Category-specific neural oscillations predict recall organization during memory search

Running Head: Oscillatory correlates of category clustering

Neal W Morton[1], Michael J. Kahana[2], Emily A. Rosenberg[2], Gordon H. Baltuch[3], Brian Litt[4], Ashwini D. Sharan[5], Michael R. Sperling[6], & Sean M. Polyn[1]

[1]Department of Psychology, Vanderbilt University, Nashville, TN 37240, [2]Department of Psychology, University of Pennsylvania, PA 19104, [3]Department of Neurosurgery, [4]Department of Neurology, Hospital of the University of Pennsylvania, Philadelphia, PA 19104, and [5]Department of Neurological Surgery, [6]Department of Neurology, Thomas Jefferson University, Philadelphia, PA 19107.

Direct correspondence to:

Dr. Sean M. Polyn
Vanderbilt University
Department of Psychology
PMB 407814
2301 Vanderbilt Place
Nashville, TN 37240-7817
Phone: 615-322-2536
Fax: 615-343-8449
E-mail: sean.polyn@vanderbilt.edu

Abstract

Retrieved-context models of human memory propose that as material is studied, retrieval cues are constructed that allow one to target particular aspects of past experience. We examined the neural predictions of these models in two free-recall experiments using electrocorticographic/depth recordings and scalp electroencephalography to characterize category-specific oscillatory activity recorded while participants studied and recalled items from distinct, neurally-discriminable categories. During study, these category-specific patterns predict whether a studied item will be recalled, as well as the order in which the items will be remembered, consistent with the proposal that a category-specific retrieval cue is constructed during the study period and deployed during memory search. These models suggest that integrative neural circuitry is critically involved in the construction and maintenance of the retrieval cue. Consistent with this, we observe category-specific patterns that rise in strength as multiple same-category items are studied sequentially, and find that individual differences in this category-specific neural integration during study predict the degree to which a participant will use category information to organize memory search. Finally, we track the deployment of this retrieval cue during memory search: Category-specific patterns are stronger when participants organize their responses according to the category of the studied material.

Keywords: category clustering, episodic memory, free recall, neural integration, pattern classification

Introduction

The electric fields of the brain, recorded with electrodes via scalp electroencephalography (EEG), intracranial electrocorticography (ECoG), and depth recordings, reveal a multitude of neurally generated signals related to human cognitive processing (Nunez and Srinivasan 2006; Jacobs and Kahana 2010). Coherent and rhythmic activation of neural populations can be detected both at the scalp and intracranially; this synchronous oscillatory activity has been related to single-unit spiking activity (Jacobs et al. 2007), and has been proposed to facilitate neural communication at both local and global spatial scales (Fries 2005; Buzsáki 2006). The spatiotemporal pattern of oscillatory activity across electrodes carries detailed information about stimulus characteristics (Freeman 1978; Jacobs and Kahana 2009) and task characteristics (Canolty et al. 2006). Furthermore, specific oscillatory components have been implicated in memory formation and retrieval, both in the local field around neurons, and at the scalp (Klimesch 1999; Sederberg et al. 2003; Summerfield and Mangels 2005; Nyhus and Curran 2010; Düzel et al. 2010; Liebe et al. 2012).

Here, we use a computational model of human memory to provide a functional interpretation of oscillatory neural signals recorded as people perform a memory task. A recent study by Manning et al. (Manning et al. 2011) reveals the promise of this approach. Using multivariate pattern analysis techniques (Duda et al. 2001), Manning et al. observed reactivation of study-period oscillatory patterns during memory search, consistent with a retrieved-context model of memory (Howard and Kahana 2002; Sederberg et al. 2008; Polyn and Kahana 2008; Polyn et al. 2009), in which a population of neural integrators (Kojima and Goldman-Rakic 1982; Fuster et al. 1982; Miller et al. 1996) is used to construct a retrieval cue while materials are being studied (Polyn and Kahana 2008; Manns et al. 2007). The retrieval cue is then deployed to allow the person to reactivate the details of recent experience.

We carried out two experiments using ECoG/depth recordings and scalp EEG, in which category-specific patterns of oscillations were characterized while participants studied and re-called items drawn from distinct taxonomic categories, allowing us to test three critical predictions of this neurocognitive account of memory search. First, the model proposes that during study, a participant constructs a category-specific retrieval cue to allow them to target items from that category during memory search. Thus, items eliciting strong category-specific neural activity at study will tend to be remembered during memory search, and will furthermore tend to be remembered in sequence with other same-category items. Second, the integrative process of retrieval cue creation suggests that category-specific patterns grow stronger as a series of same-category items are studied, and that the degree of neural integration will determine the degree to which memory search is organized by category. Third, during memory search, the retrieval cue integrates reactivated category-specific information, causing category-specific patterns to rise in strength when a participant recalls a series of items from the same category. In the reported experiments, we find evidence in support of each of these predictions.

Materials and Methods

Scalp Electroencephalography (EEG) Experiment

Participants. Forty-one paid volunteers (15 female, age 18–30 years) were recruited; 3 participants were excluded due to technical problems with the EEG recording apparatus, and 9 participants were excluded due to excessive eye movements, leaving 29 participants presented here. The research protocol was approved by the Institutional Review Board of the University of Pennsylvania.

Experimental paradigm. Stimuli consisted of color and black and white photographs of famous landmarks, celebrity faces, and common objects, with the name of the stimulus presented in text above the picture. There were 256 stimuli for each category. Stimuli were presented using pyEPL (Geller et al. 2007).

In a preliminary EEG session, participants rated their familiarity with each stimulus used in the experiment. This was done to assess participants' pre-experimental familiarity with each stimulus, to provide participants at least a minimal familiarity with each stimulus, and to provide us with category-specific oscillatory responses in the absence of the cognitive demands of a memory task. Stimuli were presented pseudorandomly, with the constraints that every group of three contained stimuli from each of the three categories, and that no two adjacent items were of the same category. Each stimulus was presented for 3500 ms, during which participants rated their familiarity with the stimulus' referent on a four-point scale. Each stimulus was followed by a blank interstimulus interval (ISI) of 1000±200 ms. Participants were given a chance to rest after each group of 48 items.

In the subsequent 3 sessions, participants were presented with 48 study-test lists. Each list was composed of 24 stimuli. There were two types of lists: *mixed-category* lists which contained 8 stimuli from each of the three categories, and *pure-category* lists which were composed of stimuli all drawn from the same category. In the mixed-category lists, items were presented in trains of same-category items, with each train containing 2–6 items. The order of category trains was pseudorandom, with the constraints that all categories appeared in each set of 3 trains, and that adjacent trains did not contain the same category. Each session contained 10 mixed-category lists and 6 pure-category lists. The pure-category lists were included to establish a baseline behavioral measure of temporal clustering, so this effect could be controlled for when examining category clustering (Polyn et al. 2009). The order of mixed-category and pure-category lists within each session was pseudorandom. Stimuli did not appear more than once within a session, and stimuli were chosen so that items from the same sub-category (e.g. stadiums, presidents) did not appear in the same list.

Each stimulus was presented for 3500 ms, during which participants made a category-specific four-point semantic judgment (celebrities: "How much do you love or hate this person?"; landmarks: "How much would you like to visit this place?"; objects: "How often do you come across this object in your daily life?"). Studied items where the participant didn't respond or responded faster than 300 ms were excluded from all analyses; 0–36 study epochs were excluded for each participant. Each stimulus was followed by a blank ISI of 1000±200 ms.

After presentation of the last stimulus, the screen was blank for 1300±100 ms, followed by presentation of a row of asterisks and a 300 ms tone signaling the start of a 90 s immediate free recall (IFR) period. Participants were instructed to recall items from the list in any order, without regard to stimulus category. Digital recordings of vocal recalls were scored using PyParse (Solway et al. 2010). Intrusions of items not in the word pool were scored to determine the category if possible (e.g. "Meryl Streep" was not in the word pool but is clearly a celebrity, while "rock" may have referred to an object or a partially recalled landmark and therefore had ambiguous category). Intrusions of ambiguous category were excluded from all analyses.

At the end of each session, there was a final free recall (FFR) period where participants were given 360 s to recall names of stimuli from any of the lists presented during the session.

Behavioral analysis. When asked to freely recall categorized materials, participants often will remember multiple same-category items sequentially, a phenomenon known as *category clustering* (see Fig. 1). We used the list-based semantic clustering index (LBC_{sem}; Stricker et al. 2002), to assess the degree of category clustering during IFR. A relabeling procedure was used to establish a baseline level of clustering expected due to the temporal contiguity of same-category items during study (Polyn et al. 2009). Each pure-category list was relabeled with a set of category labels by randomly sampling with replacement from the set of mixed-category lists for that subject. Mean LBC_{sem} was then calculated for the relabeled pure-category lists. The random relabeling procedure was repeated 10000 times to establish a null distribution of mean LBC_{sem} expected in the absence of category information. Because LBC_{sem} varies with list length, we used a different measure, the adjusted ratio of clustering (ARC) score, to compare category clustering in IFR and FFR (Roenker et al. 1971).

Scalp EEG recordings and data processing. EEG measurements were recorded using 129-channel HydroCel Geodesic Sensor Nets and a Net Amps 200 Amplifier (Electrical Geodesics, Inc.). An analog bandpass filter of 0.5–200 *Hz* was applied to recorded voltage, which was then digitized at 500 *Hz*. Recordings were initially referenced to Cz and were later converted to an average reference. In order to identify electrodes with poor contact, we first used multiple regression to remove signal related to vertical electrooculogram (VEOG) and horizontal electrooculogram (HEOG) measured using electrode pairs placed near the eye. We then created a distribution of the mean voltage for each electrode, and a distribution of the standard deviation of voltage fluctuations for each electrode. We identified an electrode as having poor contact if the absolute z-score (for either mean or standard deviation, compared to the corresponding distribution) was greater than 4. We excluded these electrodes when calculating the average reference. Line noise was removed using a Butterworth filter with zero phase distortion at 60 *Hz*.

We used a modified version of the eye motion correction procedure reported by Gratton et al. (Gratton et al. 1983) to remove blinks and eye movements. In order to better discriminate between blinks and eye movements, we identified blinks by applying a threshold to the difference between a fast and slow running average of the VEOG. Before each session of the experiment, participants were instructed to make 10 voluntary blinks and 20 eye movements (5 each of up, down, left, and right saccades) while HEOG and VEOG signals were recorded. The blink detector was applied to each participant's voluntary blinks and eye movements, and the threshold was adjusted to correctly identify at least 80% of the blinks while minimizing the number of eye movements incorrectly identified. The optimized blink detector was then applied to that participant's experimental data to identify time periods containing blinks. A buffer of 150 ms before and 500 ms after was added to each time sample identified as containing a blink to capture slower changes missed by the blink detector. Multiple linear regression was used to predict the signal at each electrode using (1) VEOG not containing blinks, (2) VEOG containing blinks, (3) HEOG not containing blinks, (4) HEOG containing blinks, and an intercept as predictors. The residuals from this regression were then used as corrected EEG. When calculating propagation factors, we did not subtract the average event-related potential (ERP) from each epoch as Gratton and colleagues (Gratton et al. 1983), because we found in an independent data set that correction performance was better when propagation factors were calculated on raw EEG rather than deviation scores (performance improved according to the metrics of variance after correction, and deviation from an estimate of the "true" ERP obtained from averaging events that passed a strict voltage threshold; Gratton et al. 1983). The EEG of 9 participants was not well-corrected by this procedure, due to large eye movement artifacts that were difficult to discriminate from blinks; these participants were excluded from the present analyses.

The EEG analyses presented below examine recordings made during familiarization, mixed-category study lists, mixed-category immediate free recall, and final free recall. In order to throughly characterize neural activity during the free-recall periods, we examined both continuous data including entire recall periods and segmented data locked to the onset of vocalized recalls.

