Neuropsychologia 143 (2020) 107471

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: http://www.elsevier.com/locate/neuropsychologia

Neural fatigue influences memory encoding in the human hippocampus

Lynn J. Lohnas [a,d,*], Lila Davachi [a,b,e], Michael J. Kahana [c]

[a] Department of Psychology, USA
[b] Center for Neural Science, New York University, 6 Washington Place, 8th Floor, New York, NY, 10003, USA
[c] Department of Psychology, University of Pennsylvania, 425 South University Avenue, Philadelphia, PA, 19104, USA
[d] Present Address: Department of Psychology, Syracuse University, 430 Huntington Hall, Syracuse, NY, 13244, USA
[e] Present Address: Department of Psychology, Columbia University, 371 Schermerhorn Hall, 1190 Amsterdam Avenue, New York, NY, 10027, USA

ARTICLE INFO

Keywords:
Memory
Episodic memory
iEEG
Temporal lobe
Encoding

ABSTRACT

Here we examine the variability underlying successful memory encoding. Successful encoding of successive study items may fatigue encoding resources, thus decreasing the ability to encode subsequent items (Tulving and Rosenbaum, 2006); alternatively, successful encoding may be persistent, leading to more successful encoding (Kahana, Aggarwal, and Phan, 2018). Analyzing intracranial electroencephalographic activity while subjects studied lists of words for subsequent free recall, we examined high-frequency activity (HFA) in hippocampus and dorsolateral prefrontal cortex (DLPFC), as HFA was greater for subsequently recalled than non-recalled items in these regions. We compared non-recalled items with good encoding history (i.e. one of the two preceding items was recalled) with non-recalled items with poor encoding history (i.e. neither prior item was recalled). In the hippocampus, good encoding history led to reduced HFA, whereas in the DLPFC, good encoding history led to enhanced HFA. Hippocampal findings appear consistent with the neural fatigue hypothesis, whereas the DLPFC results appear consistent with persistent encoding states.

1. Introduction

The ability to measure physiological activity in the human brain as people study and subsequently attempt to retrieve memoranda has uncovered a diverse network of regions whose activity predicts encoding or retrieval success. To derive biomarkers of successful encoding, researchers have compared physiological activity recorded during the study of items that are subsequently remembered to activity recorded during the study of items that are subsequently forgotten. This subsequent memory analysis has revealed increased hemodynamic and electrophysiological activity in a core network of brain regions, including the hippocampus and medial temporal lobe structures as well as in dorsolateral prefrontal cortex (DLPFC; Brewer et al., 1998; Davachi, 2006; Diana et al., 2007; Hanslmayr and Staudigl, 2013; Kim, 2011; Paller and Wagner, 2002; Rugg et al., 2012; Sederberg et al., 2007; Wagner et al., 1998). These signals are thought to reflect variability in the goodness of memory encoding arising not only from item properties but also from endogenous variation in the neurocognitive processes underlying successful memory storage (Kahana et al., 2018).

Variability is a ubiquitous feature of any complex dynamical system. However, the nature of this variability and the mechanisms that give rise to it can be characterized and explained, and such a characterization would provide further insight into the underlying system. For example, variability could be described by a stochastic process, with goodness of memory encoding starting at some value and rising or falling unpredictably over the course of a list. Such a process may by described by its persistence, with goodness of memory processes either drifting slowly around some mean value (i.e., having a high autocorrelation), or jumping unpredictably from item to item. In this formulation, the rise or fall in memory efficiency for a given item fluctuates randomly from item to item. Although the fluctuations do not depend upon the recent history of encoding efficiency, the encoding state of a given item is highly correlated with the encoding state of the preceding item.

An alternative account of variability in encoding goodness comes from the idea of "neural fatigue". According to this view, the neural processes that support effective encoding cannot be sustained indefinitely but rather depend on a latent resource that depletes during sustained periods of high activity and replenishes over periods of low activity. Tulving and Rosenbaum (2006) advanced this idea as an explanation for the well-known "law of primacy" which describes better memory for the initial items experienced within a given context (Murdock, 1962; Spurgeon et al., 2014). Neural fatigue predicts that the

* Corresponding author. Department of Psychology 430 Huntington Hall Syracuse, NY, 13244, USA.
E-mail address: ljlohnas@syr.edu (L.J. Lohnas).

https://doi.org/10.1016/j.neuropsychologia.2020.107471
Received 9 July 2019; Received in revised form 19 March 2020; Accepted 16 April 2020
Available online 22 April 2020
0028-3932/© 2020 Elsevier Ltd. All rights reserved.

encoding of a given item should be influenced by the past history of successful or unsuccessful encoding. A biomarker of the neural activity that is required to sustain good encoding should thus decline following a period of sustained successful encoding. Behaviorally, this may also lead to items being less likely to be successfully encoded.

In this study we attempt to distinguish between neural correlates of encoding goodness that are transitory, as predicted by the standard autoregressive model, and those that specifically reflect changes associated with the recent past history of encoding success, as predicted by a neural fatigue model.

2. Materials and methods

2.1. Subjects

We present novel analyses of intracranial electrophysiological recordings taken from patients undergoing invasive monitoring as treatment for drug-resistant epilepsy. Recordings from subdural grids and intraparenchymal depth electrodes were taken from patients who volunteered to participate in memory studies during their 1–3 week hospitalization. The clinical team determined the placement of electrodes entirely for purposes of seizure localization. Both the electrophysiological and the behavioral data were collected in a multi-center study from 2000 to 2017. Although portions of this dataset have been reported on previously (e.g., Burke et al., 2014; Long et al., 2017; Merkow et al., 2015), all of the analyses and results described here are novel. Data used in this report may be freely obtained from the cognitive electrophysiology portal at the University of Pennsylvania (http://memory.psych.upenn.edu/Electrophysiological_Data).

For the purpose of the present study we selected patients who had electrodes in at least one of our two regions of interest (ROIs): hippocampus and DLPFC. The clinical team determined the placement of electrodes entirely for purposes of seizure localization. In total, data were included from 223 subjects, each of whom contributed 1–8 testing sessions (96 contributed 1 session). Of these subjects, 131 had electrodes in hippocampus and 163 had electrodes in DLPFC. Our research protocol was approved by the institutional review boards at the University of Pennsylvania and our collaborating hospitals, and was carried out in accordance with the provisions of the World Medical Association Declaration of Helsinki. Informed consent was obtained from patients (or their guardians, in the case of teenage subjects). Our motivation for the sample size was simply to use every subject available from this multicenter study, producing large sample sizes for both subjects with hippocampal electrodes (N = 131) and DLPFC electrodes (N = 163).

