Behavioral/Cognitive

Signal Complexity of Human Intracranial EEG Tracks Successful Associative-Memory Formation across Individuals

©Timothy C. Sheehan,¹* ©Vishnu Sreekumar,¹* Sara K. Inati,² and Kareem A. Zaghloul¹

¹Surgical Neurology Branch and ²Office of the Clinical Director, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892

Memory performance is highly variable among individuals. Most studies examining human memory, however, have largely focused on the neural correlates of successful memory formation within individuals, rather than the differences among them. As such, what gives rise to this variability is poorly understood. Here, we examined intracranial EEG (iEEG) recordings captured from 43 participants (23 male) implanted with subdural electrodes for seizure monitoring as they performed a paired-associates verbal memory task. We identified three separate but related signatures of neural activity that tracked differences in successful memory formation across individuals. High-performing individuals consistently exhibited less broadband power, flatter power spectral density slopes, and greater complexity in their iEEG signals. Furthermore, within individuals across three separate time scales ranging from seconds to days, successful recall was positively associated with these same metrics. Our data therefore suggest that memory ability across individuals can be indexed by increased neural signal complexity.

Key words: complexity; human memory; iEEG; theta power

Significance Statement

We show that participants whose intracranial EEG exhibits less low-frequency power, flatter power spectrums, and greater sample entropy overall are better able to memorize associations, and that the same metrics track fluctuations in memory performance across time within individuals. These metrics together signify greater neural signal complexity, which may index the brain's ability to flexibly engage with information and generate separable memory representations. Critically, the current set of results provides a unique window into the neural markers of individual differences in memory performance, which have hitherto been underexplored.

Introduction

Some people consistently have better memory than others. This variability in memory performance among individuals, and even within individuals from moment to moment, is quite familiar in our daily lives. In the study of memory, however, this variability has largely been viewed as a problem that needs to be addressed through proper experimental design. As a result, the neural mechanisms that give rise to such variability have been relatively

Received Aug. 23, 2017; revised Dec. 22, 2017; accepted Dec. 30, 2017.

Author contributions: T.C.S., V.S., and K.A.Z. designed research; T.C.S., V.S., and S.K.I. performed research; T.C.S. and V.S. analyzed data; T.C.S., V.S., S.K.I., and K.A.Z. wrote the paper.

This work was supported by the Intramural Research Program of the National Institute for Neurological Disorders and Stroke. This work was also partially supported by the Defense Advanced Research Projects Agency's Restoring Active Memory Program (Cooperative Agreement N66001-14-2-4032). The views, opinions, and/or findings contained in this material are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. government. We thank John Wittig Jr. and Julio Chapeton for helpful comments on the manuscript. We are indebted to all patients who have selflessly volunteered their time to participate in this study.

*T.C.S. and V.S. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Kareem A. Zaghloul, Surgical Neurology Branch, NINDS, National Institutes of Health Building 10, Room 3D20 10 Center Drive, Bethesda, MD 20892-1414. Email: kareem.zaghloul@nih.gov.

DOI:10.1523/JNEUROSCI.2389-17.2017

Copyright © 2018 the authors 0270-6474/18/381744-12\$15.00/0

unexplored. Understanding the source of such variability among individuals can provide valuable insights into how the brain is able to successfully form and retrieve memories.

Studies of memory have typically attempted to eliminate the variability in neural activity and memory performance among individuals by regressing it out. In many paradigms evaluating memory-related changes in oscillatory activity, for example, data within each individual are normalized so as to only examine relative changes in activity when events are either successfully remembered or forgotten, yielding what has been termed the subsequent memory effect (SME). Positive and negative SMEs have been reported in different frequency bands (Hanslmayr et al., 2012; Hanslmayr and Staudigl, 2014), yet interpreting these effects has been difficult given conflicting reports of positive lowfrequency SMEs in some studies (Sederberg et al., 2003; Osipova et al., 2006; Hanslmayr et al., 2011) and negative low-frequency SMEs in others (Sederberg et al., 2006; Guderian et al., 2009; Fell et al., 2011). Hence, normalized SMEs studied in isolated frequency bands may not provide a complete description of the neural correlates of memory. Moreover, these approaches have not addressed the larger question of how variability in neural activity may be related to variability in memory performance.

An alternative and complementary approach that has emerged in response to the conflicting SME data is to describe the changes in low-frequency and high-frequency activity as arising from the same phenomenon, one that produces an overall change in the structure of the entire power spectral density (PSD; Voytek et al., 2015). Spectral power decreases linearly with frequency on a log-log scale over a broad range of frequencies (Miller et al., 2009; Milstein et al., 2009; Dehghani et al., 2010; He et al., 2010; He, 2014). Importantly, neuronal activation results in a flatter PSD slope (He, 2011), reflecting decreases in lower-frequency power and increases in higher-frequency power (Podvalny et al., 2015). These findings have led to the suggestion that flattening of the PSD slope, and the associated changes in spectral power, may therefore be a signature of increased asynchronous neuronal activity (Ray and Maunsell, 2011; Burke et al., 2015; Voytek and Knight, 2015).

Viewed from an information-coding perspective, the PSD slope and oscillatory power of a neural signal, by indicating the extent of synchrony in the underlying neural activity, may be a proxy for neural-signal complexity and underlying information content (Hanslmayr et al., 2012). Direct measures of complexity of neural signals, such as sample entropy, have supported this suggestion by demonstrating that more complex brain dynamics underlie enhanced cognitive performance (McIntosh et al., 2008), likely signifying a greater capacity to encode and process information. Indeed, several groups have advanced the notion that complexity in neural activity is functionally relevant and affords greater flexibility for cognitive processing (Stein et al., 2005; MacDonald et al., 2006; Faisal et al., 2008; Deco et al., 2009, 2013; Garrett et al., 2011, 2013; Grady and Garrett, 2014; Sleimen-Malkoun et al., 2015). Therefore, these metrics may together reflect a general capacity for processing information that may be particularly relevant for memory formation.

In this scenario, then, a possible explanation for the variability in memory performance among individuals is that different brains may exhibit differences in complexity, allowing a greater number of unique cognitive states relevant for encoding memories. We investigate this possibility here by examining changes in spectral power, PSD slope, and the sample entropy of neural signals captured from intracranial electrodes as participants perform a paired-associates verbal episodic-memory task. We were specifically interested in whether these metrics exhibit differences among individuals and changes within individuals across time that correlate with memory performance. We found that such measures of complexity and general information processing are indeed behaviorally relevant when forming and retrieving memories.

Materials and Methods

Participants. Forty-three participants [23 male; age range, 13–59 years old; 32.1 ± 11.7 (mean \pm SD) years old] with drug-resistant epilepsy underwent a surgical procedure in which platinum recording contacts were implanted subdurally on the cortical surface as well as deep within the brain parenchyma. For each participant, the clinical team determined the placement of the contacts so as to best localize epileptogenic regions (Fig. 1*C*). Preoperative clinical fMRI testing results were available for 37 participants, and 36 of these participants exhibited fMRI activity consistent with left language dominance. The Institutional Review Board approved the research protocol, and informed consent was obtained from the participants or their guardians. Data from a subset of participants were initially collected and analyzed for previous publications (Yaffe et al., 2014, 2017; Greenberg et al., 2015; Haque et al., 2015). Computational analyses were performed using custom-written Matlab (MathWorks) scripts.

Paired-associates task. Each participant performed a paired-associates verbal memory task (Fig. 1A). In the task, participants were sequentially shown a list of word pairs (encoding period) and then later cued with one

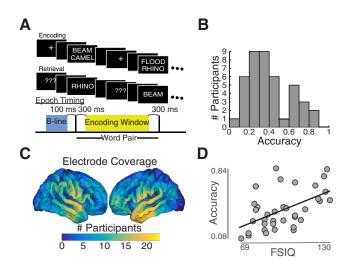


Figure 1. Paired-associates task and subject distribution. **A**, Paired-associates memory task schematic. **B**, Average performance distribution across subjects. Distribution is bimodal ranging between 0.05 and 0.84 with a median accuracy of 0.36 (N=43). **C**, Electrode coverage by spatial region of interest. Colormap reflects number of electrodes within 12.5 mm. **D**, Correlation between FSIQ and accuracy across subjects ($r_s=0.51$, p=0.0017, N=35). Line is standard least-squares regression line.

word from each pair selected at random (retrieval period), and were instructed to say the associated word into a microphone. Each participant performed one of two versions of the task that had slight differences in the experimental details. As the tasks did not differ in the fundamental objectives and performance was indistinguishable between groups, we combined the data from both sets of tasks for subsequent analyses.

A single experimental session for each participant consisted of 15 or 25 lists, where each list contained either four or six pairs of common nouns shown on the center of a laptop screen, depending on whether the participant completed the first or second version of the task respectively. Although different participants performed the task with different list lengths, the number of pairs in a list was kept constant for each participant. Words were chosen at random and without replacement from a pool of high-frequency nouns and were presented sequentially and appeared in capital letters at the center of the screen. Study word pairs were separated from their corresponding recall cue by a minimum lag of two study or test items. During the study period (encoding), each word pair was preceded by an orientation stimulus (either a '+' or a row of capital X's) that appeared on the screen for 250–300 ms followed by a blank interstimulus interval (ISI) between 500 and 750 ms. Word pairs were then presented stacked in the center of the screen for 2500 ms followed by a blank ISI of 1500 ms with a jitter of 75 ms in the first version of the task, or for 4000 ms followed by a blank ISI of 1000 ms in the second version. Following the presentation of the list of word pairs in the second version of the task, participants completed an arithmetic distractor task of the form A + B + C = ? for 20 s.

