Title: Spontaneous Pre-encoding Activation of Neural Patterns Predicts Memory

Authors: Talya Sadeh[1,4,5*], Janice Chen[2], Yonatan Goshen-Gottstein[3] and Morris Moscovitch[4,5]

Affiliations:

[1] Department of Cognitive and Brain Sciences, Ben-Gurion University of the Negev, POB 653 Beer Sheva, 84105, Israel

[2] Psychological & Brain Sciences, Johns Hopkins University, Baltimore, MD 21218

[3] School of Psychological Sciences, Tel-Aviv University, Ramat-Aviv, 69978, Israel

[4] Department of Psychology, University of Toronto, Toronto, ON M5S 3G3

[5] Rotman Research Institute, Baycrest Centre, 3560 Bathurst Street, Toronto, Ontario, Canada, M6A 2E1

*Correspondence to: Talya Sadeh

Department of Brain and Cognitive Sciences

Ben-Gurion University of the Negev

POB 653

Beer Sheva, 84105

Israel

tsadeh@bgu.ac.il

Summary: It is well-established that what determines whether information will be remembered depends not only on encoding and retrieval, but also on the neural processes that follow encoding. Extending the consideration of peri-encoding processes further in time, recent studies have indicated that memory for an experience may also be affected by processes occurring prior to that experience. We hypothesized that, in humans, spontaneous pre-encoding patterns are reinstated during succesful encoding. Individuals studied and retrieved lists of words while undergoing fMRI-scanning. Multivoxel hippocampal patterns during resting periods prior to encoding resembled hippocampal patterns at encoding most strongly for items that were subsequently remembered. Furthermore, across subjects, the magnitude of similarity correlated with a behavioural measure of episodic recall. The results indicate that the neural scaffold of a memory trace is spontaneously laid even before ever perceiving the to-be-encoded information.

Main Text:

Why are certain experiences remembered and others forgotten? It has long been established that, to answer this question, scholars should investigate not only processes occurring while individuals are learning or retrieving information, but also processes occurring following encoding (e.g., consolidation [1-4]) and during retention (e.g., decay or interference [5,6]). Can the determinants for the mnemonic fate of an experience, however, be traced back to the time *preceding* encoding? The current study is aimed at answering this question, with the hypothesis that the neural foundations of an individual memory trace are laid even before the corresponding experience occurs.

Evidence accumulated in recent years has begun to support the viability of this hypothesis. In rodents it has been shown that hippocampal place-cells firing in a particular temporal sequence while an animal is navigating a route also fire spontaneously in the same sequence during a resting period prior to the experience [7-15 i]. This striking phenomenon, referred to as "preplay", has been suggested to enhance learning via spontaneous activation of a neural pattern which future experience coopts or recruits into a memory trace. In the absence of this pre-encoding pattern, new patterns must be formed from less activated neurons, thereby reducing their chances of survival in a memory trace. Extrapolating from these studies, we hypothesized that in humans the first elements of a memory trace may be laid prior to learning.

Studies investigating processes occurring prior to encoding in humans have found that neural activity (in some studies, including the hippocampus) immediately prior to presentation of each stimulus predicted memory performance [16-19] and decision-making [20]. These findings have typically been interpreted to reflect an anticipatory state, or heightened attention during the pre-

stimulus periods. Indeed, oftentimes, pre-stimulus periods were preceded by a cue providing information about the upcoming item.

In contrast to these previous human studies, the aim of the current study was not to examine anticipation of an individual experience, or even explicit anticipation at all, but to seek pre-encoding patterns within ongoing spontaneous neural activity that support memory. Based on recent theoretical ideas [15], we hypothesized that pre-existing neural representations provide the framework for successful encoding of new information. Operationally, this entails that, in humans, spontaneous patterns of neural activity during pre-encoding rest are reinstated to support successful encoding. We focused specifically on mnemonic effects in the hippocampus and its related network.

It has been recently suggested that pre-encoding processes support learning via ***memory allocation***, whereby intrinsically excitable neurons are likely to be included in a memory trace [21-23]. Namely, the cell assembly that will form a memory trace is not arbitrarily selected at the time of memory formation, but is haphazardly activated beforehand. The high excitability of the neurons recruited to encode a certain episode is, furthermore, expected to increase the likelihood that these neurons will also be involved in encoding the subsequent episode. Consequently, these two episodes are more likely to be clustered at retrieval such that recalling the first will lead to subsequently recalling the second [24]. In humans such a clustering effect has been well-documented behaviourally [25-28]. Termed the Temporal Contiguity Effect, this phenomenon refers to the increased probability of sequentially recalling two items that were studied in close temporal contiguity. The effect is thought to arise from the largely overlapping temporal (and neural) contexts shared by neighbouring items, and is a hallmark of episodic recall [29,30]. Taken together, if memory allocation mediates pre-activation and results in temporal clustering, we

predict that the extent to which pre-encoding and encoding patterns overlap will be associated with the magnitude of temporal contiguity.