Oscillatory analysis. We measured oscillatory power using a Morlet wavelet transform with a wave-number of 6. Oscillatory power was calculated at 34 logarithmically spaced frequencies from 2 to 100 *Hz*. Power values were then log-transformed and down-sampled to 25 *Hz*. Power was z-transformed relative to the mean and standard deviation of a baseline period, separately for each frequency, electrode, and session. For study epochs, the baseline period was 500–400 ms before stimulus onset. For recall epochs locked to vocalization onset, quiet times during the recall period (where no vocalizations were being made) were used as baseline; for each list, enough 100 ms baseline epochs were randomly chosen from quiet periods to match the number of recall events on that list. For analyses examining entire recall periods, power was normalized relative to all samples in a given recall period.

Multivariate pattern analysis. We used multivariate pattern analysis (Norman et al. 2006) to decode stimulus category based on patterns of oscillatory power. Classification was carried out using penalized logistic regression (penalty parameter = 10), using L_2 regularization (Duda et al. 2001). Classification analyses were carried out using the EEG Analysis Toolbox (available at: http://code.google.com/p/eeg-analysis-toolbox) and the Princeton MVPA Toolbox (available at: http://www.pni.princeton.edu/mvpa).

Study period classification. In all of our classification analyses, the classifier was trained on one set of epochs, and applied to a distinct set. We examined two methods for training the classifier before using it to decode stimulus category during study-period epochs. Using one method, we trained and tested on the study period using a cross-validation procedure, where the classifier was trained on study epochs from all lists except one, then tested on the study epochs from the remaining list. Classifier performance was measured as the fraction of test items whose category was correctly classified. This procedure was repeated with a different list left out on each iteration, and classifier performance was averaged over iterations. In a separate set of analyses, we trained the classifier on the familiarization period. This allowed us to examine the performance of a classifier trained on activity related to stimulus presentation, in the absence of intentional episodic encoding. We trained the classifier on all epochs of the familiarization period, then applied it to all epochs of the study period, again measuring performance as the fraction of items correctly classified.

Several sets of familiarization-period and study-period patterns were created for the analyses reported below. First, for each time-bin/frequency-bin pairing from a set of 100 time-bins and 34 frequency bins (Fig. 3B), we generated an across-electrode pattern for each stimulus presentation, where the value for each feature of the pattern was the oscillatory power at that electrode/time-bin/frequency-bin combination. Separate analyses examined performance of cross-validation classification, and performance of a classifier trained on the corresponding electrode/time-bin/frequency-bin of the familiarization epochs. In order to examine the category-specificity of these patterns at particular oscillatory frequencies (without regard to time), we averaged performance within 6 frequency bands and over all times bins during 0–3500 ms after stimulus onset (Fig. 3C). The frequency bands were: delta (2–4 *Hz*), theta (4–8 *Hz*), alpha (10–14 *Hz*), beta (16–25 *Hz*), low gamma (25–55 *Hz*), and high gamma (65–100 *Hz*). We used 100 *Hz* as the upper bound of high gamma to allow comparison of the ECoG and scalp EEG signals.

In order to examine the category-specificity of these patterns over time (without regard to frequency), we created a pattern for each study epoch and familiarization epoch containing average oscillatory power in each of eight 500 ms bins swept over the stimulus presentation period (Fig. 3D). Each pattern contained a feature for each electrode/frequency-bin pairing. To obtain a measure of overall classifier performance for a given item presentation, average oscillatory power was calculated for two time-bins: 0–0.5 s post-stimulus onset (*early* time bin) and 0.5–3.5 s post-stimulus onset (*late* time bin). A pattern was created for each familiarization epoch and study epoch containing a feature for each time-bin/frequency-bin/electrode pairing.

[Figure 1 about here.]

Classifier performance and subsequent recall. To determine how oscillatory activity during study affected subsequent recall performance, we labeled study events based on how the study item was later remembered. Items were labeled as *recalled* (recalled during IFR) and *forgotten* (not recalled during IFR). Recalled items were labeled based on whether they were *subsequently clustered* (recalled as part of a sequence of 2 or more items of the same category) or *subsequently isolated* (not recalled as part of a category cluster). These conditions are illustrated in Figure 1. Analyses below report how classifier accuracy changes with subsequent memory and subsequent clustering status of a particular item.

A number of follow-up analyses were carried out to ensure the validity of analyses contrasting classifier accuracy in different conditions. The first analysis altered the classifier training sets to ensure that analyses involving unbalanced numbers of items with different labels (e.g., *recalled* vs. *forgotten*) were unbiased. When creating each training set for the classifier, we ensured that each combination of category and the conditions of interest was equally represented, by sampling randomly without replacement from the set of training patterns corresponding to each category and condition. We repeated each classification analysis 10 times to obtain a stable estimate of performance for each classified item. Classification performance was then calculated for each condition of interest, averaged over all replications of the random sampling and classification.

We measured oscillatory power using Morlet wavelets, which were convolved with the EEG to obtain instantaneous estimates of power. Although oscillatory power measured using wavelets most strongly corresponds to oscillatory power at the measured time t, it will also be influenced by surrounding time points in the interval $[t - x, t + x]$, where x depends on the frequency and wavenumber of the wavelet (Herrmann et al. 2005). For all frequencies, we used a wavenumber of 6, so the measured interval varied with frequency; since the lowest measured frequency was 2 *Hz*, the largest window for which power was affected was $t \pm 1500$ ms, raising the possibility that classifier performance for a given item might be influenced by the category identity of an adjacent item on the study list. To control for this potential influence, we divided studied items based on whether the previous and next items were of the same or a different category. We divided items into train position bins based on whether they were at the beginning of a train, in the middle of a train, at the end of a train, or presented adjacent to items of different categories (this was never the case in the scalp EEG experiment, but did occur in the ECoG experiment described below). We then examined subsequent memory and subsequent clustering contrasts while controlling for train position bin. In no case did the category identity of surrounding study items influence the conclusions of an analysis.

Integration of category-specific activity. We examined whether the category-specificity of oscillatory patterns increased as multiple same-category items were studied in sequence (Fig. 5B). Earlier analyses examined overall classifier fraction correct, where the classifier's "guess" of the category of each stimulus was based on which category was estimated as being the most probable during each item presentation. Here, we instead examined a continuous measure of category-specific activity: For each presented item, we examined the classifier's estimate of the probability of the relevant category, given the pattern of neural activity observed during presentation of the item (Kuhl et al. 2011a). We tested for evidence of neural integration of category-specific activity by examining whether classifier estimates increased with successive presentations of items in the same category. To determine whether integrative activity was related to individual differences in category clustering (Fig. 5C), we used weighted least-squares regression (weighted by the number of observations at each train position) to fit the change in classifier estimates over train positions 1–3 for each participant; we refer to this as *neural integration rate*. We then examined whether neural integration rate predicted individual differences in category clustering, by measuring the correlation between neural integration rate and LBC_{sem}.

A secondary analysis examined whether within an individual participant, differences in the amount of category clustering observed in individual lists correlated with neural integration rate for the studied items *in that particular list*. Each participant performed free recall on 30 mixed-category lists, across three experimental sessions. For each trial we calculated both the neural integration rate and the degree of category clustering. For each participant, we obtained the *t*-value of the slope of the regression of category clustering on neural integration rate. We then used a *t*-test to assess whether the regression *t*-value was significantly positive across subjects. A significantly positive slope indicates a significant relationship between list-level fluctuations in neural integration rate and category clustering.

Reactivation during recall. To examine whether patterns of oscillatory power observed during study were reactivated during recall, we trained the classifier on average power from the late time bin (see above) of item presentation, then applied the classifier to oscillatory power recorded during recall. We assessed the degree of reactivation of category-specific oscillatory patterns during the recall period using a correlation-based reactivation metric (Polyn et al. 2005).

The classifier provides an estimate of the strength $\mathbf{S}_t^i$ of each category i at each time bin t. The record of recalls during each free recall period was sampled at 25 *Hz* to match the sampling rate of the oscillatory power. Each time bin was either assigned no category (if no recalls were currently being made), or to exactly one category. The 1 s preceding onset of each vocalized recall was labeled with the category of the recalled item. When there was overlap between recalls, the earlier item took precedence. This resulted in a set of 3 vectors $\mathbf{R}^j$, where each element $\mathbf{R}_t^j$ is 1 for times t when category j is active, and 0 when category j is not active. These vectors represent the *recall record* of each recall period.

We calculated a correlation-based reactivation metric to measure reactivation of category patterns during recall. We treated all recall periods as part of one record by concatenating the recall periods together. We calculated Pearson's linear correlation between $\mathbf{S}^i$ and $\mathbf{R}^j$ for $i = 1, 2, 3$ and $j = 1, 2, 3$ to create a cross-correlation matrix. The diagonal of the cross-correlation matrix corresponds to correlations between classifier estimates and the correct recall records, while the off-diagonal entries correspond to correlation with the incorrect categories. We calculated the mean correlation in the diagonal entries and subtracted the mean correlation in the off-diagonal entries to obtain a summary index of the classifier's ability to track each subject's recall behavior, which we refer to as the reactivation metric (this measure was referred to as the OnOff metric by Polyn et al. 2005).

We used a permutation test to determine whether reactivation was statistically significant across subjects. For each subject, the columns of the cross-correlation matrix were scrambled, and the mean reactivation metric was calculated. This process was repeated 5000 times to establish a null distribution of reactivation metric scores, and reactivation was considered significant if the observed score was greater than 95% of the null distribution. We also examined reactivation at different frequencies by training and testing the classifier at each frequency individually. In order to control Type I error rate while accounting for the correlation structure of the data, we scrambled the columns of the cross-correlation matrix in the same way for each frequency, then pooled the null distributions of each frequency together to make one null distribution accounting for familywise error. This familywise null distribution was then used to set the significance threshold for all frequencies (Sederberg et al. 2003).

Category-specific activity during clustering. We also examined recall-period category-specific activity by training the classifier to identify the category of particular recalled items. The oscillatory patterns created for this analysis used a single time-bin averaged from 3 to 0.5 s before onset of vocalization, referred to as a *recall epoch* (the 0.5 s immediately before each vocalization was excluded in order to limit the influence of vocal response preparation artifacts), and all frequency-bins. Recall epochs were excluded if they overlapped with vocalizations of previous recalls (immediate free recall: 68.5% [SEM 1.3%] of epochs were excluded, leaving 44–146 epochs for each subject; final free recall: 78.0% [SEM 1.3%] of epochs were excluded, leaving 31–93 epochs for each subject). Performance was assessed using cross-validation, with one list left out on each iteration, and performance was averaged over iterations.

We compared mean classifier accuracy for clustered and isolated items, to test the hypothesis that category clustering is associated with stronger category representations. Clustered items were labeled according to the item's position within the category cluster: *initial* (preceded by an item from a different category, and followed by an item from the same category), *middle* (both preceded and followed by items from the same category), or *terminal* (preceded by an item from the same category, and followed by an item from a different category), as illustrated in Figure 1. Note that these cluster position bins are defined in a similar manner to the train position bins used for the study period (see *Classifier performance and subsequent recall*), but apply to the order in which items are recalled rather than the order in which items are presented. We used a two-way within-subjects ANOVA with previous category (same or different) and next category (same or different) as factors, to test for influences of the previous and next recalls. As with our study-period analyses, we examined both performance of a classifier trained on all events, and performance of a classifier provided with a balanced training set. For the balanced analysis, we used random sampling without replacement to create a training set with equal numbers of epochs from each combination of cluster position bin and category. We repeated the random sampling and classification 10 times to obtain a stable estimate of classifier performance.

