2012

Ensembles of Human MTL Neurons "Jump Back in Time" in Response to a Repeated Stimulus

M. W. Howard

Indre Viskontas
University of San Francisco, ivviskontas@usfca.edu

K. H. Shankar

I. Fried

Recommended Citation

Ensembles of Human MTL Neurons ''Jump Back in Time'' in Response to a Repeated Stimulus

Marc W. Howard,[1*] Indre V. Viskontas,[2] Karthik H. Shankar,[1] and Itzhak Fried[3,4]

ABSTRACT: Episodic memory, which depends critically on the integrity of the medial temporal lobe (MTL), has been described as "mental time travel" in which the rememberer "jumps back in time." The neural mechanism underlying this ability remains elusive. Mathematical and computational models of performance in episodic memory tasks provide a specific hypothesis regarding the computation that supports such a jump back in time. The models suggest that a representation of temporal context, a representation that changes gradually over macroscopic periods of time, is the cue for episodic recall. According to these models, a jump back in time corresponds to a stimulus recovering a prior state of temporal context. In vivo single-neuron recordings were taken from the human MTL while epilepsy patients distinguished novel from repeated images in a continuous recognition memory task. The firing pattern of the ensemble of MTL neurons showed robust temporal autocorrelation over macroscopic periods of time during performance of the memory task. The gradually-changing part of the ensemble state was causally affected by the visual stimulus being presented. Critically, repetition of a stimulus caused the ensemble to elicit a pattern of activity that resembled the pattern of activity present before the initial presentation of the stimulus. These findings confirm a direct prediction of this class of temporal context models and may be a signature of the mechanism that underlies the experience of episodic memory as mental time travel. © 2012 Wiley Periodicals, Inc.

KEY WORDS: episodic memory; temporal context; contiguity effect

INTRODUCTION

Theories of episodic memory postulate that the brain maintains a representation of spatiotemporal context that changes gradually from moment to moment (Tulving and Madigan, 1970; Kinsbourne and Wood, 1975; O'Keefe and Nadel, 1978; Eichenbaum et al., 2007). When an event is remembered, the state of spatiotemporal context is recovered, leading to the experience of "mental time travel." Episodic memory exhibits two extremely robust temporal effects in the laboratory: the recency effect and the contiguity effect. The recency effect describes the finding that more recently presented items are better remembered (Murdock, 1962; Glenberg et al., 1980), while the contiguity effect refers to the finding that when a stimulus from a list is remembered, memory for other stimuli nearby in the list is enhanced (Kahana, 1996; Kahana et al., 2008). Recency and contiguity are very general behavioral phenomena, having been observed in a variety of recall (e.g., Raskin and Cook, 1937; Howard et al., 2008) and recognition tasks (e.g., Ratcliff and Murdock, 1976; Schwartz et al., 2005).

There is a disconnect in the literature between cognitive theories of episodic memory and neural circuit models of the MTL. Cognitive theories focus on the experience of mental time travel, which reinstates the learning context, thereby causing the contiguity effect. Neural circuit models of the MTL have focused on forming direct inter-item associations between neurons representing different stimuli active in short-term memory (Mehta et al., 2000; Jensen and Lisman, 2005; Rolls and Kesner, 2006; Koene and Hasselmo, 2007). Direct inter-item associations can be formed between stimuli separated in time if the cells activated by the stimuli remain activated in a short-term memory after the stimulus is not longer available. Under these circumstances, local synaptic plasticity could support a contiguity effect; when one of the stimuli is repeated, enhanced synaptic connections cause the cell corresponding to the other stimulus to be reactivated. If the neurons that represent stimuli that are simultaneously active in short-term memory become associated to one another, then this would be sufficient to account for the existence of the contiguity effect (e.g., Kahana, 1996; Sirotin et al., 2005). Persistent activity in short-term memory has long been proposed as an explanation for the immediate recency effect (e.g., Atkinson and Shiffrin, 1968; Raaijmakers and Shiffrin, 1980; Davelaar et al., 2005; Grossberg and Pearson, 2008).

Another stream of work in cognitive psychology also accounts for recency and contiguity effects. Retrieved temporal context models, which have been widely applied in mathematical cognitive modeling (Dennis and Humphreys, 2001; Howard and Kahana, 2002; Howard et al., 2005; Sederberg et al., 2008, 2011; Polyn et al., 2009), postulate a memory representation that changes gradually over macroscopic periods of time. Like short-term memory this gradually-changing representation of temporal context is

caused by the previously presented stimuli and affected most strongly by the most recently presented stimuli. However, unlike short-term memory models, in retrieved temporal context models, the contiguity effect is a result of a "jump back in time." A repeated stimulus need not cause the same input as it did when it was first presented; a repeated stimulus can recover the state of context that was available before it was initially presented. The recovered state is, in turn, a mixture of the inputs caused by earlier stimuli weighted by their temporal proximity.

If it is physically implemented, the retrieved temporal context approach requires that the state of the memory system should change gradually over time and that these changes should be caused by the stimuli presented. These predictions are shared with short-term memory models. The unique predictions of the retrieved temporal context approach are manifest when one considers the comparison between the neural response caused by a repeated stimulus and the neighbors of its original presentation. If a gradually-changing memory state were a signature of a short-term memory that contains stimuli, then the input caused by the repeated stimulus would resemble the states of the short-term memory following its original presentation, when that stimulus was still active in short-term memory. The overlap between the repeated state should decrease going forward from the original presentation as that stimulus decays in short-term memory. However, if repeating a stimulus causes a jump back in time then the state caused by the repeated stimulus will be similar to neighboring states in both the forward and backward directions. That is, the signature of the jump back in time is that a repeated stimulus recovers activity that was present before the original presentation of the stimulus.

There is ample evidence for temporal autocorrelation in the MTL that would be a signature of a gradually-changing representation over various time-scales (Manns et al., 2007; Pastalkova et al., 2008; Paz et al., 2010; MacDonald et al., 2011; Naya and Suzuki, 2011). There is also ample evidence for stimulus-selectivity of MTL neurons in both human and animal studies (Heit et al., 1988; Wood et al., 1999; Hampson et al., 2004; Quiroga et al., 2005; Manns et al., 2007; Viskontas et al., 2009; Naya and Suzuki, 2011), and evidence that stimulus-selective firing is related to memory retrieval (Gelbard-Sagiv et al., 2008). In a study of the free recall task, Manning et al. (2011) found that when subjects free recalled a word from the list, the pattern of activity observed in intracranial electrodes resembled the pattern during study of preceding items. However, there are two important limitations of that study. First, it is not clear whether the context reinstatement effect generalizes to the level of units. Second, because the order of free recall is unconstrained the study leaves open the possibility that the neural state during retrieval of an item resembles the activity associated with preceding items because those preceding items were recalled shortly before the target word. There is as yet no evidence, however, that repetition of a stimulus causes recovery of prior neural ensemble states in the MTL that would be the signature of a jump back in time. This article reports results from single-unit recordings from the human MTL that address this gap using a continuous recognition experiment. Because

Number of Units for Each Recording Session

Patient number	Session number	No. units
371	1	44
	2	45
	3	22
373	1	45
	2	29
378	1	38
380	1	24
	2	23

the sequence of repetitions is under experimental control rather than a consequence of the subject's own retrieval strategy it provides a better opportunity to assess the causal relationship between repetition of a stimulus and the jump back in time.

METHODS

Patients

Four patients with pharmacologically resistant epilepsy for whom extensive noninvasive evaluation failed to yield a single epileptogenic zone participated in our study. To obtain localizing information for potential curative resection, patients were stereotactically implanted with 6 to 14 depth electrodes from a lateral orthogonal approach aiming at targets selected using clinical criteria. Following implantation, patients remained on the ward for 1 to 2 weeks, and were monitored for spontaneous seizures. The four patients participated in a total of eight recording sessions (see Table 1) during which they performed a continuous recognition task. All patients provided informed consent and every recording session conformed to the guidelines of the Medical Institutional Review Board at UCLA. Single-unit analyses of the repetition effect observed for a subset of these data have already been reported (Viskontas et al., 2006).

