others within the same clusters, they consistently underperformed in relation to when they were instead paired with regions from the other cluster.* ***c.*** *The clustering analysis revealed that delta did not possess a suitable clustering solution, as evidenced by no systematic pairwise performance variations following a 2-cluster fitting.* ***d.*** *Mean delta-predictive performance of each region based on its pair category.*

The smaller of the two clusters of alpha information (Fig. 5a, yellow), included visual cortex (both the ventral and dorsal visual streams), prefrontal regions associated with visual cognition, and three subcortical regions that are responsive to contents of visual information (the accumbens, amygdala and hippocampus). Alpha rhythms have previously been proposed to be involved in visual information flow from lower-level perceptual areas to higher-order cognitive regions [15,36]; this cluster we identified includes key regions linked to visual information transfer.

The larger cluster of alpha information (Fig. 5a, red) included the thalamus, the dorsal striatum (putamen and caudate) and the pallidum—which were also the four significantly predictive subcortical regions from the single-region analysis (Fig. 3a). These regions are involved in arousal control [37], and the dorsal striatum projects to the pallidum, which in turn projects to the thalamus [38], suggesting that these regions work in tandem to mediate alpha activity. Finally, several cortical regions fell into this same cluster: given that the subcortical regions outlined above are implicated in arousal, and that arousal variations influence the fMRI signal and are detected to some degree across all brain regions [39,40], this red cluster could reflect arousal fluctuations. Finally, we emphasize that, while information is shared within each network, it is not shared equally—some regions express this information more strongly than others, as highlighted by our single-region analysis (Fig. 3a).

In contrast, the gap statistic revealed that delta results did not present a suitable clustering solution (Extended Data 5). The lack of a suitable clustering solution for delta predictions highlights its diffuse spatial property, with no separable groups that share distinct information from each other. To further demonstrate this, we performed a two-cluster analysis on the pairwise delta predictions (Fig. 5c), showing that this forced clustering solution did not exhibit the systematic cluster differences present on the two-cluster alpha analysis (Fig. 5b, 5d). Delta-predictive information, therefore, was not expressed as separable networks; instead, it was represented as a large, distributed network, with predictive performance benefiting from a greater number of regions (see Fig. 4). It is likely that subtle differences in fMRI cortical activity come together to predict variations in EEG delta power—as highlighted both by the coarse-grained analysis (Fig. 2b), which showed the conjunction of cortical areas to be highly predictive of delta power, and by the spatial scales analyses (Fig. 4f), where the regions preserved by the model were primarily cortical. Finally, this distributed nature of delta-predictive information, along with our finding that this information has a significant non-gray matter component, may explain why traditional univariate correlations between fMRI activity and continuous delta power have not previously been reported. This highlights our method's ability to make inferences about relationships involving multiple brain regions, including accounting for non-neural contributions.

DISCUSSION

While delta and alpha rhythms are reliable indicators of brain arousal states with established relationships to cognition, the brainwide activity underlying their variations has been unclear. We used machine learning to predict continuous fluctuations in EEG power in the alpha and delta bands from fMRI data, and achieved robust performance in held-out subjects, with a mean correlation value of r=0.49 between predictions (on held-out subjects) and ground truth following feature selection (Fig. 4g-h). This prediction-based approach demonstrated reproducibility across datasets, and was able to uncover EEG-predictive information in individual fMRI regions. We discovered that distinct network patterns were

coupled to each rhythm, with strikingly differentiated spatial scales: two focused, separable networks predicted alpha rhythms from moment to moment, whereas a highly diffuse network of primarily cortical regions provided information about delta oscillations.

Alpha rhythms have been widely studied in the EEG-fMRI literature, with focus on identifying its specific generators by using mass-univariate approaches to correlate individual fMRI regions to EEG oscillations. Our findings replicate prior univariate relationships between the thalamus/V1 and alpha power [17–24], and owing to our predictive approach unconstrained by a pre-defined HRF, we also identified novel relationships between alpha power and fMRI activity in the dorsal striatum and pallidum. Given that alpha is not only tied to arousal states but also to shifts in high-level cognition [11,15], a key question was not which focal regions generate alpha directly, but instead which broader networks underlie the dynamic variations in alpha power. Our results demonstrate that separable networks are involved with alpha variations, suggesting they may reflect two distinct potential mechanisms of occipital alpha power expression: the dynamics within the visual system itself, versus the control from the subcortical arousal system. The thalamus has often been viewed as the primary orchestrator of alpha rhythms [18–20,24], and our results indicate that this thalamic control is likely part of a circuit involving other key subcortical regions: the dorsal striatum and the pallidum. Meanwhile, the separable alpha-predictive information identified in the visual system points to a distinct and dissociable mechanism of alpha expression, which may reflect that cortical neurons can intrinsically fire in the alpha range [41] independent of thalamic input.