As for the study-period analyses, we also carried out follow-up analyses to rule out the possibility that the wavelet-based power estimates were influenced by the neural signal related to adjacent events in the recall sequence. Since our wavelet estimates of instantaneous oscillatory power are influenced by oscillations within an extended interval, classification of items recalled as part of a cluster may be improved by the influence of oscillatory power related to adjacent recalls of same-category items. If clustered items are better classified due to influence of nearby recalls on power estimates, this difference should only appear for time bins that are less than 1500 ms from the closest recall event. Therefore, we focused on the period from 1500 to 500 ms before vocalization onset (using 500-ms time bins), which cannot be influenced by adjacent recalls (based on our criteria for creating recall epochs, see above). We averaged classifier performance over this interval for items following a recall of the same category (middle/terminal) and items following a recall of a different category (isolated/initial) to determine whether the category of the previous recall has an effect on classifier performance. Similarly, we compared classifier performance during IFR and FFR over the critical period from 1500 to 500 ms before vocalization onset. In no case did the category identity of surrounding recall events influence the conclusions of our analysis.

[Figure 2 about here.]

[Table 1 about here.]

Electrocorticography (ECoG) Experiment

Participants We tested 11 patients (3 female; age 18–44, mean 35.5, SD 8.2) with medication-resistant epilepsy who were undergoing invasive ECoG and depth electrode monitoring to determine the location of epileptogenic foci for subsequent resection. See Table 1 for detailed patient information. The patients had a total of 864 surface and depth electrodes (Fig. 2); electrode placement was determined by the clinical team.

The research protocol was approved by the relevant institutional review boards, and informed consent was obtained from all participants. To limit the effects of seizures and medication on task performance and brain activity, we refrained from testing when patients were on high doses of pain medications or anti-epileptic drugs, and during the 6 hour period following any clinically significant seizure. For 2 sessions, patient 7 was implanted with one set of electrodes. This patient underwent another surgery prior to her remaining 8 sessions. During this surgery, some electrodes were added, and some were removed to create a second set of electrodes. In the reported classification analyses, we treated these 2 sets of sessions as coming from distinct participants. The number of electrodes that overlapped between the 2 sets of sessions was relatively small (24 electrodes; 26.1% of the first set of electrodes, and 27.9% of the second set), precluding our ability to combine these 2 sets of data for pattern classification analysis.

Materials The word pool consisted of the 216 items from the scalp EEG experiment that were the most recognizable (as judged by the experimenters). In addition to the original picture used in the scalp EEG experiment, 4 additional pictures were found for each item.

Procedure Participants were presented with lists of 9 items, with 3 items from each category. Category order was pseudorandom within each list. Before each item, a text cue indicating the category of the upcoming item was shown for 1000 ms. There was a 200–500 ms ISI before presentation of the item, which lasted for 3500 ms. While the item was on the screen, participants made a category-specific judgment, as in the scalp EEG experiment. The ISI between each item and the next category cue was 800–1200 ms. After presentation of the last stimulus, the screen was blank for 1200–1400 ms, followed by presentation of a row of asterisks and a 300 ms tone signaling the start of a 60 s IFR period. If 60 s had not passed yet, but the participant indicated that he or she had finished recall, the experimenter pressed a button to end the recall period. Each item had 5 distinct pictures, which all appeared during the session (but never in the same list). Participants were told that the same item might appear multiple times, but to simply focus on remembering the items from the current list. Participants were presented with 20 lists in each session. There was a 240 s final free recall test at the end of each session. Each participant completed 1–10 sessions (see Table 1 for the number of sessions completed by each participant).

ECoG recordings and data processing ECoG was recorded using a Grass Telefactor or Nicolet digital video-EEG system. ECoG was sampled at 400 or 512 *Hz*. A digital Butterworth notch filter with zero phase distortion at 60 *Hz* was used to remove electrical noise. Synchronization pulses controlled by the computer presenting the stimuli were sent to the EEG monitoring system, and later used to align electrophysiological data to events in the experiment (precision < 4 ms).

Oscillatory power was measured at 37 logarithmically spaced frequencies from 2 to 128 *Hz*. Power was log-transformed and down-sampled to 16 *Hz*. Power was normalized using similar techniques as used in the scalp EEG experiment, except power measured during study epochs was normalized relative to 500–400 ms before onset of the category cue (rather than the onset of the stimulus itself). Epochs examined during the recall period consisted of data from 2000 ms before to 1000 ms after vocalization onset. Epochs were only included if they did not contain previous vocalizations (immediate free recall: 70.8% [SEM 4.5%] of epochs were excluded, leaving 5-744 epochs for each subject; final free recall: 67.6% [SEM 3.9%] of epochs were excluded, leaving 3-90 epochs for each subject). Power was normalized relative to periods with no vocalizations; for each recall period, enough 100 ms baseline epochs were randomly chosen from quiet periods to match the number of recalls during that recall period. We also examined continuous data including entire recall periods; power was z-transformed based on the mean and standard deviation of power over each recall period, separately for each electrode and frequency.

The locations of the intracranial electrodes were determined using an indirect stereotactic technique based on co-registered post-operative computed tomography and pre- or post-operative magnetic resonance imaging, and converted into Montreal Neurological Institute coordinates. The Talairach Atlas was used to determine the anatomical location of each electrode (Talairach and Tournoux 1988; Lancaster et al. 2000). Electrodes were divided into 7 regions of interest (ROIs; see Fig. 2): frontal (220 electrodes), prefrontal (188), temporal (532), medial temporal (76), hippocampus (22), occipital (56), and parietal (57). The prefrontal ROI is a subset of electrodes in the frontal ROI; similarly, the hippocampal ROI is a subset of the medial temporal ROI, which is a subset of the temporal ROI. We used brain images from the WFU Pick-Atlas for data visualization (Maldjian et al. 2003).

Multivariate pattern analysis Pattern analysis methods were the same as in the scalp EEG experiment, except that classification analyses were carried out separately for each ROI. Significance of classifier performance during study was assessed using a permutation test. The labels corresponding to each category were permuted 5000 times, and the mean fraction correct over participants was calculated for each permutation. The same permutations were used across all ROIs. The permuted distribution of fraction correct scores was pooled over all ROIs to create one null distribution, which was used to establish a significance threshold that controls familywise Type I error at $\alpha < 0.05$ (Sederberg et al. 2003). A similar method was used to assess significance of reactivation during recall; in this case, columns of the cross-correlation matrix for each participant were permuted to calculate a null distribution of reactivation metrics, which was pooled over all ROIs to set familywise Type I error at $\alpha < 0.05$.

To assess the time-course of reactivation during recall, we divided recall epochs into 500 ms time bins, and ran a separate classification analysis for each bin. Significance of classifier performance was assessed using a permutation test. The labels corresponding to each category were permuted 5000 times, and the mean fraction correct over participants was calculated for each permutation. The same permutations were used across all time bins and ROIs. The permuted distribution of fraction correct scores was pooled over all time bins and ROIs to create one null distribution, which was used to establish a significance threshold that controls familywise Type I error at $\alpha < 0.05$ (Sederberg et al. 2003).

We also examined whether reactivation of category-specific oscillations during recall was restricted to particular frequency bands. We ran a separate classification analysis for each frequency to determine the degree of reactivation of each frequency of oscillation. Significance was assessed using a permutation test where the columns of each participant's cross-correlation matrix were permuted 5000 times; the significance threshold was determined by pooling all frequencies and ROIs into one null distribution (familywise Type I error < 0.05; Sederberg et al. 2003).

Results

An overview of the modeling framework

When participants freely recall studied material, the order of their responses reveals the associative structure of their stored memories (Puff 1979; Polyn et al. 2009). Retrieved-context models of memory (Howard and Kahana 2002; Sederberg et al. 2008; Polyn et al. 2009) explain these organizational phenomena (as well as a vast array of other behavioral phenomena) in terms of the interactions between a representation of the studied material, and an internal retrieval cue. The retrieval cue is characterized as a population of integrative elements; the persistent activity of these integrators causes the retrieval cue to slowly change its state. One characteristic of these models is that the retrieval cue is both always active, and ever-changing. During study, the system (1) integrates details of the studied items into the retrieval cue itself, and (2) engages associative processes that directly link the cue with the neural representation of the studied material. These characteristics of the system allow it to more accurately target the studied material during a later search attempt. In these experiments, participants studied items from categories associated with distinct neural representations (Polyn et al. 2005), allowing us to track this category-specific neural activity as it is integrated into the retrieval cue, and deployed during memory search. Using these category-specific oscillatory patterns, we are able to predict which studied items will be subsequently remembered (Kuhl et al. 2011b), as well as the order in which they will be remembered. Furthermore, we are able to track category-specific activity as memory search unfolds. In the following sections, we use the retrieved-context framework to interpret the dynamics of category-specific patterns of oscillatory neural activity, recorded both at the scalp, and with ECoG/depth electrodes.

[Figure 3 about here.]

Category-specific oscillatory patterns during encoding

Central to modern cognitive neuroscientific theory is the hypothesis that the characteristics of a particular study event are reflected in a distributed, attribute-based representation that spans multiple brain areas (Haxby et al. 2001; Polyn et al. 2005; Martin 2007). Both the local oscillatory signals picked up by the intracranial electrodes, and the more global signals picked up by the scalp electrodes reveal distinct patterns of oscillatory activity associated with the category identity of a studied item. Overall classifier percent correct (allowing the classifier to use 2 time-bins, 34 frequency-bins, and all electrodes to decode stimulus category using a cross-validation procedure; see Materials and Methods) was 58.9% (SEM 1.0%) and 82.4% (SEM 2.6%) for the scalp EEG and ECoG experiments, respectively. As we describe in this section, the category-specific oscillatory patterns recorded by each of these techniques are quite similar in terms of their time-course and frequency profile, though the ECoG/depth recordings contain significantly more category-specific high-gamma activity.

Figure 3 depicts a number of analyses characterizing the time-course and frequency profile of these category-specific patterns, for each of these datasets. As described above (*Study Period Classification methods*) we created two sets of category-specific patterns: one corresponding to the study events in the free-recall sessions, and one corresponding to the stimuli presented during an independent familiarization session (this latter set was only collected as part of the scalp EEG experiment). Here, we characterize the category specificity of the patterns recorded during the free-recall sessions (as this was part of both experiments), and in later sections we contrast the free recall study events with the category-specific patterns collected during the familiarization session in the scalp EEG experiment.

We first conducted a series of classification analyses separately for each time-bin/frequency-bin pairing, relative to item onset. This allowed us to map out category-specificity of the neural signal for both intracranially implanted (Fig. 3A) and scalp EEG–monitored (Fig. 3B) participants. The time-frequency distribution of category-specific neural signals is remarkably similar across the two groups of participants, suggesting that the scalp electrodes are sensitive to the same category-specific patterns characterized by the intracranial electrodes. Category specificity at widespread frequencies is seen in the first 500 ms after item presentation, and longer-lasting category differences are observed in the delta (2–4 *Hz*), theta (4–8 *Hz*), and alpha (10–14 *Hz*) frequency bands. Sustained high gamma (65–128 *Hz*) category differences are observed in the recordings from intracranial electrodes (Fig. 3A). Gamma-band (30–100 *Hz*) activity at the scalp is also somewhat sensitive to stimulus category (Fig. 3B).

Figure 3C presents the category-specific response at different frequency bands during the study period, for each of these datasets. We averaged classifier cross-validation performance over the entire stimulus presentation period (0–3500 ms post-stimulus onset) for 6 frequency bands: delta, theta, alpha, beta (16–25 *Hz*), low gamma (25–55 *Hz*), and high gamma (65–100 *Hz*). Here, we used 100 *Hz* as the upper bound of high gamma to allow comparison of the ECoG/depth and scalp EEG signals. Classifier performance for the intracranial experiment was greater than scalp EEG ($F(1, 234) = 188.66$, $p < 0.0001$). There was also a main effect of frequency ($F(5, 234) = 15.45$, $p < 0.0001$), and a significant interaction ($F(5, 234) = 3.51$, $p < 0.005$). There was an interaction between ECoG and scalp EEG in the low and high gamma bands ($F(1, 78) = 10.88$, $p < 0.002$), with the ECoG data showing a greater advantage for high gamma over low gamma (Fig. 3C). The lack of an increase in classifier performance for high gamma (over low gamma) in the scalp EEG experiment may reflect attenuation of high-frequency oscillations by the skull (Nunez and Srinivasan 2006).