Behavioral Task

Patients were shown pictures of unknown faces and places, each stimulus repeated once, and asked to indicate whether or not the picture had been previously presented: via button press, patients were instructed to respond "no" to the first presentation of a given image and "yes" to the second. Pictures were shown in four alternating blocks of 20 nonfamous faces taken from the Stirling database (available at http://pics.stir.ac.uk) and 20 unfamiliar indoor and outdoor places. All pictures were black and white photographs and were presented for 2 s. The distances between repeated stimuli varied quasi-randomly with the constraint that the repetition of a stimulus took place within the same block of 20 stimulus presentations in which the first presentation of that stimulus took place. The patients
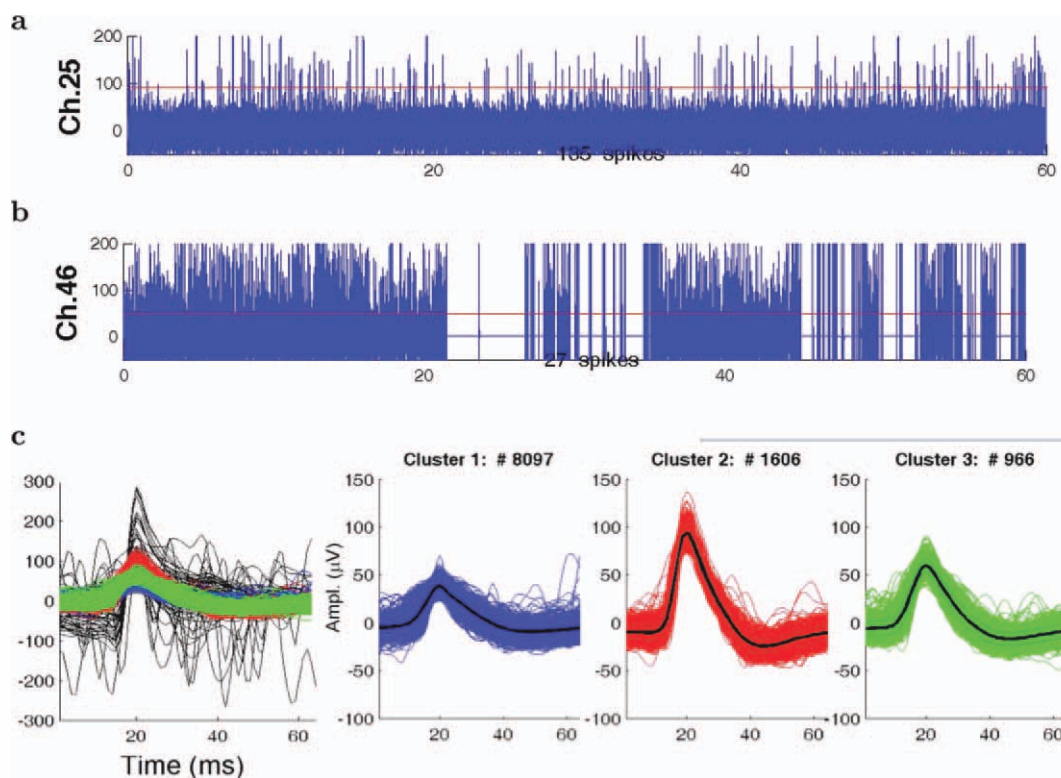
FIGURE 1. Extracellular recording. a and b. Selection of channels. Voltage is shown as a function of time over the course of the recording session for two channels. a. Channel 25 was considered viable. b. Channel 46 was not considered viable. It is likely that channel 46 was contaminated by electrical noise that was not biological in origin. c. Example of spike clustering. In this case, three clusters were isolated from the channel shown on the left. The panel on the far left shows all the waveforms for all of the spikes recorded by that channel. The three clusters to the right of that figure show the results of the algorithm. Cluster 1 was classified as a multiunit (8,097 spikes), whereas clusters 2 and 3 were classified to be single units. Note that the figures include all of the instances of spiking throughout the recording session. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

were able to reliably discriminate new from previously seen stimuli and showed no trend toward a response bias (hit rate $= 0.82 \pm 0.07$; correct rejection rate $= 0.83 \pm 0.07$). Patients participated in the task up to three times, with each recording session separated by a minimum of 48 h. The image sets were not randomized; that is, the first session for every patient contained the same images, and each session following also used the same stimulus set, albeit in a different order.

Stimuli within a block were presented in a quasi-random order. It would be impossible to completely control all possible relevant variables. For instance, it is necessarily the case that the first stimulus shown in a block is new and the last stimulus in the block is old. There is a general tendency for there to be more old stimuli later in the block than earlier in the block. In addition, the recencies that can be sampled at any one time depend on the position within the block and the stimuli that have been previously presented.

Recording Methods

At the tip of each intracranial depth electrode was a set of nine 40 µm platinum–iridium microwires: the ninth microwire served as a reference. It was stripped of insulation to allow sam-

pling over a wider anatomical region and had a lower impedance than the other eight microwires, which provided possible cellular signals. Anatomical locations of electrodes were verified via postplacement MRI scans and images created by fusing CT scans taken while electrodes were implanted with high-resolution MRI scans taken immediately before implantation (Fried et al., 1999). Signals from each microwire were amplified (gain $= 10,000$), digitally sampled at 27.8 kHz and bandpass filtered between 1 and 6 kHz (Neuralynx, Tucson, AZ). Using the spike separation algorithm wave_clus (Quiroga et al., 2004), we isolated single- and multiunit activity during microwire recordings. Single units were defined as waveforms with clear refractory periods, that were of high amplitude (>100 microvolts), and had less than 1% of spikes occurring at less than 3 ms. As an additional check for noise, we plotted the power spectral density using the times when spikes occurred for that unit; putative cell activity showing significant amounts of 60 Hz power-line activity was excluded from further analysis. We took steps to eliminate channels with gross changes in the distribution of voltages or signal-to-noise ratio over the course of the recording session. All channels were visually inspected for such changes. Figures 1a,b give an example of a channel that was included and a channel that was rejected on the basis of
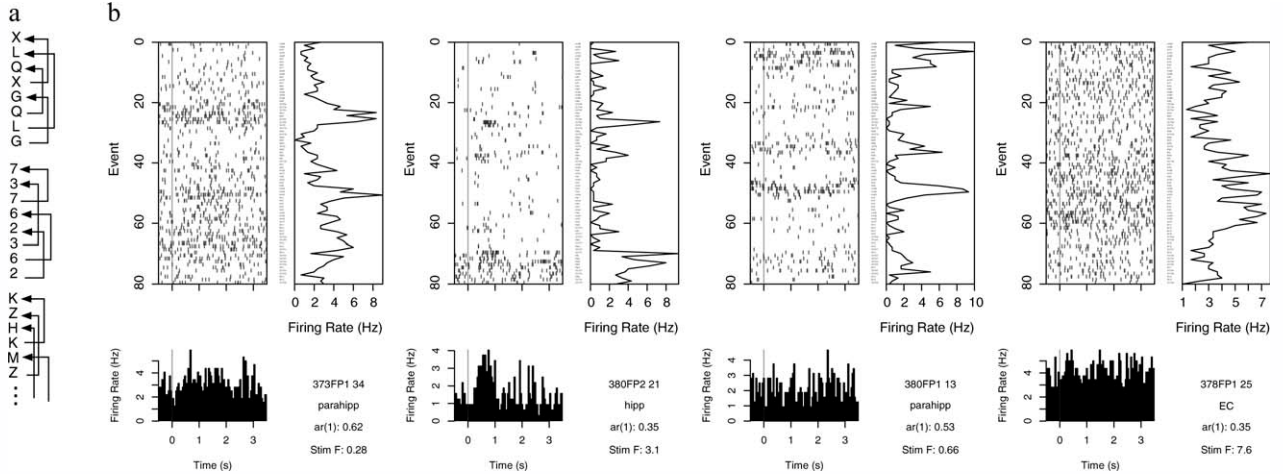
FIGURE 2.   MTL units show autocorrelated firing over macroscopic periods of time. a. Schematic of the behavioral task. Subjects were presented with a series of pictures. For each picture, the subject's task was to indicate whether or not the stimulus had been seen previously. There were four blocks of 20 stimuli. Stimulus repetitions were within the same block and all stimuli within a block were from the same category. The stimuli in blocks 1 and 3 were faces (schematically represented by letters); in blocks 2 and 4 were places (schematically represented by numbers). b. Representative units showing autocorrelated firing. For each unit, the upper left shows a raster plot from each of the 80 stimulus events aligned on stimulus presentation. The bottom left panel shows a peristimulus time histogram. The upper right shows the firing rate averaged across the 3 s post-stimulus for each event. The bottom right shows some properties of the units, including the session and unit, the brain region the unit was recorded from (hipp: hippocampus; parahipp: parahippocampus; EC: entorhinal cortex). The number ar(1) gives the magnitude of the unit's lag one regression coefficient. Stim F gives the F statistic for an ANOVA evaluating stimulus type. With 3 and 74 degrees of freedom, values of Stim F above 2.73 are significant at the 0.05 level. See text for details. The Appendix shows similar plots for an additional 24 units that were classified as temporally-autocorrelated.

these criteria. Figure 1c shows an example of a channel that was clustered into three units. These results were typical of the units in our experiment. Each and every unit was observed and evaluated by a human observer in addition to the clustering algorithm.

Neural data Analyses

We computed the average firing rate for each unit in the 3 s interval following each stimulus presentation event. The minimum amount of time between presentation of stimuli, including the time for the behavioral response, was 3.10 s, so there is no chance that a single spike is counted as part of two separate events. Let us refer to the firing rate on the jth event as f_j. To calculate the lag one regression coefficient for each cell and estimate its responsiveness to the stimulus categories, we fit the linear model

$$f_j = c + \phi f_{j-1} + \boldsymbol{\beta}' \vec{X_j} + \boldsymbol{\epsilon}_j \qquad (1)$$

where X_j is a vector describing the category of the stimulus presented at event j and $\boldsymbol{\beta}'$ are the coefficients estimated from that fit. We refer to the fitted value of ϕ as the lag one regression coefficient (see e.g., Fig. 3a). Because information about subcategory—male or female face, indoor or outdoor place—was recorded we used all four categories in this regression. Using just the broader categories of face versus place did not affect the results of the analyses examined here in any significant way.

To compare ensemble similarities of event vectors, we z-transformed each unit's activity across events before taking a weighted inner product. Referring to one ensemble vector across units as $\vec{v}$, and another as $\vec{u}$,

$$\text{Sim}(\vec{v}, \vec{u}) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{\sigma_k^2} (v_k - \mu_k)(u_k - \mu_k), \qquad (2)$$

where k indexes the units, μ_k is the mean of the kth unit across events, σ_k is the standard deviation of the kth unit across events and N is the number of units in the ensemble for that session. In addition to comparing event vectors, we also used the similarity function, Eq. (2), to compare the subset of the ensemble that was temporally autocorrelated.

Because of the variability in the number of units across recording sessions (Table 1), we do not simply report means averaged across sessions. Rather, the mean values report the average across units. That is, for the purposes of doing statistics, each unit was treated as a one-cell ensemble. Because recording sessions from a particular patient were separated by at least 48 h, we treated each unit as an independent observation.