We found a distinct network pattern linked to delta power, which indexes several sleep oscillatory dynamics, including slow oscillations and slow waves. While EEG-fMRI studies have identified fMRI activity locked to individual slow wave events [42–44], the power variations in the delta band we examine here represent distinct information, as they index the frequency and amplitude of a continuous spectrum of oscillatory dynamics, and are a metric of brain state. However, past work that analyzed univariate correlations between delta power and fMRI activity had failed to identify significant relationships [21]. Our work explains this null result by showing that delta power predictions rely on combined information across a large number of fMRI regions; this can be contrasted with alpha, which is readily predicted by a small network, and unlike delta displays a large number of significant univariate relationships (see Fig. 3).

Delta predictions not only relied on a large network, but represented an accumulation of diffuse and unified activity distributed throughout the cortex. What dynamics within this broad cortical fMRI network could be predicting delta power? Prior work has shown that arousal fluctuations are associated with cortical infra-slow propagating fMRI activity [40], and that infra-slow fMRI is coupled to delta power [45]; it is therefore possible that this propagating infra-slow fMRI activity—which is global in nature—was core to the delta predictions. The varying phase across brain regions may explain why delta-predictive information required a large number of regions, yet was not separable in different networks. In addition, delta power itself exhibits complex propagating spatiotemporal dynamics [46], and can be both local and global in nature [4,47], suggesting that this diffuse fMRI predictive network may also represent a collection of diverse local slow wave dynamics. Future work could explore these questions by identifying traveling wave events in simultaneous EEG-fMRI and investigating whether distinct spatiotemporal features of infra-slow fMRI oscillations predict local or global delta power.

The strong role of the putamen in predicting variations in delta power was unexpected: it was the only region which contained enough unique predictive information by itself to surpass non-gray matter regions. What aspect of its activity is contributing to delta predictions? Activity in the putamen is linked to memory consolidation, which is facilitated by slow waves [3,48]. Past work has identified experience-dependent activity in putamen during sleep [49] and increased putamen activity during spindle-slow wave complexes [50], which play a causal role in memory consolidation [1,6–8]. While slow waves have well-established direct cortical generators [51,52], the subcortex still exerts a role in controlling or initiating cortical slow waves [53–55]. Our results point to a potentially unique involvement of the putamen in slow wave activity that warrants future investigation as neuroscientists continue to decipher the network activity that underlies slow waves.

Past EEG-fMRI analyses of alpha and delta rhythms have primarily relied on univariate correlations, considering each brain region or voxel separately [17–21,23], yet fMRI activity is strongly characterized by network-level activity, with variations in multiple brain regions coming together to generate distinct spatiotemporal patterns [56,57]. The differences between our single-region analysis and our multi-region analyses highlight that distinct patterns can be discovered when going beyond

univariate correlations: as an example, had our work been focused exclusively on univariate relationships, we would have concluded that the putamen is the only region predictive of delta variations (Fig. 3b), but our network-level analyses revealed the importance of a large array of cortical regions for capturing delta variations (Fig. 2b, Fig. 4d-f).

Predictive machine learning has supported neuroscientific insights in event-related classification analyses across a range of neuroimaging modalities [58–61]. Our method aimed to extend this framework to a challenging task: achieving continuous, dynamic predictions in the absence of external stimuli and task conditions. Our approach can be generalized to other continuous crossmodal analyses in neuroimaging data, to identify how spontaneous brain dynamics are coupled in a broad range of questions. It has several advantages, such as correcting for non-neuronal effects (by including data from the white matter and ventricles as predictors) without requiring regression that can introduce errors [62–64], ensuring that effects are generalizable across subjects through cross-validation, and allowing for the identification of multi-region relationships that represent network-scale information. This approach could be broadly useful for identifying distributed dynamics underlying brain rhythms and states.

In conclusion, we identified the fMRI network correlates of EEG delta and alpha rhythms across sleep and wakefulness, going beyond mass univariate relationships by extending a predictive approach to continuous cross-modal data. We surpassed our expected performance ceiling, successfully predicting alpha and delta fluctuations on held-out subjects. Alpha and delta were captured by distinct fMRI components, with alpha predictions being driven by two separable networks—one weighted heavily toward the visual system, and the other by arousal-regulatory subcortical regions—and delta being represented by a diffuse array of primarily cortical regions with unique contributions from the putamen. Our work expands our understanding of networks involved with human arousal states, identifying the large-scale activity patterns underlying the EEG oscillations that define sleep and wakefulness. Furthermore, this approach establishes a framework for predictive analyses of EEG-fMRI data that could be broadly employed to identify brainwide network states linked to neuronal oscillatory dynamics.

ACKNOWLEDGEMENTS

This study was funded by National Institutes of Health grants R01-AG070135, U19-NS128613, and U19-NS123717; the Sloan Fellowship, the McKnight Scholar Award, the Pew Biomedical Scholar Award, the Simons Collaboration on Plasticity in the Aging Brain (811231), and the One Mind Rising Star Award. Resources were provided by NSF Major Research Instrumentation grant BCS-1625552.