Although gamma-band oscillations were attenuated at the scalp electrodes, classifier performance for frequencies in the gamma band was still reliably above chance during the study period, consistent with research suggesting that induced gamma activity is involved in perceptual binding during object perception (Tallon-Baudry et al. 1996, 1997). Recent work raises the possibility that some high-frequency EEG activity measured at scalp electrodes is related to miniature saccades, rather than brain activity (Yuval-Greenberg et al. 2008). Voltage potentials related to these miniature saccades do not differ in polarity on different sides of the eye, so our regression procedure for subtracting the influence of eye movements (which relied on difference potentials to measure eye movements) would not be effective at removing these signals (Yuval-Greenberg et al. 2008). If participants made distinct micro-saccadic activity for each of the three categories, these signals could affect our analyses. Thus, we carried out a second series of analyses on the scalp EEG signal from the study period, in which oscillations of frequency higher than 30 *Hz* were excluded, and found that overall classifier performance was very similar, and the conclusions from all reported analyses were unchanged. This suggests, for scalp EEG at least, either that the information contained by the high-frequency category-specific activity is redundant with the low-frequency information, or that the most important category-specific activity is carried by lower frequencies.

Figure 3D shows the time-course of category-specific oscillatory patterns relative to item onset. For each time-bin, the classifier is trained on oscillatory information from all frequencies. Classifier performance peaks at 500–1000 ms for both datasets, but category-specific information persists for at least as long as the item remains on-screen. We found some evidence in the scalp EEG study that category-specific information related to the previous study item persists during the inter-item interval. When the classification analysis for this baseline period includes all study events classifier performance is significantly above chance (mean 35.81%, SEM 0.45%, $t(28) = 5.53$, $p < 0.0001$). In this experiment 69% of items are preceded by an item of the same category, raising the possibility that persistent category-specific activity could influence these estimates. To obtain a clean baseline estimate of category discriminability, we restricted our analysis (for this bin) to items presented immediately after a category switch; the other time bins show performance averaged over all items. This change caused baseline classification to drop to chance levels. In the ECoG/depth experiment, each item was preceded by a cue indicating the category of the upcoming item. Classifier performance in the 500 ms before stimulus onset was significantly above chance (mean 37.81%, SEM 1.53%, $t(11) = 24.7$, $p < 0.0001$; see Fig. 3D); this may be due to activity related to anticipation of the stimulus category or preparation for the category-specific judgment to be made about the item.

Category-specific information predicts subsequent memory

While this category-specific neural activity could reflect a neural representation of the semantic characteristics of the studied item, there are potentially many reasons why neural activity might reflect category identity. Thus, it is important to determine whether these category-specific activation patterns tell us anything about the memorability of the stimulus itself. The next set of analyses show that, in line with this hypothesis, the strength of the oscillatory patterns elicited for a particular studied item provide information about whether that item will be recalled. This is consistent with the hypothesis that a substantial component of these category-specific patterns relates to the representation of the studied item, the representation of a category-based retrieval cue, or some combination of the two.

[Figure 4 about here.]

In the scalp EEG dataset, we found that subsequently recalled items were classified more accurately than subsequently forgotten items. This was the case regardless of the training set used to train the classifier. When the classifier was trained on activity observed during the study period of free-recall trials (using a cross-validation procedure), classifier performance was greater for subsequently recalled items (percent correct: mean 60.4%, SEM 1.2%) than for subsequently forgotten items (percent correct: mean 57.5%, SEM 0.9%); this difference was significant ($t(28) = 3.60$, $p < 0.0001$). Similarly, when the classifier was trained on the independent familiarization session (in which there were no demands regarding a later memory test), classifier performance was greater for subsequently recalled items (percent correct: mean 54.8%, SEM 1.0%) than for subsequently forgotten items (percent correct: mean 52.0%, SEM 0.9%); this difference was significant ($t(28) = 3.15$, $p < 0.005$).

In the intracranial experiment, during the study period, every brain region which had substantial electrode coverage showed reliable category-specific differences in oscillatory power (Fig. 4A; $p < 0.05$, permutation test). While temporal and occipital regions showed the strongest category-specific activity, the uneven electrode coverage of different regions of interest across patients makes it difficult to draw strong conclusions about the relative contributions of different areas. However, we can examine how category-specific activity within a particular region of interest changes under different experimental conditions. In the ECoG/depth dataset, we found that there was no difference in classifier performance between recalled and forgotten items at any ROI (all $p > 0.05$, Bonferroni corrected). Since performance at temporal electrodes is near ceiling (mean 79.22%, SEM 2.96%), we also examined whether the raw classifier estimates predicted subsequent memory status. Classifier estimates are free to vary continuously, so they may be more sensitive in some cases than fraction correct, which is binary for each classified item (Kuhl et al. 2011b). We found that classifier estimates at temporal electrodes were significantly greater for subsequently recalled than forgotten items (Fig. 4B; $t(11) = 8.50$, $p < 0.0005$, Bonferroni corrected). While occipital electrodes also showed strong category-related oscillatory signal (Fig. 4A), the fidelity of this signal did not predict whether an item would be remembered ($t(9) = 2.91$, $p > 0.05$, Bonferroni corrected).

In each of these datasets, we carried out a follow-up analysis of variance to examine whether classifier performance was influenced by the sensitivity of the wavelet-based power estimates to the category identity of surrounding studied items. As described in the methods, we included a factor in the analysis of variance relating to the category identity of the surrounding items (*Classifier performance and subsequent recall*). In the scalp dataset, this revealed a significant main effect of subsequent memory ($F(1, 28) = 9.98$, $p < 0.005$), no effect of surrounding category identity ($F(2, 56) = 1.56$, $p = 0.22$), and no interaction between these factors ($F(2, 56) < 1$). Similarly, in the ECoG/depth dataset, this revealed a significant main effect of subsequent memory ($F(1, 11) = 12.55$, $p < 0.005$) at temporal electrodes, but no effect of surrounding category identity ($F(3, 33) < 1$), and no interaction between these factors ($F(3, 33) < 1$). We also carried out a follow-up analysis in which the number of recalled and forgotten items (as well as the number of items from the different categories) were balanced within the training set (see Materials and Methods). We found that, even with a balanced training set, classifier estimates in temporal electrodes were significantly greater for subsequently recalled items compared to forgotten items ($t(11) = 6.98$, $p < 0.0005$, one-sided test). This suggests that classifier estimates are greater for subsequently recalled items because they are associated with higher-fidelity category activity, and not merely because there are more subsequently recalled items in the training set. The other ROIs showed no significant differences (all $p > 0.05$, Bonferroni corrected).

Category-specific activity predicts subsequent recall organization

Finding that category-specific activity is related to the memorability of particular stimuli suggests that this activity is, at the very least, involved in creating a neural representation of the stimulus (encoding). For example, if a person's attention wanes for a particular stimulus, then the neural representation of that stimulus will be weak or not present, which will cause that stimulus to be poorly remembered. Retrieved-context models suggest that if some component of the observed category-specific oscillatory activity is related to the operation of a category-based retrieval cue, then an item with a strong category response during study should not only be better recalled, but there should also be an increased likelihood of that item being recalled in succession with other items from the same category (i.e., the item should be clustered with same-category items during recall).

We first present results from our behavioral analyses documenting the presence and reliability of category-related organization of responses during the free-recall periods. Participants exhibited reliable category clustering in both experiments. In the scalp EEG experiment, LBC_{sem} in immediate free recall (IFR) was 3.66 (SEM 0.25); this exceeded the amount of category clustering expected given temporal influences on recall, calculated using a relabeling procedure (mean 0.808, SD 0.061; $p < 0.0002$; Polyn et al. 2009). In the ECoG/depth experiment, because items from each category were randomly placed in the list, LBC_{sem} expected by chance is 0. LBC_{sem} in IFR was 1.19 (SEM 0.16), which was significantly greater than chance ($t(11) = 7.39$, $p < 0.0001$). While both experiments showed evidence that participants were organizing their memory search by category, only the scalp EEG experiment yielded reliable evidence that the category-specific oscillatory responses were related to this category organization. This may be due to the global estimate of category-specific neural response provided by scalp EEG, or may be due to the larger number of participants and longer study lists of the scalp EEG experiment, which leads to a larger set of study and recall events to examine.

We found that item-level fluctuations in classifier performance predict subsequent clustering by category; we refer to this as the *subsequent clustering effect*. When the classifier was trained on the familiarization period, then used to identify the category corresponding to each study item, the subsequently clustered items were identified more reliably than the items that would be forgotten ($t(28) = 3.26$, $p < 0.005$), and more importantly, than the subsequently isolated items ($t(28) = 2.39$, $p < 0.05$; Fig. 5A). As above, a follow-up analysis of variance showed that this effect was not influenced by the category identity of surrounding items (we found a main effect of subsequent clustering [$F(1, 28) = 5.20$, $p < 0.05$], no effect of surrounding category identity [$F(2, 56) < 1$], and no interaction [$F(2, 56) = 1.42$, $p = 0.25$]).

We found that while the classifier trained on the familiarization session was sensitive to subsequent clustering, the classifier trained on the free-recall sessions was not. We carried out an analysis to examine this differential sensitivity. We ran an analysis of variance on the classifier performance to examine whether there was an interaction between training period (familiarization or study) and subsequent organization (isolated or clustered). There was a significant main effect of training period ($F(1,28) = 32.42$, $p < 0.0001$; accuracy was better when the classifier was trained on the study period), no main effect of subsequent organization ($F(1,28) = 2.53$, $p = 0.13$), and no interaction ($F(1,28) = 1.75$, $p = 0.20$). Therefore, although there is a significant difference in classifier performance between clustered and isolated items when the classifier is trained on the familiarization period, and no difference when the classifier is trained on the study period, the magnitude of the difference in classifier performance between subsequently clustered and subsequently isolated items does not significantly interact with training period. This raises the possibility that it is not necessarily a fundamental difference in the familiarization vs. free-recall category patterns that is causing the difference in sensitivity to subsequent clustering; it may be a matter of statistical power. Along these same lines, we also found that the difference in classifier performance between recalled and forgotten items does not depend on the training period. Another analysis of variance revealed a significant main effect of training period ($F(1,28) = 53.78$, $p < 0.0001$), a significant main effect of subsequent memory ($F(1,28) = 16.87$, $p < 0.0005$), and no interaction ($F(1,28) < 1$).

[Figure 5 about here.]

Integration of category-specific information

According to retrieved-context models of memory, organizational effects arise through the interaction of a retrieval cue with the contents of memory. When many items on the study list have similar characteristics (e.g., they are from the same category), one can construct a retrieval cue that contains those common characteristics to effectively retrieve those items during memory search. According to these models, the processes that construct the retrieval cue are integrative: they create a representation that changes its state slowly over time. This allows the retrieval cue to synthesize the properties of a particular episode, and serve as an effective cue for the events occurring over a rather large temporal interval (Howard and Kahana 2002). Thus, for a particular neural signal to be a candidate for being part of a retrieval cue, one would predict that it would be sensitive to the category identity of previous stimuli.

As in the above analysis of subsequent clustering, we found that the more global scalp EEG signal contained evidence for integrative activity, while the EGoG/depth signal did not. This may be due to the design of the ECoG/depth experiment, where same-category items did not often occur in sequence. For a classifier trained on the scalp EEG study period (using a cross-validation procedure), we found that the fidelity of the observed category-specific neural activity increased as multiple items from the same category were presented in succession (Fig. 5B). The classifier estimate for the category corresponding to the studied item increased for the first three positions of a same-category train of items and leveled off beyond that. A weighted least-squares regression (weighted by the number of observations at each train position) was used to fit the change in classifier estimates over train positions 1–3 for each participant. The mean slope was 0.0078 (SEM 0.002), which was significantly positive ($t(28) = 3.68$, $p < 0.001$).