2.2. Procedure

Subjects studied lists of common nouns chosen at random and without replacement from a pool of high-frequency words. Fourteen subjects received 20-item lists; 145 received 12-items lists; 64 received 15-item lists. Following an orienting signal the computer displayed each list item. For subjects with 12-item lists, the orienting signal was watching the screen for 10s as centrally-placed numbers counted down from 10, 9, 8, …to 1; for other subjects the orienting signal was a + sign on the screen for 1600 ms followed by an 800–1200 ms blank interstimulus interval (ISI). Following the orienting signal, each list item was displayed in capital letters for 1600 ms, followed by an 800–1200 ms blank ISI. The variation in the duration of the ISI served to decorrelate the physiological responses from successive word presentations (Sederberg et al., 2007). Following the final list item, subjects were shown a series of arithmetic problems of the form A + B + C = ? where A, B, and C were randomly chosen digits in the set {1, …,9}. Subjects responded by typing the answer on a computer keyboard. Immediate feedback was given in the form of a high-pitched tone for correct entries and a low-pitched tone for incorrect answers. After performing this distractor task for 20 s a row of asterisks accompanied by a 300 ms tone signaled the start of the recall period. Responses to the arithmetic problems were self-paced, and so a subject may have been in the middle of a problem when the distractor task ended. Following the distractor task, subjects were given a fixed amount of time to recall items aloud from the current list in any order: subjects with 12-items lists were given 30s for recall; all others were given 45s for recall. Vocal responses were digitally recorded and scored for analysis following each session (Solway et al., 2010).

Subjects completed 6–25 lists per session. After excluding lists with no recalls (see Data Analysis), one session had 5 included lists, and 90% of included sessions had at least 14 lists.

2.3. Electrophysiology

Intracranial electroencephalography (EEG) activity was recorded from subdural arrays (grids of 3 mm diameter contacts spaced 1 cm apart) or depth probes with 1 mm collar electrodes spaced 8 mm apart. EEG recordings were sampled at 256 – 1 KHz depending on the clinical center. Physiological and behavioral measures were synchronized using optically isolated pulses received from the testing computer on an additional recording channel. Signals were converted to a bipolar montage by differencing the signals between each pair of immediately adjacent electrodes on grids and depth electrodes (Burke et al., 2013). The resulting bipolar signals were treated as new virtual electrodes and are referred to as such in the remainder of the text.

Images with the electrode placements on individual patient brains were created by coregistering a computed tomography (CT) scan with a preoperative magnetic resonance image (MRI). The electrodes were manually identified using the postoperative CT scans. These images were then normalized to a standardized brain in MNI space (Maldjian et al., 2003). The MNI coordinates were then transformed to Tailarach space (Lancaster et al., 2000), and Tailarach coordinates were used to determine the side of each electrode as well as the Brodmann areas of cortical electrodes. Cortical electrodes in Brodmann areas 9 and 46 defined the DLPFC region of interest. For depth electrodes implanted in the temporal lobe, an experienced clinician reviewing CT scans and MRIs labeled anatomic locations.

Electrodes were positioned by clinical teams to identify seizure foci and functional brain regions to guide potential resective surgery. As a result, most electrodes were usually placed in temporal cortex, but many electrodes were also placed in the hippocampi and frontal, occipital, and parietal cortices. Because the clinical procedure of identifying seizure foci entails placing electrodes in any region that is potentially epileptogenic, the majority of recordings come from brain regions outside the area that is eventually determined to be involved in seizures (Jacobs and Kahana, 2010).

2.4. Data analysis

To eliminate electrical line and equipment noise, data were notch-filtered on-line at 50 or 60 Hz with a Butterworth filter with zero phase distortion. Power was calculated during each word presentation (0–1600 ms post-onset) with a 1000 ms buffer on either side. The Morlet wavelet transform with a wave number of 6 was used to compute spectral power as a function of time. Following previous studies, we defined wavelets at 46 log-spaced frequencies between 2 and 100 Hz, but here only report results for HFA, 44–100 Hz (Long et al., 2014; Long and Kahana, 2015). After calculating power values at each electrode and frequency value, we then log transformed the power values. Next, we Z-scored the values across all events and mean time within a session, separately at each frequency and electrode. Finally, we took the mean across time points and frequency values contained within the HFA range.

To ensure reliable statistics from each subject, we required that each subject contribute at least 10 observations per session for each of the four behavioral conditions of interest: good encoding versus poor

encoding (i.e. whether an item was successfully recalled), each considered with good versus poor encoding history (i.e. whether at least one of the prior two items was recalled). In addition, any lists with no remembered items were excluded from analysis. These criteria led to the exclusion of 19 subjects. Requiring a minimum number of observations per session also led to the exclusion of individual sessions for individual subjects: 33 subjects had 1 session excluded, 6 had 2 sessions excluded, 3 had 3 sessions excluded, and 1 subject had 6 sessions excluded. In the behavioral analysis, we include the same sessions that were included for subjects, but nonetheless include recall from lists with no recalls, to highlight that memory performance was not at floor. Observations were collapsed across session for each subject, such that the number of observations per subject per condition were: good history, poor encoding: 20–564 (mean 130); poor history, poor encoding: 19–1197 (mean 210); good encoding: 21–760 (mean 141); poor encoding: 39–1734 (mean 340).

To determine significance for a pairwise comparison, for each subject and electrode, a *t*-statistic was generated through an unpaired *t*-test comparing the Z-scored HFA power between the two conditions. These t-statistics were averaged across electrodes within an ROI, creating a single t-statistic for each subject and ROI. The distribution of subject t-statistics was compared to zero using an unpaired *t*-test. To determine significance between pairwise comparisons, we extracted the distribution of t-statistics from each comparison, then compared them using a paired *t*-test.