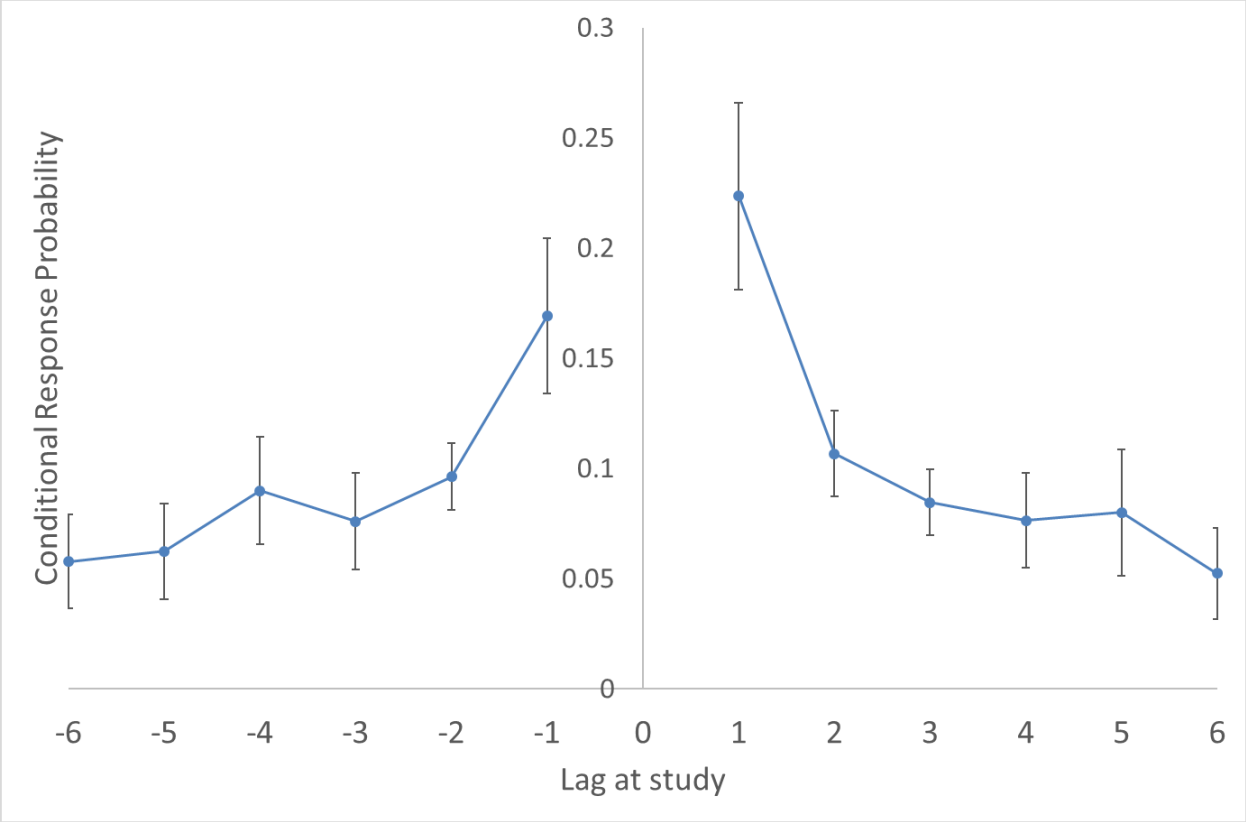
Our operationalization of the notion of pre-activation in humans entails that an item is more likely to be remembered the more similar the neural representation of its memory trace is to a representation spontaneously elicited during pre-encoding rest. Our design thus included 15-20 s resting periods prior to presentation of the study materials. We used a word-list free recall task, in which participants (n=23) studied and immediately recalled 24 lists of 12 words. During all phases, neural activity was measured using functional magnetic resonance imaging (fMRI). From these data, neural activity patterns for each of the timepoints (TRs) were extracted. This procedure was implemented both for the pre-encoding resting timepoints and for each of the timepoints corresponding to studying a word. A similarity score could then be calculated between each resting timepoint and each study timepoint.

Results

Behavioural Results

Participants correctly recalled a mean of 6.34 words per list (SEM = 0.36). As illustrated in Figure 1, a typical Temporal Contiguity Effect [31] was observed, demonstrating that the closer two items had been presented at study, the higher their probability of being recalled consecutively. Because we excluded the first three words recalled from each list to eliminate possible effects of working memory, whose span is 3-4 items [32-35], the magnitude of the Temporal Contiguity Effect is comparable to that of delayed free-recall paradigms [27,31].

Fig. 1. The Temporal Contiguity Effect. The Temporal Contiguity Effect for all but the first three words recalled (which may reflect reliance on working memory rather than episodic memory; see Materials and Methods). Conditional Response Probability (CRP) measures the probability that a transition between two successively-recalled items would be made across a certain lag [31]. The probabilities are conditional on the event that a transition of a certain lag would yield a studied item that has not already been retrieved. Lag refers to the distance, at study, between the serial-positions of two successively recalled words. Error bars reflect 95% confidence intervals for within-subject designs [36].