We found that individual differences in the slope of classifier estimates over train position significantly correlated with participants' tendency to engage in category clustering during memory search as measured by $\mathrm{LBC}_{\mathrm{sem}}$ (Fig. 5C; $r = 0.421$, $p < 0.05$; with 2 outliers removed, $r = 0.500$, $p < 0.01$). In contrast, individual differences in overall discriminability of category patterns at study did not correlate with category clustering during recall ($r = 0.268$, $p = 0.18$; with 2 outliers removed, $r = 0.251$, $p = 0.19$).

We also found that fluctuations in the slope of these category-specific estimates on a trial-by-trial basis were related to trial-by-trial fluctuations in category clustering behavior within a given subject. For each trial we calculated both the slope of classifier estimates (the *neural integration rate*) and the degree of category clustering, and performed a regression on these two measures. We found that these two measures were reliably positively related to one another (t-value of the slope of the regression: mean 0.476, SEM 0.160; $t(28) = 2.98$, $p < 0.01$), indicating a significant relationship between list-level fluctuations in neural integration rate and category clustering.

The training period is important for determining whether the classifier is sensitive to effects of integration of category representations over multiple item presentations. When the classifier was trained on the familiarization period, and tested on the study periods from the free-recall sessions, there was no increase in classifier estimate with train position (slope over train positions 1–3, based on weighted least-squares regression: mean 0.0010, SEM 0.0021, $t(28) = 0.51$, $p = 0.31$, one-sided test compared to 0). Slope was significantly greater when the classifier was trained on the study period ($t(28) = 3.12$, $p < 0.005$). Furthermore, when the classifier was trained on the familiarization period, the slope of classifier estimates did not correlate with LBC_{sem} ($r = 0.00062$, $p = 0.997$). A dependent correlations test showed that LBC correlates significantly better with slope for study-session cross-validation than for familiarization-to-study classification ($t(26) = 2.30$, $p < 0.05$). Similarly, the slope of classifier estimates is more sensitive to differences in the amount of clustering on individual lists when the classifier is trained on the study period, compared to when it is trained on the familiarization period (t for the slope of the regression of LBC_{sem} on slope of classifier estimates: mean 0.190, SEM 0.159). This difference is marginally nonsignificant ($t(28) = 1.96$, $p = 0.061$).

These results suggest that the classifier trained on the familiarization session stimuli (where there are no memory demands) is not sensitive to this integrative category-specific neural activity. This is consistent with the retrieved context framework, which suggests that there are two kinds of category-specific activity: the item representations themselves (which are not integrative) and the category retrieval cue (which has integrative properties). We return to this point in the discussion.

Reactivation of category-specific patterns during memory search

According to retrieved context models (indeed, many models of memory), when one remembers a past event, the memory system reactivates the pattern of neural activity that prevailed when that event occurred. This reactivation process has been used to describe remembering and reminiscence as a form of *mental time travel* (Tulving 1993; Wheeler et al. 1995; Polyn et al. 2005; Danker and Anderson 2010). Both the scalp EEG and the ECoG/depth studies showed evidence for reactivation of category-specific oscillatory patterns during memory search, although the patterns recorded with ECoG/depth electrodes were of substantially higher fidelity.

These ECoG category-specific oscillatory patterns reactivate during memory search, and this reactivation tracks which category is being recalled by the participant on a moment-to-moment basis. There was reliable reactivation in frontal, prefrontal, temporal, medial temporal, hippocampal, and occipital electrodes (Fig. 4C; $p < 0.05$, permutation test). Notably, the category-specific occipital patterns observed while the stimulus was presented visually were only weakly reactivated during memory search. Classifier accuracy peaks during the 1 s before onset of vocalization, then drops during vocalization of the recalled item. In temporal electrodes, oscillatory patterns at all frequencies were reactivated; in medial temporal electrodes, delta, theta, alpha, and beta patterns were reactivated; and in frontal electrodes, delta, theta, and alpha patterns were reactivated.

The ECoG experiment also revealed reliable reactivation of category-specific patterns during the final free recall period at frontal, prefrontal, temporal, medial temporal, hippocampal, and parietal electrodes (Fig. 4D; $p < 0.05$, permutation test). Occipital category-specific patterns were significantly less reactivated in final free recall compared to immediate free recall ($t(9) = 3.76$, $p < 0.05$, Bonferroni corrected). Reactivation follows a similar time-course as in IFR, with classifier performance peaking around 1 s before onset of vocalization, then decreasing after vocalization. Temporal electrodes demonstrated reactivation in all frequency bands except high gamma; in frontal electrodes, theta, beta, and high gamma power was reactivated; and in medial temporal electrodes, beta power was reactivated.

The global category-specific patterns observed at the scalp showed marginally nonsignificant reactivation during immediate free recall (reactivation metric: mean = 0.0029, SEM = 0.0012; $p = 0.057$, permutation test). However, there was significant reactivation in scalp EEG during final free recall (reactivation metric: mean = 0.0130, SEM = 0.0041; $p < 0.005$, permutation test). There was significant reactivation in delta and theta power. We attempted to characterize the time-course of reactivation (relative to recall onset) by examining reactivation in the segmented recall data; however, this less-sensitive analysis revealed no significant reactivation.

Pattern fidelity during retrieval correlates with category clustering

Retrieved-context models describe how, during memory search, the retrieval cue is constantly updated by the information that is retrieved from memory. These models predict that when a participant recalls an item from a particular category, the category-specific information that is retrieved is integrated into the retrieval cue, making it a better match for other memories from the same category. This context retrieval operation leads to the prediction that category-specific oscillatory activity observed during recall should increase in fidelity as multiple items are recalled from the same category.

To test this prediction, we examined the dynamics of category-specific activity during memory search. A classifier was trained to identify the category associated with particular recalled items, using patterns of oscillatory activity recorded prior to the vocalization of that item. The classifier was then tested on the neural patterns preceding a left-out set of recalled items. Again, the larger scalp EEG dataset affords us a closer examination of the nuanced dynamics of these category-related patterns, though at the cost of anatomical localization of the signal. The scalp EEG–monitored participants showed reliable category-specific activity during recall: mean classifier performance was significantly above chance for recall tests that were administered immediately after the list (mean 36.6%, SEM 1.3%, $t(28) = 2.47$, $p < 0.01$), as well as during the final free recall test administered at the end of the session (Fig. 6A; mean 42.7%, SEM 2.1%, $t(28) = 4.44$, $p < 0.0001$). During final free recall (FFR), classifier accuracy was significantly greater than it was during immediate free recall (IFR; $t(28) = 2.55$, $p < 0.05$).

Because LBC_{sem} varies with list length, we used a different measure of semantic clustering, the adjusted ratio of clustering (ARC) score (Roenker et al. 1971), to compare category clustering in IFR and FFR. In the scalp EEG experiment, ARC score for IFR was 0.60 (SEM 0.02); ARC score for FFR was 0.88 (SEM 0.02), and was significantly greater than IFR ($t(28) = 14.93$; $p < 0.0001$). A similar difference was observed in the ECoG experiment (IFR: mean 0.62, SEM 0.25; FFR: mean 0.85, SEM 0.03; $t(11) = 3.65$, $p < 0.002$). In other words, the later recall period gave rise to stronger category-related organization of responses during memory search. Given that category clustering was greater during FFR than during IFR, this suggests that the strength of category-specific activity during recall was related to the degree of category clustering. This proposed effect was observed for the scalp EEG experiment (as described above) but was not statistically reliable in the ECoG/depth experiment (although there was a trend towards greater classifier performance during FFR in the parietal region of interest).

As with our analyses of the study-period data, we examined whether the category identity of neighboring items in the recall sequence could influence classifier performance. In order to account for this, we examined the period from 1500 to 500 ms before vocalization onset, which cannot be influenced by adjacent recalls given the parameters of the wavelets we used to measure oscillatory power, and the criteria used to choose recall epochs. We averaged classifier performance over this interval. Again, FFR classifier performance was significantly greater than IFR classifier performance during this critical interval ($t(28) = 1.98$, $p < 0.03$, one-sided test). This increased classifier performance for the FFR period was robust to the training set used to train the classifier: The same effect was observed when the classifier was trained on category-specific patterns from the study period, and was then applied to the IFR and FFR recall periods. Classifier performance was greater for FFR (reactivation metric: mean 0.0130, SEM 0.0041) than IFR (mean 0.0029, SEM 0.0012); this difference is significant ($t(28) = 2.45$, $p < 0.05$).

In FFR, when significantly more same-category items are recalled successively, the category-specific patterns are of higher fidelity than in IFR. This is in line with the predictions of retrieved context models, which suggest that the retrieval cue will be reliably more category-specific during periods of category clustering. The model predicts that we should see similar variability in the fidelity of category-specific patterns during IFR. We confirmed this in a final series of analyses.

During IFR, the strength of the category-specific patterns rise and fall as a function of whether a participant is producing a cluster of same-category responses or is transitioning from category to category (Fig. 6B). Clustered items were classified with greater accuracy than isolated items ($t(28) = 2.51$, $p < 0.05$). To further examine the influence of adjacent recalls on classifier performance, we used a two-way within-subjects analysis of variance with category of the previous recall (same as the current category or different) and category of the next recall (same or different) as factors. The category of the previous recall is important, as the classifier can better identify the category of a middle or terminal item in a sequence of same-category recalls, as compared to an initial or isolated item ($F(1, 28) = 10.89$, $p < 0.005$). To rule out effects of adjacent recalls on our wavelet-based power measures, we examined the critical time interval of 1500 to 500 ms before vocalization onset, and found a significant difference between middle/terminal items and isolated/initial items ($t(28) = 2.82$, $p < 0.005$, one-sided test). These results are consistent with retrieved-context models, which propose that category-specific activity is integrated over time during recall, and will therefore be stronger when the previous recall was from the same category as the current recall. Classifier performance was also greater for recalls in the initial or middle position of a sequence of same-category responses, as compared to isolated responses from a given category, and terminal responses from a sequence of same-category responses ($F(1, 28) = 4.38$, $p < 0.05$). In other words, a recall associated with higher-fidelity category-specific activity will tend to be followed by a recall from the same category. This is consistent with the proposal that retrieved category-specific patterns are used to guide memory search. The influences of category of the previous recall and the category of the next recall did not interact ($F(1, 28) < 1$).

[Figure 6 about here.]

As with our study-period analyses, we ran a secondary analysis to control for effects of training set imbalances. Similar results were obtained when random sampling without replacement was used to obtain a training set with an equal number of each combination of category and cluster position bin. Classifier performance was greater for recalls preceded by an item of the same category (i.e. middle and terminal items), compared to items preceded by a recall of a different category (i.e. isolated and initial items; $F(1, 28) = 12.54, p < 0.005$). Classifier performance was also significantly greater for recalls that were followed by an item of the same category (i.e. initial and middle items), compared to items that were followed by an item of a different category (i.e. isolated and terminal items; $F(1, 28) = 8.86$, $p < 0.01$). There was no interaction between previous category and next category ($F(1, 28) = 2.87$, $p = 0.1$).