METHODS

EEG-fMRI acquisition

Both datasets analyzed in this work were collected as part of other studies. Collection for Dataset A was ongoing while our analyses were conducted; the reported number of subjects for this dataset is the number that had been collected at the time our analyses began. Subsequently collected subjects were not counted or considered for the purposes of this study. Dataset B is from a previously published study [12].

Dataset A

We obtained written informed consent from 15 healthy young adults; 12 were analyzed for the purposes of this study. Subject inclusion was based on sufficient EEG data quality to track continuous power throughout the scan (as determined by visual examination of EEG spectrograms). The protocol was approved by the Boston University Charles River Campus Institutional Review Board. Subjects were scanned on a 3T Siemens Prisma scanner with a 64-channel head and neck coil. Anatomical references were acquired using a 1 mm isotropic T1-weighted multi-echo MPRAGE [65]. Functional runs were acquired using TR = 0.378 s, TE = 31 ms, 2.5 mm isotropic voxels, 40 slices, Multiband factor=8, blipped CAIPI shift = FOV/4, flip angle=37°, no in-plane acceleration. Three successive runs of 25 minutes each were obtained per subject.

EEG was acquired using MR-compatible 32-channel EEG caps fitted with 4 carbon wire loops (BrainProducts GmbH, Germany) at a sampling rate of 5000 Hz. EEG acquisition was synchronized to the scanner 10 MHz clock to reduce aliasing of high-frequency gradient artifacts. Additional sensors were used to record systemic physiology: respiration was measured simultaneously using an MRI-safe pneumatic respiration transducer belt around the abdomen and pulse was measured with a photoplethysmogram (PPG) transducer (BIOPAC Systems, Inc., Goleta, CA, USA). Physiological signals were acquired at 2000 Hz using Acqknowledge software and were aligned with MRI data using triggers sent by the MRI scanner.

Dataset B

We used a previously published dataset which had collected data from 13 healthy young adults; 9 were analyzed for the purposes of this study. Subject inclusion was based on sufficient EEG data quality to track continuous power throughout the scan (as determined by visual examination of EEG spectrograms). The protocol was approved by the Massachusetts General Hospital Institutional Review Board. Subjects were scanned on a 3T Siemens Prisma scanner with a 64-channel head-and-neck coil. Anatomical references were acquired using a 1 mm isotropic T1-weighted multi-echo MPRAGE [65]. Functional runs acquired 40 interleaved BOLD-weighted EPI slices with 2.5 mm^3 isotropic voxels. fMRI protocols consisted of a single-shot gradient echo SMS-EPI [66] with MultiBand factor=8, matrix=92×92, blipped CAIPI shift=FOV/4, TR=367 ms, TE=30-32 ms (changed midway through the study due to a software upgrade) nominal echo-spacing=0.53 ms, flip angle=32-37°, and no in-plane acceleration. VERSE factor was set between 1 and 1.5 depending on individual subject SAR constraints. Individual runs could last up to 2 hours. If runs ended earlier, subsequent runs would be started up to a maximum total scan duration of 2.5 hours as long as subjects were still comfortable and sleeping.

EEG was acquired using MR-compatible 256-channel geodesic nets and a NA410 amplifier (Electrical Geodesics, Inc., Eugene, OR USA) at a sampling rate of 1000 Hz. EEG acquisition was synchronized to the scanner 10 MHz clock to reduce aliasing of high-frequency gradient artifacts. The scanner cryopump was temporarily shut off during EEG acquisition to reduce vibrational artifact. To acquire reference signals to be used for EEG noise removal [67], subjects wore a reference layer cap composed of an isolating vinyl layer and conductive satin layer on the head, with grommets inserted to allow electrodes to pass through and make contact with the scalp, while other electrodes remained isolated from the scalp and recorded the noise, resulting in a total of 30–36 EEG electrodes per subject. Physiological signals were simultaneously acquired using a Physio16 device (Electrical Geodesics, Inc., Eugene, OR USA). ECG was measured through two disposable electrodes placed on the chest diagonally across the heart, with an MR-compatible lead (InVivo Corp, Philips). Respiration was measured through a piezoelectrical belt (UFI systems, Morro Bay, CA USA) around the chest.

EEG pre-processing

Dataset A

Gradient artifacts were removed using average artifact subtraction [68], with moving average of the previous 20 TRs. Ballistocardiogram artifacts were removed from each EEG channel using signals from the 4 carbon wire loops using the sliding Hanning window regression method from the EEGlab CWL toolbox [69] with the following parameters: window = 25 s and a lag = 0.09 ms. Clean EEG signals were re-referenced to the average of EEG channels.