Controlling for distance from pre-encoding between Remembered and Forgotten items

Our analysis of pre-activation effects was aimed at demonstrating that, as compared to Forgotten items, the neural patterns elicited by Remembered items exhibit enhanced similarity with spontaneous patterns elicited during the pre-encoding phase. It was essential to rule out that pre-activation effects are driven by differential distances from pre-encoding between Remembered and Forgotten items: if Remembered items are closer to the pre-encoding phase than Forgotten items, differences between the two conditions could be due to trivial effects of temporal autocorrelation, rather than pre-activation effects.

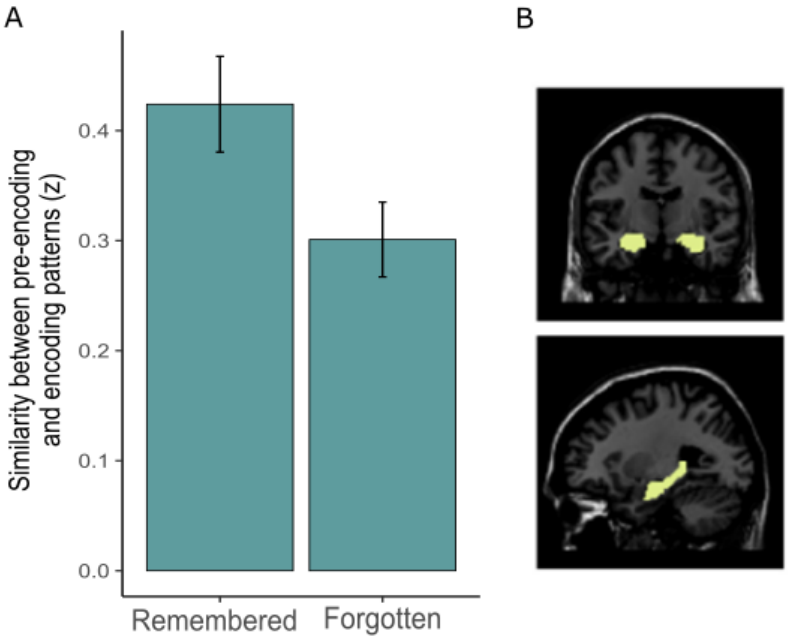
For each encoded item, we calculated its distance in TRs from the pre-encoding phase. As might be expected considering the primacy effect, Remembered items were closer to the pre-encoding phase than Forgotten items (Mean Remembered = 7.4 TRs, Mean Forgotten = 8.4 TRs; paired-sample t-test: $t(22) = 3.63$, $p = .001$). To control for the effects of distance from encoding, for participants whose Remembered items were, on average, closer to the pre-encoding phase than the Forgotten items ($n = 16$), we randomly excluded ~25% of the closest Remembered items and ~25% of the farthest Forgotten items. With this exclusion, the distances from pre-encoding were greater (though not significantly; $t(22) < 1$, $p = .9$) for Remembered than Forgotten trials (Mean Remembered = 7.64 TRs, Mean Forgotten = 7.62 TRs). To examine whether, as compared to Forgotten items, Remembered items still included a greater proportion of items which were close to the pre-encoding stage, we examined the distribution of output positions within each participant, separately for Remembered and Forgotten items. If a larger proportion of Remembered items were close to the pre-encoding stage, their output positions' variance should be greater than those of the Forgotten items. In contrast, the data showed that the variance of Remembered trials was, in fact, lower than that of the Forgotten trials (Mean standard deviations for Remembered = 10.56, for Forgotten = 12.81). All subsequently-reported

analyses were conducted on the data controlling for the differences in distances between Remembered and Forgotten items as described above (Mean number of Remembered trials per block = 2.43, S.D. = 0.82; Mean number of Forgotten trials per block = 3.3, S.D. = 1.25).

Pre-activation Effects in the Hippocampus

Fisher-transformed Pearson correlation coefficients indexing similarity scores were calculated between the multi-voxel pattern of each pre-encoding TR and of each study trial (see Methods for further details). For each participant, similarity scores were then averaged across trials for each of the two conditions (Remembered and Forgotten). Figure 2 depicts the results of this analysis for our a-priori region of interest (ROI): the bilateral hippocampi. As predicted, the mean similarity scores in the Remembered condition (Mean Fisher's Z = 0.42) were greater than in the Forgotten condition (Mean Fisher's Z = 0.3; paired-sample t-test: $t(22) = 2.59$, $p = .017$, Cohen's $d = 0.54$). Thus, a significant pre-activation effect was demonstrated.

Fig. 2. Pre-activation Effect in the Bilateral Hippocampi. Difference between Remembered and Forgotten words with regard to similarity between pre-encoding and encoding patterns ($t(22) = 2.59$, $p = .017$, Cohen's $d = 0.54$). Error bars denote standard error of the mean.