RESULTS

Our analyses revealed a variety of findings. First we characterize the tendency of many units to change their firing gradually over macroscopic periods of time, on the order of tens of seconds. Second, we describe the stimulus- and category-specificity of recorded units. Finally we describe a conti-
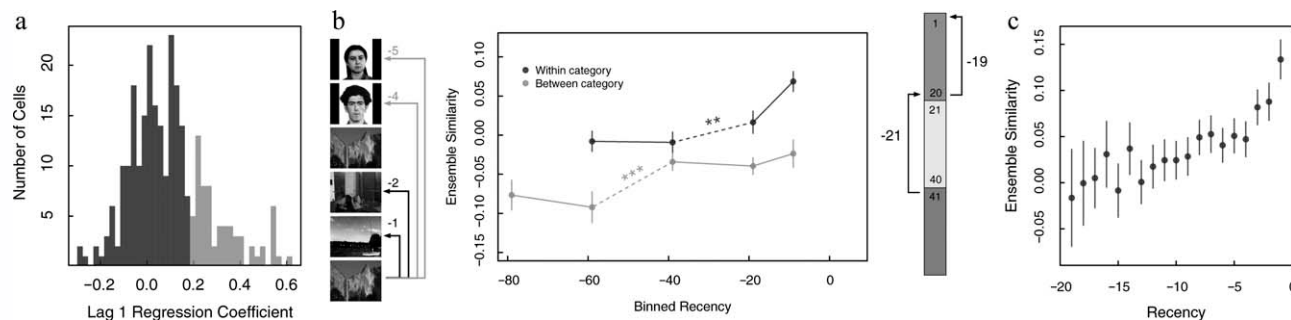
FIGURE 3. Quantification of the temporal autocorrelation of MTL ensembles. a. Distribution of lag one regression coefficients across cells. The lag one regression coefficient indicates the degree to which the firing of the cell at a stimulus event is predicted by its firing on the previous event, after controlling for stimulus identity. The distribution is significantly skewed in the positive direction. Light-shaded region indicates cells categorized as temporally-auto-correlated cells. b. MTL ensemble similarity reflected stimulus category, recency within block, as well as block structure. Left: event comparisons between different stimuli were aggregated as a function of recency separately for comparisons involving events from the same stimulus category (dark) or different categories (light). Middle: Ensemble similarity for all units was aggregated as a function of recency collapsed into five bins. The bins were chosen such that the largest absolute value of recency in each bin was 10, 20, 40, 60, or 80. The average ensemble similarity in each bin is shown as a function of the least recent event in each bin for within-category comparisons (dark) and between-category comparisons (light). The dashed lines connect recency bins that are separated by an entire block of 20 stimuli from one category. **$P < 0.005$; ***$P < 0.001$. Right: For example, the largest recency possible within a block of stimuli from the same category is 19. A recency of 20 is not possible for comparisons of stimuli from the same category. A recency of 21 for a same category comparison requires that a block of stimuli from the other category intervened between the two events. Similarly, an across-category comparison with a lag of 40 is not possible. c. Similarity of the ensemble for each event to preceding events within block as a function of the recency of the preceding event. For instance, comparing stimulus event 10 to stimulus event 3 yields a recency of 7. Pairs of events corresponding to the same stimulus or to stimuli from different categories are excluded from this analysis. All units were included in this analysis. In all figures, error bars are 95% confidence intervals calculated across units.

guity analysis performed on the gradually-changing part of the ensemble. To anticipate the results, units will show temporal autocorrelation over tens of seconds. The contiguity analysis will both rule out the possibility that the temporal autocorrelation is solely a recording artifact and also demonstrate both forward and backward contiguity effects, confirming a key prediction of the retrieved context approach.

MTL Ensembles Change Gradually Over Time

A memory representation that changes gradually over time should show evidence for temporal autocorrelation in unit firing. Informal observation of the firing of units across time revealed evidence of strong temporal autocorrelation (Fig. 2b). For many units, firing persisted for long periods of time. For instance, the second unit shown starts firing at an elevated rate after presentation of the stimulus on the 69th event and fires at an elevated rate for the remainder of the session—more than 30 s. Similarly, the fourth unit shown appears to show a gradual increase in firing rate from about event 22 to about event 60, events that are separated by more than 100 s, followed by a gradual decrease in firing rate that takes about 10 events, or more than 30 s. Autocorrelated units showed elevated firing in the time before stimulus presentation suggesting they were not simply responding to the just-presented stimulus.

Analysis of Single-Unit Autocorrelation

To quantify the autocorrelation across the behavioral session, we first averaged the firing rate of each unit over the 3 s following

presentation of each stimulus. We then regressed firing rate across events onto the firing rate from the previous event and the category of the presented stimulus. The value of the lag one regression coefficient provides a measure of the unit's autocorrelation across events. The distribution of lag one regression coefficients across units was significantly greater than zero, mean 0.094 ± 0.01, $t(269) = 9.31$, $P < .001$ (see Fig. 3a). We refer to units that had a lag one regression coefficient significantly greater than zero ($P < .05$) as temporally-autocorrelated units. Almost one quarter of all units recorded (0.248) were classified as temporally-autocorrelated units. The proportion of temporally-autocorrelated units in the hippocampus, parahippocampal region, and entorhinal cortex far exceeded the Type I error rate (0.05) and did not reliably differ among those regions, $\chi^2 (2) = 4.87$, $P > 0.05$ (see Table 2). The number of units in the amygdala classified as temporally-autocorrelated did not exceed the Type I error rate. The proportions of autocorrelated units did not differ across sessions $\chi^2 (7) = 7.03$, $P > 0.4$. We did not identify any other reliable differences between brain regions in these analyses. Note, however, that we averaged over 3 s intervals so that differences between regions that manifest on shorter time scales would be invisible to us.

Ensemble Changes Across Blocks of Stimuli

To characterize the temporal sensitivity of the ensemble of units, we constructed event vectors by taking the average firing rate for each event across units. We then aggregated the ensemble similarity between events as a function of the recency from each event to its predecessors. For instance, comparison of the

Number of Units and Number of Units That Showed Significant Temporal Autocorrelation (t-Units) as a Function of the Anatomical Region in Which They Were Located

Region	*t*-Units	Total	Prop.
EC	19	98	0.19
PHG	25	73	0.34
Hipp.	20	72	0.28
Amyg.	3	27	0.11

EC, entorhinal cortex; PHG: parahippocampal gyrus; Hipp.: hippocampus; Amyg.: amygdala.

event corresponding to the presentation of the 10th stimulus to the event corresponding to the 7th stimulus yields a recency of -3. Similarly, comparison of the event corresponding to presentation of the 10th stimulus to the event corresponding to the third stimulus yields a recency of -7. For this analysis, we excluded comparisons in which the events corresponded to presentation of the same stimulus.

There was a large effect of stimulus category and block of the experiment on the ensemble response that coexisted with the robust temporal autocorrelation effect.

Figure 3b shows similarity averaged across all units for each of five recency bins. The first bin included recencies greater than -10. The second bin included $-20 > r \geq -10$. The borders between the subsequent bins were -40 and -60; the last bin contained $r < -60$. Ensemble similarity is broken out for comparisons between pairs of event of the same category excluding events with the same stimulus (darker symbols), and among pairs of events from different categories (lighter symbols). Both within- and across-category comparisons showed a recency effect. The ensemble similarity for across-category comparisons in the first recency bin, including recencies -1 to -9 was significantly greater than in the second recency bin, which contained recencies -10 to -19, $t(269) = 2.03$, $P < 0.05$. Even when controlling for the recency effect, there was a dramatic difference among within- and between-category comparisons. In addition, there was also an effect of category block. The dashed lines connect points corresponding to bins that are separated by a complete block of stimuli from one category. That is, because of the blocked structure of the session, a recency of -19 for a within-category comparison (the last event in a block compared with the first event in that block) means that only events from the same category intervene. A recency of -21 for a within-category comparison (the first event of one block and the last event of the previous block with the same category), however, requires that a block of 20 events from the other category intervene. For the within-category comparisons, there was a highly significant difference between the bin with recencies -10 to -19 and the bin containing recencies -21 to -40, $t(269) = 2.88$, $P < 0.005$, but no difference between bins 4 and 5, $t(269) = 0.19$. Similarly, for the between-category comparisons, the difference between bins three and four was highly significant, $t(269) = 5.49$, $P < 0.001$.

Hippocampus

Within-Block Temporal Autocorrelation

To isolate temporal autocorrelation from the effects of stimulus category and block (which were also autocorrelated in time due to the structure of the experiment) we restricted our attention to pairs of events within the same block. Because of the blocked structure of the stimulus presentation, pairs of events from the same block are necessarily from the same stimulus category. As before, we excluded pairs of events corresponding to the same stimulus. Figure 3c illustrates the results of this analysis. The ensemble similarity of pairs of event vectors decreased gradually as the time between the events increased. A linear regression showed a significant effect of recency on ensemble similarity, regression coefficient 0.006 ± 0.001, $F_{(1,17)} = 53.3$, $P < 0.001$. The change over time was not simply due to an elevated level of ensemble similarity for adjacent events. A regression excluding recencies -8 to -1, focusing on just recencies -19 to -9, still found a significant regression coefficient, 0.003 ± 0.001, $F_{(1,9)} = 5.64$, $P < 0.05$. This analysis sets a lower limit on the time scale of the change over time. Because the similarity continued to decrease after recency -9, this means that the ensemble continued changing reliably after nine intervening stimuli, or a delay of more than 27 s.