Dataset B

Gradient artifacts were removed using average artifact subtraction [68], with moving average of the previous 20 TRs. Electrodes were then re-referenced to the common average, computing this separately for electrodes contacting the head, and those placed on the reference layer. Channels on the cheeks and borders of the reference cap were excluded from the common average. Ballistocardiogram artifacts were removed using regression of reference signals from the isolated EEG electrodes [70]. Since there was a larger number of noise electrodes than signal electrodes, the regression was performed after subsampling the noise electrodes, using only every fourth isolated electrode. Because the position of and physiological noise influences on the electrodes can vary over the long recording times used here, we implemented a dynamic time-varying regression of the reference signals. Beta coefficients for the best-fit regression within 30 s sliding time windows were fit using least-squares; these beta values were then linearly interpolated over the nonoverlapping

windows. The resulting interpolated beta value at every time point was then used for a local subtraction of the reference signals from the modeled EEG recording. This regression was performed individually for each EEG channel.

fMRI pre-processing

The cortical surface was reconstructed from the MEMPRAGE volume using Freesurfer version 6 [71]. All functional runs were slice-time corrected using FSL version 6 (slicetimer; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki) [72] and motion corrected to the middle frame using AFNI (3dvolreg; https://afni.nimh.nih.gov/). Each motion-corrected run was then registered to the anatomical data using boundary-based registration (bbregister) [73].

For Dataset A, physiological noise was removed using HRAN, open-source software implementing a statistical model of harmonic regression with autoregressive noise [74].

For Dataset B, Physiological noise removal was performed using dynamic regression based on the concept of RETROICOR [75] and adapted for fast fMRI as follows. The respiratory trace was bandpass filtered between 0.16–0.4 Hz using a finite impulse response filter and the instantaneous phase was computed as the angle of the Hilbert transform. The cardiac peaks were detected automatically using the FASST toolbox (http://www.montefiore.ulg.ac.be/~phillips/FASST.html) and the phase was modeled as varying linearly between each identified peak. Sine and cosine basis functions using the phase of the signal and its second harmonic were generated as regressors for physiological noise. This regression was performed over 1000 s windows sliding every 400 s to enable high-quality physiological noise removal as the heart rate and respiratory rate varied throughout the scan. No spatial smoothing was applied.

Data preparation

Data preparation, model training, and statistical analyses were conducted in MATLAB unless noted otherwise. EEG power was extracted from occipital electrodes via the multitaper spectrogram method [76,77]. Spectrograms were obtained using a 5 second window, sliding every 0.1 s. The power time series was then extracted for the frequencies of interest: 1.1-4 Hz for delta (chosen to minimize contamination from low-frequency artifacts in EEG-fMRI) and 8.5-12 Hz for alpha (chosen to minimize contamination from theta and spindle power while still covering subject variations in alpha range). Power in the 40-52 Hz range was also extracted for the purpose of artifact detection (see below).

For dataset A, power was extracted from three occipital electrodes (O1, O2, and Oz) and then averaged. For dataset B, which used a reference layer and therefore did not have identical channel layouts across subjects, power was extracted from the single occipital electrode closest to Oz which had good data quality based on visual inspection of the spectrogram. Power time series were then interpolated to match fMRI TR timings.

fMRI segmentation was performed in FreeSurfer based on the Desikan-Killiany atlas [78]. The following regions were obtained: 31 bilateral cortical regions (the entire cortex minus the temporal pole, frontal pole, and entorhinal regions, which were not covered by the acquisition volume in all subjects), 7 bilateral subcortical regions (thalamus, pallidum, amygdala, caudate, putamen, accumbens, hippocampus), and 8 non-grey matter regions (left- and right- total white matter as an index of global blood flow, left- and right- lateral and inferior lateral ventricles, third ventricle, and fourth ventricle). Anatomical labels were interpolated to functional space, and functional voxels with 70% or above of overlap with the label were averaged to generate the mean region time series. The first 20 TRs of each run were discarded, and the fMRI timeseries of each region (separately for each subject) was z-scored.

To mark data points unsuitable for analyses, the following criteria were used: 1) TRs with motion above 0.3 mm; 2) z-scored EEG power in the 40-52Hz range with a value above 1 (adjusted to be lower for runs with particularly noisy data); 3) detected local outliers (distance larger than 6 local scaled median absolute deviation (MAD)) within a sliding window of 1000 TRs (for dataset B, a threshold of 3 MAD was used instead due to noisier data; thresholds were also stricter for subjects with particularly noisy data). The outlier detection method was applied to delta power and power in the 40-52Hz range (but not to alpha power, as its naturally fast fluctuations would be erroneously detected as outliers; however, three subjects with noisy data underwent alpha band outlier removal based on a smaller window). Datapoints that fulfilled any

of these criteria were then removed, and the remaining data was saved as separate segments to avoid creation of spurious continuities (segments below 60 TRs were discarded, as 60 TRs of continuous fMRI are required to predict each EEG point). EEG data for each subject was then rescaled to 0-1.

Model training and testing

Following the above, fMRI data was split into overlapping 60-TR sections, with the goal of predicting the EEG point at the center (equivalent to the position of the 30th fMRI point). Given that no fMRI points would be available for EEG points at the edges of the data segments following artifact removal, the first 29 and the last 30 EEG points of each segment were discarded.