In both task versions, one word was randomly chosen during the test period (retrieval) from each of the presented pairs and presented in random order, and the participant was asked to recall the other word from the pair by voicing a response. Each cue word was preceded by an orientation stimulus (a row of question marks) that appeared on the screen for 250-300 ms followed by a blank ISI of 500-750 ms. Cue words were then presented on the screen for 3000 ms followed by a blank ISI of 4500 ms in the first version of the task, or for 4000 ms followed by a blank ISI of 1000 ms in the second version. Participants could voice their response any time during the recall period after cue presentation. We manually designated each recorded response as correct, intrusion, or pass. A response was designated as pass when no vocal response was made or when the participant said "pass." We defined all intrusion and pass trials as incorrect trials. A single experimental session contained 60, 100, or 150 word pairs or trials, depending on the task version. We included only participants who engaged in ≥2 separate sessions of the paired-associates

task such that each participant completed between two and five sessions taking 31.1 \pm 1.7 (mean \pm SEM) min each with a median of 24.8 h between sessions.

Intracranial EEG recordings. Intracranial EEG (iEEG) signals were referenced to a common electrode and were resampled at 1000 Hz. We applied a fourth-order 2 Hz stopband Butterworth notch filter at 60 Hz to eliminate electrical-line noise. The testing laptop sent synchronization pulses via an optical isolator into a pair of open lines on the clinical recording system to synchronize the iEEG recordings with behavioral events.

We collected electrophysiological data from 3756 subdural and depth recording contacts (PMT; Ad-Tech Medical Instrument). Subdural contacts were arranged in both grid and strip configurations with an intercontact spacing of 5 or 10 mm. Contact localization was accomplished by coregistering the postoperative CTs with the postoperative MRIs using both FSL [FMRIB (Oxford Centre for Functional MRI of the Brain) Software Library] Brain Extraction Tool (BET) and FLIRT (FMRIB's Linear Image Registration Tool) software packages and mapped to both MNI and Talairach space using an indirect stereotactic technique and OsiriX Imaging Software DICOM (Digital Imaging and Communications in Medicine) viewer package. The resulting contact locations were subsequently projected to the cortical surface of a population-average brain. Preoperative MRIs were used when postoperative MRI images were not available.

We took several steps to reduce the influence of pathologic activity on our results. First, we excluded from further analysis 414 electrodes identified clinically as having prominent interictal or ictal activity based on the evaluation of a board-certified epileptologist. To minimize the extent of transient epileptic activity (interictal discharges) in the remaining electrodes, we then performed an iterative cleaning procedure on the common averaged electrode signals to eliminate both electrodes and trials with a kurtosis of >2.8 SDs or a variance of >2.2 SDs from the persistent sample's mean. This procedure eliminated an additional 698 electrodes from further examination as well as 1047 of 12,650 individual trials. The remaining 2644 electrode contacts recorded over 11,576 trials were used to generate our dataset.

We analyzed iEEG data using bipolar referencing to reduce volume conduction and confounding interactions between adjacent electrodes (Nunez and Srinivasan, 2006). We defined the bipolar montage in our dataset based on the geometry of iEEG electrode arrangements. For every grid, strip, and depth probe, we isolated all pairs of contacts positioned immediately adjacent to one another; bipolar signals were then found by finding the difference in the signals between each pair of immediately adjacent contacts. The resulting bipolar signals were treated as new virtual electrodes (referred to as electrodes throughout the text) originating from the midpoint between each contact pair. All subsequent analyses were performed using these derived bipolar signals. Our dataset consisted of 2347 bipolar referenced electrodes derived from the set of original monopolar electrodes that remained following our cleaning procedure as described above.

Data analyses and spectral power. We computed the spectral power for each bipolar electrode at every time point during the experimental session by convolving the raw iEEG signal with complex valued Morlet wavelets (wavelet number, 6) to obtain the magnitude of the signal at each of 30 logarithmically spaced frequencies ranging from 3 to 180 Hz. We squared and log-transformed the magnitude of the continuous-time wavelet transform to generate a continuous measure of instantaneous power. During every trial, we convolved each wavelet with two separate time windows: a baseline period extending from 600 to 100 ms before word-pair presentation and an encoding period from 300 ms after wordpair presentation until 300 ms before the offset of the word pair from the display screen (Fig. 1A). In addition, we computed power for ten 2000 ms windows from the beginning of the clinical recording segment before task-specific activity began for each session and averaged those to get an extra-task window. We included an additional 1000 ms buffer on either side of each time window to minimize any edge effects. This buffer was not subsequently analyzed.

To examine the relation between overall raw power and performance across participants, we used the above measures of raw spectral power. To examine how changes in power on individual trials affected performance, we z-scored each session's power values independently to remove the effects of across-participant and session-level variations.

Calculating spectral slope. To understand how spectral power changes as a function of frequency, we calculated spectral slope. For each participant, we computed an average PSD across all trials and electrodes and computed slope in log–log space across the broadband range of 10 to 100 Hz (Podvalny et al., 2015). To identify the general $1/f^{\alpha}$ slope of the spectrum and avoid contamination of narrowband oscillations, we used a robust fitting algorithm with bisquare weighting (Matlab robustfit.m function). Additionally, we computed slope over a range of frequency values and spectral widths as described in Results. We defined spectral width using units of octaves such that the spectral width of a given slope was equal to the \log_2 of the ratio of the highest frequency to the lowest frequency.

Calculating sample entropy. We used a metric of sample entropy to measure the complexity of the iEEG signal. Sample entropy, by construction, is a measure of predictability. Specifically, the sample entropy (SampEn) of a time series is the negative natural logarithm of the conditional probability that any two subsequences of length m within the series and are similar within a tolerance r remain similar at length m+1 (Richman and Moorman, 2000). Two patterns that are close together in m-dimensional space and that remain close together in m+1-dimensional space indicate fewer irregularities or less complexity in the signal. Similarity is measured using the Chebyshev distance between the two subsequences. A smaller value of SampEn denotes greater repetitiveness and less complexity in a given signal.

For an embedding dimension m and a tolerance r, the formal equations for the calculation of sample entropy for a given time series of total length N are as follows (Eqs. 1–3; Vakorin and McIntosh, 2012; Sokunbi et al., 2013):

$$SampEn = -\ln \frac{U^{m+1}}{U^m} \tag{1}$$

where,
$$U^m = [N-m]^{-1} \sum_{i=1}^{N-m} \frac{B_i}{N-m-1}$$
 (2)

and,
$$B_i = \sum_{j \neq i}^{N-m} H[r - ||x_m(i) - x_m(j)||]$$
 (3)

where \parallel refers to the maximum norm, $x_m(i)$ is a vector $\{xi, xi+1, \ldots, xi+m-1\}$ within our time series, and H is the Heaviside step function. B_i is the number of m-dimensional vectors that are within a tolerance of r from a given template $x_m(i)$, excluding self-matches. B_i is normalized by the number of possible matches, N-m-1, and averaged over the N-m possible template vectors to get U^m , the probability of any two m-dimensional vectors in a series being within a Chebyshev distance r.

An embedding dimension m of 2 and a tolerance r of 0.2 * std(x(t)) were used in all analyses. Of note, the number of three-element matching template sequences is necessarily less than or equal to the number of two-element matching template sequences, implying that the ratio $\frac{U^{m+1}}{U^m}$ in Equation 1 is bounded between 0 and 1. Therefore, the range of SampEn is $[0, \infty)$. For computational considerations, we down-sampled all

in Equation 1 is bounded between 0 and 1. Therefore, the range of SampEn is $[0, \infty)$. For computational considerations, we down-sampled all iEEG signals to 250 Hz for this analysis, making our sampling period in between points in x 4 ms. To avoid infinite values, we excluded the few trials with no matching three-element templates.

Commonality analysis. To understand whether the metrics of power, spectral slope, and sample entropy uniquely account for variance in memory performance across participants or are redundant, we performed a commonality analysis (Nimon et al., 2008; Ray-Mukherjee et al., 2014), which partitions variance (R^2) into parts unique to each predictor variable and parts shared among all possible combinations of the predictors. To remain consistent with our rank-based analyses used throughout the text and to remain sensitive to nonlinear relationships, commonality analysis was performed on the ranks of our neural and performance measures. The unique contribution of a predictor is calculated as the proportion of variance attributed to it when it is entered last

in a regression analysis. For example, consider a hypothetical case where dependent-variable y is explained by two predictors i and j. Here, the total variance in y explained jointly by both predictors is $R_{y,ij}^2$, while the variance in y explained by i is $R_{y,i}^2$, and the variance explained by j is $R_{y,j}^2$. The unique contribution of a given variable is obtained by subtracting the contribution of the other variable from the joint contribution $R_{y,ij}^2$. Therefore, the variances uniquely explained by i and j respectively are as follows (Eqs. 4 and 5):

$$U_{(i)} = R_{y.ij}^2 - R_{y.j}^2 \tag{4}$$

and

$$U_{(i)} = R_{v,i}^2 - R_{v,i}^2 \tag{5}$$

The common variance in y explained by i and j is equal to the total variance explained jointly less the unique contributions of i and j, expressed as follows (Eq. 6):

$$C_{(i,j)} = R_{y.ij}^2 - U_{(i)} - U_{(j)}$$

simplifying to:

$$C_{(i,i)} = R_{v,i}^2 + R_{v,i}^2 - R_{v,i}^2 \tag{6}$$

Commonality analysis decomposes explained variance into 2^k-1 independent effects for k predictor variables. Therefore, the number of effects increases exponentially with the number of predictors. We used the R package yhat (Nimon et al., 2013) to perform commonality analysis.