MTL Ensembles Coded for Stimulus and Category

Before turning to the contiguity analyses, we first characterize the sensitivity of MTL units to stimulus identity and category. MTL ensembles exhibited robust stimulus and category selectivity. Both of these properties were exhibited both by units that did and that did not exhibit temporal autocorrelation. The linear model used to assess lag one autocorrelation also provided information about the responsiveness of the units to the category of stimulus at each event. We categorized units that showed an effect at the 0.05 level, $F_{(3,74)} > 2.73$, as category-selective. A total of 44/270 units were category-selective. This greatly exceeded the Type I error rate, $\chi^2(1) = 70.1$, $P < 0.001$. Of the 44 category-selective units, 12 were also classified as temporally-autocorrelated. This proportion (0.27) did not differ from the proportion of autocorrelated units among units that were not category-selective (0.25), $\chi^2(1) = 0.05$.

Because of the blocked structure of this experiment, the category of the stimulus is confounded with the event number. As a consequence, comparisons that are recent are also likely to be stimuli from the same category. In order to isolate the effects of stimulus and category in relative isolation of recency, ensemble similarities were aggregated as a function of recency separately for (a) comparisons between presentations of the same stimulus, (b) pairs in which both stimuli were from the same category (faces or scenes) but were different stimuli, and (c) pairs in which the stimuli were from different categories. In order to assess the effect of these variables independently of recency, we performed comparisons across pairs of similarities with the same recency. These analyses yielded several findings.
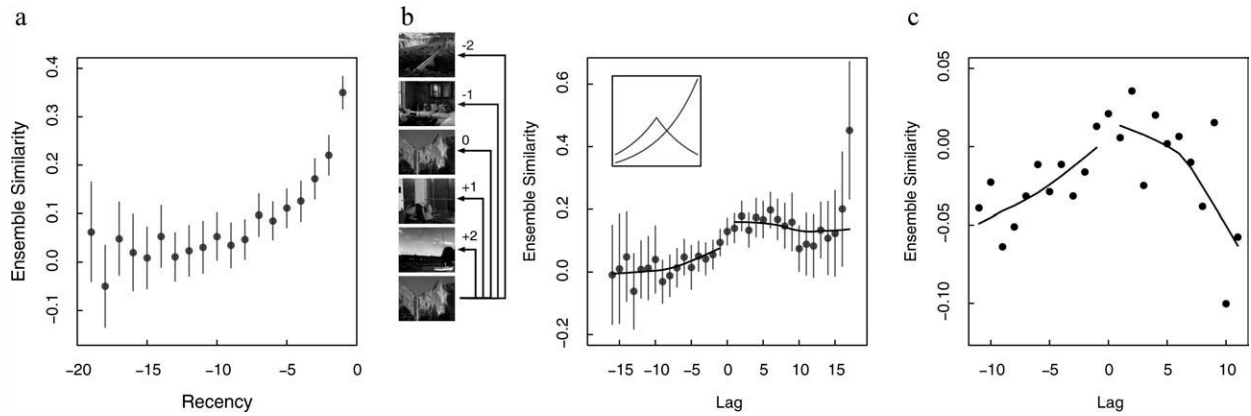
FIGURE 4. MTL ensembles show the neural signature of a jump back in time. a. Recency analysis, as in Figure 3b, but restricted to cells showing a significant lag one autocorrelation function. b. Contiguity analyses. Left: Similarity of the ensemble representation during presentation of a repeated stimulus was compared with preceding stimulus events. As with the recency analysis in a, comparisons to stimuli from the other stimulus category were excluded. Rather than aggregating the results as a function of recency, results were aggregated as a function of lag, the difference in serial position between the original presentation of the stimulus and its neighbors. Note the comparison is always between an event and another event that preceded it. Right: Contiguity analysis using only the temporally-correlated units. Rather than a smoothly decreasing function, as one would expect from a pure recency effect, the results appear to correspond to a peak around zero superimposed on a recency effect. Inset: a recency effect is shown with a contiguity effect. c. Analysis in (b) repeated after subtracting out the part attributable to the recency effect. The results show a robust contiguity effect, with significant effects of |lag| in both the forward and backward directions. For all panels, error bars are 95% confidence intervals calculated across units and solid lines are the result of a LOESS locally-weighted least squares spline regression.

First, MTL ensembles distinguished individual stimuli. Same-stimulus pairs were more similar than same category-pairs, paired t-test across recencies, $t(15) = 3.18$, $P < 0.01$. Second, MTL ensembles captured information about properties of the stimuli other than stimulus identity. The ensemble similarity between same-category pairs was greater than for different-category pairs, $t(56) = 13.9$, $P < 0.001$. Because this comparison is across recencies, the difference cannot be attributed to a confounding of recency and category structure.

We repeated these analyses for the units categorized as temporally-autocorrelated, which we refer to as the temporally-autocorrelated part of the ensemble, and the units that were not categorized as temporally-autocorrelated. Robust stimulus coding was not observed using this method in the temporally-autocorrelated part of the ensemble, $t(15) = 0.46$.

Note that a contiguity effect caused by the stimuli in the temporally-autocorrelated part of the ensemble would make an effect of stimulus identity harder to identify. If the effect of a stimulus is spread across many events, the difference between the event when the stimulus was presented and neighboring events would be reduced. However, there was robust category-specificity in the temporally-autocorrelated part of the ensemble; similarity was greater for comparisons between same-category event pairs than between different-category pairs, $t(56) = 5.02$, $P < 0.001$. The temporally-autocorrelated part of the ensemble also showed a large effect of block structure comparable to that shown in Figure 3b, with a significant effect between bins 2 and 3 for the within-category comparisons, $t(66) = 3.32$, $P < 0.005$, and between bins 3 and 4 for the between-category comparisons, $t(66) = 3.67$, $P < 0.001$. The non-auto-correlated part of the ensemble showed robust stimulus and category selectivity. Across recencies, same-stimulus comparisons showed greater similarity than did same-category comparison, $t(15) = 4.55$, $P < 0.001$. Similarly, within-category comparisons were more similar to one another than between-category comparisons, $t(56) = 14.68$, $P < 0.001$.

A Neural Contiguity Effect in the MTL Ensemble

Artifactual accounts of the temporal autocorrelation require that there be no causal relationship between the stimuli being presented and the gradually-changing ensemble response. To take one example, although the shape of action potentials recorded extracellularly might change over time, leading to an artifactual temporal autocorrelation in the ensemble response, there is no reason to suspect that the specific identity of the picture presented to the patient affects the shape of the action potential. We undertook a contiguity analysis of the temporally-autocorrelated ensemble. The analysis allows us to diagnose a causal relationship between the stimulus presented and the gradually-changing response, ruling out artifactual accounts of the temporal autocorrelation. It also enables us to look for the neural signature of a jump back in time.

It does not make sense to ask about whether the neural signature of mental time travel is observed in units that do not show temporal autocorrelation. Because our interest here is in the temporally-autocorrelated part of the ensemble response, we restricted our attention to the units classified as temporally-correlated (lighter distribution in Fig. 3a). Not surprisingly, this

temporally-autocorrelated ensemble showed robust autocorrelation as revealed by the recency analysis (Fig. 4a). Because the experiment includes repeated presentations of the same stimulus, we can examine the degree to which the ensemble responds to that particular stimulus. Because we are interested not in stimulus-selectivity per se but the degree to which the temporally-autocorrelated activity is caused by the stimuli, we examined the degree to which the response to a repeated stimulus resembles the neighbors of its original presentation. In order to avoid artifactual category-specific contiguity effects, we restricted our attention to pairs of events within the same block—these were necessarily from the same category. The left part of Figure 4b describes this schematically. For each repeated stimulus, we compared the ensemble activity for that event to preceding events from the same block and aggregated the results, not as a function of the recency of the comparison, but as a function of the lag between the earlier event and the initial presentation of the repeated stimulus.

Figure 4b shows the results of the contiguity analysis conducted on the temporally-autocorrelated part of the ensemble. Visually, the results are just as we would expect if it were the combination of a recency effect due to ongoing activity with a contiguity effect superimposed on top of it (inset). The recency effect is visible in the overall tendency for the curve to rise from left to right. The apparent boost in the similarity to the neighbors of the original presentation of the repeated item around lag = 0 is consistent with a contiguity effect. To statistically remove the recency effect, we estimated the results of the contiguity analysis we would have obtained if recency were the only factor causing ensemble similarity.

The strategy of this analysis was to directly estimate the way the ensemble similarity as a function of lag (Fig. 4b) would look if it were due only to the empirically-observed recency gradient and then subtract that estimate from the actual curve. The difference between the actual curve and the one expected on the basis of recency ought to be attributable to the input to the ensemble caused by the most recently presented stimulus. The method thus enables a comparison of the input caused by the repeated stimulus to the neighbors of its original presentation.

The method starts by estimating the recency gradient isolated as far as possible from contiguity effects. To do this, we took the comparison between all pairs of events where the later event was a stimulus presented for the first time and the previous event was from the same block of stimuli. We then aggregated these ensemble similarities as a function of recency to compute a recency gradient. This recency gradient closely resembled Figure 4a (see Fig. A1a). The next step used this recency gradient to estimate the part of the contiguity effect solely attributable to recency. To do so, we took the same pairs of events as used in the contiguity analysis. However, rather than aggregating the observed ensemble similarity between those events as a function of lag, we aggregated the value of the recency gradient corresponding to the recency between the two events being compared as a function of lag. This measure

uses the cleanest estimate of recency independent of stimulus repetitions available to us and controls for the actual distribution of recencies sampled in the contiguity analysis. Because the ensemble similarities get very noisy around the edges (note the size of the error bars for extreme values of lag in Fig. 4b) we focused our attention on the central region of the contiguity analysis.