At each training fold, all data for one subject was held-out as test data. The model was then trained on the remaining subjects using the Matlab *fitrlinear* function. The function fits a weight value for each input feature (each TR of each input fMRI region), along with a bias value, and a lambda regularization term. The regularization term multiples the sum of squares of all model weights to penalize large weights and prevent the model from overfitting (L2 regularization). During training, the model adjusts its weights and bias using stochastic gradient descent.

In each fold, the lambda regularization term was adjusted through Bayesian hyperparameter optimization by partitioning the training data (i.e. the held-out subject was not used) in 5 folds over 50 iterations, with repartitioning at every iteration. The best lambda value was then selected, and the model was trained on all training data at once. While the model frequently chose very similar lambda values in each held-out subject fold, this additional computational effort resulted in a strict training regimen where no data from the held-out subject was ever used to tune the model. In earlier explorations of the model, the Bayesian hyperparameter optimization was also allowed to select between L1 and L2 regularization, but given that L2 consistently outperformed L1, we fixed L2 as the regularization option to reduce computational processing time.

Once the optimal lambda value was determined and the weights and biases were fit on the training data, the model was fixed and used to generate predictions on the held-out subject. We highlight that normalization (see 'Data preparation' above) was done separately within each subject so that at no point did the held-out subject data affect training data. A Pearson's linear correlation coefficient was then calculated between the predictions made for the held-out subject and the ground truth.

Statistical analyses

Major Anatomical groups analysis

The model underwent the training and testing routine described above using different groups of fMRI regions as input: only the subcortical regions, only the cortical regions, only the non-gray matter regions, and all of them together. Finally, the model also went through the same training routine using all the regions as input, but with a 2000 TR circular shift of the EEG in relation to the fMRI (as a control condition). Once the correlation performance was obtained for all subjects and conditions, an ANOVA was conducted, with subject ID as a random variable, and condition as a fixed variable. Correction for multiple comparisons was performed with Tukey's HSD (honest significant difference). Multiple comparison correction was done within each dataset (A/B/combined) and analyses type (alpha/delta) only.

Bootstrapping for single-region analysis

To assess whether prediction performance differences between individual regions were not the result of the ordering of random noise, we shuffled the region labels within each subject, then averaged across the shuffled regions, ordered the means, and fit a line to the ordered means. This was repeated 1000 times; the resulting lines were then averaged, and a 95% confidence interval was calculated. Verification that the line fit to the real (non-shuffled) data was outside these bounds was taken to mean that region differences (i.e. spatial pattern of predictive performance across fMRI regions) were significant and not due to ordering of random noise. Following this analysis, paired t-tests were conducted to assess at which point the best-performing regions were no longer superior to control (or to non-gray matter regions in the case of the delta analyses).

Spatial scales analysis

After training and testing of the model using all 84 parcellated regions, the model's beta weights were extracted, and weights (across all TRs) for each region were squared and summed. The regions that displayed weights whose sum of squares were below the bottom 10% (as distribution percentiles) were then removed from the model, which was then once again trained and tested. This was repeated up until 5 regions were left in the model (or 9 plus the non-gray matter regions, for the delta analysis where non-gray matter regions were ineligible to be removed).

Given the large number of regions present in the delta model at its peak performance, and our access to two datasets, we examined which regions (regardless of hemisphere) were commonly shared by the model's peak iteration across the two datasets. The model peaked at a similar number of regions (33 for dataset A and 27 for dataset B, for a mean of 30, not counting the 8 non-gray matter regions which were ineligible to be dropped). Observing the iteration with 30 regions of the model for each dataset, dataset A displayed 22 unique regions (collapsing across hemispheres), while dataset B displayed 28. Within these, 18 regions were shared between the two datasets (with A being the ceiling at 22; this amounts to 82% of them being shared).

Clustering of pairwise performance

The model was trained and tested on each pair of gray matter regions. For delta, the non-gray matter regions were also added to each gray matter pair. For each pair, the performance values from the single-region analysis (correlation between predictions and ground truth) were noted, and the highest performance was subtracted from the pair's performance. This value represented the performance gain (or loss) of the pair in relation to its best individual component. These values were then used for clustering.

To assess the optimal clustering solution (including the null hypothesis of no suitable clustering), we used the gap statistic [79]. The gap value is defined as the difference between the within-cluster dispersion obtained by Monte Carlo sampling from a reference distribution (generated from a uniform distribution over a box aligned with the principal components of the data) and the actual within-cluster dispersion obtained by clustering the real data. The ideal number of clusters is defined as the smallest number of clusters that satisfies $\mathrm{Gap}(k) \geq GAPMAX - \mathrm{SE}(GAPMAX)$, where Gap(k) is the gap value for the k-clusters solution, GAPMAX is the largest gap value, and SE(GAPMAX) is the corresponding standard error.