Anatomic visualization. To visualize how the relation between spectral power and task performance is spatially distributed, we created 1441 regions of interest (ROIs) evenly spaced across a 1 × 1 cm grid covering the pial surface of a population-average brain. In each participant, we identified all electrodes located within 12.5 mm of each ROI. We designated the raw power for each ROI in each participant as the average raw power across all electrodes assigned to that ROI. For each ROI that included electrodes from ≥6 participants, we determined the Spearman's correlation between raw power and task performance across the participants with electrodes contributing to that ROI. We therefore generated a value for the correlation between raw power and task performance for each ROI. Any ROI that contained electrodes from <6 participants was excluded from statistical analyses.

We generated cortical topographic plots of the anatomic distribution of these correlations by assigning each vertex in the 3D-rendered image of the standard brain a weighted average of the mean value of each ROI that includes that vertex. Weighted values for each vertex were assigned by convolving a 3D Gaussian kernel (radius, 12.5 mm; $\sigma=4.17$ mm) with center weight 1 with the values of surrounding ROIs. We projected these vertex values onto the standard brain. Intensity varied as a function of the statistic metric in question, either Fisher-transformed correlation or t score, in each ROI and with the SD of the Gaussian kernel, which was used purely as a visualization technique.

Statistical analysis. All statistical tests were assessed for significance using two-tailed distributions. As most of our distributions, including accuracy and raw power, were not normally distributed, we used Spearman's rank correlation when evaluating the monotonic relationship between two variables. Spearman's correlation uses only the order of data points and is thus not biased by outliers as with Pearson's correlation. We made an exception, however, when examining the relation between sessions within individual participants. Because we analyzed session counts as low as three, Spearman's correlation is prone to produce extreme values of ± 1 , which cannot be analyzed with cohort-level statistics, necessitating the use of Pearson's correlation in this instance.

To compare correlations across participants, we used a Fisher z transformation on the correlation coefficients calculated for each participant. The transformation stabilizes the variance of these correlations, reduces bias toward lower correlations, and results in a normalized distribution of coefficients. For each correlation, we therefore calculated the Fisher

z transform as follows: $z=\frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$, where r is the correlation coefficient. We used the mathematically equivalent formula, z=arctanh(r) in our calculations.

To determine whether any anatomic region exhibited a significant correlation across participants, we used a nonparametric spatial clustering procedure (Maris and Oostenveld, 2007). This procedure identifies contiguous ROIs where the distribution of correlation coefficients across participants significantly deviates from chance correlation while controlling for the familywise error rate. Briefly, for each ROI, we calculated the true Fisher-transformed correlation coefficient between memory performance and raw spectral power across participants. We then generated 1000 permuted values for each ROI. In each permutation, we randomly assigned each participant a level of task performance drawn from the original distribution of task performance across participants without replacement. In this manner, each permutation involves a random pairing between task performance and raw spectral power. We then determined a z score for each true value and each permuted value in each ROI by comparing that value to the distribution of permuted values. For the true data and for each permutation, we identified contiguous spatial clusters of ROIs, exhibiting z scores with a magnitude >1.96 (corresponding to a two-tailed p value of <0.05). For each cluster, we computed the cluster statistic as the sum of all z scores in that cluster. In this manner, large-magnitude cluster statistics can arise from large deviations in the distributions of correlation coefficients across participants extending over a small spatial region, or moderate deviations that extend over larger regions. We then calculated the exact two-tailed p value for each cluster observed in the true dataset by comparing its cluster statistic to the distribution of largest cluster statistics drawn from each permutation. Clusters were determined to be significant and corrected for multiple comparisons if their p value calculated in this manner was < 0.05.

To assess whether the relation between sample entropy and performance at different time scales was significantly different from zero when summarizing across participants, we used a similar permutation procedure. In this case, for every ROI, we used a two-tailed t test to compare the distribution of values to zero. This generates a t statistic for the true data. Then, during every permutation, we randomly inverted the sign of the metric and produced a permuted distribution of t statistics. We compared the true t statistic to the permuted distribution to generate a t value and t score for every ROI. As above, we used a clustering procedure to identify contiguous ROIs with t 0.05, assigned each contiguous cluster a cluster statistic based on the sum of the corresponding t statistics, and then calculated the exact two-tailed t value for each cluster observed in the true dataset by comparing its cluster statistic to the distribution of the largest cluster statistics drawn from each permutation.

Manual inspection for artifacts. To evaluate the influence of pathological activity on our results, a board-certified clinical epileptologist evaluated a subset of our recordings for the presence of interictal epileptiform discharges (IEDs), allowing us to examine the effects of this pathological activity on theta power in a given electrode or during a given trial, as well as on the average theta power for a given participant. For each participant, we selected and analyzed the five electrodes and 10 trials exhibiting both the highest-magnitude and the lowest-magnitude theta power. Using custom viewing software, and blinded to the method of selecting trials or electrodes, the epileptologist was asked to evaluate whether a given trial did or did not contain epileptiform activity, and subsequently to identify the number of IEDs of any amplitude present in a specified bipolar electrode channel in a given 2 min sample. To determine whether IEDs were more likely during high theta power events or trials, we compared the two groups within each participant. There was no significant difference in the number of IEDs observed in a 2 min period between low theta power (1.12 \pm 1.39 IEDs) and high theta power (0.88 \pm 1.72 IEDs) electrodes across 10 participants ($t_{(9)} = 0.707$, p = 0.50, paired t test). There was also no significant difference between the percentage of events exhibiting IEDs anywhere between low theta power (45.0 \pm 31.0%) and high theta power (61.0 \pm 31.8) trials ($t_{(9)} = -0.97$, p = 0.36). Last, to determine whether IEDs were biased with respect to average power for each participant, we correlated total number of IEDs in our examined

electrodes with average theta power and found there was not a significant correlation ($r_s = 0.40$, p = 0.26, N = 10).

Results

Forty-three participants with drug-resistant epilepsy who underwent surgery for placement of intracranial electrodes for seizure monitoring participated in a verbal pairedassociates task (Fig. 1A). Participants studied 294.2 \pm 20.0 (mean \pm SEM) word pairs, split across multiple experimental sessions, and successfully recalled 40.1 ± 3.2% $(mean \pm SEM)$ words with a mean response time of 1837 ± 65 ms. Response accuracy across participants exhibited a bimodal distribution (Fig. 1B). On 14.9 \pm 1.7% of trials, participants responded with an incorrect word (intrusions) with a mean response time of 2687 \pm 83 ms. For the remaining 44.9 ± 2.6% of trials, participants either made no response to the cue word or said "pass" with a mean response time of 3494 \pm 176 ms. We designated all trials in which a participant successfully said the correct word as correct, and all other trials as incorrect. A participants accuracy was calculated as (# correct trials)/(# trials attempted). Recordings were included from all electrode contacts (number of participants with contacts in each cortical location shown in Fig. 1C).

We measured full-scale IQ (FSIQ) in 35 participants before electrode implantation as part of the routine clinical preoperative evaluation. Participants had an average preoperative FSIQ of 98.5 \pm 2.9 (mean \pm SEM). Across all sessions for each participant, we found that preoperative FSIQ significantly correlated with accuracy during the task ($r_s = 0.51$, p = 0.0017, N = 35; Fig. 1D), suggesting that task performance is related to normal psychometric measurements.

Raw power is negatively correlated with performance

Raw iEEG power can reflect the extent of overall neural activity in each participant's brain and has occasionally been shown to relate to a participant's abilities (Hanslmayr et al., 2007). We were therefore interested in examining whether the raw overall power in each participant as captured by iEEG was related to their task performance. As typical spectral analysis involves examining changes in *z*-scored power relative to an individual's baseline activity, this relation between raw power and task performance would be unexplored in most planned analyses.

In each participant, we extracted the raw spectral power contained in the signal during a baseline time window before wordpair presentation and during the encoding period. To generate an overall level of broadband power for each participant, we averaged the extracted spectral power over all frequencies between 3 and 180 Hz (broadband power), over all trials, and over all electrodes for each time window in each participant. We found that the average raw broadband power during the encoding period demonstrated a significant negative correlation with accuracy during the task ($r_s = -0.39$, p = 0.01109, N = 43; Fig. 2A). This was unchanged if we z-scored each frequency band across subjects to equalize contributions across bands. As with task performance,

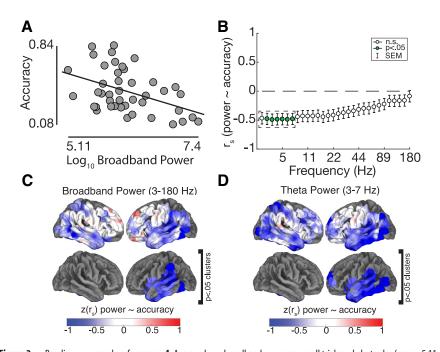


Figure 2. Baseline power and performance. **A**, Average \log_{10} broadband power across all trials and electrodes (range, 5.11–7.40 A.U.) is negatively correlated with performance ($r_s = -0.39$, p = 0.011, N = 43). Line is standard least-squares regression line. **B**, Power–accuracy correlation by frequency band. The negative correlation between power and accuracy exists across all bands and is significantly negatively correlated at all frequency bands between 3.5 and 9 Hz (p < 0.05, Bonferroni correct for 30 frequency bands). The error bars indicate SEM for Spearman's correlation $\left(\frac{0.6325}{\sqrt{n-1}}\right)$. Theta power spectral region of interest is inside of dashed box. **C**, Broadband (Fisher transformed) correlation across spatial ROIs. Bottom shows significant regions (p < 0.05) compared with a permuted distribution through a clustering procedure. **D**, Same as **C** for theta-band power.

broadband power was also negatively correlated with FSIQ across participants ($r_s = -0.408$, p = 0.0348, N = 35).