Figure 4c shows the difference between the observed ensemble similarity (Fig. 4b) and the estimate due to the recency gradient (the results of this intermediate step can be found in Fig. A1b). First we note that if the temporal autocorrelation in the ensemble response were solely caused by uncontrolled variables or recording artifacts, this would require that the curve in Figure 4c be flat. This prediction was falsified. An ANOVA with |lag| as a regressor and direction (forward vs. backward) as a categorical variable yielded a main effect of the absolute value of lag, $F_{(1,18)} = 13.1$, MSe = 0.0081, $P < 0.002$, definitively demonstrating that the autocorrelation in the ensemble cannot be solely caused by random variation independent of the identity of the stimuli.

Figure 4c also enables us to examine the similarity of the input caused by the repeated item to the neighbors of the original presentation in the region around lag = 0. Consider first what we would expect if the temporal autocorrelation was a consequence of a short-term memory acting as a container to holding recently-presented stimuli for an extended period of time. In this case, the input caused by the stimulus would be the same across presentations. We would expect to see a boost in similarity falling off as a decreasing function of lag in the forward direction but no effect of lag in the backward direction. In contrast, if the repeated stimulus caused a "jump back in time" by retrieving information available before the original presentation of the stimulus, then ensemble similarity will decrease with increasing |lag| in both the forward and backward directions. The decrease in the backward direction would imply that the repeated item caused a recovery of information available before its original presentation.

The ANOVA did not show a main effect of direction, $F_{(1,18)} = 1.64$, MSe = 0.001, nor an interaction between contiguity and direction, $F_{(1,18)} = 1.18$, MSe = 7×10^{-4}. To analyze the effect of contiguity separately in the forward and backward directions, we conducted linear regressions. There was a significant decrease in similarity as a function of |lag| in the forward direction, 0.008 ± 0.003, $F_{(1,9)} = 7.04$, $P < 0.03$, consistent with both short-term memory and retrieved context accounts. There was also a significant effect in the backward direction. The regression of similarity onto |lag| produced a value, significantly less than zero, 0.0042 ± 0.0015, $F_{(1,9)} = 7.48$, $P < 0.03$. The results of the contiguity analysis are inconsistent with what one would expect if the repeated stimulus were simply matching ongoing activity that was induced by its original presentation. They are consistent with the hypothesis that repetition of a stimulus causes a jump back in time by recovering information that was available prior to the original presentation of the stimulus.

DISCUSSION

The contiguity effect is an extremely robust feature of episodic memory experiments (Kahana et al., 2008). One hypothesis for the contiguity effect in episodic memory as mental time travel requires that the brain maintains a representation of the "now" that changes gradually from moment to moment. In this view episodic memory for an event should be accompanied by a jump back in time in which the previous event is recovered. This account of the contiguity effect is distinguishable from persistent stimulus activation coupled with direct inter-item associations by a qualitatively different neural contiguity profile. Our results show strong evidence that the MTL maintains a representation that changes gradually over time that is not attributable to recording artifacts nor to uncontrolled experimental variables. These accounts would have predicted a flat curve in Figure 4c. The backward effect in Figure 4c is not predicted by persistent stimulus activation in short-term memory but is predicted by a jump back in time.

Is it Possible These Findings are Recording Artifacts?

The unit isolation algorithms used here are consistent with many prior studies (Ekstrom et al., 2003; Gelbard-Sagiv et al., 2008; Viskontas et al., 2006, 2009). Nonetheless, the problem of sorting spikes from many individual neurons recorded on a single recording wire is a difficult one. There are two types of unit isolation errors we should consider. In one type of error, a single neuron could have spikes sorted into multiple units. If whatever factors that favor one unit over the other are autocorrelated, then this type of error could result in an artifactual autocorrelation. This artifact, however, could not account for the stimulus-specificity, sensitivity of the ensemble to the block structure, or the fact that the autocorrelated signal retains information about preceding stimuli. In the other type of error, spikes from multiple neurons are counted as belonging to one unit. This type of error could lead to artifactual conjunctions in the properties of units. For instance, if spikes from a stimulus-specific neuron were counted as part of the same unit as an autocorrelated neuron, we might incorrectly conclude that there was an autocorrelated stimulus-specific cell. Because the identity of the stimulus affects the response to the neighbors of the original presentation of the stimulus, the results of our contiguity analysis is not subject to this type of artifact either. At least one of the neurons contributing to the unit would have to have the property that its activity is sensitive to the recent history of stimulus presentation.

Although artifacts may exist, the contiguity analysis provides very strong evidence against the hypothesis that all of the temporal autocorrelation in ensemble activity is due to artifactual sources. The fact that the contiguity analyses resulted in curves that are peaked around zero and fall off gradually, in both the forward and the backward direction, means that there is a causal relationship between the stimulus being presented to the patient and autocorrelated ensemble activity. If this result were an artifact, this would require that there is a causal relationship between the particular stimulus being shown to the patient and the shape of the spikes being recorded.

The contiguity analysis also rules out the possibility that the temporal autocorrelation results from uncontrolled environmental stimuli that are extended in time (Estes, 1955; Mensink and Raaijmakers, 1988; Murdock, 1997). For instance, perhaps a cell fired for part of the recording session because there was a fan or other environmental sound that was present only during that part of the session. A number of mathematical and computational models of memory postulate a memory representation that changes gradually from moment to moment in a way that is more or less independent of the stimuli presented (e.g., Estes, 1955; Mensink and Raaijmakers, 1988; Murdock, 1997; Burgess and Hitch, 1999; Brown et al., 2000).

How Does the MTL Accomplish a Jump Back in Time?

The present results imply that the MTL responds to stimuli over long periods of time and that repetition of a stimulus enables the MTL to partially recover its previous state. While necessarily speculative, it is worth thinking about how these two functions might be accomplished.

The long-term temporal autocorrelation could result from either network properties or intrinsic cellular properties that lead to persistent firing in response to a transient input. An extensive and growing body of work from slice preparations suggests that neurons in many parts of the MTL are equipped with a variety of intrinsic mechanisms that enable them to show persistent firing over long time scales in the absence of synaptic input. Such cells have been observed in the entorhinal cortex (Egorov et al., 2002; Fransén et al., 2006; Tahvildari et al., 2007), lateral amygdala (Egorov et al., 2006), postsubiculum (Yoshida and Hasselmo, 2009), perirhinal cortex (Navaroli et al., in press), and hippocampus (Sheffield et al., 2011). Similar phenomena have been observed outside of the MTL in the anterior cingulate cortex (Zhang and Seguela, 2010) and prefrontal cortex (Sidiropoulou et al., 2009). Some of these mechanisms appear dependent on acetylcholine (Egorov et al., 2002; Tahvildari et al., 2007; Yoshida and Hasselmo, 2009; Navaroli et al., in press); others do not (Yoshida et al., 2008; Sheeld et al., 2011). These findings from slice preparations indicate a variety of intracellular mechanisms in a number of different MTL structures that lead to autocorrelation over long periods of time. This suggests that autocorrelation over macroscopic periods of time is an essential aspect of MTL function.

If there is a gradually-changing representation of recent experience accessible to the MTL, this naturally raises the question of how it could be recovered to implement a jump back in time. For decades, researchers have noted that the structure of the hippocampus, in particular the recurrent structure of CA3, could be used for autoassociative pattern completion (Marr, 1971; McNaughton and Morris, 1987; Hasselmo and Wyble, 1997; Norman and O'Reilly, 2003). The basic idea is that during encoding, recurrent synapses are enhanced between coactive

neurons. When a subset of the active neurons are reactivated during retrieval, these strengthened connections enable reactivation of the remainder of the neurons that were active during encoding. Of course, the pattern that is completed depends on the nature of the representation that provides input to the hippocampal region. If the representation in extrahippocampal MTL regions is temporally-autocorrelated, then pattern completion of such a gradually-changing pattern would seem to be sufficient to account for neural contiguity effect we observed here. At each moment, the current state of the gradually-changing representation is updated with input caused by the present stimulus. At any moment, inputs caused by multiple preceding stimuli are available. Suppose that when a particular stimulus is repeated, this result in pattern completion of the gradually-changing pattern that was present when it was initially experienced. This recovered pattern resembles the states that preceded the original presentation of the stimulus in a way that falls off with distance from the original presentation. In addition, because the recovered state—and any stimulus-specific pattern that is recovered—will also resemble states that followed the initial presentation of the stimulus. The degree of this resemblance will fall off with the distance from the original presentation of the stimulus.

The present result, coupled with extensive behavioral modeling of recency and contiguity effects in episodic memory tasks (e.g., Howard et al., 2009a,b; Polyn et al., 2009; Sederberg et al., 2008, 2011), suggest a proximal target for computational neuroscience models that aspire to describe MTL-dependent cognition. This work suggests that understanding the cellular and network mechanisms that support temporal autocorrelation and the recovery of gradually-changing MTL representations is central to understanding the cognitive function of the MTL.

What is Being Retrieved?

Retrieved temporal context models hypothesize that the input caused by a stimulus changes across presentations to reflect the temporal contexts in which it has been experienced. While we did observe strong evidence that the autocorrelated ensemble activity is causally affected by the stimuli and reflects the category membership of the stimuli, we cannot determine with certainty from these data whether the change in the ensemble activity caused by the initial presentation of a stimulus had any relationship to the particular stimulus shown. A statistically reliable asymmetry in Figure 4c would have been evidence that the input to the autocorrelated ensemble during the initial presentation was caused by the stimulus. Because we did not observe such an asymmetry, it is possible that during the initial presentation of a stimulus MTL activity changed due to random processes; this random state could have been subsequently recovered by the repeated item.