Discussion

Category-specific topographic patterns of oscillatory activity, recorded both at the scalp and intracranially, allow us to examine how people create and search through the structures of memory. Using pattern classification techniques, we characterized the dynamics of category-specific activity during both study and memory search, allowing us to relate neural measures of oscillatory power to the strength of attribute-based cognitive representations characterized by retrieved-context models of human memory (Howard and Kahana 2002; Howard 2004; Sederberg et al. 2008; Polyn et al. 2009). The subsequent memory effect, in which subsequently remembered items elicit a stronger neural response in certain critical brain regions than subsequently forgotten items (Paller and Wagner 2002), is used by researchers to implicate particular brain regions or neural signals in memory-related processes. A recent neuroimaging study showed that the strength of category-specific patterns elicited during study predicted whether an item would be remembered during a later paired-associates memory test, thus extending the subsequent memory effect to category-specific patterns of neural activity (Kuhl et al. 2011b). The current results extend the Kuhl et al. finding from the domain of cued recall to the domain of free recall. The self-directed nature of this task allows us to examine the structures formed in memory, by examining the order with which the participant discovers the studied items during memory search. A recent study by Long et al. (Long et al. 2010) found a region of ventrolateral prefrontal cortex whose activity levels were sensitive to whether an item would be subsequently clustered according to its semantic category. Here, we show that a global category-specific response recorded with scalp EEG predicts which items are more likely to be subsequently clustered; furthermore, we show that these category-specific patterns show integrative dynamics, as predicted by retrieved-context models of memory.

A separate session in the scalp EEG experiment allowed us to assess the influence of memory-related cognitive processes on category-specific neural activity. We propose that the category-specific patterns recorded during a familiarization session, where there was no demand for the participant to memorize the stimuli, reveal item-specific category representations, as there is no demand for the participant to create an integrative retrieval cue. In contrast, the category-specific patterns recorded during the study periods of the free-recall sessions, where participants knew there was a memory test following each list, reflect both item-specific and retrieval cue–related category representations. Regardless of whether the classifier was trained on the familiarization session or the study period, we found that classifier performance for a study item predicted whether that item would be subsequently recalled. However, we found that these different training periods revealed two distinct subsequent clustering effects. First, we found that a classifier trained to discriminate the neural representations of the three categories using the familiarization session data identified an item-level subsequent clustering effect when applied to the study periods of the free-recall experiment. Particular studied items that would be later remembered in a sequence with other same-category items elicited a stronger category-specific representation. This is consistent with a model in which the category-specific activity reflects a strong (or canonical) item representation, which would be well-targeted by a category-specific retrieval cue during memory search.

In contrast, a classifier trained to discriminate the neural representations of the three categories using the study period data revealed an integrative subsequent clustering effect, at both the level of individual participants, as well as at the level of individual lists. Participants who tended to organize their memory search according to category showed a large increase in classifier performance with successively presented same-category items during study. This is consistent with a model in which the category-specific activity reflects a retrieval cue that integrates category-related information from each studied item. By this model, participants that tend to integrate more category-specific information (as opposed to, for example, idiosyncratic item features not broadly associated with the category) will tend to exhibit more category clustering. Future work will be needed to determine whether this form of integrative neural activity is dependent on an intent to memorize the studied material, as items from the same category were never presented adjacent to one another during the familiarization session. These results raise the possibility that these category-specific patterns reflect both item-specific, and retrieval cue–specific category information. Thus, an important direction for future work will be to tease apart the contributions of these distinct cognitive processes. Retrieved-context models provide a direction forward, as they make distinct predictions regarding the functional contributions of each model component to memory search.

A second set of analyses revealed category-specific neural activity during memory search in both the ECoG/depth and scalp EEG data-sets. Only a few studies have examined neural activity during free recall, using any neurorecording modality (Polyn et al. 2005; Gelbard-Sagiv et al. 2008; Sederberg et al. 2007; Long et al. 2010; Manning et al. 2011; Polyn et al. 2011). Polyn et al. (Polyn et al. 2005) showed, using fMRI, that brain-wide patterns of category-specific activity were reactivated when participants searched memory for studied material, and that the rise and fall of this category-specific hemodynamic activity predicted the category identity of recalled items. In the ECoG/depth experiment, we found evidence for reactivation of the same category-specific oscillatory patterns characterized during the study period. In the scalp EEG experiment, these reactivation effects were reliable, but very weak. However, we found strong category-specific patterns in the recall periods of the scalp EEG experiment that did not match the patterns observed during study. How these recall-period category-specific patterns relate to those observed in the ECoG/depth experiment is a question for future work. However, the recall-period category-specific patterns characterized with scalp EEG showed dynamics consistent with retrieved-context models of human memory. First, we found that these category-specific patterns were increased in strength during a final recall period characterized by strong category clustering, as compared to an immediate recall period characterized by more modest category clustering. Second, we found that even during the immediate recall period, these category-specific patterns increased in strength during periods of category clustering, as compared to periods where the participant was shifting between categories. These results are consistent with the idea from retrieved-context models that when an item is remembered, the retrieval cue integrates the reactivated information. Thus, each time an item from a particular category is remembered, the model predicts that the retrieval cue will become more category-specific, in line with both of these observations. Furthermore, we found that the strength of category-specific activity during recall of an item predicted the category of the next recalled item: Recalls associated with strong category-specific patterns were more likely to be followed by a recall from the same category, suggesting that this category-specific activity is used to guide memory search.

In this study we used retrieved-context models of human memory as a framework to interpret the functional importance of various category-specific neural signals observed during study and memory search. However, this work only begins to tap the potential for computational modeling to inform neural investigations. In a number of cognitive domains, computational models are being used to bridge between the neural signals recorded while a participant performs a task, and the behavioral measures characterizing that performance (Purcell et al. 2010; Ratcliff et al. 2009; Polyn et al. 2011; Davis et al. 2012). Polyn et al. (Polyn et al. 2011) used the Context Maintenance and Retrieval (CMR) model of memory search (a retrieved-context model; Polyn et al. 2009) to interpret task-specific patterns of hemodynamic activity recorded as participants performed a free-recall task. They found that the discriminability of task-specific patterns of neural activity was related to the magnitude of the recency effect, and showed that a particular model parameter controlling integration rate could be used to explain individual differences in both the neural data (classifier performance in identifying task identity of a studied item) and the behavioral data (the tendency of the participant to initiate recall with the final studied item). Applying the model more closely to the data from the current study will allow us to better understand the similarities and differences between task organization and category organization. For example, in the current study, we did not observe a reliable relationship between category discriminability and the recency effect. This may be due to differences in how task and category information are processed by the neurocognitive system; task information represents a rapidly formed association between a study item and the circumstances in which it is encountered, whereas category information represents longstanding semantic associations between all of the members of a category. Computational models such as CMR provide a common framework for understanding both how these different forms of information are processed by the brain, and how they relate to the neural measures recorded during study and free recall.

Funding

This work was supported by the National Institute of Mental Health (MH055687 and MH061975); the National Institutes of Health (NS067316); and the Dana Foundation.

Acknowledgements

We thank Kenneth Norman, Joshua Jacobs, Gordon Logan, Geoff Woodman, John Burke, Ashwin Ramayya, James Kragel, and Zachary Cohen for useful discussions. We also thank Jonathan Miller, Richard Lawrence, John Burke, and Matthew Mollison for help collecting data, and Evan Stein for helping to organize testing logs.

References

Buzsáki G. 2006. Rhythms of the Brain. New York (NY): Oxford University Press.

Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT. 2006. High gamma power is phase-locked to theta oscillations in human neocortex. Science. 313:1626–1628.

Danker JF, Anderson JR. 2010. The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. Psychological Bulletin. 136:87–102.

Davis T, Love BC, Preston AR. 2012. Learning the exception to the rule: Model-based fMRI reveals specialized representations for surprising category members. Cerebral Cortex. 22:260–273.

Duda RO, Hart PE, Stork DG. 2001. Pattern classification. 2nd ed. New York (NY): Wiley.

Düzel E, Penny W, Burgess N. 2010. Brain oscillations and memory. Current opinion in neurobiology. 20:143–149.

Freeman WJ. 1978. Spatial properties of an EEG event in the olfactory bulb and cortex. Electroencephalogr Clin Neurophysiol. 44:586 – 605.

Fries P. 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends in Cognitive Sciences. 9:474–480.

Fuster J, Bauer R, Jervey J. 1982. Cellular discharge in the dorsolateral prefrontal cortex of the monkey in cognitive tasks. Experimental Neurology. 77:679–694.

Gelbard-Sagiv H, Mukamel R, Harel M, Malach R, Fried I. 2008. Internally generated reactivation of single neurons in human hippocampus during free recall. Science. 3:96–101.

Geller AS, Schleifer IK, Sederberg PB, Jacobs J, Kahana MJ. 2007. PyEPL: A cross-platform experiment-programming library. Behavior Research Methods. 39:950–958.

Gratton G, Coles MGH, Donchin E. 1983. A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology. 55:468–484.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 293:2425–2429.

Herrmann CS, Grigutsch M, Busch NA. 2005. EEG oscillations and wavelet analysis. In: Handy T, editor. Event-Related Potentials: A Methods Handbook. Cambridge (MA): The MIT Press. p 229–260.

Howard MW. 2004. Scaling behavior in the temporal context model. Journal of Mathematical Psychology. 48:230–238.

Howard MW, Kahana MJ. 2002. A distributed representation of temporal context. Journal of Mathematical Psychology. 46:269–299.

Jacobs J, Kahana MJ. 2009. Neural representations of individual stimuli in humans revealed by gamma-band ECoG activity. Journal of Neuroscience. 29:10203–10214.

Jacobs J, Kahana MJ. 2010. Direct brain recordings fuel advances in cognitive electrophysiology. Trends in Cognitive Sciences. 14:162–171.

Jacobs J, Kahana MJ, Ekstrom AD, Fried I. 2007. Brain oscillations control timing of single-neuron activity in humans. Journal of Neuroscience. 27:3839–3844.

Klimesch W. 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Research Reviews. 29:169–195.

Kojima S, Goldman-Rakic PS. 1982. Delay-related activity of prefrontal neurons in rhesus monkeys performing delayed response. Brain Research. 248:43–49.

Kuhl BA, Rissman J, Chun MM, Wagner AD. 2011a. Fidelity of neural reactivation reveals competition between memories. Proceedings of the National Academy of Sciences of the United States of America. 108:5903–5908.

Kuhl BA, Rissman J, Wagner AD. 2011b. Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. Neuropsychologia. doi:10.1016/j.neuropsychologia.2011.09.002.

Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT. 2000. Automated Talairach atlas labels for functional brain mapping. Hum Brain Mapp. 10:120–131.

Liebe S, Hoerzer G, Logothetis N, Rainer G. 2012. Theta coupling between v4 and prefrontal cortex predicts visual short-term memory performance. Nature Neuroscience. 15:456–464.

Loftus GR, Masson MEJ. 1994. Using confidence intervals in within-subject designs. Psychonomic Bulletin & Review. 1:476–490.

Long NM, Oztekin I, Badre D. 2010. Separable prefrontal cortex contributions to free recall. Journal of Neuroscience. 30:10967–10976.

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 19:1233–1239.

Manning JR, Polyn SM, Baltuch G, Litt B, Kahana MJ. 2011. Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. Proceedings of the National Academy of Sciences, USA. 108:12893–12897.

Manns JR, Howard MW, Eichenbaum H. 2007. Gradual changes in hippocampal activity support remembering the order of events. Neuron. 56:530–540.

Martin A. 2007. The representation of object concepts in the brain. Annual Review of Psychology. 58:25–45.

Miller EK, Erickson CA, Desimone R. 1996. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. Journal of Neuroscience. 16:5154.

Norman KA, Polyn SM, Detre GJ, Haxby JV. 2006. Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. Trends in Cognitive Sciences. 10:424–430.

Nunez PL, Srinivasan R. 2006. Electric fields of the brain: The neurophysics of EEG. 2nd ed. New York (NY): Oxford University Press.

Nyhus E, Curran T. 2010. Functional role of gamma and theta oscillations in episodic memory. Neuroscience & Biobehavioral Reviews. 34:1023–1035.

Paller KA, Wagner AD. 2002. Observing the transformation of experience into memory. Trends in Cognitive Sciences. 6:93–102.