REFERENCES

1. Geva-Sagiv, M. *et al.* Augmenting hippocampal–prefrontal neuronal synchrony during sleep enhances memory consolidation in humans. *Nat Neurosci* **26**, 1100–1110 (2023).

2. Helfrich, R. F., Mander, B. A., Jagust, W. J., Knight, R. T. & Walker, M. P. Old Brains Come Uncoupled in Sleep: Slow Wave-Spindle Synchrony, Brain Atrophy, and Forgetting. *Neuron* **97**, 221-230.e4 (2018).

3. Holz, J. *et al.* EEG sigma and slow-wave activity during NREM sleep correlate with overnight declarative and procedural memory consolidation. *J Sleep Res* **21**, 612–619 (2012).

4. Huber, R., Ghilardi, M. F., Massimini, M. & Tononi, G. Local sleep and learning. **430**, 4–7 (2004).

5. Latchoumane, C. F. V., Ngo, H. V. V., Born, J. & Shin, H. S. Thalamic Spindles Promote Memory Formation during Sleep through Triple Phase-Locking of Cortical, Thalamic, and Hippocampal Rhythms. *Neuron* **95**, 424-435.e6 (2017).

6. Maingret, N., Girardeau, G., Todorova, R., Goutierre, M. & Zugaro, M. Hippocampo-cortical coupling mediates memory consolidation during sleep. *Nat Neurosci* **19**, 959–964 (2016).

7. Marshall, L., Helgadóttir, H., Mölle, M. & Born, J. Boosting slow oscillations during sleep potentiates memory. *Nature* **444**, 610–613 (2006).

8. Ngo, H. V. V., Martinetz, T., Born, J. & Mölle, M. Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. *Neuron* **78**, 545–553 (2013).

9. Schreiner, T., Petzka, M., Staudigl, T. & Staresina, B. P. Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nat Commun* **12**, (2021).

10. Benwell, C. S. Y. *et al.* Frequency and power of human alpha oscillations drift systematically with time-on-task. *Neuroimage* **192**, 101–114 (2019).

11. Schneider, D., Herbst, S. K., Klatt, L. I. & Wöstmann, M. Target enhancement or distractor suppression? Functionally distinct alpha oscillations form the basis of attention. *European Journal of Neuroscience* **55**, 3256–3265 (2022).

12. Fultz, N. E. *et al.* Coupled electrophysiological, hemodynamic, and cerebrospinal fluid oscillations in human sleep. *Science (1979)* **366**, 628–631 (2019).

13. Adamantidis, A. R., Gutierrez Herrera, C. & Gent, T. C. Oscillating circuitries in the sleeping brain. *Nat Rev Neurosci* **20**, 746–762 (2019).

14. Pfurtscheller, G., Stancák, A. & Neuper, Ch. Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology* **24**, 39–46 (1996).

15. Clayton, M. S., Yeung, N. & Cohen Kadosh, R. The many characters of visual alpha oscillations. *European Journal of Neuroscience* **48**, 2498–2508 (2018).

16. Xie, L. *et al.* Sleep Drives Metabolite Clearance from the Adult Brain. *Science (1979)* **342**, 373–377 (2013).

17. Martínez-Montes, E., Valdés-Sosa, P. A., Miwakeichi, F., Goldman, R. I. & Cohen, M. S. Concurrent EEG/fMRI analysis by multiway Partial Least Squares. *Neuroimage* **22**, 1023–1034 (2004).

18. de Munck, J. C. *et al.* The hemodynamic response of the alpha rhythm: An EEG/fMRI study. *Neuroimage* **35**, 1142–1151 (2007).

19. Feige, B. *et al.* Cortical and subcortical correlates of electroencephalographic alpha rhythm modulation. *J Neurophysiol* **93**, 2864–2872 (2005).

20. Goldman, R., Stern, J., Engel, J. & Cohen, M. S. Simultaneous EEG and fMRI of the alpha rhythm. *Neuroreport* 2487–2492 (2002) doi:10.1097/01.wnr.0000047685.08940.d0.

21. Tyvaert, L., LeVan, P., Grova, C., Dubeau, F. & Gotman, J. Effects of fluctuating physiological rhythms during prolonged EEG-fMRI studies. *Clinical Neurophysiology* **119**, 2762–2774 (2008).

22. Moosmann, M. *et al.* Correlates of alpha rhythm in functional magnetic resonance imaging and near infrared spectroscopy. *Neuroimage* **20**, 145–158 (2003).

23. Laufs, H. *et al.* EEG-correlated fMRI of human alpha activity. *Neuroimage* **19**, 1463–1476 (2003).

24. Liu, Z. *et al.* Finding thalamic BOLD correlates to posterior alpha EEG. *Neuroimage* **63**, 1060–1069 (2012).

25. Dang-Vu, T. T. *et al.* Cerebral correlates of delta waves during non-REM sleep revisited. *Neuroimage* **28**, 14–21 (2005).

26. Handwerker, D. A., Ollinger, J. M. & D'Esposito, M. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* **21**, 1639–1651 (2004).