At minimum our results show that the input caused by a repeated item reflects the temporal context in which it has been presented. This finding is consistent with neurophysiological findings showing that stimulus representations in the MTL and inferotemporal cortex do in fact change over time to reflect the temporal cooccurrence of stimuli (Miyashita, 1988; Naya et al., 2001). Indeed, a recent fMRI study showed that activation in the parahippocampal cortex responds not only to the visual category of the scene shown, but also to whether it is presented in the same temporal context in which it was previously experienced (Turk-Browne et al., unpublished data). Carried to the extreme with many presentations of a stimulus in a variety of temporal contexts over a statistically rich learning environment, this process could result in the learning of useful semantic representations (Rao and Howard, 2008; Howard et al., 2011).

Time-Scale of Autocorrelation

Recency and contiguity effects have been observed over time scales considerably longer than the tens of seconds we were able to measure reliable autocorrelation in this experiment (Howard et al., 2008; Unsworth, 2008; Moreton and Ward, 2010). The results from this experiment, however, do not place a strong constraint on the upper limit of the time-scale of the autocorrelation of the ensemble representation. The experimental methods impose two time scales. If there was autocorrelation caused by the stimuli at a time scale faster than 3 s, the presentation rate of the stimuli, it would have been invisible to these analyses. Similarly, autocorrelation on the scale of the blocks of stimuli (on the order of a minute) could have easily been confused with responsiveness to the stimulus categories. Nonetheless, the present analyses suggest that the MTL maintains information over at least a few dozen seconds. Although our data analyses assume that the ensemble changes over time they do not uniquely support a particular mathematical form of the change. It is also possible that the MTL is maintaining a representation of the recent past in which cells are responding to prior history at a delay (Grossberg and Merrill, 1992; Hasselmo, 2007; Shankar and Howard, 2012). Indeed, MacDonald et al. (2011) have shown evidence that hippocampal pyramidal neurons encode the time since presentation of a nonspatial stimulus.

Jumping Back in Time in Other Brain Regions

While the present study showed evidence for the neural signature of a jump back in time in the medial temporal lobe, this does not preclude the possibility that similar results could be observed in other brain regions. In particular, it has been suggested that the neural representation of a gradually-changing state of temporal context should reside in prefrontal cortex (Polyn and Kahana, 2008). Given the strong reciprocal functional connections between the MTL and the PFC (Simons and Spiers, 2003; Siapas et al., 2005; Anderson et al., 2010), this is not at all contradictory to the hypothesis that the MTL maintains and recovers a gradually-changing representation of temporal context (Howard et al., 2005).

A recent study that examined evidence for a neural contiguity effect in electrodes distributed throughout the brain found strong evidence for a neural contiguity effect in the temporal lobe but did not find evidence for a neural contiguity effect in other regions. Manning et al. (2011) recorded from patients implanted with intracranial electrodes while they freely-recalled

from lists of high-frequency words. They compared the oscillatory components in the time period just before recalling an item to the states during the study of the neighbors of that item, demonstrating a neural contiguity effect not unlike the one observed in the present study. Notably, when considering electrodes from the entire brain, they established a significant association between the behavioral contiguity effect and the oscillatory contiguity effect. These effects were also observed when they restricted their attention to the temporal lobe (including MTL sites). Frontal lobe electrodes showed a trend towards an effect but it did not reach significance.

There is other evidence that suggests that there is a special role for the prefrontal cortex in temporal context memory. Neuropsychological findings indicate that frontal lesions preferentially affect temporal order judgments (Shimamura et al., 1990; McAndrews and Milner, 1991). Jenkins and Ranganath (2010) presented fMRI results suggesting that there is a gradually-changing representation in prefrontal regions that is associated with success on a judgment of recency task that depends on subjects' ability to remember the time at which an item was presented. That study also found that activation in the parahippocampal cortex and bilateral hippocampus was correlated with success on temporal ordering tasks. One possible way to reconcile these findings with the present result is that a gradually-changing representation of the past is maintained in both PFC and the MTL but recovery of prior states depends on the MTL. Perhaps performance in judgment of recency tasks requires detailed examination of the present state rather than recovery of a past state. Perhaps the ability to scan through current and recovered memory states (Hacker, 1980; Chan et al., 2009) depends critically on subregions of the PFC.

Accommodating the Present Results in Neural and Cognitive Models of Memory

Although the results of the contiguity analysis are not predicted by identification of the MTL ensemble with persistent stimulus-specific activation, our findings could be accommodated within existing mathematical cognitive models of short-term memory. For instance, in the search of associative memory model (SAM, Raaijmakers and Shiffrin, 1980), associations in long-term memory are formed between items that co-occur in short-term store. If the rules describing the contents of short-term store were altered to reflect those associations, such that repetition of a stimulus caused not only reactivation of that stimulus in short-term memory, but also activated associated stimuli, then the contents of short-term memory would show the backward contiguity profile seen in Figure 4c. It is possible that this conceptual change could represent a significant neurocomputational challenge to neural circuit models of short-term memory. If persistent activity in short-term memory reflects reverberatory activity via recurrent connections (Durstewitz et al., 2000; Wang, 2001), then directly changing inter-item weights within short-term memory could also affect the stability of those attractor states thus affecting the stability of short-term memory. In any event, by modifying the conception of short-

term memory in this way, our neural results could be accommodated, but at the cost of being able to conceptualize short-term store as a container that holds representations of stimuli. Put another way, the present results could be accommodated within models of short-term memory by making short-term memory less like a container that holds recently-presented stimuli and more like a gradually-changing memory representation that jumps back in time when a stimulus is repeated.

It is also possible to account for the results of our analyses by arguing that they reflect the recovery of information from long-term memory rather than persistent activation of stimuli in short-term memory (which may take place outside of the MTL). This position is not falsified by our findings, but there are several considerations worth mentioning. First, this position would nonetheless have strong implications for neural circuit models of the MTL that focus on direct inter-item associations rather than a jump back in time (Lisman, 1999; Mehta et al., 2000; Rolls and Kesner, 2006; Koene and Hasselmo, 2007). Second, there is no task requirement for retrieval of associations from long-term memory in item recognition. Although contiguity effects can be observed in item recognition (Schwartz et al., 2005), they are considerably more subtle than in recall tasks that explicitly tap the associations made between study words. If it turns out that effortful retrieval of associations from long-term memory is not necessary to observe a neural contiguity effect, then this would argue against this view.

Recent mathematical and neurocomputational models of item recognition in long-term memory (e.g., Shiffrin and Steyvers, 1997; Norman and O'Reilly, 2003) would also not account for the neural contiguity effect without significant modification. For instance in the retrieving effectively from memory (REM) model (Shiffrin and Steyvers, 1997) of item recognition, a probe item is compared with each trace in long-term memory in leading to a recognition decision. Even if the model were elaborated such that traces included the contents of short-term memory, then the match of an old probe to neighboring study traces would not exhibit the bidirectional contiguity profile we observed (Fig. 4c), showing instead only a forward effect. However, if an old probe caused reactivation of the trace that was formed when that stimulus was originally presented, this would account for the backward effect. This is so because the study trace would include information from recently-presented stimuli; the recovered study trace and would thus overlap with traces that preceded initial presentation of the stimulus. This elaboration of the model would exhibit the bidirectional contiguity profile we observed. Similarly, the Norman and O'Reilly (2003) model as published would not account for the neural contiguity effect but could be elaborated to do so. According to Norman and O'Reilly (2003) the recollective process in recognition memory results in recovery of a specific trace, enabling the recovery of temporally-defined associations. However, in their simulations of item recognition, only a single item was available in the entorhinal cortex input to the hippocampus at one time. The Norman and O'Reilly (2003) model could be altered to account for our neural contiguity findings by assuming that the representation in entorhinal

cortex that provides input to the hippocampus changes gradually over time reflecting information from several recently-presented stimuli. This entorhinal representation would change gradually over time and the input caused by a stimulus would change from one presentation of the stimulus to another. Notably, in each of these cases, the modifications necessary to account for our findings would move them closer to the conception of episodic memory as a jump back in time to a previous state of a gradually-changing now.

CONCLUSIONS

Work from the behavioral human memory literature suggest that the contiguity effect is a major component of episodic memory (Kahana et al., 2008; Sederberg et al., 2010). Modeling work suggests that episodic memory is cued by a gradually-changing representation of temporal context. In these models, the contiguity effect is caused by a "jump back in time" by which a repeated stimulus recovers the state of temporal context that was available before its original presentation. We found evidence for both of these predictions. The evidence for nonartifactual temporal autocorrelation is strong. The ensemble response changed gradually over tens of seconds and was causally affected by the identity of the stimulus presented. The contiguity analysis showed a statistically reliable backward effect (Fig. 4c) confirming a prediction of models that postulate that the contiguity effect results from a jump back in time of a neural representation.

We sought evidence for a jump back in time by comparing the ensemble state of a repeated stimulus to the states neighboring the original presentation of that stimulus. Manning et al. (2011) used a similar strategy to analyze human ECoG during performance of a free recall task. In principle at least, the same gambit could be applied to data from a variety of tasks and dependent measures, including fMRI data. The complexity of the contiguity analyses used here was necessary to disentangle a large recency effect from the contiguity effect. Future research should disentangle recency from contiguity by, for instance, presenting repeated stimuli after a long delay. In addition, behavioral measures distinguishing whether the repeated stimulus evoked an episodic memory would also be extremely useful establishing a relationship between the neural jump back in time and the cognitive experience of episodic memory.

Acknowledgments

REFERENCES

Anderson K, Rajagovindan R, Ghacibeh G, Meador K, Ding M. 2010. Theta oscillations mediate interaction between prefrontal cortex and medial temporal lobe in human memory. Cereb Cortex 20:1604.

Atkinson RC, Shiffrin RM. 1968. Human memory: A proposed system and its control processes. In: Spence KW, Spence JT, editors. The Psychology of Learning and Motivation, Vol. 2. New York: Academic Press. pp89–105.