We found that this relation between raw overall broadband power and task performance was robust and independent of the specific task periods. For example, raw broadband power during the baseline period before word presentation was also inversely correlated with task accuracy ($r_s = -0.38$, p = 0.0125, N = 43). Moreover, we also found a significant relation between raw broadband power and task accuracy when we examined power separately during only correct or only incorrect trials ($r_s = -0.38$, p = 0.0132, and $r_s = -0.38$, p = 0.0114, respectively), suggesting that this effect reflects each participant's overall baseline neural activity rather than simply the proportion of trials that featured successful encoding in each participant. Finally, to determine whether this relation reflects each participant's underlying physiology or is dependent on a task-evoked state, we also examined this relation during an epoch recorded before the beginning of the task when the participant was awake, at rest, and under no instruction. We found the negative correlation between overall broadband power and task performance was also preserved during this extra-task period, suggesting that this effect is not task-dependent but is related to baseline cognitive behavior ($r_s = -0.36$, p = 0.0198). This finding departs from most memory studies in that we claim that our result does not depend on the fact that the subject is undertaking a memory task at the time, allowing us to generalize our electrophysiological correlates to normal daily activities.

We next examined whether the inverse correlation between raw power and task performance was specific to individual frequency bands by separately computing correlations between narrow-band frequencies and task performance (Fig. 2B). We found power at every frequency band between 3 and 180 Hz was

negatively correlated with performance. All frequencies between 3.5 and 9 Hz had a significant negative correlation between overall raw power and accuracy when corrected for multiple comparisons across frequencies (Fig. 2B; p < 0.05, Bonferroni-corrected for 30 frequency bands). This suggests that this effect is spectrally broad but driven by low-frequency activity. We therefore restricted subsequent power analyses to power averaged across the theta band, which had a correlation of $r_s = -0.50$, p = 0.0008(3–7 Hz; Fig. 2B, dashed box). As performance was shown to be strongly correlated with IQ, it is not immediately clear whether the relationship between power and accuracy is simply a manifestation of a relationship between power and general ability or whether power explains additional variance in task performance not explained by IQ. While theta power is negatively correlated with IQ ($r_s = -0.44$, p = 0.0074, N = 35), it is also correlated with performance after regressing out the variance in accuracy explained by IQ ($r_s = -0.42$, p = 0.014, N = 35), indicating it is independently predictive of task-specific performance.

We were also interested in whether the relation between raw power and task accuracy varied across brain regions (ROIs; see Materials and Methods). For every ROI, we determined the correlation between both average raw broadband and theta power in all electrodes within that ROI and task performance across participants. The inverse correlation that we found between cortically distributed raw power and task performance localized to regions of the temporal and parietal lobes in both hemispheres (Fig. 2C,D, top). Using a nonparametric clustering algorithm, we found that spatially contiguous regions exhibited a significant correlation across participants within the left temporal lobe for both broadband and theta band power (p < 0.05, permutation procedure; see Materials and Methods; Fig. 2C,D, bottom). These data suggest that individuals with less broadband and lowfrequency power in the temporal lobe have greater ability to encode associative memories.

Assessing cortical activation through PSD slope

The PSD of iEEG signals falls off with frequency following a power law distribution. The slope of the PSD in log–log space has been shown to flatten in response to task activation (Podvalny et al., 2015), and the extent to which the slope flattens has been related to cognitive effort (Churchill et al., 2016). As such, we examined the overall raw PSD in each participant to determine whether the observed changes in broadband power with task performance may be related to changes in the slope of the PSD.

We first divided the participants into three terciles based on task performance (low, medium, or high accuracy) to visualize the average raw PSD in each cohort (Fig. 3A). Below 30 Hz, the average PSDs of the three populations easily separate, with the lowest-performing participants exhibiting the largest low-frequency raw power. As suggested by our analysis examining the correlation between raw power and task performance, dividing participants into these terciles yielded a significant effect of performance tercile on low-frequency power (ANOVA using average raw power <30 Hz; $F_{(2,40)} = 4.40$, p = 0.019). At higher frequencies (>30 Hz), however, the distinction between participant groups was negligible ($F_{(2,40)} = 0.913$, p = 0.409).

We next calculated the slope of the average PSD in log–log space between 10 and 100 Hz for each participant (Fig. 3*B*, inset). We chose this frequency range to avoid the low-frequency knee and the effects of action potential contamination at higher frequencies (Podvalny et al., 2015). Across participants, PSD slopes [range, -3.36 to -2.07; -2.67 ± 0.05 (mean \pm SEM); Fig. 3*B*] are in the range of those reported by others using similar metrics

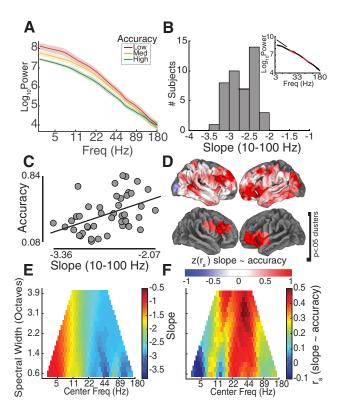


Figure 3. Spectral slope and performance. **A**, Average PSD across tertials of subjects sorted by performance. Shading shows SEM. **B**, Distribution of average spectral slopes across subjects (2.67 \pm 0.05; mean \pm SEM). Inset shows example subject, red is range of frequencies slope is calculated over (10 – 100 Hz), and dashed line shows robust fit line. **C**, Spectral slope is positively correlated with accuracy across subjects ($r_s = 0.48$, p = 0.0014, N = 43). Line is standard least-squares regression line. **D**, Correlation of spectral slope and accuracy across spatial ROIs. Bottom shows significant regions (p < 0.05) compared with a permuted distribution through a clustering procedure. **E**, Average spectral slope as a function of center frequency and spectral width. **F**, Average correlation as a function of center frequency and spectral width as in **E**.

(Podvalny et al., 2015). Across participants, we found that PSD slope was positively correlated with task performance, such that participants with flatter slopes performed better ($r_s = 0.48$, p = 0.0014, N = 43; Fig. 3C). Slope was not significantly correlated with IQ ($r_s = 0.17$, p = 0.32), and was still correlated with performance after regressing out the effects of IQ ($r_s = 0.49$, p = 0.0032, N = 35). We examined the anatomic regions that demonstrated a significant relation between PSD slope and task performance (p < 0.05, permutation procedure) and localized them to the left and right frontal lobes (Fig. 3D).

As with raw broadband and theta power, this relation between PSD slope and task performance was robust and independent of when during the task the calculation of PSD was made. We found participants with greater recall accuracy had flatter slopes when examining recordings from the baseline period ($r_s = 0.45$, p = 0.0026), during correct trials only ($r_s = 0.47$, p = 0.0018), or during incorrect trials only ($r_s = 0.47$, p = 0.0015). We found that like broadband power, the significant relationship between slope and accuracy was preserved when examining extra-task epochs, during which participants were awake and at rest ($r_s = 0.44$, p = 0.0034), indicating that as with raw power, this relation is not task-evoked.

Although several studies have identified measures of broadband power or spectral slope as a proxy for spike rate (Manning et al., 2009), cortical activation (Podvalny et al., 2015), or the balance between cortical excitation and inhibition (Gao et al., 2017), there remains no consensus regarding the frequency range over which one should calculate the PSD slope to identify the nonoscillatory components of spectral power. In our initial analysis, we used a range of 10 to 100 Hz. However, other groups have used different frequency ranges and it is possible our findings are sensitive to this parameter.

To ensure that the observed relation between PSD slope and task performance was not specific to the range of frequencies we used to calculate PSD slope, we iterated through a library of different frequency windows, each comprising a center frequency and a spectral width, to compute PSD slopes. We examined PSD slope using every possible frequency window between 3 and 180 Hz. We found that the slope was largely unaffected by spectral width, and that although the slope increased as a function of center frequency, average slope beyond a center frequency of 20 Hz stabilized to an overall mean across participants of ~ 2.7 (Fig. 3E). Above 18 Hz, the PSD slopes ranged between -2 and -3, while at lower center frequencies we observed PSD slopes as flat as -1. We examined how varying our measure of PSD slope affected the relation between PSD slope and task performance. We correlated each PSD slope calculated with different center frequencies and spectral widths with task performance and confirmed that PSD slopes were positively correlated with task accuracy for most center frequencies regardless of spectral width (Fig. 3F). This suggests that when examining the role of spectral slope, most ranges centered between 20 and 50 Hz should give congruent results.

Assessing information content through sample entropy

The relation between spectral slope and accuracy may be partially explained by the complexity of the underlying iEEG signal, which may in turn suggest a higher capacity for processing information. However, while spectral slope is related to signal complexity (Keshner, 1982), it is not a direct measure. We therefore calculated sample entropy to quantify signal complexity of the iEEG trace, a measure that has previously been successfully used for discerning differences among EEG signals (Fig. 4A; see Materials and Methods; Mizuno et al., 2010; Catarino et al., 2011; Vaz et al., 2017). Sample entropy measures the predictability of a signal, is robust to low-level noise and artifacts, and has been found to be more robust for shorter data lengths than other measures of entropy, such as approximate entropy (Yentes et al., 2013; Sokunbi, 2014). Indeed, the complexity of two example iEEG signals is visible in the raw recording and reflected in the measured sample entropy (Fig. 4B).