Polyn SM, Kahana MJ. 2008. Memory search and the neural representation of context. Trends in Cognitive Sciences. 12:24–30.

Polyn SM, Kragel JE, Morton NW, McCluey JD, Cohen ZD. 2011. The neural dynamics of task context in free recall. Neuropsychologia. doi:10.1016/j.neuropsychologia.2011.08.025.

Polyn SM, Natu VS, Cohen JD, Norman KA. 2005. Category-specific cortical activity precedes retrieval during memory search. Science. 310:1963–1966.

Polyn SM, Norman KA, Kahana MJ. 2009. A context maintenance and retrieval model of organizational processes in free recall. Psychological Review. 116:129–156.

Puff CR. 1979. Memory organization research and theory: The state of the art. In: Puff CR, editor. Memory Organization and Structure. New York (NY): Academic Press. p 3–17.

Purcell BA, Heitz RP, Cohen JY, Schall JD, Logan GD, Palmeri TJ. 2010. Neurally constrained modeling of perceptual decision making. Psychological Review. 117:1113–1143.

Ratcliff R, Philiastides MG, Sajda P. 2009. Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. Proceedings of the National Academy of Sciences of the United States of America. 106:6539–6544.

Roenker DL, Thompson CP, Brown SC. 1971. Comparison of measures for the estimation of clustering in free recall. Psychological Bulletin. 76:45–48.

Sederberg PB, Howard MW, Kahana MJ. 2008. A context-based theory of recency and contiguity in free recall. Psychological Review. 115:893–912.

Sederberg PB, Kahana MJ, Howard MW, Donner EJ, Madsen JR. 2003. Theta and gamma oscillations during encoding predict subsequent recall. Journal of Neuroscience. 23:10809–10814.

Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, Litt B, Brandt A, Kahana MJ. 2007. Gamma oscillations distinguish true from false memories. Psychological Science. 18:927–932.

Solway A, Geller AS, Sederberg PB, Kahana MJ. 2010. Pyparse: A semiautomated system for scoring spoken recall data. Behavior Research Methods. 42:141–147.

Stricker JL, Brown GG, Wixted JT, Baldo JV, Delis DC. 2002. New semantic and serial clustering indices for the california verbal learning test–second edition: Background, rationale, and formulae. Journal of the International Neuropsychological Society. 8:425–435.

Summerfield C, Mangels JA. 2005. Coherent theta-band EEG activity predicts item-context binding during encoding. NeuroImage. 24:692–703.

Talairach J, Tournoux P. 1988. Co-planar stereotaxic atlas of the human brain. Stuttgart: Verlag.

Tallon-Baudry C, Bertrand O, Delpuech C, Permier J. 1997. Oscillatory gamma-band (30-70 hz) activity induced by a visual search task in humans. Journal of Neuroscience. 17:722–734.

Tallon-Baudry C, Bertrand O, Delpuech C, Perner J. 1996. Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. Journal of Neuroscience. 16:4240–4249.

Tulving E. 1993. What is episodic memory? Current Directions in Psychological Science. 2:67–70.

Wheeler MA, Stuss DT, Tulving E. 1995. Frontal lobe damage produces episodic memory impairment. Journal of the International Neuropsychological Society. 1:525–536.

Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY. 2008. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. Neuron. 58:429–441.

ID	HOSP	AGE	SEX	HAND	ELC	SES
1	UP	18	M	A	100	6
2	UP	39	M	L	77	1
3	UP	40	M	R	38	2
4	TJ	25	M	R	35	3
5	TJ	40	F	R	82	10
6	TJ	39	M	L	52	4
7	TJ	34	F	R	92	2
7	TJ	34	F	R	86	8
8	TJ	39	F	R	85	1
9	TJ	44	M	R	124	4
10	TJ	29	M	R	36	1
11	TJ	43	M	R	57	5

Table 1: This table provides the hospital (HOSP) at which each patient's data were collected, as well as each patient's age in years (AGE), sex (SEX), handedness (HAND), number of implanted electrodes (ELC), and number of testing sessions (SES). Patient 7 underwent invasive monitoring with 2 partially overlapping sets of electrodes (see text for details). A, ambidextrous; F, female; L, left; M, male; R, right; TJ, Thomas Jefferson Hospital (Philadelphia, PA); UP, Hospital of the University of Pennsylvania (Philadelphia, PA).

Figure 1: Illustration of category clustering in a sample mixed-category list. A *cluster* is defined as a sequence of two or more same-category recalls. Recalled items are labeled according to their position in a category cluster (Init: initial, Mid: middle, Term: terminal); items not recalled in a category cluster are labeled Iso (isolated). Study items are labeled according to their subsequent recall organization. SC: subsequently clustered (i.e. recalled as an initial, middle, or terminal item in a category cluster); SI: subsequently isolated.

Figure 2: Right and left sagittal views of electrode coverage in intracranially implanted patients. Temporal, parietal, and occipital regions are denoted by blue, yellow, and green dots, respectively. Prefrontal electrodes are shown in orange; other frontal electrodes are shown in red. Not shown: 76 medial temporal electrodes.

Figure 3: Category-specific oscillations during study. (**A**) Oscillations in ECoG and depth electrodes at widespread frequencies show category-specificity during study. The colorbar indicates classifier performance (using a cross-validation procedure) at each oscillatory frequency and time bin relative to stimulus onset (at time = 0). Dark blue corresponds to chance performance (1/3). (**B**) Scalp EEG oscillatory activity shows a similar time-frequency distribution of category specificity, although category-specific high gamma oscillatory activity is attenuated. (**C**) Classifier performance as a function of frequency band, averaged over the stimulus presentation interval. D: delta, 2–4 *Hz*, T: theta, 4–8 *Hz*, A: alpha, 10–14 *Hz*, B: beta, 16–25 *Hz*, LG: low gamma, 25–55 *Hz*, HG: high gamma, 65–100 *Hz*. The dotted line indicates chance performance (1/3). Error bars represent standard error of the mean. (**D**) Performance of a classifier provided with information from all frequencies, plotted against time after stimulus onset. Category-specific patterns peak at about 500 ms after stimulus onset, and persist throughout stimulus presentation, both for intracranially implanted and scalp EEG-monitored participants. Scalp classifier performance during the 500 ms before stimulus onset is averaged over items that followed an item of a different category; all other time bins show performance averaged over all items. Error bars represent 95% confidence intervals based on within-subject error (Loftus and Masson 1994). The dotted line indicates chance performance (1/3).

Figure 4: Category-specific oscillations during study and reactivation during recall. (**A**) Classifier performance during study is significantly above chance for all regions of interest (ROIs). (**B**) The fidelity of category-specific patterns in ECoG recorded from temporal electrodes predicts subsequent memory. The difference in classifier category strength estimates between recalled and forgotten items is shown for each ROI. For temporal electrodes, classifier estimates were greater for subsequently recalled items, compared to subsequently forgotten items. Error bars indicate 95% confidence intervals corresponding to a one-tailed paired t-test; * indicates $p < 0.05$, Bonferroni corrected. (**C**) Reactivation of category-specific information during immediate free recall is observed in frontal, prefrontal, temporal, medial temporal, hippocampal, and occipital electrodes. (**D**) Reactivation during final free recall is observed in frontal, prefrontal, temporal, medial temporal, hippocampal, and parietal electrodes. FR: frontal lobe, PFC: prefrontal cortex, Temp: temporal lobe, MTL: medial temporal lobe, Hipp: hippocampus, Occ: occipital lobe, Par: parietal lobe. The dotted lines indicate significance thresholds for permutation tests comparing performance to chance (familywise Type I error rate < 0.05).

Figure 5: Dynamics of category-specific scalp EEG activity during study. (**A**) The fidelity of category-specific oscillatory patterns predicts subsequent memory and recall organization. When a classifier was trained on item presentations during a familiarity judgment task and applied to the study period of free recall lists, subsequently clustered (SC) items were classified more accurately than both subsequently isolated (SI) items and forgotten items. Error bars represent 95% confidence intervals based on within-subject error (Loftus and Masson 1994). (**B**) When the classifier is trained on a left-out portion of the study period, the persistence of category-related neural patterns is seen in the increased fidelity of category patterns when multiple same-category items are presented in succession. The classifier's estimate of the strength of the current category is plotted as a function of position within a train of same-category item presentations. On average, classifier estimates rose with successive same-category stimuli. Error bars represent 95% confidence intervals based on within-subject error (Loftus and Masson 1994). (**B**) The slope of the regression of classifier estimate on train position was correlated with individual differences in category clustering as measured by LBC_{sem} ($r = 0.500$, $p < 0.01$). Two outliers have been removed from the plot; with them included, the correlation is still significant ($r = 0.421$, $p < 0.05$).

Figure 6: Category-specific patterns of neural activity during recall. (**A**) In the scalp EEG experiment, fidelity of category-specific patterns (measured using cross-validation of recall events) was greater during the final free recall period (FFR), which exhibited greater category clustering than the immediate free recall period (IFR). Error bars indicate 95% confidence intervals corresponding to a one-tailed *t*-test vs. chance (1/3; indicated by the dotted line). (**B**) During IFR, classifier performance was significantly higher for positions in the recall sequence where the previous item was from the same category as the current item (middle, terminal), as compared to positions in the recall sequence where the previous item was from a different category (isolated, initial). Classifier performance was also greater when the next item was the same category (initial, middle), compared to when the next item was a different category (isolated, terminal). Error bars indicate 95% confidence intervals corresponding to a one-tailed *t*-test vs. chance (1/3; indicated by the dotted line).

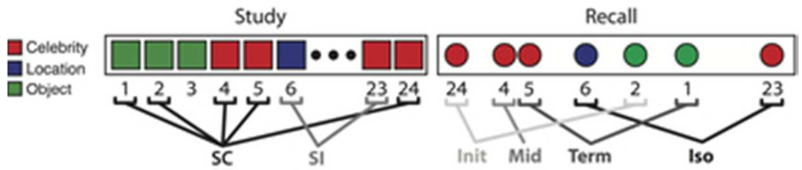

Figure 1: Illustration of category clustering in a sample mixed-category list. A cluster is defined as a sequence of two or more same-category recalls. Recalled items are labeled according to their position in a category cluster (Init: initial, Mid: middle, Term: terminal); items not recalled in a category cluster are labeled Iso (isolated). Study items are labeled according to their subsequent recall organization. SC: subsequently clustered (i.e. recalled as an initial, middle, or terminal item in a category cluster); SI: subsequently isolated.
33x6mm (300 x 300 DPI)

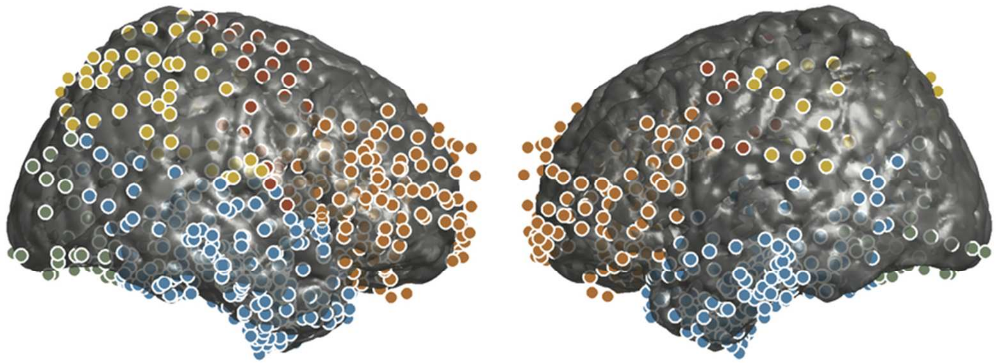

Figure 2: Right and left sagittal views of electrode coverage in intracranially implanted patients. Temporal, parietal, and occipital regions are denoted by blue, yellow, and green dots, respectively. Prefrontal electrodes are shown in orange; other frontal electrodes are shown in red. Not shown: 76 medial temporal electrodes.
74x27mm (300 x 300 DPI)