Brown GD, Preece T, Hulme C. 2000. Oscillator-based memory for serial order. Psychol Rev 107:127–181.

Burgess N, Hitch GJ. 1999. Memory for serial order: A network model of the phonological loop and its timing. Psychol Rev 106:551–581.

Chan M, Ross B, Earle G, Caplan JB. 2009. Precise instructions determine participants' memory search strategy in judgments of relative order in short lists. Psychon Bull Rev 16:945–951.

Davelaar EJ, Goshen-Gottstein Y, Ashkenazi A, Haarmann HJ, Usher M. 2005. The demise of short-term memory revisited: Empirical and computational investigations of recency effects. Psychol Rev 112:3–42.

Dennis S, Humphreys MS. 2001. A context noise model of episodic word recognition. Psychol Rev 108:452–478.

Durstewitz D, Seamans JK, Sejnowski TJ. 2000. Neurocomputational models of working memory. Nat Neurosci 3:1184–1191.

Egorov AV, Hamam BN, Fransén E, Hasselmo ME, Alonso AA. 2002. Graded persistent activity in entorhinal cortex neurons. Nature 420:173–178.

Egorov AV, Unsicker K, von Bohlen und, Halbach O. 2006. Muscarinic control of graded persistent activity in lateral amygdala neurons. Eur J Neurosci 24:3183–3194.

Eichenbaum H, Yonelinas A, Ranganath C. 2007. The medial temporal lobe and recognition memory. Ann Rev Neurosci 30:123–152.

Ekstrom AD, Kahana MJ, Caplan JB, Fields TA, Isham EA, Newman EL, Fried I. 2003. Cellular networks underlying human spatial navigation. Nature 425:184–188.

Estes WK. 1955. Statistical theory of spontaneous recovery and regression. Psychol Rev 62:145–154.

Fransén E, Tahvildari B, Egorov AV, Hasselmo ME, Alonso AA. 2006. Mechanism of graded persistent cellular activity of entorhinal cortex layer V neurons. Neuron 49:735–746.

Fried I, Wilson CL, Maidment NT, Engel J, Behnke E, Fields TA, MacDonald KA, Morrow JW, Ackerson L. 1999. Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients. Technical note. J Neurosurg 91:697–705.

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 322:96–101.

Glenberg AM, Bradley MM, Stevenson JA, Kraus TA, Tkachuk MJ, Gretz AL. 1980. A two-process account of long-term serial position effects. J Exp Psychol 6:355–369.

Grossberg S, Merrill J. 1992. A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. Cogn Brain Res 1:3–38.

Grossberg S, Pearson LR. 2008. Laminar cortical dynamics of cognitive and motor working memory, sequence learning and performance: Toward a unified theory of how the cerebral cortex works. Psychol Rev 115:677–732.

Hacker MJ. 1980. Speed and accuracy of recency judgments for events in short-term memory. J Exp Psychol 15:846–858.

Hampson RE, Pons TP, Stanford TR, Deadwyler SA. 2004. Categorization in the monkey hippocampus: a possible mechanism for encoding information into memory. Proc Natl Acad Sci USA 101:3184–3189.

Hasselmo ME. 2007. Arc length coding by interference of theta frequency oscillations may underlie context-dependent hippocampal unit data and episodic memory function. Learn Mem 14:782–794.

Hasselmo ME, Wyble BP. 1997. Free recall and recognition in a network model of the hippocampus: Simulating effects of scopolamine on human memory function. Behav Brain Res 89:1–34.

Heit G, Smith ME, Halgren E. 1988. Neural encoding of individual words and faces by the human hippocampus and amygdala. Nature 333:773–775.

Howard MW, Fotedar MS, Datey AV, Hasselmo ME. 2005. The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. Psychol Rev 112:75–116.

Howard MW, Jing B, Rao VA, Provyn JP, Datey AV(2009a) Bridging the gap: Transitive associations between items presented in similar temporal contexts. J Exp Psychol 35:391–407.

Howard MW, Kahana MJ. 2002. A distributed representation of temporal context. J Math Psychol 46:269–299.

Howard MW, Sederberg PB, Kahana MJ. 2009b. Reply to Farrell and Lewandowsky: Recency-contiguity interactions predicted by TCM. Psychon Bull Rev 16:973–984.

Howard MW, Shankar KH, Jagadisan UKK. 2011. Constructing semantic representations from a gradually-changing representation of temporal context. Top Cogn Sci 3:48–73.

Howard MW, Youker TE, Venkatadass V. 2008. The persistence of memory: Contiguity effects across several minutes. Psychon Bull Rev 15:58–63.

Jenkins LJ, Ranganath C. 2010. Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. J Neurosci 30:15558–15565.

Jensen O, Lisman JE. 2005. Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. Trends Neurosci 28:67–72.

Kahana MJ. 1996. Associative retrieval processes in free recall. Memory Cognit 24:103–109.

Kahana MJ, Howard M, Polyn S. 2008. Associative processes in episodic memory. In Roediger III HL, editor. Cognitive Psychology of Memory, Vol. 2. Learning and Memory—A Comprehensive Reference (Byrne J, Editor). Elsevier, Oxford. pp 476–490.

Kinsbourne M, Wood F. 1975. Short-term memory and the amnesic syndrome. In: Deutsch D, Deutsch JA, editors. Short Term Memory. New York: Academic Press. pp 257–291.

Koene RA, Hasselmo ME. 2007. First-in-first-out item replacement in a model of short-term memory based on persistent spiking. Cereb Cortex 17:1766–1781.

Lisman JE. 1999. Relating hippocampal circuitry to function: Recall of memory sequences by reciprocal dentate-CA3 interactions. Neuron 22:233–242.

MacDonald CJ, Lepage KQ, Eden UT, Eichenbaum H. 2011. Hippocampal "time cells" bridge the gap in memory for discontiguous events. Neuron 71:737–749.

Manning JR, Polyn SM, Litt B, Baltuch G, Kahana MJ. 2011. Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. Proc Natl Acad Sci USA 108:12893–12897.

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

Marr D. 1971. Simple memory: A theory for archicortex. Philos Trans R Soc B 262:23–81.

McAndrews M, Milner B. 1991. The frontal cortex and memory for temporal order. Neuropsychologia 29:849–859.

McNaughton BL, Morris RG. 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. Trends Neurosci 10:408–415.

Mehta MR, Quirk MC, Wilson MA. 2000. Experience-dependent asymmetric shape of hippocampal receptive fields. Neuron 25:707–715.

Mensink GJM, Raaijmakers JGW. 1988. A model for interference and forgetting. Psychol Rev 95:434–455.

Miyashita Y. 1988. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. Nature 335:817–820.

Moreton BJ, Ward G. 2010. Time scale similarity and long-term memory for autobiographical events. Psychon Bull Rev 17:510–515.

Murdock BB. 1962. The serial position effect of free recall. J Exp Psychol 64:482–488.

Murdock BB. 1997. Context and mediators in a theory of distributed associative memory (TODAM2). Psychol Rev 104:839–862.

Navaroli VL, Zhao Y, Boguszewski P, Brown TH. Muscarinic receptor activation enables persistent firing in pyramidal neurons from superficial layers of dorsal perirhinal cortex. Hippocampus (in press).

Naya Y, Suzuki W. 2011. Integrating what and when across the primate medial temporal lobe. Science 333:773–776.

Naya Y, Yoshida M, Miyashita Y. 2001. Backward spreading of memory-retrieval signal in the primate temporal cortex. Science 291:661–664.

Norman KA, O'Reilly RC. 2003. Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. Psychol Rev 110:611–646.

O'Keefe J, Nadel L. 1978. The Hippocampus as a Cognitive Map. New York: Oxford University Press.

Pastalkova E, Itskov V, Amarasingham A, Buzsaki G. 2008. Internally generated cell assembly sequences in the rat hippocampus. Science 321:1322–1327.

Paz R, Gelbard-Sagiv H, Mukamel R, Harel M, Malach R, Fried I. 2010. A neural substrate in the human hippocampus for linking successive events. Proc Natl Acad Sci USA 107:6046–6051.

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

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

Quiroga RQ, Nadasdy Z, Ben-Shaul Y. 2004. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Computation 16:1661–1687.

Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I. 2005. Invariant visual representation by single neurons in the human brain. Nature 435:1102–1107.

Raaijmakers JGW, Shiffrin RM. 1980. SAM: A theory of probabilistic search of associative memory. In: Bower GH, editor. The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 14. New York: Academic Press. pp207–262.

Rao VA, Howard MW. 2008. Retrieved context and the discovery of semantic structure. In: Platt J, Koller D, Singer Y, Roweis S, editors, Advances in Neural Information Processing Systems 20. Cambridge, MA: MIT Press. pp1193–1200.

Raskin E, Cook SW. 1937. The strength and direction of associations formed in the learning of nonsense syllables. J Exp Psychol 20:381–395.

Ratcliff R, Murdock BB. 1976. Retrieval processes in recognition memory. Psychol Rev 83:190–214.

Rolls ET, Kesner RP. 2006. A computational theory of hippocampal function, and empirical tests of the theory. Prog Neurobiol 79:1–48.

Schwartz G, Howard MW, Jing B, Kahana MJ. 2005. Shadows of the past: Temporal retrieval effects in recognition memory. Psychol Sci 16:898–904.

Sederberg PB, Gershman SJ, Polyn SM, Norman KA. 2011. Human memory reconsolidation can be explained using the temporal context model. Psychon Bull Rev 18:455–468.

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

Sederberg PB, Miller JF, Howard MW, Kahana MJ. 2010. The temporal contiguity effect predicts episodic memory performance. Mem Cognit 38:689–699.