Pre-activation Effects in the Hippocampal Network

Our a-priori assumption was that pre-activation reflects the laying down of a scaffold for a memory trace of a future episode. On the neural level, the memory trace consists of an entire hippocampal-neocortical ensemble, in which the hippocampus acts as a pointer to a cortical network representing the perceptual and semantic details of the episode [37,38]. We, therefore, hypothesized that the pre-activation effect extends to the network of regions which, together with the hippocampus, constitutes the memory trace.

To examine this possibility, we next turned to investigate pre-activation effects within the network of regions that co-fluctuated with the hippocampus throughout the experiment. To this end, the timecourses for all voxels within the bilateral hippocampi ROI were averaged and used as the seed in a functional connectivity analysis. This analysis aimed to identify voxels whose

activity correlates with the hippocampus throughout the experiment. The correlations were calculated separately for each run, then averaged across runs. For each participant, a Hippocampal Network ROI was defined that included all voxels for which the correlation coefficient surpassed a threshold of 0.3[1]. Using the same threshold, a group Hippocampal Network ROI was created by averaging together all of the individual-subjects' Hippocampal Network ROIs (see Figure 3 for the Group Hippocampal Network ROI). This group-level functional-connectivity analysis revealed regions primarily in bilateral posterior areas. These included the parahippocampal gyrus, occipital-temporal regions, the posterior cingulate cortex, lingual gyrus and precuneus. In addition, large clusters were detected in the thalamus, the striatum and regions in the cerebellum.

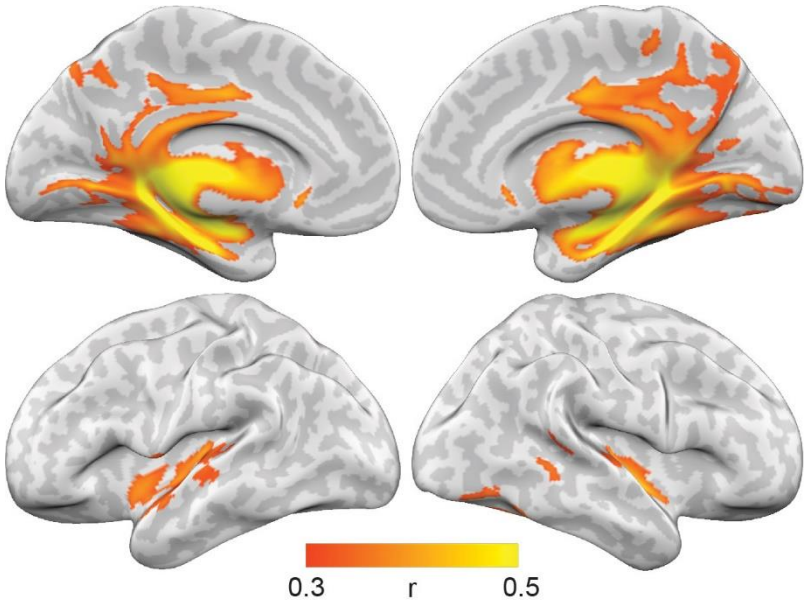
The analysis comparing pre-activation effects between the Remembered and Forgotten conditions was repeated for these ROIs. As in the analysis of the hippocampus, here, too, the mean similarity score in the Remembered condition (Mean Fisher's Z = 0.19) was significantly greater than in the Forgotten condition (Mean Fisher's Z = 0.096), as revealed in a paired-sample t-test ($t(22) = 3.32$, $p = .003$, Cohen's $d = 0.69$). In the Group Hippocampal Network ROI, the effect was also significant (Mean Fisher's Z for Remembered = 0.14; Mean Fisher's Z for Forgotten = 0.038; $t(22) = 2.97$, $p = .007$, Cohen's $d = 0.62$).

Importantly, though these analyses reveal that the pre-activation effect extends to brain regions which are functionally-coupled with the hippocampus, it is unlikely to be a whole-brain effect, as revealed by a follow-up analysis. This analysis explored 90 individual AAL ROIs spanning the entire brain and found that only a subset of 26 regions—20 of which overlapped

[1] The current analysis is not redundant with the one finding a preactivation effect in the hippocampus. While the hippocampal preactivation effect concerns spatial correlation and is calculated on an item-by-item basis, the current analysis concerns temporal correlation and is calculated across the whole session.

with the Hippocampal Network ROI—showed a significant effect at $p < .05$ (for the full list see Table S1). Furthermore, none of these regions passed correction for multiple comparisons ($q(FDR) < .05$ [39]).

Fig. 3. Group Hippocampal Network ROI. Regions co-fluctuating (functionally connected) with the hippocampus during the experiment (voxels for which the mean correlation coefficient surpassed a threshold of 0.3, averaged across participants).