For the final analysis the interaction of encoding history with serial position, we quantified the primacy effect in the aggregate data by comparing the change in recall performance across successive serial positions. This analysis showed that recall performance significantly decreases with each successive pair of SPs ($t(222)>3$, $p < .001$) until positions 5 vs. 6 ($t(222) = 1.15$, $p > .2$). Thus, we compared HFA collapsed across early list positions 1–4 to HFA collapsed across the same number of mid-list positions, i.e. 5-8. We included sessions with at least 10 items in each considered condition. This analysis included a subset of the primary analyses: 126 subjects with DLPFC electrodes and 109 subjects with hippocampal electrodes.

3. Results

According to a standard view of variability in encoding efficacy, periods of good encoding will tend to be persistent, with encoding states that lead to successful subsequent recall tending to be followed by further good encoding states. In contrast, according to a neural fatigue account, extended periods of good encoding will tend to deplete neural resources thus making subsequent epochs more likely to be poorly encoded. To evaluate these hypotheses, we examined HFA in the local-field potential across two regions of interest within the core verbal memory network: hippocampus and DLPFC. We chose to examine HFA because of its strong correlation with the firing rates of individual neurons in both human and non-human animals (Hirabayashi et al., 2014; Manning et al., 2009). Furthermore, previous research has established hippocampus and prefrontal cortical HFA as biomarkers of

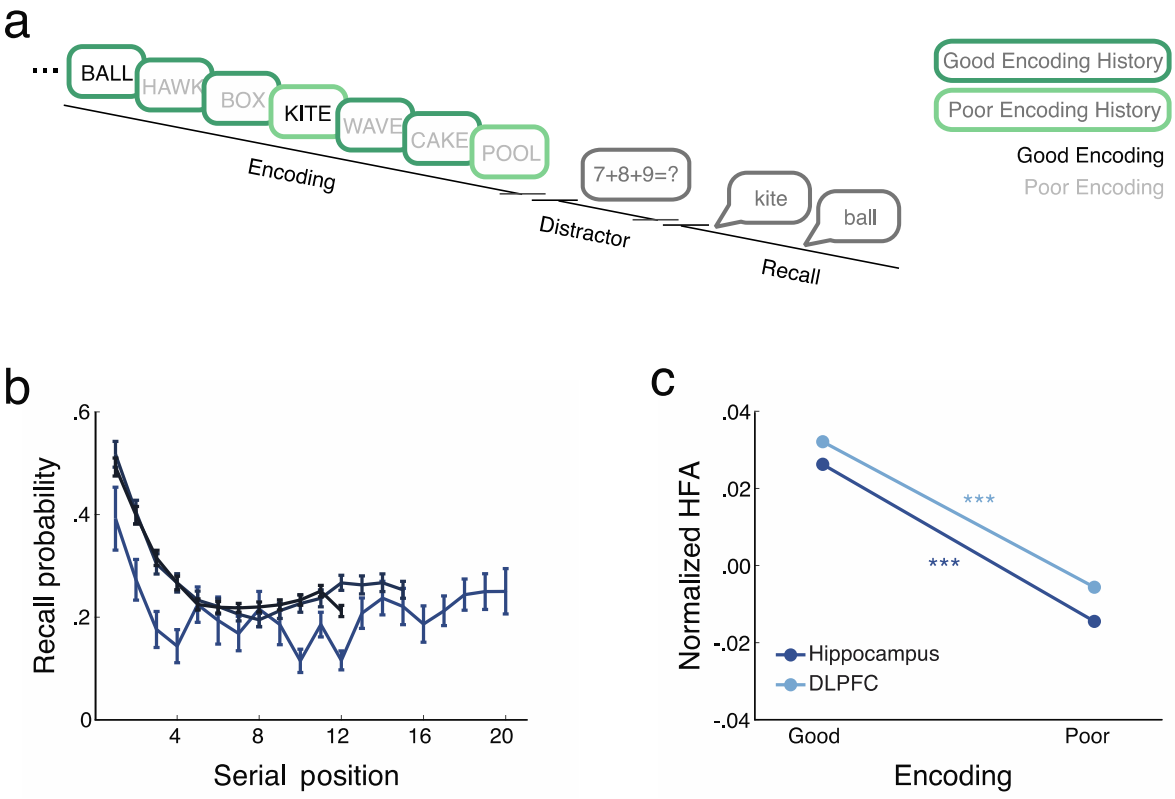

Fig. 1. Experiment design, behavioral performance, and high frequency activity (HFA) as a biomarker of good encoding. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

a. Subjects studied lists of common nouns which they subsequently freely recalled after a distractor-filled delay interval. Each item was retroactively labeled according to two variables: 1) successful encoding, based on whether the item was subsequently recalled (black for good and light gray for poor); 2) successful recent encoding history, based on whether either of the two preceding items were recalled (dark green for good encoding history and light green for poor encoding history).
b. Probability of recall as a function of serial position, partitioned on the basis of list-length: 12 (N = 145), 15 (N = 64), 20 (N = 14). Error bars indicate ±1 SEM.
c. HFA in hippocampus and dorsolateral prefrontal cortex (DLPFC). HFA is a biomarker of good encoding, being significantly greater for items that were subsequently recalled (good encoding) in comparison to items that were subsequently not recalled (poor encoding). Asterisks indicate significance between groups, with the color corresponding to the region referenced in the figure legend (***$p < .001$).

good memory encoding in humans (Long et al., 2014; Long and Kahana, 2015).

Here we compared HFA as subjects encoded lists of items for subsequent free recall (Fig. 1a and b). We found significantly higher HFA during encoding of items that were subsequently recalled than during encoding of items that were not recalled (Fig. 1c), both in the hippocampus ($t(130) = 4.74$, $p < .001$) and in DLPFC ($t(162) = 4.10$, $p < .001$). We thus interpret greater HFA in these regions as a biomarker of good memory encoding.

Having established HFA in hippocampus and DLPFC as biomarkers of good memory encoding, we next sought to characterize how these biomarkers vary over time during encoding. For each item we defined two variables: whether it was subsequently recalled and whether at least one of the prior two words was subsequently recalled (Fig. 1a). The former measure is a surrogate for successful current encoding, while the latter measure is a surrogate for successful encoding history. To condition memory encoding on a *history* of prior successful encoding one would ideally like to use a recency-weighted measure of encoding success, such that more recent encoding success indicates greater neural fatigue. For instance, one could argue that the neural resources required to encode a given item *i* would be more depleted if only item *i*-1 was encoded successfully than if only item *i*-2 was encoded successfully. However, under the assumption that the encoding success of both of the prior items should influence encoding of the current item, combined with short lists of items and discrete observations, a multi-item window approach is more parsimonious. That is, although a larger time window of encoding history (e.g. considering the history of the prior 3 items) confers the advantage of more precisely estimating encoding goodness by averaging over more items, it is more challenging to argue that a more distantly encoded item (e.g. *i*-3) would influence encoding success of the current item. A smaller window (e.g. considering just the prior 1 item) restricts the focus on the most relevant epoch just preceding the target item. Thus, we define encoding history based on encoding success of the prior two items.