Shankar KH, Howard MW. 2012. A scale-invariant representation of time. Neural Comput 24:134–193.

Sheffield ME, Best TK, Mensh BD, Kath WL, Spruston N. 2011. Slow integration leads to persistent action potential firing in distal axons of coupled interneurons. Nat Neurosci 14:200–207.

Shiffrin RM, Steyvers M. 1997. A model for recognition memory: REM—Retrieving effectively from memory. Psychon Bull Rev 4:145–166.

Shimamura A, Janowsky J, Squire L. 1990. Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. Neuropsychologia 28:803–813.

Siapas AG, Lubenov EV, Wilson MA. 2005. Prefrontal phase locking to hippocampal theta oscillations. Neuron 46:141–151.

Sidiropoulou K, Lu FM, Fowler MA, Xiao R, Phillips C, Ozkan ED, Zhu MX, White FJ, Cooper DC. 2009. Dopamine modulates an mGluR5-mediated depolarization underlying prefrontal persistent activity. Nat Neurosci 12:190–199.

Simons JS, Spiers HJ. 2003. Prefrontal and medial temporal lobe interactions in long-term memory. Nat Rev Neurosci 4:637–648.

Sirotin YB, Kimball DR, Kahana MJ. 2005. Going beyond a single list: Modeling the effects of prior experience on episodic free recall. Psychon Bull Rev 12:787–805.

Tahvildari B, Fransén E, Alonso AA, Hasselmo ME. 2007. Switching between "On" and "Off" states of persistent activity in lateral entorhinal layer III neurons. Hippocampus 17:257–263.

Tulving E, Madigan SA. 1970. Memory and verbal learning. Annu Rev Psychol 21:437–484.

Unsworth N. 2008. Exploring the retrieval dynamics of delayed and final free recall: Further evidence for temporal-contextual search. J Mem Lang 59:223–236.

Viskontas IV, Knowlton BJ, Steinmetz PN, Fried I. 2006. Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. J Cognit Neurosci 18:1654–1662.

Viskontas IV, Quiroga RQ, Fried I. 2009. Human medial temporal lobe neurons respond preferentially to personally relevant images. Proc Natl Acad Sci USA 106:21329–21334.

Wang X. 2001. Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci 24:455–463.

Wood ER, Dudchenko PA, Eichenbaum H. 1999. The global record of memory in hippocampal neuronal activity. Nature 397:613–616.

Yoshida M, Fransén E, Hasselmo ME. 2008. mGluR-dependent persistent firing in entorhinal cortex layer III neurons. Eur J Neurosci 28:1116–1126.

Yoshida M, Hasselmo ME. 2009. Persistent firing supported by an intrinsic cellular mechanism in a component of the head direction system. J Neurosci 29:4945–4952.

Zhang Z, Séguéla P. 2010. Metabotropic induction of persistent activity in layers II/III of anterior cingulate cortex. Cereb Cortex 20:2948–2957.

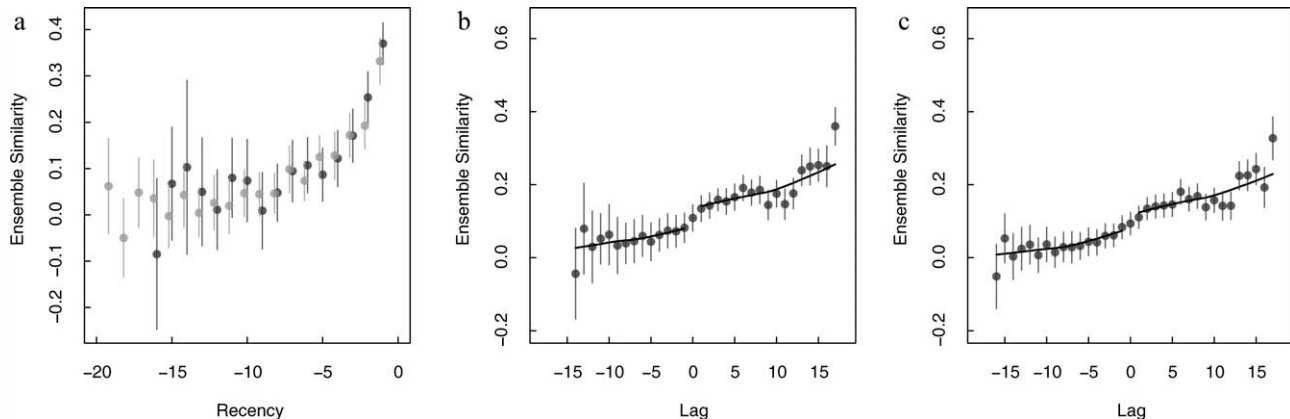## APPENDIX: ADDITIONAL EXAMPLES OF TEMPORALLY-AUTOCORRELATED UNITS AND INTERMEDIATE STAGES OF THE CONTIGUITY ANALYSIS

FIGURE A1. Intermediate stages of the contiguity analysis shown in Figure 4. a. The recency gradient (an analog of Fig. 4a) calculated separately according to whether the second event in the pair corresponded to a stimulus presented for the first time (dark) or a repeated stimulus (light). Note the slight change of scale compared to Figure 4a. The location on the *x*-axis of the lighter points is slightly offset for clarity. b. The estimate of the contiguity effect (Fig. 4b) that results from using the recency gradient computed when the second item is presented for the first time. Figure 4c is computed by taking the difference between the observed ensemble similarity (Fig. 4b) and this result. c. Same as b, but computed with the recency gradient when the second event corresponds to a repeated stimulus (light line in a). The choice of whether to use the analysis in b or c did not have a dramatic effect on the results of the contiguity analysis. The slight shoulder around zero in b and c is a consequence of the non-uniform sampling of recencies.
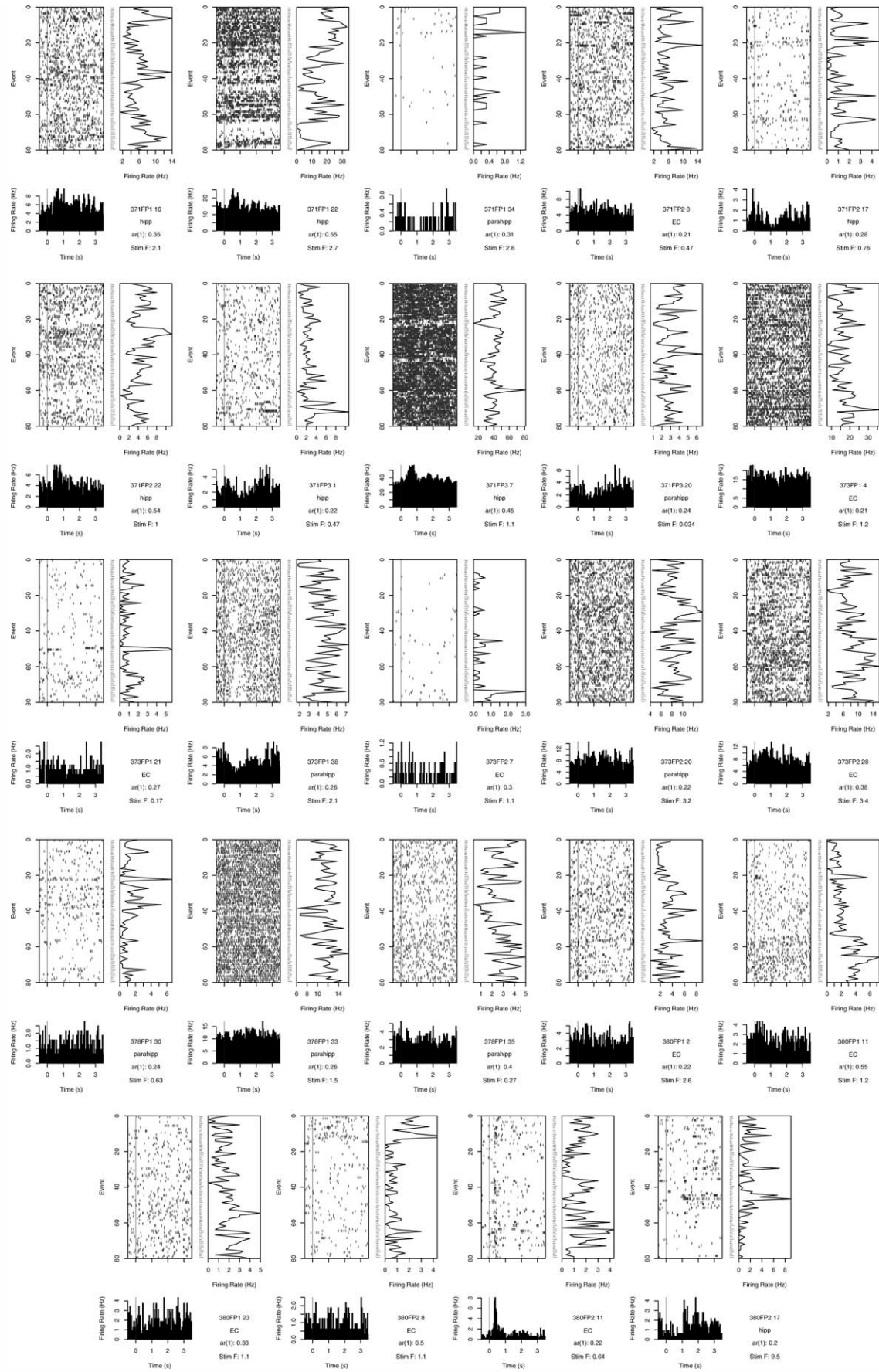
FIGURE A2. Additional examples of autocorrelated units. Format follows Figure 2b. Unit numbers, brain regions and various statistics can be found in each panel.