Based on the observed changes in raw power and spectral slope, and the theoretical suggestion that increased information content involves signal desynchronization (Hanslmayr et al., 2012), we hypothesized that participants with greater complexity in their iEEG signal, and therefore higher sample entropy, would exhibit better task performance. We calculated the average sample entropy during the encoding period across all trials and all electrodes and found that participants with greater sample entropy performed significantly better on the task (Fig. 4C; $r_s =$ 0.51, p = 0.00065, N = 43), suggesting that the observed relation between task performance and PSD slope is related to the complexity of the underlying neural signal. Similar to slope, sample entropy was not correlated with IQ ($r_s = 0.26$, p = 0.14) and was correlated with accuracy after regressing out the effects of IQ (r_s = 0.52, p = 0.0015, N = 35). The relation between sample entropy and accuracy was distributed across the cortex but was particularly localized to the left temporal lobe (Fig. 4D).

As with power and spectral slope, we found that this relation was preserved when looking at correct trials ($r_s = 0.48$, p =

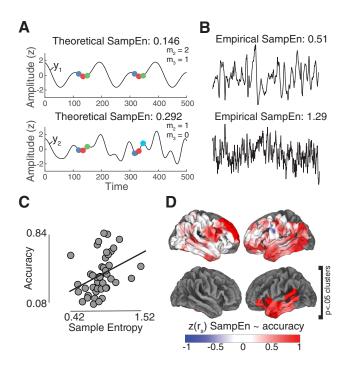


Figure 4. Sample entropy and performance. **A**, Sample entropy schematic for theoretical signals. Color of dots superimposed on signals indicate discretized voltage bin. Signal y_2 is more complex than y_1 , making subsequent points relatively more difficult to predict. **B**, Example epochs from two participants with low and high entropy. The upper signal is from participant with an average sampEn of 0.51. This epoch has a measured sampEn of 0.67. The lower signal is from participant with an average sample entropy of 1.29. This epoch has a measured sample entropy of 1.51. **C**, Sample entropy is positively correlated with performance across participants ($r_s = 0.51$, p = 0.0007). Line is standard least-squares regression line. **D**, Sample entropy correlation across spatial ROIs. Bottom shows significant regions (p < 0.05) compared with a permuted distribution through a clustering procedure.

0.0012), incorrect trials ($r_s = 0.52$, p = 0.00049), the baseline period ($r_s = 0.53$, p = 0.00028), or an extra-task epoch ($r_s = 0.34$, p = 0.025). Moreover, participants with flatter PSD slopes and less theta power had greater sample entropy ($r_s = 0.63$, $p = 1.2 \times 10^{-5}$, and $r_s = -0.41$, p = 0.0067, respectively) demonstrating that spectral slope and low-frequency power are strong indicators of signal complexity.

Signal complexity across time scales

Our data demonstrate that related measures of signal complexity—low-frequency power, spectral slope, and sample entropy—show strong relations with overall performance during an associative-memory task across participants. However, most studies of human memory have focused on SMEs in which differences between correct and incorrect trials are assessed within individuals. We were therefore interested in whether the observed changes in neural signal complexity across participants would also be observed across different time scales within participants. We specifically investigated changes in sample entropy during individual sessions and trials to index changes in brain-state complexity at the time scales of hours and seconds, respectively.

We first examined the relation between sample entropy and performance during individual sessions for each electrode in each participant who completed ≥ 3 sessions. In individual participants, we found that sample entropy correlated with performance on a session-by-session basis (Fig. 5*A*). Across all participants, we found that this relation was consistent, although the distribution of correlation coefficients was not significantly different from zero (two-tailed t test of Fisher-transformed correlation coefficients, $t_{(21)}$ =

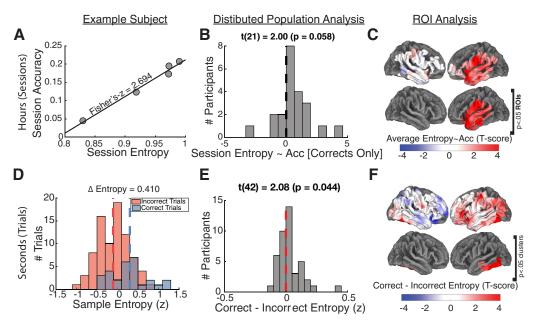


Figure 5. Sample entropy across time scales. **A**, Example subject with positive session level correlation of sample entropy to accuracy. **B**, Distribution of Fisher-transformed ρ values across subjects trends toward positive correlations ($t_{(21)} = 2.00$, p = 0.058). **C**, Population average session correlation by ROI (t score). Bottom shows independently significant ROIs (p < 0.05, uncorrected). **D**, Example subject-level distribution of sample entropy values for correct versus incorrect trials. **E**, Distribution of correct-incorrect sample entropy across subjects is significantly > 0 ($t_{(42)} = 2.08$, p = 0.044). **F**, Population average change in sample entropy by item (t score) across ROIs. Bottom shows significant regions (p < 0.05) compared with a permuted distribution through a clustering procedure.

2.00, p = 0.058; Fig. 5*B*). However, while this relationship was not significant on a global level, we did find that individual ROIs throughout the left temporal and parietal lobes were significant (Fig. 5*C*). In addition, we found that theta power and slope also showed strong relationships with session accuracy that were similar to those found across participants ($t_{(21)} = -2.37$, p = 0.028, and $t_{(21)} = 1.83$, p = 0.081, respectively).

Next, as is routine in most memory studies, we examined differences between correct and incorrect trials to understand the relation between sample entropy and memory encoding at the time scale of seconds. We first z-scored sample entropy within each session to eliminate any session-level variance. Hence, both the session-level and item-level effects are calculated such that they are completely independent from one another and the previously explored participant-level effects. We found that participants exhibited significantly higher sample entropy for correct compared with incorrect trials (Fig. 5D, E; $t_{(42)} = 2.08$, p =0.0044). The item-level changes in sample entropy localized to the left inferior temporal lobe (Fig. 5E). The differences between correct and incorrect theta power and slope were also both significantly biased in the same direction as their across-subject and across-session effects ($t_{(42)} = -2.57$, p = 0.014 and $t_{(42)} = 2.06$, p = 0.046, respectively). To complement our analysis across participants, we also examined whether individual SMEs were correlated with raw power measurements. While slope and sample entropy showed no relationship (p > 0.90), raw theta power was significantly and negatively correlated with the difference between correct and incorrect theta power. Participants with greater raw theta power had more negative SMEs, while those with less raw theta power had more positive SMEs ($r_s = -0.33$, p = 0.031). This suggests that, while by far the overriding trend in our cohort is that less theta power is better for encoding across all time scales, for a subset of participants with overall low-theta power, this is not the case.

In evaluating across subject SMEs, it is unclear over what time scale the changes in sample entropy are occurring. These changes

may be related to word-pair adaptation, or they may reflect a more slowly fluctuating dynamic. To explore this, we made three additional comparisons. We compared the sample entropy during the baseline period between correct and incorrect trials, we compared the sample entropy during correct trials between the encoding and baseline periods, and we compared the sample entropy during incorrect trials between the encoding and baseline periods. Interestingly, while we found minimal difference in sample entropy during the baseline periods between correct and incorrect trials ($t_{(42)} = -1.71$, p = 0.095), we also found that sample entropy significantly increased from baseline during correct trials, and significantly decreased from baseline on incorrect trials ($t_{(42)} = 2.41, p = 0.020$ and $t_{(42)} = -2.36, p = 0.023$). These data suggest that relatively fast changes in the sample entropy, and therefore complexity, of the signal contribute to subsequent remembering and subsequent forgetting along with changes over much longer time scales.

Are theta power, spectral slope, and sample entropy redundant?

Sample entropy was shown earlier to be positively correlated with spectral slope and negatively correlated with theta-band power. Spectral slope and theta-band power are inversely correlated with each other as well ($r_s = -0.35$, p = 0.02). It is unclear, given the high level of collinearity between these variables, whether they are describing unique underlying processes or are really redundant factors. To determine the proportions of variance in memory performance across participants that are uniquely attributed to these metrics as well as those that are common among all possible combinations of these metrics, we performed a commonality analysis (see Materials and Methods). The commonality analysis (Table 1) showed that spectral power in the theta band uniquely accounted for 22.28% of the total variance explained by the predictors, with slope and entropy uniquely explaining 6.98 and 8.64%, and jointly explaining 20.04% of the total variance explained. The total variance accounted for by power, PSD slope,

Table 1. Commonality analysis output describing unique and common contributions of the three predictor variables (theta-band power, spectral slope, and sample entropy) to the regression effect explaining memory performance across participants

	Power	Slope	Entropy	Percentage of total
Unique to power	0.0851			22.28
Unique to slope		0.0267		6.98
Unique to entropy			0.0330	8.64
Common to power and slope	0.0151	0.0151		3.96
Common to power and entropy	0.0381		0.0381	9.97
Common to slope and entropy		0.0766	0.0766	20.04
Common to power, slope, and entropy	0.1075	0.1075	0.1075	28.12

Total $R^2 = 0.3821$. Unique + Common = 100% of R^2 .

and sample entropy through both unique and shared contributions are 24.58, 22.58, and 25.52% respectively with 10.75% of the variance being common to all three. This analysis illustrates that, while all three metrics capture properties of neural activity relevant for memory performance, theta power may capture relatively distinct features from those captured by spectral slope and sample entropy.