Having established measures of successful encoding and successful encoding history, these two variables enable us to distinguish between the autocorrelated and the neural fatigue accounts of encoding success. The two aforementioned accounts make differential predictions regarding items with good encoding history yet poor encoding: According to the neural fatigue account, this reflects a depletion in neural resources, thus hindering the encoding of the current item; according to the autocorrelated account, this situation arises due to fluctuations in good or poor encoding states. Thus, the neural fatigue account predicts that a good encoding history should reduce HFA for non-recalled items. In contrast, the autocorrelated account predicts that HFA for non-recalled items should be greater with a good encoding history, as these items with good encoding should have higher HFA. Fig. 2 presents representative signals recorded during individual word lists. In these example lists, filled markers denote subsequently remembered items whereas open markers denote subsequently forgotten items; larger markers denote good encoding history whereas smaller markers denote poor encoding history. In the hippocampal electrode examples, when there is a transition from a good to poor encoding state (such that an item has poor current encoding but is preceded by a good encoding history), this coincides with a sharp decrease in HFA (green arrows). This is consistent with a neural fatigue account, whereby a good encoding state, reflected both by successful subsequent recall and greater HFA, cannot be maintained for extended periods of time. In contrast, fluctuations in DLPFC HFA from good to poor encoding states are not as drastic, and thus are more consistent with an autocorrelated process.

Going beyond single examples, we next quantified the extent of the changes in HFA for non-recalled items as a function of their encoding history. In the hippocampus, HFA for non-recalled items was significantly lower for items with a good than a poor encoding history (Fig. 3; $t(130) = 3.18$, $p = .002$). This is consistent with the neural fatigue hypothesis which predicts that a recent history of successful encoding should deplete cognitive resources, thus leading the current item to be in a poor encoding state. As such, the item will not be recalled, and hippocampal HFA will be lower. In contrast, HFA in DLPFC trended towards being significantly greater for good vs. poor encoding history ($t(162) = 1.90$, $p = .059$). Although we cannot consider this result definitive, it aligns most closely with the autocorrelated account of encoding variability. If the autocorrelation was strong, then the prior encoding state should persist into the current encoding state. As a result, even if an item were not recalled, it would nonetheless exhibit greater HFA if it were preceded by a good encoding history than a poor encoding history. In the case of weaker autocorrelation, where successful memory encoding is highly variable from item to item, encoding history may have little influence on HFA of the current item, and thus HFA would not differ with encoding history.

To further query HFA differences in encoding history between hippocampus and DLPFC, we asked whether there was an interaction between these regions. As a conservative estimate of this interaction, we

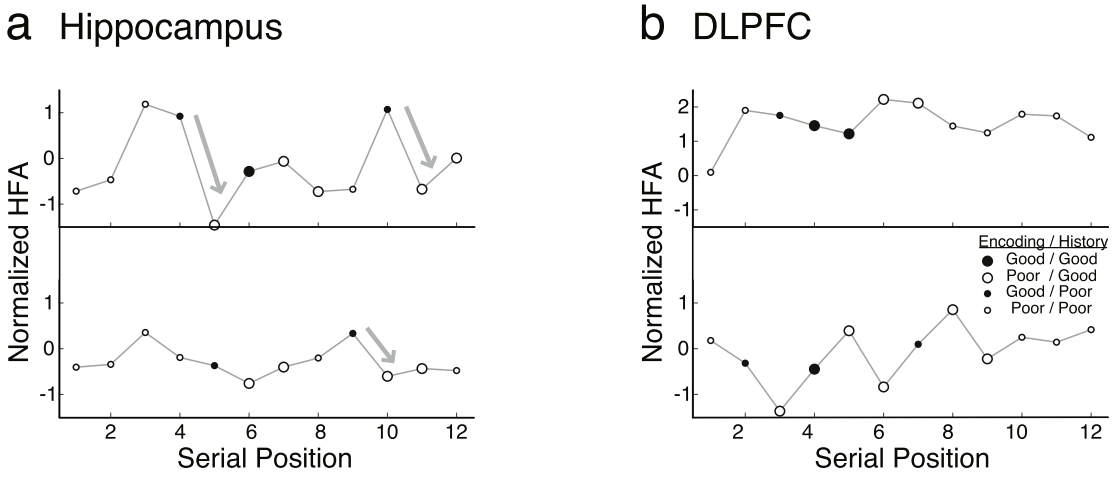

Fig. 2. Modulations of high-frequency activity (HFA) by encoding state. a. Examples of HFA in single hippocampal electrodes exhibiting activity consistent with a neural fatigue process in a single list. The top panel is from an electrode from left hippocampus, and the bottom panel is from an electrode in right hippocampus. b. Examples of HFA in a single DLPFC electrodes exhibiting activity consistent with an autocorrelated process in a single list. The top panel is an electrode from right Brodmann Area (BA) 9, and the bottom panel is an electrode from left BA9. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