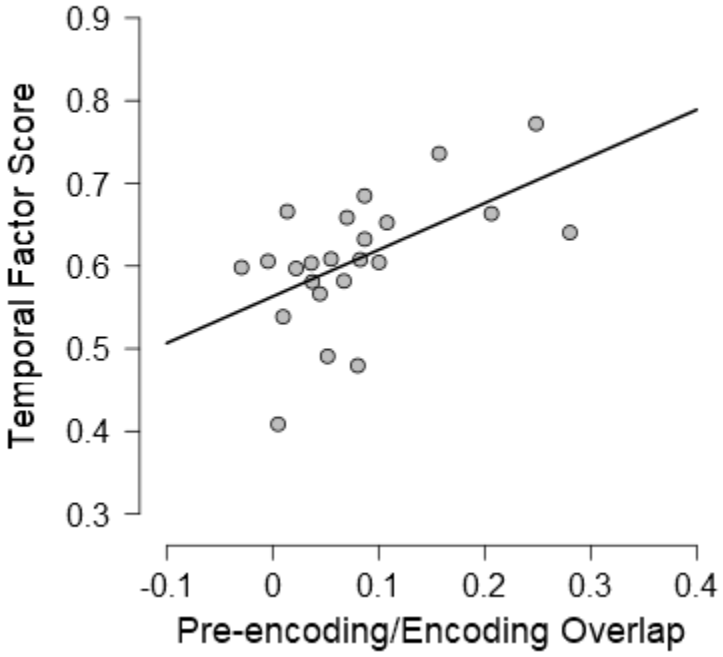
Across-subject Correlation between Pre-encoding/Encoding Overlap and the Temporal Contiguity Effect

Next, we examined the across-subject correlation between (1) the magnitude of overlap between pre-encoding and encoding representations and (2) the temporal contiguity effect—a hallmark of

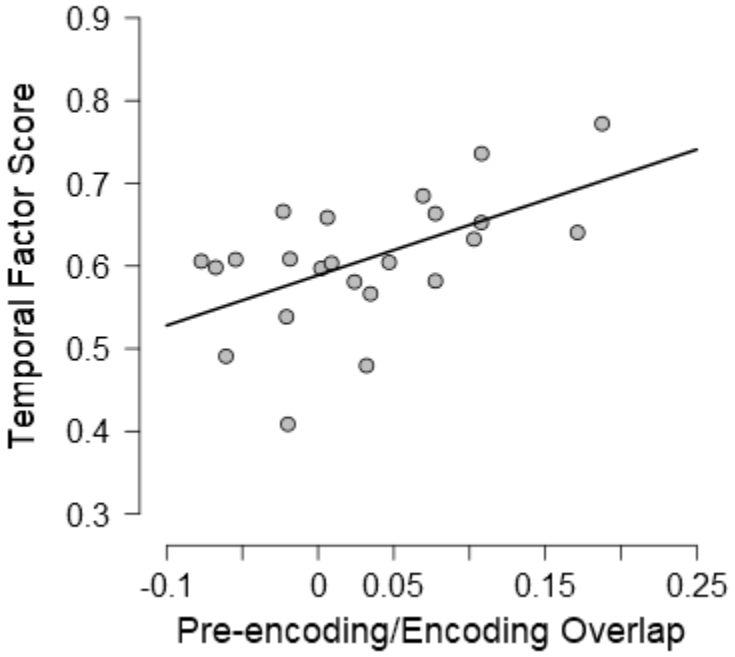
episodic recall [25,26,31]. The magnitude of pre-encoding/encoding overlap per participant was indexed by the mean correlation coefficients between pre-encoding trials and subsequent encoding trials. With regard to the behavioural measure, for each participant, a temporal factor score was calculated. This score represents the tendency of a participant to successively retrieve items in short temporal lags, or to rely on temporal context at recall (see "Behavioural Analysis"). A significant positive correlation was found both for the bilateral hippocampus and for the Hippocampal Network ROI: the greater the pre-encoding/encoding overlap, the stronger the reliance on temporal context (for bilateral hippocampus: $r = .56$, $p = .006$; for Hippocampal Network ROI: $r = .56$, $p = .006$). Results for the bilateral hippocampus and the Hippocampal Network are illustrated in Figure 4.

Fig. 4. Correlation between pre-encoding/encoding overlap and the Temporal Contiguity Effect.

Bilateral Hippocampus



Hippocampal Network ROI



Waxing and waning of a general encoding state?