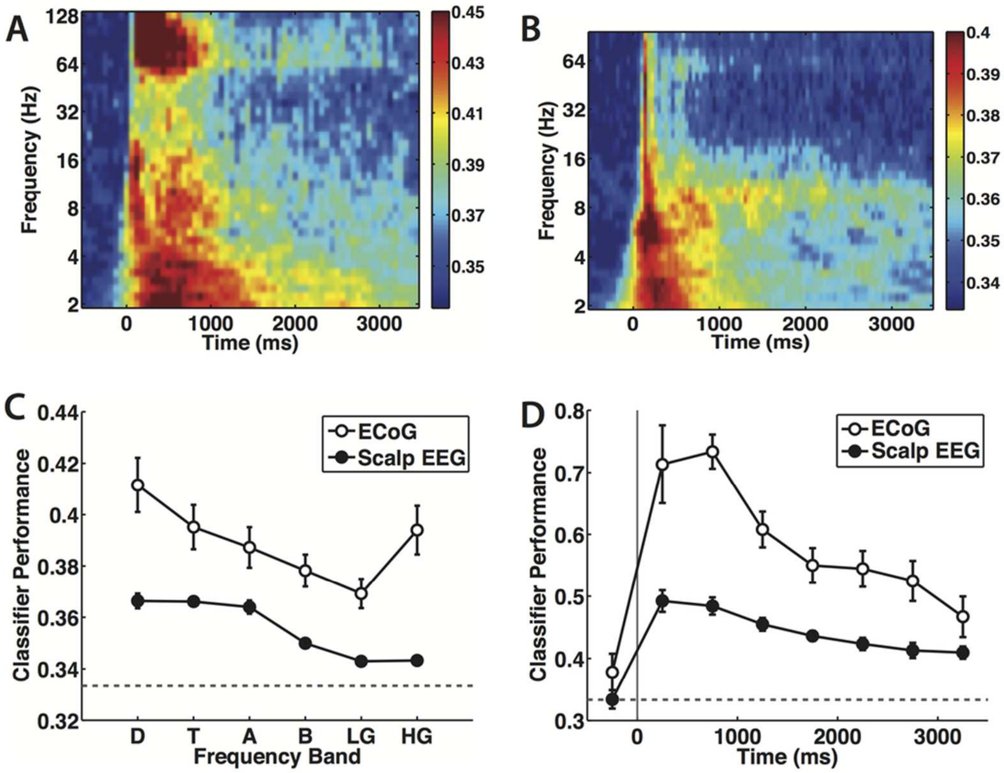

Figure 3: Category-specific oscillations during study. (A) Oscillations in ECoG and depth electrodes at widespread frequencies show category-specificity during study. The colorbar indicates classifier performance (using a cross-validation procedure) at each oscillatory frequency and time bin relative to stimulus onset (at time = 0). Dark blue corresponds to chance performance (1/3). (B) Scalp EEG oscillatory activity shows a similar time-frequency distribution of category specificity, although category-specific high gamma oscillatory activity is attenuated. (C) Classifier performance as a function of frequency band, averaged over the stimulus presentation interval. D: delta, 2–4 Hz, T: theta, 4–8 Hz, A: alpha, 10–14 Hz, B: beta, 16–25 Hz, LG: low gamma, 25–55 Hz, HG: high gamma, 65–100 Hz. The dotted line indicates chance performance (1/3). Error bars represent standard error of the mean. (D) Performance of a classifier provided with information from all frequencies, plotted against time after stimulus onset. Category-specific patterns peak at about 500 ms after stimulus onset, and persist throughout stimulus presentation, both for intracranially implanted and scalp EEG-monitored participants. Scalp classifier performance during the 500 ms before stimulus onset is averaged over items that followed an item of a different category; all other time bins show performance averaged over all items. Error bars represent 95% confidence intervals based on within-subject error (Loftus and Masson 1994). The dotted line indicates chance performance (1/3).
87x67mm (300 x 300 DPI)

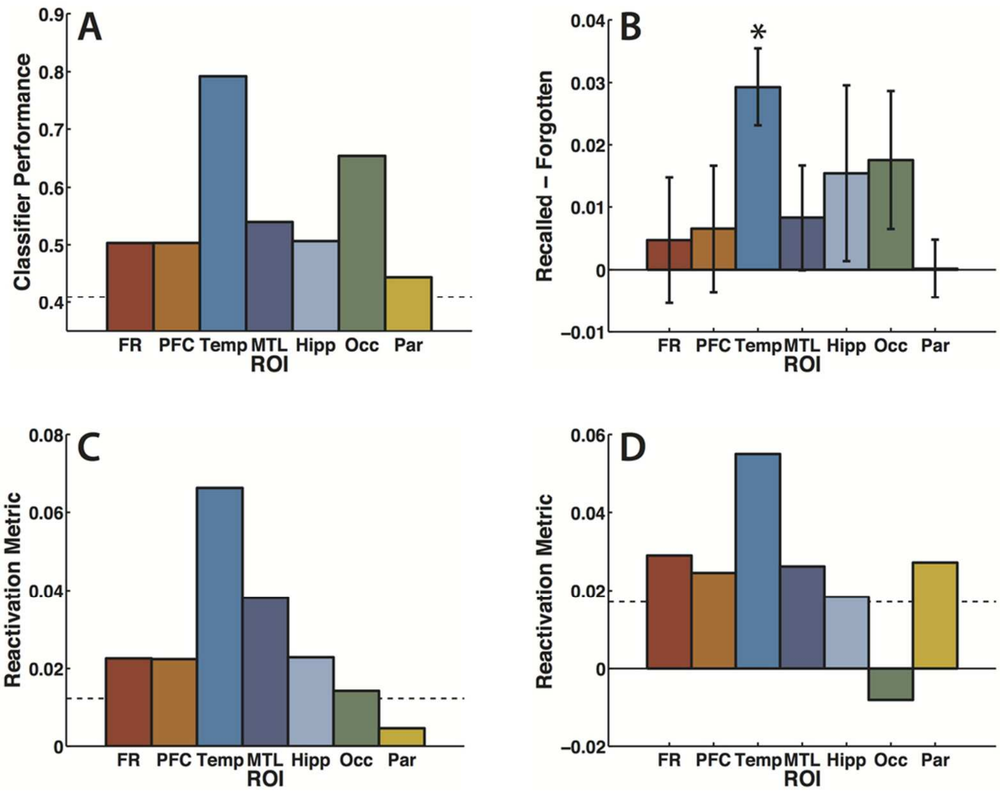

Figure 4: Category-specific oscillations during study and reactivation during recall. (A) Classifier performance during study is significantly above chance for all regions of interest (ROIs). (B) The fidelity of category-specific patterns in ECoG recorded from temporal electrodes predicts subsequent memory. The difference in classifier category strength estimates between recalled and forgotten items is shown for each ROI. For temporal electrodes, classifier estimates were greater for subsequently recalled items, compared to subsequently forgotten items. Error bars indicate 95% confidence intervals corresponding to a one-tailed paired t-test; * indicates $p < 0.05$, Bonferroni corrected. (C) Reactivation of category-specific information during immediate free recall is observed in frontal, prefrontal, temporal, medial temporal, hippocampal, and occipital electrodes. (D) Reactivation during final free recall is observed in frontal, prefrontal, temporal, medial temporal, hippocampal, and parietal electrodes. FR: frontal lobe, PFC: prefrontal cortex, Temp: temporal lobe, MTL: medial temporal lobe, Hipp: hippocampus, Occ: occipital lobe, Par: parietal lobe. The dotted lines indicate significance thresholds for permutation tests comparing performance to chance (familywise Type I error rate < 0.05).
90x72mm (300 x 300 DPI)

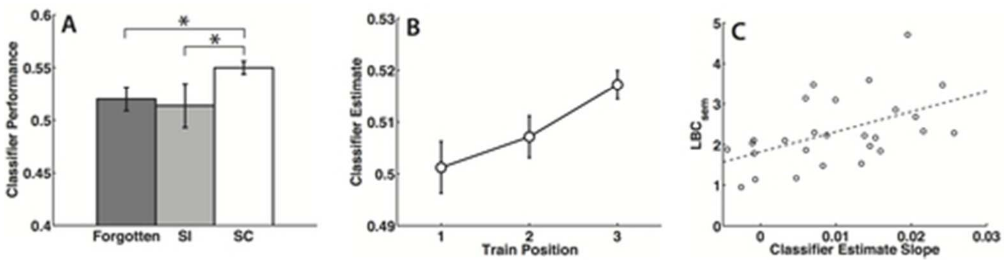

Figure 5: Dynamics of category-specific scalp EEG activity during study. (A) The fidelity of category-specific oscillatory patterns predicts subsequent memory and recall organization. When a classifier was trained on item presentations during a familiarity judgment task and applied to the study period of free recall lists, subsequently clustered (SC) items were classified more accurately than both subsequently isolated (SI) items and forgotten items. Error bars represent 95% confidence intervals based on within-subject error (Loftus and Masson 1994). (B) When the classifier is trained on a left-out portion of the study period, the persistence of category-related neural patterns is seen in the increased fidelity of category patterns when multiple same-category items are presented in succession. The classifier's estimate of the strength of the current category is plotted as a function of position within a train of same-category item presentations. On average, classifier estimates rose with successive same-category stimuli. Error bars represent 95% confidence intervals based on within-subject error (Loftus and Masson 1994). (B) The slope of the regression of classifier estimate on train position was correlated with individual differences in category clustering as measured by LBCsem ($r = 0.500$, $p < 0.01$). Two outliers have been removed from the plot; with them included, the correlation is still significant ($r = 0.421$, $p < 0.05$).

43x11mm (300 x 300 DPI)

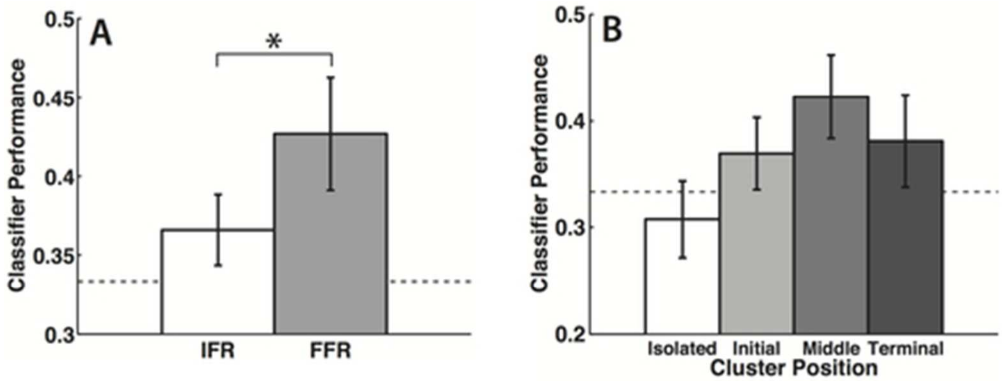

Figure 6: Category-specific patterns of neural activity during recall. (A) In the scalp EEG experiment, fidelity of category-specific patterns (measured using cross-validation of recall events) was greater during the final free recall period (FFR), which exhibited greater category clustering than the immediate free recall period (IFR). Error bars indicate 95% confidence intervals corresponding to a one-tailed t-test vs. chance (1/3; indicated by the dotted line). (B) During IFR, classifier performance was significantly higher for positions in the recall sequence where the previous item was from the same category as the current item (middle, terminal), as compared to positions in the recall sequence where the previous item was from a different category (isolated, initial). Classifier performance was also greater when the next item was the same category (initial, middle), compared to when the next item was a different category (isolated, terminal). Error bars indicate 95% confidence intervals corresponding to a one-tailed t-test vs. chance (1/3; indicated by the dotted line).
42x16mm (300 x 300 DPI)