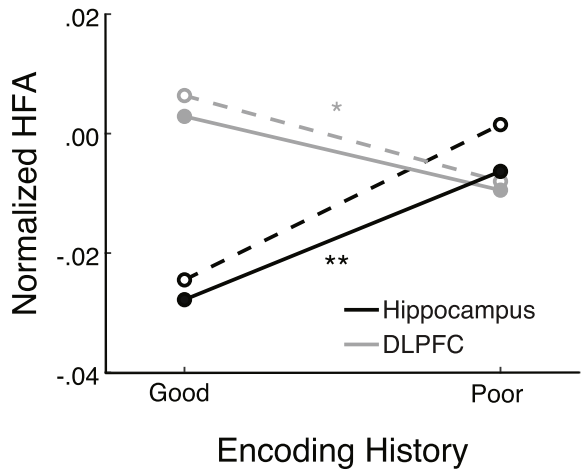

Fig. 3. Modulations of high-frequency activity (HFA) by encoding history for nonrecalled items. In the hippocampus, HFA is significantly greater for items with poor encoding history than good encoding history (N = 131). The dorsolateral prefrontal cortex (DLPFC) shows a trend toward greater HFA for items with good than poor encoding history (N = 163). Solid lines and filled circles indicate values across all subjects who had electrodes in either region. Dashed lines and open circles indicate values in the same regions when calculated for the subset of subjects who had electrodes in both regions (N = 71). Asterisks indicate significance between conditions for each brain region, for all subjects, with the color corresponding to the region referenced in the figure legend (*p < .06, **p < .005). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

only considered those subjects who had electrodes in both hippocampus and DLPFC. In this subgroup ($N = 71$), we observed main effects of region with greater hippocampal HFA for items with a poor than good encoding history ($t(70) = 3.34$, $p = .001$), and in DLPFC there was a trend towards greater HFA for items with a good encoding history ($t(70) = 1.52$, $p = .13$). We also found a significant interaction between regions ($t(70) = 3.31$, $p = .001$), due to greater HFA for nonrecalled items with a poor encoding history in hippocampus but not in DLPFC. These results indicate that neural activity in the human hippocampus aligns most closely with the predictions of the neural fatigue account, whereas neural activity in the human DLPFC aligns most closely with predictions of the autocorrelated encoding account.

The neural fatigue account helps to explain why encoding may be unsuccessful, and in particular why encoding is generally worse following the first few items in a new context, termed the "law of primacy" or primacy effect (Murdock, 1962; Spurgeon et al., 2014). In the current study of free recall, the primacy effect manifests as greater recall probability for items in early serial positions (Fig. 1b). Although the above HFA analyses collapse across serial positions, we examined HFA differences in encoding history as a function of serial position as recommended by one of our paper's referees. Prior work has revealed neural correlates of primacy, with several brain regions exhibiting greater HFA for early serial positions irrespective of encoding success (Sederberg et al., 2006; Serruya et al., 2014). Thus, we examined whether there was an interaction of encoding history with serial position; in other words, we examined whether the differences in HFA by encoding history for nonrecalled items was greater for early serial positions (1–4) versus mid-list serial positions (5–8). We did not find a significant interaction in either brain region (hippocampus: $t(108) = 1.25$, $p > .2$; DLPFC: $t(125) = 0.63$, $p > .2$) and thus tentatively conclude that our reported differences in HFA appear consistent throughout the encoding period. However, future work remains to characterize the state of HFA for early list items, a point we consider further in the Discussion.

4. Discussion

Our ability to form new memories varies over time, with periods of good encoding being interrupted by periods of poor encoding. By recording neural activity during the learning process we can identify biomarkers of this variability. Research using both non-invasive and invasive measures of human brain activity have identified reliable biomarkers of good memory encoding (Davachi, 2006; Paller and Wagner, 2002; Sederberg et al., 2003; Wagner et al., 1998). Intracranial EEG studies in particular have shown that in both hippocampus and DLPFC, HFA increases during the encoding of words that will be subsequently remembered as compared with those that are forgotten (Long et al., 2014; Long and Kahana, 2015). Although there are several possible mechanisms that could give rise to these dynamics, we sought to test the hypothesis that sustained successful memory encoding will induce fatigue in the core memory network, and that this fatigue will appear as a subsequent drop in the good encoding biomarkers. An alternative to this neural fatigue hypothesis is the idea that goodness of encoding varies stochastically, being equally likely to rise or fall independent of the recent history of memory encoding. This prediction would arise if the biomarker followed an autoregressive process.

To test the neural fatigue and autoregressive hypotheses, we analyzed hippocampal (N = 131) and DLPFC (N = 163) HFA measured as neurosurgical patients studied lists of items for a subsequent recall test. We classified an item as having good encoding history if either one or both of the two preceding items was successfully recalled. We expected to observe higher HFA for recalled than for non-recalled items. The neural fatigue hypothesis predicts that HFA is more likely to decline than it is to rise following a sustained period of good encoding, and this would result in a failure to recall the subsequent item. This implies that non-recalled items with good encoding history would have lower HFA than those with a poor encoding history. Hippocampal HFA dynamics appeared consistent with this neural fatigue hypothesis, whereas HFA in the DLPFC did not align with these predictions, but rather followed the predictions of an autocorrelated goodness=of=encoding account (Fig. 3).

According to Tulving and Rosenbaum's (2006) neural fatigue account, the networks engaged during an item's first presentation fatigue when an item or components of that item are repeated. Thus, the primacy effect observed in free recall and other memory tasks arises from the lack of neural fatigue for the first few list items (Murdock, 1962; Spurgeon et al., 2014). Our primary analyses considered encoding history irrespective of serial position, and we did not find an interaction of encoding history with serial position. This suggests that primacy items alone do not drive encoding history effects. As another approach to query neural activity of early list items, it may also be informative to examine the period immediately preceding list onset. Although during this time, presumably neural encoding resources have not yet been depleted, other differences in the cognitive operations confound this comparison.

The neural fatigue hypothesis also provides an explanation for the build-up and release of proactive interference exhibited in lists of items from the same semantic category (Loess, 1967; Wickens, 1970); items drawn from the same category share many stimulus features, thus imposing similar demands on encoding and fatiguing encoding processes. Here we examined neural fatigue within each list yet not across lists, as the across-list analyses are beyond the scope of the present study and design. However, examining predictions of neural fatigue across time scales would be a valuable future direction.

The neural fatigue hypothesis also provides intuition in the more extreme case in which an item itself is repeated within a list (Tulving, 2008), as the neural resources needed for a repeated item's encoding may have been depleted from the item's first presentation. This explanation is consistent with findings that, following presentation of an item, repeating the item shortly thereafter leads to reduced neural activity in medial temporal regions including hippocampus (Stern et al., 1996;

Kirchhoff et al., 2000; Yassa and Stark, 2011; Lohnas et al., 2018). More broadly, this phenomenon of reduced or suppressed activity for repeated items, termed repetition suppression, has been hypothesized to reflect fatigue of neuronal processes (for a review see Grill-Spector et al., 2006). Relevant to our measure of HFA, Merzagora et al. (2014) examined HFA during the encoding of repeated items in a Sternberg short-term item recognition task. They found that HFA during the second occurrence of an item was markedly attenuated, consistent with our findings that decreases in HFA may reflect a depletion of resources to devote to a particular stimulus.