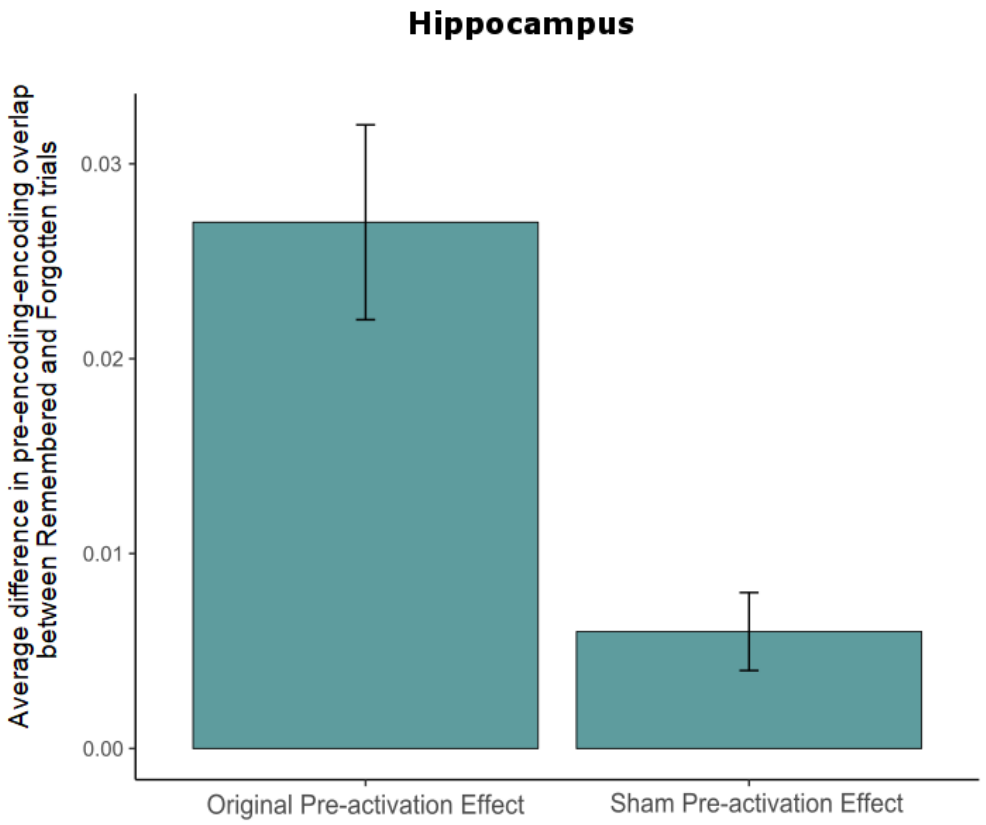
We maintain that the pre-activation effects demonstrated here reflect correspondence between specific pre-encoding neural patterns and neural patterns of subsequently-experienced events which form the memory trace. However, an alternative interpretation of our findings is that the pre-activation effects are a result of a general encoding state (or even an attentional state; see Discussion) that waxes and wanes over time. According to this interpretation, the pre-

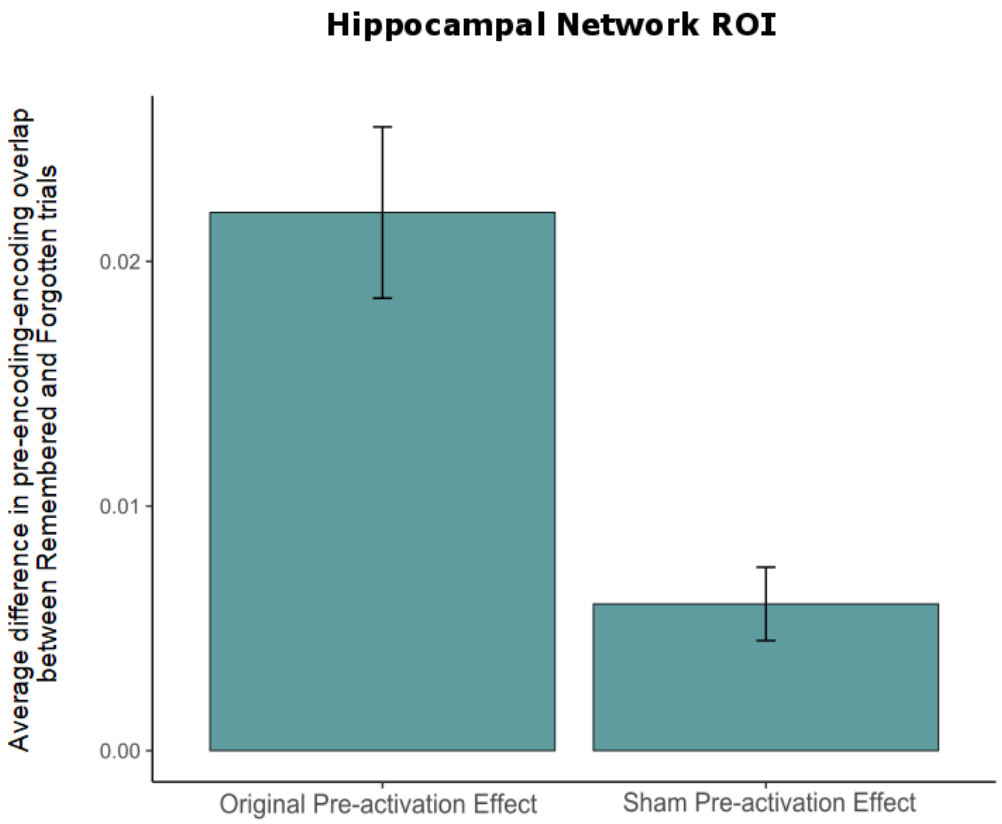
activation effects reflect a general similarity between pre-encoding and encoding neural states that is not item-specific.