27. Lewis, L. D., Setsompop, K., Rosen, B. R. & Polimeni, J. R. Fast fMRI can detect oscillatory neural activity in humans. *Proceedings of the National Academy of Sciences* **113**, E6679–E6685 (2016).

28. Polimeni, J. R. & Lewis, L. D. Imaging faster neural dynamics with fast fMRI: A need for updated models of the hemodynamic response. *Progress in Neurobiology* vol. 207 Preprint at https://doi.org/10.1016/j.pneurobio.2021.102174 (2021).

29. Lindquist, M. A., Meng Loh, J., Atlas, L. Y. & Wager, T. D. Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling. *Neuroimage* **45**, S187–S198 (2009).

30. Shen, X. *et al.* Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat Protoc* **12**, 506–518 (2017).

31. Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E. & Gallant, J. L. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* **532**, 453–458 (2016).

32. Finn, E. S. & Rosenberg, M. D. Beyond fingerprinting: Choosing predictive connectomes over reliable connectomes. *Neuroimage* **239**, 118254 (2021).

33. Lake, E. M. R. *et al.* Simultaneous cortex-wide fluorescence Ca2+ imaging and whole-brain fMRI. *Nat Methods* **17**, 1262–1271 (2020).

34. Chang, C. *et al.* Tracking brain arousal fluctuations with fMRI. *Proc Natl Acad Sci U S A* **113**, 4518–4523 (2016).

35. Falahpour, M., Chang, C., Wong, C. W. & Liu, T. T. Template-based prediction of vigilance fluctuations in resting-state fMRI. *Neuroimage* **174**, 317–327 (2018).

36. Zumer, J. M., Scheeringa, R., Schoffelen, J. M., Norris, D. G. & Jensen, O. Occipital Alpha Activity during Stimulus Processing Gates the Information Flow to Object-Selective Cortex. *PLoS Biol* **12**, (2014).

37. Lazarus, M., Chen, J. F., Urade, Y. & Huang, Z. L. Role of the basal ganglia in the control of sleep and wakefulness. *Curr Opin Neurobiol* **23**, 780–785 (2013).

38. Haber, S. N. & Calzavara, R. The cortico-basal ganglia integrative network: The role of the thalamus. *Brain Res Bull* **78**, 69–74 (2009).

39. Setzer, B. *et al.* A temporal sequence of thalamic activity unfolds at transitions in behavioral arousal state. *Nat Commun* **13**, (2022).

40. Raut, R. V *et al.* Global waves synchronize the brain's functional systems with fluctuating arousal. *Sci Adv* **7**, 1–16 (2021).

41. Silva, L. R., Amitai, Y. & Connors, B. W. Intrinsic Oscillations of Neocortex Generated by Layer 5 Pyramidal Neurons. *Science (1979)* **251**, 432–435 (1991).

42. Betta, M. *et al.* Cortical and subcortical hemodynamic changes during sleep slow waves in human light sleep. *Neuroimage* **236**, 118117 (2021).

43. Ilhan-Bayrakcı, M., Cabral-Calderin, Y., Bergmann, T. O., Tüscher, O. & Stroh, A. Individual slow wave events give rise to macroscopic fMRI signatures and drive the strength of the BOLD signal in human resting-state EEG-fMRI recordings. *Cereb Cortex* **32**, 4782–4796 (2022).

44. Dang-Vu, T. T. *et al.* Spontaneous neural activity during human slow wave sleep. *Proc Natl Acad Sci U S A* **105**, 15160–15165 (2008).

45. Mitra, A. *et al.* Human cortical-hippocampal dialogue in wake and slow-wave sleep. *Proc Natl Acad Sci U S A* **113**, E6868–E6876 (2016).

46. Massimini, M., Huber, R., Ferrarelli, F., Hill, S. & Tononi, G. The sleep slow oscillation as a traveling wave. *Journal of Neuroscience* **24**, 6862–6870 (2004).

47. Malerba, P., Whitehurst, L. N., Simons, S. B. & Mednick, S. C. Spatio-temporal structure of sleep slow oscillations on the electrode manifold and its relation to spindles. *Sleep* **42**, 1–14 (2019).

48. Skelin, I. *et al.* Coupling between slow waves and sharp-wave ripples engages distributed neural activity during sleep in humans. *Proceedings of the National Academy of Sciences* **118**, e2012075118 (2021).

49. Ribeiro, S. *et al.* Long-lasting novelty-induced neuronal reverberation during slow-wave sleep in multiple forebrain areas. *PLoS Biol* **2**, 126–137 (2004).

50. Baena, D. *et al.* Functional differences in cerebral activation between slow wave-coupled and uncoupled sleep spindles. *Front Neurosci* **16**, 1–15 (2023).

51. Timofeev, I., Grenier, F., Bazhenov, M., Sejnowski, T. J. & Steriade, M. Origin of Slow Cortical Oscillations in Deafferented Cortical Slabs. *Cerebral Cortex* **10**, 1185–1199 (2000).