Although the current study did not involve item or category repetition, all of the items within a list share contextual features representing their overlapping temporal attributes and encoding task context (Manning et al., 2011; Polyn et al., 2009). Further, subjects may automatically segment subsequences of items into meaningful "chunks" or events (Clewett and Davachi, 2017; Farrell, 2012; Zacks et al., 2001). Such structure, whether endogenously created or exogenously imposed, has been shown to modulate both encoding success (Ezzyat and Davachi, 2011; Heusser et al., 2018; Kurby and Zacks, 2008; Speer and Zacks, 2005) as well as hippocampal activity (DuBrow and Davachi, 2014; Ezzyat and Davachi, 2011, 2014). These findings provide support for our interpretation of hippocampal HFA dynamics as reflecting neural fatigue.

Our findings are also consistent with evidence suggesting that, during successful encoding, hippocampal activity is more likely to be in a good attentional state (Aly and Turk-Browne, 2016; Uncapher and Rugg, 2009). In a complementary way, several studies have found that hippocampal activity is associated with guiding attention more efficiently based on previously encoded information (Chun et al., 2011; Goldfarb et al., 2016; Summerfield et al., 2006). Taken together, this suggests that hippocampal HFA not only reflects the current encoding state but also reflects the influence of prior encoding states on the current encoding state.

Here we examined biomarkers of good and poor encoding during a free recall task. In line with studies that define good encoding based on successful subsequent memory performance (Kim, 2011; Paller and Wagner, 2002), we operationalized encoding history as successful memory of the preceding two items. However, the current encoding state may be influenced by other factors of encoding history, such as the duration of encoding history (i.e. the number of prior encoded items) or the time elapsed since that good encoding state took place. We did not wish to make strong assumptions of how these variables influenced the current encoding state, as if these assumptions were false, it would be challenging to discern whether they falsify the neural fatigue hypothesis or simply falsify our operationalization of encoding history. As noted in the Results section, we considered two items to be a reasonable tradeoff between considering a longer history of good encoding in lists with few recalled items and the narrower history considered by only the prior item. However, it will be important for future work to characterize neural fatigue according to different assumptions of encoding history.

Regardless of the assumptions of encoding history, the neural fatigue account assumes that periods of good encoding states fatigue neural resources, leading to worse encoding of an item following a good encoding state. This may lead to the intuition that if a given item is encoded successfully and recalled, then recall of the following item is *less* likely. Such an intuition may seem to contradict the finding in free recall that subjects are *more* likely to recall items successively from nearby serial positions, termed the temporal contiguity effect (Healey et al., 2019; Kahana, 1996). Whereas the contiguity effect relates temporal order at study with temporal order at retrieval, our analysis relates the nature of autocorrelated encoding states with subsequent memory irrespective of an item's position in the recall sequence. Indeed, an item's probability of being recalled may be influenced by prior recalled items (e.g. Lohnas and Kahana, 2014), whereas nonrecalled items do not have these complications. In other words, the temporal contiguity effect can only be characterized using sequences of recalled items, yet our analyses only considered the role of encoding history on nonrecalled items. In this way, we avoided issues of retrieval dynamics, and restricted our comparisons to the conditions where the neural fatigue account makes the strongest predictions: when neural resources are depleted, leading to poor encoding.

Measures of neural activity in the core neural memory network provide a window into the cognitive processes underlying successful encoding. By recording intracranial activity during memory encoding for subsequent free recall, we found that brain regions established as being modulated by encoding success are also modulated by a recent history of encoding success. Although both DLPFC and hippocampus exhibit standard increases in HFA that mark successful encoding, the hippocampus exhibits a unique signature consistent with Tulving's neural fatigue hypothesis. By defining an item's recent encoding history as at least one of the prior two items being recalled, hippocampal HFA of non-recalled items with a good encoding history was significantly lower than HFA for non-recalled items with a poor encoding history. According to the neural fatigue account, sustained periods of successful encoding and greater hippocampal HFA eventually fatigue local neural networks, thus leading to poor encoding and lower hippocampal HFA. These results provide the first neural evidence that the biomarkers of good memory encoding exhibit fatigue following successful encoding of prior items.

5. Declarations of interest

None.

Declaration of competing interest

The authors have no competing interests to declare.

CRediT authorship contribution statement

Lynn J. Lohnas: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Lila Davachi:** Funding acquisition, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - review & editing. **Michael J. Kahana:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

Acknowledgments

We would like to especially thank Marijke Beulen, Joshua Jacobs and Mijail Serruya who carried out important preliminary analyses that informed our work. We also owe a special debt of gratitude to Endel Tulving, who suggested that we pursue this question and provided valuable discussion related to this work. We thank Blackrock Microsystems for providing neural recording and stimulation equipment. This work was supported by the DARPA Restoring Active Memory (RAM) program [Cooperative Agreement N66001-14-2-4032]; and National Institutes of Health [grant numbers MH106266, MH55687]. The views, opinions, and/or findings contained in this material are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

References

Aly, M., Turk-Browne, N.B., 2016. Attention promotes episodic encoding by stabilizing hippocampal representations. Proc. Natl. Acad. Sci. Unit. States Am. 113 (4), E420–E429. https://doi.org/10.1073/pnas.1518931113.