To examine this possibility, we ran a sham analysis in which we aggregated across all lists to create two correlation matrices per subject: one including correlation coefficients between all pre-encoding trials (across all lists) and all subsequently-remembered encoding trials, and one between all pre-encoding trials and all subsequently-forgotten trials. An average correlation coefficient was then calculated for each of these two matrices. A sham pre-activation effect was defined as the difference between these two averages—namely, between the subsequently-remembered average correlation coefficient and the subsequently-forgotten average correlation coefficient. If the alternative, "encoding state", interpretation is true, it is expected that the sham pre-activation effects would be of a similar magnitude as the original pre-activation effects. Results of the sham analysis argue against this interpretation. For both the hippocampus and the Hippocampal Network ROI, the original pre-activation effects were significantly greater than the sham pre-activation effects (for the hippocampus: $t(22) = 2.29$, $p = .032$, Cohen's $d = 0.48$; for the Hippocampal Network ROI: $t(22) = 2.7$, $p = .013$, Cohen's $d = 0.56$; Figure 5). This indicates that the patterns of remembered items are more similar to the pre-encoding phase which immediately preceded them as compared to the pre-encoding phases of the other lists.

Fig. 5. Comparison of Original and Sham Pre-activation Effects. Preactivation effects are the difference between Remembered and Forgotten trials with regard to similarity between pre-encoding and encoding patterns. The original pre-activation effect refers to similarity between the pre-encoding trials preceding a certain list and the encoding trials of that specific list. The sham pre-activation effect refers to similarity between all pre-encoding trials (across all lists) and

all encoding trials. The top panel depicts pre-activation effects in the bilateral hippocampus and the bottom panel depicts pre-activation effects in the Hippocampal Network ROI. Error bars denote standard error of the mean.


Hippocampus
Average difference in pre-encoding-encoding overlap between Remembered and Forgotten trials
0.03
0.02
0.01
0.00
Original Pre-activation Effect
Sham Pre-activation Effect


Hippocampal Network ROI
Average difference in pre-encoding-encoding overlap between Remembered and Forgotten trials
0.02
0.01
0.00
Original Pre-activation Effect
Sham Pre-activation Effect

Discussion

Our results demonstrate that spontaneous neural patterns elicited during an ongoing pre-encoding resting period are reinstated during successful encoding of individual items. The pre-activation effects we found are not merely a result of temporal autocorrelation, as revealed by our controls for temporal distance. Neither were these effects a result of a general encoding or attentional state [40], as revealed by the sham analysis. As predicted, these effects were found in item-specific spatial activity patterns of the hippocampus, as well as in its functionally-coupled network.

Furthermore, we found that individual differences in the magnitude of pre-encoding/encoding overlap correlated with reinstatement of temporal context at retrieval, as indexed by the temporal contiguity effect. This result was predicted based on the notion of memory allocation [21] proposed in the rodent literature. According to the notion of memory allocation, neurons with intrinsically high excitability are more likely to be allocated to a memory trace of an item. These neurons are also more likely to be included in the memory traces of items studied adjacently to the given item. Consequently, neighbouring items at study are also more likely to cue each other during test and thus be clustered together at retrieval. The behavioural manifestation of this clustering phenomenon, described in the context of the memory allocation theory, is exactly that which is documented in the seminal Temporal Contiguity Effect [26,31]. Importantly, our findings demonstrate, for the first time, that reliance on temporal context to

guide retrieval—as captured by the Temporal Contiguity Effect—can be predicted by neural representations prior to learning—as captured by the pre-encoding/encoding overlap.

Are the pre-encoding patterns merely random variations in neural firing that are fortuitously co-opted by neural events at encoding? Results of the sham analysis suggest that they are not. Namely, it is unlikely that the pre-activation effects reported here reflect a general encoding (or attentional) state that waxes and wanes over time [41]. Instead, these effects more likely reflect the laying down of the scaffold for an item-specific memory trace.

Whereas we conceive of the pre-encoding neural patterns as random fluctuations in neural activation, it is possible that on some occasion, these fluctuations reflect thoughts that emerge spontaneously and are idiosyncratically associated with the to-be-studied items. This idea does not, of-course, entail any form of pre-cognition [42]. Namely, we do not claim that during pre-encoding rest participants spontaneously thought of the exact same words that subsequently appeared during the study phase. Rather, it is possible that spontaneous idiosyncratic thoughts during pre-encoding became associated with (at least some of) the study words and constituted part of their memory traces, perhaps due to shared contextual features between the spontaneous thoughts and the memory traces of the study words.