To see how these metrics relate across time, we looked at their correlations across trials for individual participants. Across trials, z-scored by session, theta power and sample entropy showed a strong negative correlation (median $\rho=-0.77$). This aligns with our interpretation that, from an information-coding perspective, predictable oscillations contain less information than less predictable stochastic dynamics. As expected, entropy and slope were positively correlated on a trial level (median $\rho=0.59$) and slope and theta power were inversely correlated (median $\rho=-0.60$). Notably, these relationships were all highly significant on a population level ($|t_{42}| > 17$, $p < 10^{-20}$).

Discussion

Our analyses demonstrate that low-frequency power and complexity of cortical activity track variability in memory performance at the level of individuals, days, and events. Specifically, improved memory performance at any given scale is related to decreased low-frequency power and increased signal complexity as measured by the PSD slope and sample entropy. Notably, while these metrics show collinearity on both an individual and population level, commonality analysis revealed substantial unique contributions of each metric in explaining memory performance. These findings suggest that the complexity of brain activity may reflect an individual's ability to occupy variable cognitive states and the extent to which information can be coded in their brain signals, which is evident in associative-memory performance.

The suggestion that cognitive flexibility may improve task performance appears intuitive. Indeed, the ability to explore the brain's dynamic repertoire during rest is thought to be a marker of healthy brain function and may underlie introspection and rehearsal (Ghosh et al., 2008). Therefore, it seems likely that a high-performing brain is one that engages with the world by assuming a variety of functional configurations. Whether such variability and flexibility may be relevant for associative-memory performance has, until now, not been directly established. We establish this link here by demonstrating that memory performance is significantly correlated with signal complexity both across and within individuals. Cognitive flexibility lends neural systems the ability to explore their state space (Deco and Jirsa, 2012), which may lead to separable memory representations less susceptible to interference. Consistent with the idea that increased complexity may lead to increased separability of events, prestimulus weighted permutation entropy of scalp EEG can bias participants' perception of identical auditory stimuli by changing the fidelity with which the stimulus was encoded (Waschke et al., 2017). Furthermore, multiscale entropy of brain signals (scalp EEG) correlate with participants' ratings of famous-face familiarity and increase with learning over multiple exposures to previously unfamiliar faces (Heisz et al., 2012). Hence, the observed correlations between entropy and associative-memory performance here suggests that neural signal complexity reflects the capacity to successfully encode associative memories by flexibly engaging with the presented material.

The paired-associates memory task used here requires participants to form associations between unrelated words that constitute individual episodes or experiences that are subsequently recalled. Encoding these associations draws upon the meanings of the words to form a conceptual and semantic link between them (Kahana et al., 2008; Madan et al., 2010; Jang et al., 2017). Therefore, forming these associations should engage cortical regions, such as the anterior temporal lobe, that are involved in semantic processing (Binder et al., 2009; Ralph et al., 2017). In our data, we observe strong correlations between memory performance and low-frequency power and entropy in these same left temporal lobe regions. The relationship between cognitive flexibility, as assessed by these metrics in the temporal lobe, and verbal associative memory performance across individuals may therefore emerge because of the involvement of the temporal lobe in helping encode verbal associations.

Our approach here differs from that of earlier studies of human memory encoding and retrieval by specifically asking whether there are systematic differences in neural activity across participants that may predict individual memory performance. Most previous studies of human episodic memory have focused on relative changes in neural oscillatory activity between correctly and incorrectly encoded events (Sederberg et al., 2003, 2007; Burke et al., 2014; Long et al., 2014; Greenberg et al., 2015). While these studies have significantly advanced our understanding of the neural correlates of human memory, an unresolved question has been why different studies have demonstrated conflicting results, particularly with respect to low-frequency oscillatory power (Hanslmayr and Staudigl, 2014). In our data, we tracked memory performance using low-frequency power that was not normalized relative to any baseline and found that it was inversely correlated with overall memory performance. Moreover, within each individual, fluctuations in neural activity were predictive of how well they performed at any given moment. Interestingly, we found that these fluctuations were dependent on overall power measurements for each participant. Our data therefore may provide some insight into the conflicting data observed in previous studies. These conflicts have been previously attributed to differences in task design and electrode coverage. However, because of the variability in baseline power between individuals, these conflicts may also be affected by where each cohort of participants sits in this range of baseline power and how that may affect the changes in power observed over shorter time scales.

As examining the structure of the full PSD across all frequencies can often yield a more complete picture of neural activity (Podvalny et al., 2015), our analyses of PSD slope changes complement the observed changes in low-frequency power. The slope of the PSD has been hypothesized to reflect the balance between excitation and inhibition, and computational modeling of neural activity has demonstrated that reducing the excitation/inhibition ratio results in a steeper PSD slope (Gao et al., 2017). Both *in vitro*

and in vivo cortical networks show maximal dynamic range under balanced excitation-inhibition conditions (Shew et al., 2009, 2011). An increased dynamical range of neuronal responses may improve adaptability and efficiency of neural systems in service of memory. Another possibility is that a shallower PSD slope may emerge due to the infusion of noise into the neural signal via asynchronous firing activity (Usher et al., 1995; Pozzorini et al., 2013; Podvalny et al., 2015; Voytek and Knight, 2015; Voytek et al., 2015). Whether such noise is beneficial is unclear, as the effect of noise on information coding depends on whether noise is correlated between neurons (Averbeck et al., 2006). Nevertheless, our finding that flatter PSD slopes and increased sample entropy relate to better memory performance suggests that in our data, more complex brain signals reflect more informationally rich signals as posited by others (Mitchell et al., 2009; Schneidman et al., 2011; Hanslmayr et al., 2012, 2016).

Ultimately, the slope of the PSD and the low-frequency power contributing to that slope should be related to the underlying complexity of the neural signal, which can be directly assessed using measures of entropy as we do here. Although greater signal complexity does not always reflect greater information content, entropy of the EEG signal may increase with healthy aging (McIntosh et al., 2008; Waschke et al., 2017), and higher entropy is also associated with greater task efficiency and network efficiency (Mišić et al., 2010, 2011). Entropy of resting-state brain signals can distinguish children at high risk for autism spectrum disorder from normal-developing children (Bosl et al., 2011), and healthy from epileptogenic neural tissues (Protzner et al., 2010). Here, we directly show that the complexity of the neural signal captured using iEEG tracks associative-memory performance across individuals, providing further support to the proposition that brainsignal variability is functionally relevant. Moreover, we show that within individuals, variability in neural-signal complexity across time scales also tracks memory performance for the individual participant. The variability we experience in our daily lives with memory performance is likely therefore influenced by these changing levels of neural-signal complexity.

Of note, participants in our study were also neurosurgical patients with drug-resistant epilepsy. In most cases, their seizure activity localized to the temporal lobes, raising the possibility that the observed effects in this brain region may also be related to the underlying pathology of the disorder itself. Greater disruptions of normal temporal lobe function could result in less signal complexity in this brain region, which could then lead to worse memory performance on this paired-associates task. We took several precautions to mitigate the effects of epileptic activity on our study, including removing electrodes identified as ictal or interictal, and removing electrodes and individual trials that showed higher variance or kurtosis relative to the rest of the population. In addition, we visually checked for both high-amplitude and low-amplitude IEDs in electrodes and trials chosen from a subset of participants. While we found that IEDs were approximately equally present in both high-power and low-power trials and electrodes, it is also clear that our data were not devoid of these artifacts. Hence, it is possible that pathological activity may contribute to some of the observed relationships between complexity metrics and memory performance. In this scenario, however, the interpretation of our data does not change, since decreased neural-signal complexity, regardless of whether it can be attributed to normal or pathologic variability, would still be related to decreased memory ability.

Previous studies have indeed shown that IEDs during encoding and retrieval can impair memory performance (Horak et al.,

2017), and IED rates decrease from baseline during correct, but not incorrect, encoding trials (Matsumoto et al., 2013). Critically, however, increases in IEDs during rest or distractor periods in these studies do not appear to reduce memory performance, and the overall IED rates do not relate to recall performance across participants (Horak et al., 2017). This is in contrast to the effects of overall power on memory performance that we report here, which are observed during both rest and task periods. Moreover, controlling for the same level of overall pathology within individuals, we find the same metrics were related to memory performance across multiple timescales. It is difficult to explain how pathologic activity alone would consistently predict memory performance at every different timescale, or even why most effects in our data also extend to generally nonpathologic frontal lobe clusters in our dataset.

In addition to the brief disruptions in temporal lobe function caused by IEDs, transient neurologic dysfunction can be observed after a seizure lasting minutes to hours. In the case of left temporal lobe epilepsy, postictal impairment can be seen in verbal and visual recognition memory. However, postictal effects are unlikely to play a significant role in our findings, since we avoided administering cognitive testing for several hours following a seizure episode. In addition, if patients were still symptomatic following a seizure, testing was usually deferred by the study team or by the participant until they had regained their baseline function. The longer-term effects of such pathologic activity may, however, contribute to changes in IQ, which could in turn mediate the across-participant relationship between signal complexity and memory. However, we found that only theta power is significantly correlated with IQ. Furthermore, theta power, PSD slope, and sample entropy are all significantly correlated with performance even after the effects of IQ are regressed out, suggesting that signal complexity is indeed specifically relevant for memory formation. Therefore, although the participants' underlying disorder may affect normal neural information processing and overall cognitive ability, our data suggest that the individual differences in neural-signal complexity that relate to differences in memory performance are unlikely to be driven by pathology alone.