52. Steriade, M., Nunez, A. & Amzica, F. Intracellular analysis of relations between the slow (< 1 Hz) neocortical oscillation and other sleep rhythms of the electroencephalogram. *The Journal of Neuroscience* **13**, 3266–3283 (1993).

53. Schreiner, T., Kaufmann, E., Noachtar, S., Mehrkens, J.-H. & Staudigl, T. The human thalamus orchestrates neocortical oscillations during NREM sleep. *Nat Commun* **13**, 5231 (2022).

54. David, F. *et al.* Essential Thalamic Contribution to Slow Waves of Natural Sleep. *The Journal of Neuroscience* **33**, 19599–19610 (2013).

55. Lewis, L. D. *et al.* Thalamic reticular nucleus induces fast and local modulation of arousal state. *Elife* **4**, (2015).

56. Bolt, T. *et al.* A parsimonious description of global functional brain organization in three spatiotemporal patterns. *Nat Neurosci* **25**, 1093–1103 (2022).

57. Shine, J. M. *et al.* Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nat Neurosci* **22**, 289–296 (2019).

58. Jacob, L. P. L. & Huber, D. E. Neural Habituation Enhances Novelty Detection: an EEG Study of Rapidly Presented Words. *Comput Brain Behav* **3**, 208–227 (2020).

59. Sarma, A., Masse, N. Y., Wang, X.-J. & Freedman, D. J. Task-specific versus generalized mnemonic representations in parietal and prefrontal cortices. *Nat Neurosci* **19**, 143–149 (2016).

60. Spellman, T. *et al.* Hippocampal–prefrontal input supports spatial encoding in working memory. *Nature* **522**, 309–314 (2015).

61. Chang, L. J., Gianaros, P. J., Manuck, S. B., Krishnan, A. & Wager, T. D. A sensitive and specific neural signature for picture-induced negative affect. *PLoS Biol* **13**, 1–28 (2015).

62. Lindquist, M. A., Geuter, S., Wager, T. D. & Caffo, B. S. Modular preprocessing pipelines can reintroduce artifacts into fMRI data. *Hum Brain Mapp* **40**, 2358–2376 (2019).

63. Chen, J. E., Jahanian, H. & Glover, G. H. Nuisance Regression of High-Frequency Functional Magnetic Resonance Imaging Data: Denoising Can Be Noisy. *Brain Connect* **7**, 13–24 (2017).

64. Bright, M. G., Tench, C. R. & Murphy, K. Potential pitfalls when denoising resting state fMRI data using nuisance regression. *Neuroimage* **154**, 159–168 (2017).

65. van der Kouwe, A. J. W., Benner, T., Salat, D. H. & Fischl, B. Brain morphometry with multiecho MPRAGE. *Neuroimage* **40**, 559–569 (2008).

66. Setsompop, K. *et al.* Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn Reson Med* **67**, 1210–1224 (2012).

67. Luo, Q., Huang, X. & Glover, G. H. Ballistocardiogram artifact removal with a reference layer and standard EEG cap. *J Neurosci Methods* **233**, 137–149 (2014).

68. Allen, P. J., Josephs, O. & Turner, R. A Method for Removing Imaging Artifact from Continuous EEG Recorded during Functional MRI. *Neuroimage* **12**, 230–239 (2000).

69. van der Meer, J. N. *et al.* Carbon-wire loop based artifact correction outperforms post-processing EEG/fMRI corrections—A validation of a real-time simultaneous EEG/fMRI correction method. *Neuroimage* **125**, 880–894 (2016).

70. Chowdhury, M. E. H., Mullinger, K. J., Glover, P. & Bowtell, R. Reference layer artefact subtraction (RLAS): A novel method of minimizing EEG artefacts during simultaneous fMRI. *Neuroimage* **84**, 307–319 (2014).

71. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–781 (2012).

72. Woolrich, M. W. *et al.* Bayesian analysis of neuroimaging data in FSL. *Neuroimage* **45**, S173–S186 (2009).

73. Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* **48**, 63–72 (2009).

74. Agrawal, U., Brown, E. N. & Lewis, L. D. Model-based physiological noise removal in fast fMRI. *Neuroimage* **205**, (2020).

75. Glover, G. H., Li, T. Q. & Ress, D. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med* **44**, 162–167 (2000).

76. Thomson, D. J. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE* **70**, 1055–1096 (1982).

77. Prerau, M. J., Brown, R. E., Bianchi, M. T., Ellenbogen, J. M. & Purdon, P. L. Sleep Neurophysiological Dynamics Through the Lens of Multitaper Spectral Analysis. (2017) doi:10.1152/physiol.00062.2015.

78. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).

79. Tibshirani, R., Walther, G. & Hastie, T. Estimating the number of clusters in a data set via the gap statistic. *J R Stat Soc Series B Stat Methodol* **63**, 411–423 (2001).