Brewer, J.B., Zhao, Z., Desmond, J.E., Glover, G.H., Gabrieli, J.D., 1998. Making memories: brain activity that predicts how well visual experience will be

remembered. Science 281, 1185–1187. https://doi.org/10.1126/science.281.5380.1185.
Burke, J.F., Zaghloul, K.A., Jacobs, J., Williams, R.B., Sperling, M.R., Sharan, A.D., Kahana, M.J., 2013. Synchronous and asynchronous theta and gamma activity during episodic memory formation. J. Neurosci. 33, 292–304. https://doi.org/10.1523/JNEUROSCI.2057-12.2013.
Burke, J.F., Long, N.M., Zaghloul, K.A., Sharan, A.D., Sperling, M.R., Kahana, M.J., 2014. Human intracranial high-frequency activity maps episodic memory formation in space and time. Neuroimage 85, 834–843. https://doi.org/10.1016/j.neuroimage.2013.06.067.
Chun, M.M., Golomb, J.D., Turk-Browne, N.B., 2011. A taxonomy of external and internal attention. Annu. Rev. Psychol. 62, 73–101. https://doi.org/10.1146/annurev.psych.093008.100427.
Clewett, D., Davachi, L., 2017. The ebb and flow of experience determines the temporal structure of memory. Curr. Opin. Behav. Sci. 17, 186–193. https://doi.org/10.1016/j.cobeha.2017.08.013.
Davachi, L., 2006. Item, context and relational episodic encoding in humans. Curr. Opin. Neurobiol. 16, 693—700. https://doi.org/10.1016/j.conb.2006.10.012.
Diana, R.A., Yonelinas, A.P., Ranganath, C., 2007. Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends Cognit. Sci. 11, 379–386. https://doi.org/10.1016/j.tics.2007.08.001.
DuBrow, S., Davachi, L., 2014. Temporal memory is shaped by encoding stability and intervening item reactivation. J. Neurosci. 34, 13998–14005. https://doi.org/10.1523/JNEUROSCI.2535-14.2014.
Ezzyat, Y., Davachi, L., 2011. What constitutes an episode in episodic memory? Psychol. Sci. 22, 243–252. https://doi.org/10.1177/0956797610393742.
Ezzyat, Y., Davachi, L., 2014. Similarity breeds proximity: pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. Neuron 81, 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042.
Farrell, S., 2012. Temporal clustering and sequencing in short-term memory and episodic memory. Psychol. Rev. 119, 223–271. https://doi.org/10.1037/a0027371.
Goldfarb, E.V., Chun, M.M., Phelps, E.A., 2016. Memory-guided attention: independent contributions of the hippocampus and striatum. Neuron 89, 1–8. https://doi.org/10.1016/j.neuron.2015.12.014.
Grill-Spector, K, Henson, R, Martin, A, 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cognit. Sci. 10 (1), 14–23. https://doi.org/10.1016/j.tics.2005.11.006.
Hanslmayr, S., Staudigl, T., 2013. How brain oscillations form memories – a processing based perspective on oscillatory subsequent memory effects. Neuroimage 85, 648–655. https://doi.org/10.1016/j.neuroimage.2013.05.121.
Healey, M.K., Long, N.M., Kahana, M.J., 2019. Contiguity in episodic memory. Psychonomic Bull. Rev. 26, 699–720. https://doi.org/10.3758/s13423-018-1537-3.
Heusser, A.C., Ezzyat, Y., Shiff, I., Davachi, L., 2018. Perceptual boundaries cause trade-offs between local and boundary processing and across-trial associative binding. J. Exp. Psychol. Learn. Mem. Cognit. 44 (7), 1075–1090. https://doi.org/10.1037/xlm0000503.
Hirabayashi, T., Tamura, K., Takeuchi, D., Takeda, M., Koyano, K., Miyashita, Y., 2014. Distinct neuronal interactions in anterior inferotemporal areas of macaque monkeys during retrieval of object association memory. J. Neurosci. 34, 9377–9388. https://doi.org/10.1523/JNEUROSCI.0600-14.2014.
Jacobs, J., Kahana, M.J., 2010. Direct brain recordings fuel advances in cognitive electrophysiology. Trends Cognit. Sci. 14 (4), 162–171. https://doi.org/10.1016/j.tics.2010.01.005.
Kahana, M.J., 1996. Associative retrieval processes in free recall. Mem. Cognit. 24, 103–109.
Kahana, M.J., Aggarwal, E.V., Phan, T.D., 2018. The variability puzzle in human memory. J. Exp. Psychol. Learn. Mem. Cognit. 44 (12), 1857–1863. https://doi.org/10.1016/j.tics.2010.01.005.
Kim, H., 2011. Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. Neuroimage 54, 2446–2461. https://doi.org/10.1016/j.neuroimage.2010.09.045.
Kirchhoff, B.A., Wagner, A.D., Maril, A., Stern, C.E., 2000. Prefrontal-temporal circuitry for episodic encoding and subsequent memory. J. Neurosci. 20, 6173–6180. https://doi.org/10.1523/JNEUROSCI.20-16-06173.2000.
Kurby, C.A., Zacks, J.M., 2008. Segmentation in the perception and memory of events. Trends Cognit. Sci. 12 https://doi.org/10.1016/j.tics.2007.11.004.
Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T., 2000. Automated Talairach atlas labels for functional brain mapping. Hum. Brain Mapp. 10, 120–131, 10.1002/1097-0193(200007)10:3<120::Aid-Hbm30>3.0.Co;2-8.
Loess, H., 1967. Short-term memory, word class, and sequence of items. J. Exp. Psychol. 74, 556–561. https://doi.org/10.1037/h0024787.
Lohnas, L.J., Kahana, M.J., 2014. Compound cuing in free recall. J. Exp. Psychol. Learn. Mem. Cognit. 12–24. https://doi.org/10.1037/a0033698.
Lohnas, L.J., Duncan, K., Doyle, W.K., Thesen, T., Devinsky, O., Davachi, L., 2018. Time-resolved neural reinstatement and pattern separation during memory decisions in human hippocampus. Proc. Natl. Acad. Sci. Unit. States Am. 115, E7418–E7427. https://doi.org/10.1073/pnas.1717088115.
Long, N.M., Kahana, M.J., 2015. Successful memory formation is driven by contextual encoding in the core memory network. Neuroimage 119, 332–337. https://doi.org/10.1016/j.neuroimage.2015.06.073.
Long, N.M., Burke, J.F., Kahana, M.J., 2014. Subsequent memory effect in intracranial and scalp EEG. Neuroimage 84, 488–494. https://doi.org/10.1016/j.neuroimage.2013.08.052.