It is also possible that the changes in neural complexity that we interpret to denote cognitive flexibility in fact simply capture changing levels of attention. For example, patients may feel drowsier in some experimental sessions than in others and these differences in levels of engagement may be captured by our complexity metrics. However, we note that changes in sample entropy from baseline to encoding states also occur over shorter timescales within an individual experimental session. These fine-grained temporal changes consistently capture differences between successful and unsuccessful associative-memory encoding trials even though the baseline entropy levels are not different between the two conditions. Moreover, at the other extreme of time scales, participant-level complexity metrics correlate with memory performance as well as IQ. Therefore, it is unlikely that drowsiness explains all the observed relationships found here between neural complexity and memory performance at multiple scales. Attention may indeed play a direct role in determining the extent to which neural state space is explored during a task. However, the possibility that changing levels of attention may explain our results is still consistent with the interpretation that theta power, spectral slope, and sample entropy ultimately reflect cognitive flexibility and a capacity to encode information.

Aside from the immediate effects of interictal and ictal epileptiform activity, it is also possible that some of these relationships are affected by the influence of antiepileptic drugs (AEDs). All participants were chronically taking AEDs, which were weaned at

a variable rate following surgery at the discretion of the treating clinicians. Because participants were on varying AEDs at varying doses with varying pharmacokinetics, we did not explicitly control for AEDs as a factor in our study. In general, since testing began several days postoperatively, AEDs were at a significantly lower level than at baseline for a given participant. AEDs are known to reduce attention and vigilance, but other studies have suggested that the cognitive impacts of AEDs may be overrated when compared with the pathological and psychosocial effects of epilepsy itself (Meador, 2002; Park and Kwon, 2008). Studies exploring electrophysiological changes related to AED use have found highly heterogenous results across participants and medications, with some evidence for short-term reductions in gamma power (Arzy et al., 2010) and long-term slowing of EEG rhythms (Salinsky et al., 1994), measures that have minimal overlap with our metrics. Hence, while AEDs may certainly be a relevant factor, they are unlikely to be the primary driving force of the reported effects that persist across timescales and individuals.

Together, our data therefore provide insight into why memory performance may be variable both between and within individuals. Our data suggest that how well one can encode and retrieve memories is related to the flexibility in their cognitive processing. Such flexibility is captured directly by measuring the sample entropy of the neural signal, and corroborated by our measures of low-frequency power and the PSD slope. People with better memory have neural signals that exhibit greater complexity, and therefore are capable of exhibiting more flexible behavior, which is beneficial for memory formation.

References

- Arzy S, Allali G, Brunet D, Michel CM, Kaplan PW, Seeck M (2010) Antiepileptic drugs modify power of high EEG frequencies and their neural generators. Eur J Neurol 17:1308–1312. CrossRef Medline
- Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nat Rev Neurosci 7:358–366. CrossRef Medline
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? a critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex 19:2767–2796. CrossRef Medline
- Bosl W, Tierney A, Tager-Flusberg H, Nelson C (2011) EEG complexity as a biomarker for autism spectrum disorder risk. BMC Med 9:18. CrossRef Medline
- Burke JF, Long NM, Zaghloul KA, Sharan AD, Sperling MR, Kahana MJ (2014) Human intracranial high-frequency activity maps episodic memory formation in space and time. Neuroimage 85:834–843. CrossRef Medline
- Burke JF, Ramayya AG, Kahana MJ (2015) Human intracranial high-frequency activity during memory processing: neural oscillations or stochastic volatility? Curr Opin Neurobiol 31:104–110. CrossRef Medline
- Catarino A, Churches O, Baron-Cohen S, Andrade A, Ring H (2011) Atypical EEG complexity in autism spectrum conditions: a multiscale entropy analysis. Clin Neurophysiol 122:2375–2383. CrossRef Medline
- Churchill NW, Spring R, Grady C, Cimprich B, Askren MK, Reuter-Lorenz PA, Jung MS, Peltier S, Strother SC, Berman MG (2016) The suppression of scale-free fMRI brain dynamics across three different sources of effort: aging, task novelty and task difficulty. Sci Rep 6:30895. CrossRef Medline
- Deco G, Jirsa VK (2012) Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. J Neurosci 32:3366–3375. CrossRef Medline
- Deco G, Jirsa V, McIntosh AR, Sporns O, Kötter R (2009) Key role of coupling, delay, and noise in resting brain fluctuations. Proc Natl Acad Sci U S A 106:10302–10307. CrossRef Medline
- Deco G, Jirsa VK, McIntosh AR (2013) Resting brains never rest: computational insights into potential cognitive architectures. Trends Neurosci 36:268–274. CrossRef Medline
- Dehghani N, Bédard C, Cash SS, Halgren E, Destexhe A (2010) Comparative power spectral analysis of simultaneous electroencephalographic and

- magnetoencephalographic recordings in humans suggests non-resistive extracellular media. J Comput Neurosci 29:405–421. CrossRef Medline
- Faisal AA, Selen LP, Wolpert DM (2008) Noise in the nervous system. Nat Rev Neurosci 9:292–303. CrossRef Medline
- Fell J, Ludowig E, Staresina BP, Wagner T, Kranz T, Elger CE, Axmacher N (2011) Medial temporal theta/alpha power enhancement precedes successful memory encoding: evidence based on intracranial EEG. J Neurosci 31:5392–5397. CrossRef Medline
- Gao R, Peterson EJ, Voytek B (2017) Inferring synaptic excitation/inhibition balance from field potentials. Neuroimage 158:70–78. CrossRef Medline
- Garrett DD, Kovacevic N, McIntosh AR, Grady CL (2011) The importance of being variable. J Neurosci 31:4496–4503. CrossRef Medline
- Garrett DD, Samanez-Larkin GR, MacDonald SW, Lindenberger U, McIntosh AR, Grady CL (2013) Moment-to-moment brain signal variability: a next frontier in human brain mapping? Neurosci Biobehav Rev 37:610–624. CrossRef Medline
- Ghosh A, Rho Y, McIntosh AR, Kötter R, Jirsa VK (2008) Noise during rest enables the exploration of the brain's dynamic repertoire. PLoS Comput Biol 4:e1000196. CrossRef Medline
- Grady CL, Garrett DD (2014) Understanding variability in the BOLD signal and why it matters for aging. Brain Imaging Behav 8:274–283. CrossRef Medline
- Greenberg JA, Burke JF, Haque R, Kahana MJ, Zaghloul KA (2015) Decreases in theta and increases in high frequency activity underlie associative memory encoding. Neuroimage 114:257–263. CrossRef Medline
- Guderian S, Schott BH, Richardson-Klavehn A, Düzel E (2009) Medial temporal theta state before an event predicts episodic encoding success in humans. Proc Natl Acad Sci U S A 106:5365–5370. CrossRef Medline
- Hanslmayr S, Staudigl T (2014) How brain oscillations form memories—a processing based perspective on oscillatory subsequent memory effects. Neuroimage 85:648–655. CrossRef Medline
- Hanslmayr S, Aslan A, Staudigl T, Klimesch W, Herrmann CS, Bäuml, KH (2007) Prestimulus oscillations predict visual perception performance between and within subjects. Neuroimage 37:1465–1473. CrossRef Medline
- Hanslmayr S, Volberg G, Wimber M, Raabe M, Greenlee MW, Bäumel KH (2011) The relationship between brain oscillations and bold signal during memory formation: a combined EEG-fMRI study. J Neurosci 31: 15674–15680. CrossRef Medline
- Hanslmayr S, Staudigl T, Fellner MC (2012) Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. Front Hum Neurosci 6:74. CrossRef Medline
- Hanslmayr S, Staresina BP, Bowman H (2016) Oscillations and episodic memory: addressing the synchronization/desynchronization conundrum. Trends Neurosci 39:16–25. CrossRef Medline
- Haque RU, Wittig JH Jr, Damera SR, Inati SK, Zaghloul KA (2015) Cortical low-frequency power and progressive phase synchrony precede successful memory encoding. J Neurosci 35:13577–13586. CrossRef Medline
- He BJ (2011) Scale-free properties of the functional magnetic resonance imaging signal during rest and task. J Neurosci 31:13786–13795. CrossRef Medline
- He BJ (2014) Scale-free brain activity: past, present, and future. Trends Cogn Sci 18:480–487. CrossRef Medline
- He BJ, Zempel JM, Snyder AZ, Raichle ME (2010) The temporal structures and functional significance of scale-free brain activity. Neuron 66:353–369. CrossRef Medline
- Heisz JJ, Shedden JM, McIntosh AR (2012) Relating brain signal variability to knowledge representation. Neuroimage 63:1384–1392. CrossRef Medline
- Horak PC, Meisenhelter S, Song Y, Testorf ME, Kahana MJ, Viles WD, Bujarski KA, Connolly AC, Robbins AA, Sperling MR, Sharan AD, Worrell GA, Miller LR, Gross RE, Davis KA, Roberts DW, Lega B, Sheth SA, Zaghloul KA, Stein JM, et al. (2017) Interictal epileptiform discharges impair word recall in multiple brain areas. Epilepsia 58:373–380. CrossRef Medline
- Jang AI, Wittig JH Jr, Inati SK, Zaghloul KA (2017) Human cortical neurons in the anterior temporal lobe reinstate spiking activity during verbal memory retrieval. Curr Biol 27:1700–1705.e5. CrossRef Medline
- Kahana MJ, Howard MW, Polyn SM (2008) Associative retrieval processes in episodic memory. In: Cognitive psychology of memory (Roediger HL III, ed). Oxford, UK: Elsevier.
- Keshner MS (1982) 1/f noise. Proc IEEE 70:212-218. CrossRef
- Long NM, Burke JF, Kahana MJ (2014) Subsequent memory effect in intracranial and scalp EEG. Neuroimage 84:488–494. CrossRef Medline

- MacDonald SW, Nyberg L, Bäckman L (2006) Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. Trends Neurosci 29:474–480. CrossRef Medline
- Madan C, Glaholt M, Caplan J (2010) The influence of item properties on association-memory. J Mem Lang 63:46–63. CrossRef
- Manning JR, Jacobs J, Fried I, Kahana MJ (2009) Broadband shifts in LFP power spectra are correlated with single-neuron spiking in humans. J Neurosci 29:13613–13620. CrossRef Medline
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190. CrossRef Medline
- Matsumoto JY, Stead M, Kucewicz MT, Matsumoto AJ, Peters PA, Brinkmann BH, Danstrom JC, Goerss SJ, Marsh WR, Meyer FB, Worrell GA (2013) Network oscillations modulate interictal epileptiform spike rate during human memory. Brain 136:2444–2456. CrossRef Medline
- McIntosh AR, Kovacevic N, Itier RJ (2008) Increased brain signal variability accompanies lower behavioral variability in development. PLOS computational biology 4:e1000106. CrossRef Medline
- Meador KJ (2002) Cognitive outcomes and predictive factors in epilepsy. Neurology 58 [8 Suppl 5]:S21–S26. Medline
- Miller KJ, Sorensen LB, Ojemann JG, den Nijs M, Sporns O (2009) Powerlaw scaling in the brain surface electric potential. PLoS Comput Biol 5:e1000609. CrossRef Medline
- Milstein J, Mormann F, Fried I, Koch C (2009) Neuronal shot noise and Brownian $1/f^2$ behavior in the local field potential. PLoS One 4:e4338. CrossRef Medline
- Mišić B, Mills T, Taylor MJ, McIntosh AR (2010) Brain noise is task dependent and region specific. J Neurophysiol 104:2667–2676. CrossRef Medline
- MišićB, Vakorin VA, Paus T, McIntosh AR (2011) Functional embedding predicts the variability of neural activity. Front Syst Neurosci 5:90. CrossRef Medline
- Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in macaque area v4. Neuron 63:879–888. CrossRef Medline
- Mizuno T, Takahashi T, Cho RY, Kikuchi M, Murata T, Takahashi K, Wada Y (2010) Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy. Clin Neurophysiol 121:1438–1446. CrossRef Medline
- Nimon K, Lewis M, Kane R, Haynes RM (2008) An R package to compute commonality coefficients in the multiple regression case: an introduction to the package and a practical example. Behav Res Methods 40:457–466. CrossRef Medline
- Nimon K, Oswald F, Roberts JK (2013) yhat: interpreting regression effects, r package version 2.0-0 edition. Vienna: R Foundation. Available at: https://rdrr.io/cran/yhat/
- Nunez PL, Srinivasan R (2006) Electric fields of the brain. New York: Oxford UP.
- Osipova D, Takashima A, Oostenveld R, Fernández G, Maris E, Jensen O (2006) Theta and gamma oscillations predict encoding and retrieval of declarative memory. J Neurosci 26:7523–7531. CrossRef Medline
- Park, SP, Kwon SH (2008) Cognitive effects of antiepileptic drugs. J Clin Neurol 4:99–106. CrossRef Medline
- Podvalny E, Noy N, Harel M, Bickel S, Chechik G, Schroeder CE, Mehta AD, Tsodyks M, Malach R (2015) A unifying principle underlying the extracellular field potential spectral responses in the human cortex. J Neurophysiol 114:505–519. CrossRef Medline
- Pozzorini C, Naud R, Mensi S, Gerstner W (2013) Temporal whitening by power-law adaptation in neocortical neurons. Nat Neurosci 16:942–948. CrossRef Medline
- Protzner AB, Valiante TA, Kovacevic N, McCormick C, McAndrews MP (2010) Hippocampal signal complexity in mesial temporal lobe epilepsy: a noisy brain is a healthy brain. Arch Ital Biol 148:289–297. Medline
- Ralph MA, Jefferies E, Patterson K, Rogers TT (2017) The neural and computational bases of semantic cognition. Nat Rev Neurosci 18:42–55. CrossRef Medline
- Ray S, Maunsell JH (2011) Different origins of gamma rhythm and highgamma activity in macaque visual cortex. PLoS Biol 9:e1000610. CrossRef Medline
- Ray-Mukherjee J, Nimon K, Mukherjee S, Morris DW, Slotow R, Hamer M (2014) Using commonality analysis in multiple regressions: a tool to

- decompose regression effects in the face of multicollinearity. Methods Ecol Evol 5:320-328. CrossRef
- Richman JS, Moorman JR (2000) Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 278:H2039–H2049. CrossRef Medline
- Salinsky MC, Oken BS, Morehead L (1994) Intraindividual analysis of antiepileptic drug effects on EEG background rhythms. Electroencephalogr Clin Neurophysiol 90:186–193. CrossRef Medline
- Schneidman E, Puchalla JL, Segev R, Harris RA, Bialek W, Berry MJ 2nd (2011) Synergy from silence in a combinatorial neural code. J Neurosci 31:15732–15741. CrossRef Medline
- Sederberg PB, Kahana MJ, Howard MW, Donner EJ, Madsen JR (2003) Theta and gamma oscillations during encoding predict subsequent recall. J Neurosci 23:10809–10814. Medline
- Sederberg PB, Gauthier LV, Terushkin V, Miller JF, Barnathan JA, Kahana MJ (2006) Oscillatory correlates of the primacy effect in episodic memory. Neuroimage 32:1422–1431. CrossRef Medline
- Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, McCarthy DC, Brandt A, Tully MS, Kahana MJ (2007) Hippocampal and neocortical gamma oscillations predict memory formation in humans. Cereb Cortex 17:1190–1196. CrossRef Medline
- Shew WL, Yang H, Petermann T, Roy R, Plenz D (2009) Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. J Neurosci 29:15595–15600. CrossRef Medline
- Shew WL, Yang H, Yu S, Roy R, Plenz D (2011) Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. J Neurosci 31:55–63. CrossRef Medline
- Sleimen-Malkoun R, Perdikis D, Müller V, Blanc JL, Huys R, Temprado JJ, Jirsa VK (2015) Brain dynamics of aging: multiscale variability of EEG signals at rest and during an auditory oddball task(1,2,3). eNeuro 2:ENEURO.0067-14.2015. CrossRef Medline
- Sokunbi MO (2014) Sample entropy reveals high discriminative power between young and elderly adults in short fMRI datasets. Front Neuroinform 8. CrossRef
- Sokunbi MO, Fung W, Sawlani V, Choppin S, Linden DE, Thome J (2013) Resting state fMRI entropy probes complexity of brain activity in adults with ADHD. Psychiatry Res 214:341–348. CrossRef Medline
- Stein RB, Gossen ER, Jones KE (2005) Neuronal variability: noise or part of the signal? Nat Rev Neurosci 6:389–397. CrossRef Medline
- Usher M, Stemmler M, Olami Z (1995) Dynamic pattern formation leads to 1/f noise in neural populations. Phys Rev Lett 74:326–329. CrossRef Medline
- Vakorin V, McIntosh AR (2012) Mapping the multiscale information content of complex brain signals. In: Principles of brain dynamics: global state interactions (Rabinovich MI, Friston KJ, Varona P, eds), pp 184–208. Cambridge, MA: MIT.
- Vaz AP, Yaffe RB Jr, Wittig JH Jr, Inati SK, Zaghloul KA (2017) Dual origins of measured phase-amplitude coupling reveal distinct neural mechanisms underlying episodic memory in the human cortex. Neuroimage 148:148–159. CrossRef Medline
- Voytek B, Knight RT (2015) Dynamic network communication as a unifying neural basis for cognition, development, aging, and disease. Biol Psychiatry 77:1089–1097. CrossRef Medline
- Voytek B, Kayser AS, Badre D, Fegen D, Chang EF, Crone NE, Parvizi J, Knight RT, D'Esposito M (2015) Oscillatory dynamics in human frontal networks in support of goal maintenance. Nat Neurosci 18:1318–1324. CrossRef Medline
- Waschke L, Wöstmann M, Obleser J (2017) States and traits of neural irregularity in the age-varying human brain. Sci Rep 7:17381. CrossRef Medline
- Yaffe RB, Kerr MS, Damera S, Sarma SV, Inati SK, Zaghloul KA (2014) Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. Proc Natl Acad Sci U S A 111:18727–18732. CrossRef Medline
- Yaffe RB, Shaikhouni A, Arai J, Inati SK, Zaghloul KA (2017) Cued memory retrieval exhibits reinstatement of high gamma power on a faster time-scale in the left temporal lobe and prefrontal cortex. J Neurosci 37:4472–4480. CrossRef Medline
- Yentes JM, Hunt N, Schmid KK, Kaipust JP, McGrath D, Stergiou N (2013) The appropriate use of approximate entropy and sample entropy with short data sets. Ann Biomed Eng 41:349–365. CrossRef Medline