Long, N.M., Sperling, M.R., Worrell, G.A., Davis, K.A., Gross, R.E., Lega, B.C., Jobst, B.A., Sheth, S.A., Zaghloul, K., Stein, J.M., Kahana, M.J., 2017. Contextually mediated spontaneous retrieval is specific to the hippocampus. Curr. Biol. 27, 1–6. https://doi.org/10.1016/j.cub.2017.02.054.
Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19, 1233–1239. https://doi.org/10.1016/S1053-8119(03)00169-1.
Manning, J.R., Jacobs, J., Fried, I., Kahana, M.J., 2009. Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. J. Neurosci. 29, 13613–13620. https://doi.org/10.1523/JNEUROSCI.2041-09.2009.
Manning, J.R., Polyn, S.M., Baltuch, G., Litt, B., Kahana, M.J., 2011. Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. Proc. Natl. Acad. Sci. Unit. States Am. 108, 12893–12897. https://doi.org/10.1073/pnas.1015174108. USA.
Merkow, M.B., Burke, J.F., Kahana, M.J., 2015. The human hippocampus contributes to both the recollection and familiarity components of recognition memory. Proc. Natl. Acad. Sci. Unit. States Am. 112, 14378–14383. https://doi.org/10.1073/pnas.1513145112.
Merzagora, A.R., Coffey, T.J., Sperling, M.R., Sharan, A., Litt, B., Baltuch, G., Jacobs, J., 2014. Repeated stimuli elicit diminished high-gamma electrocorticographic responses. Neuroimage 85, 844–852. https://doi.org/10.1016/j.neuroimage.2013.07.006.
Murdock Jr., B.B., 1962. The serial position effect of free recall. J. Exp. Psychol. 64 (5), 482–488, 10.1037.h0045106.
Paller, K.A., Wagner, A.D., 2002. Observing the transformation of experience into memory. Trends Cognit. Sci. 6, 93–102. https://doi.org/10.1016/S1364-6613(00)01845-3.
Polyn, S.M., Norman, K.A., Kahana, M.J., 2009. A context maintenance and retrieval model of organizational processes in free recall. Psychol. Rev. 116, 129–156. https://doi.org/10.1037/a0014420.
Rugg, M.D., Vilberg, K.L., Mattson, J.T., Yu, S.S., Johnson, J.D., Suzuki, M., 2012. Item memory, context memory and the hippocampus: fMRI evidence. Neuropsychologia 50, 3070–3079. https://doi.org/10.1016/j.neuropsychologia.2012.06.004.
Sederberg, P.B., Kahana, M.J., Howard, M.W., Donner, E.J., Madsen, J.R., 2003. Theta and gamma oscillations during encoding predict subsequent recall. J. Neurosci. 23, 10809–10814. https://doi.org/10.1523/JNEUROSCI.23-34-10809.2003.
Sederberg, P.B., Gauthier, L.V., Terushkin, V., Miller, J.F., Barnathan, J.A., Kahana, M.J., 2006. Oscillatory correlates of the primacy effect in episodic memory. Neuroimage 32, 1422–1431. https://doi.org/10.1016/j.neuroimage.2006.04.223.
Sederberg, P.B., Schulze-Bonhage, A., Madsen, J., Bromfield, E.B., McCarthy, D.C., Brandt, A., Kahana, M.J., 2007. Hippocampal and neocortical gamma oscillations predict memory formation in humans. Cerebr. Cortex 17, 1190–1196. https://doi.org/10.1093/cercor/bhl030.
Serruya, M.D., Sederberg, P.B., Kahana, M.J., 2014. Power shifts track serial position and modulate encoding in human episodic memory. Cerebr. Cortex 24, 403–413. https://doi.org/10.1093/cercor/bhs318.
Solway, A., Geller, A.S., Sederberg, P.B., Kahana, M.J., 2010. Pyparse: a semi-automated system for scoring spoken recall data. Behav. Res. Methods 42, 141–147. https://doi.org/10.3758/BRM.42.1.141.
Speer, N.K., Zacks, J.M., 2005. Temporal changes as event boundaries: processing and memory consequences of narrative time shifts. J. Mem. Lang. 53, 125–140. https://doi.org/10.1016/j.jml.2005.02.009.
Spurgeon, J., Ward, G., Matthews, W.J., 2014. Why do participants initiate free recall of short lists of words with the first list item? Toward a general episodic memory explanation. J. Exp. Psychol. Learn. Mem. Cognit. 40 (6), 1551–1567. https://doi.org/10.1037/xlm0000028.
Stern, C.E., Corkin, S., González, R.G., Guimaraes, A.R., Baker, J.R., Jennings, P.J., Carr, C.A., Sugiura, R.M., Vedantham, V., Rosen, B.R., 1996. The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. Proc. Natl. Acad. Sci. Unit. States Am. 93, 8660–8665. https://doi.org/10.1073/pnas.93.16.8660.
Summerfield, J.J., Lepsien, J., Gitelman, D.R., Mesulam, M.M., Nobre, A., 2006. Orienting attention based on long-term memory experience. Neuron 49, 905–916. https://doi.org/10.1016/j.neuron.2006.01.021.
Tulving, E., 2008. On the law of primacy. In: Gluck, M.A., Anderson, J.R., Kosslyn, S.M. (Eds.), Memory and Mind: A Festschrift for Gordon H. Bower. Erlbaum, Mahwah, New Jersey, pp. 31–48.
Tulving, E., Rosenbaum, R.S., 2006. What do explanations of the distinctiveness effect need to explain? In: Hunt, R.R., Worthen, J.B. (Eds.), Distinctiveness and Memory. Oxford University Press, New York, pp. 407–423. https://doi.org/10.1093/acprof:oso/9780195169669.003.0018.
Uncapher, M.R., Rugg, M.D., 2009. Selecting for memory? the influence of selective attention on the mnemonic binding of contextual information. J. Neurosci. 29, 8270–8279. https://doi.org/10.1523/JNEUROSCI.1043-09.2009.
Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B.R., Buckner, R.L., 1998. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. Science 281, 1188–1191. https://doi.org/10.1126/science.281.5380.1188.
Wickens, D.D., 1970. Encoding categories of words: an empirical approach to meaning. Psychol. Rev. 77, 1–15. https://doi.org/10.1037/h0028569.
Yassa, M.A., Stark, C.E.L., 2011. Pattern separation in the hippocampus. Trends Neurosci. 34, 515–525. https://doi.org/10.1016/j.tins.2011.06.006.
Zacks, J.M., Braver, T.S., Sheridan, M.A., Donaldson, D.I., Snyder, A.Z., Ollinger, J.M., Buckner, R.L., Raichle, M.E., 2001. Human brain activity time-locked to perceptual event boundaries. Nature 4, 651–655. https://doi.org/10.1038/88486.