Is the pre-activation effect driven by fluctuations in attention such that participants are more attentive prior to- and during encoding some of the lists than others? This would entail that the number of remembered items significantly varies between lists. However, an analysis of the difference in the number of words recalled per list revealed that this number does not vary significantly within subjects. The standard deviation of the number of items recalled per list, averaged across all participants is low—only 1.56 items. Therefore, it seems unlikely that the current pre-activation effects are driven by fluctuations in attention between lists.

All pre-activation effects reported so far were found both for the a-priori bilateral hippocampal ROI and for the network of regions functionally coupled with the hippocampus, (referred to in the Results sections as “Hippocampal Network ROIs”). This latter result held both when the functional network was defined per subject and when defined at the group level. An additional analysis (see Supplementary Information) sought to identify whether the network results reflect only patterns within individual ROIs (or subsets of the network) or also a global pattern of amplitudes across the different ROIs in the network—namely, a network effect. The analysis, which we term the “Global-pattern ROI”, also revealed a significant pre-activation effect. This finding suggests that the current pre-activation effects may be driven in part by a global pattern of amplitudes (namely, by network-wide effects).

Demonstrating the existence of a mechanism of pre-activation provides a crucial contribution to our understanding of why certain experiences are remembered and others are forgotten. The mnemonic fate of studied-items greatly depends on the extent to which spontaneous neural representations, within an ongoing pre-encoding period, are reinstated during encoding. Because the individual's neural state at the resting time preceding encoding likely differs from one occasion to the next—and correspondingly so do the idiosyncratic thoughts/autobiographical associations which are prominent during rest [43,44]—so will the mnemonic outcome of studied items be different, even if all external stimuli and conditions are kept the same in all occasions.

Whatever the specific nature of pre-encoding patterns might be, our findings illustrate that they play a crucial role in memory for subsequently-presented information. With that we provide a novel demonstration in humans that the determinants for the mnemonic fate of an experience can be traced back to spontaneously-elicited neural patterns prior to the experience.

Thereby, we establish that pre-activation constitutes a fundamental aspect of the neurocognitive basis of human memory. These findings have important implications for memory-interventions in healthy and in memory-impaired people from early development to aging, by extending the focus of such interventions to processes occurring prior to presentation of memoranda.

Methods

Participants

Participants were 28 neurologically-intact native Hebrew speakers (17 women), right-handed and with normal or corrected-to-normal vision. Data from four participants were excluded due to excessive motion inside the scanner (over 4 mm). Data from an additional participant were excluded due to low performance on the behavioural memory task (mean number of recalled items was over two standard deviations below the group average). All reported analyses thus include data of 23 participants (15 women; ages 21-32 years, mean = 24.5). Participants gave their informed consent prior to the experiment and were compensated for their time monetarily or with course credit. All experimental procedures were approved by the Tel-Aviv Medical Center's Clinical Investigation committee.

Materials

The stimuli consisted of 336 Hebrew nouns. All nouns were 5-6 letters long (mean = 5.46). The nouns were divided into 28 lists of 12 words; 24 lists were used for the experiment and 4 lists for the practice phases (see below).

Behavioral Procedure

Prior to entering the scanner, participants were given detailed instructions and two practice lists of the free-recall task, which included overt pronunciation of the encoded words. A primary goal of this practice session was to ensure that participants verbalized the words clearly

yet softly enough to avoid head motion[2]. Two additional practice lists were presented within the scanner.

The experiment consisted of four identical free-recall runs, each lasting 7:25 minutes. The set of words was randomly divided into four, with a quarter of the words (i.e., 72 words, divided into six lists of 12 words each) assigned to each run. The order of the runs was counterbalanced across participants. In addition, three runs of a semantic fluency task were interleaved in between the free-recall runs. The purpose of the semantic fluency task was extraneous to the current endeavor, and therefore we do not discuss it further.

Figure S1 illustrates the design of a free-recall run. In each of the four runs, six lists of 12 words each were presented for study followed by a free-recall test. The order in which the words were presented was random. Presentation of three of the six lists (randomly selected) was preceded by a fixation block, which we term the "Pre-encoding Phase". During this phase, participants were instructed to rest while fixating on a cross in the middle of the screen. The durations of the pre-encoding phases were 15, 17.5 and 20 s and their order was counterbalanced across runs. Only lists which included a pre-encoding phase were included in the current analyses.

Presentation of each of the study lists was preceded by a 2.5 s display in which the word PREPARE appeared at the center of the screen, signaling participants to prepare for the upcoming list. The study phase then began, with each of the 12 words in the list presented sequentially for 1750 msec in the center of the screen followed by a 750-4,750 msec fixation cross. For four of the six lists (whose order was randomly assigned and counterbalanced across sessions), a fixation trial of 0.5 to 8.25 s was presented at the offset of the study phase. The aim

[2] We were able to record overt responses during the fMRI scanning sessions by using adaptive noise cancelling microphone and headphones (FOMRI-III; Optoacoustics, Israel; see also Sadeh et al., 2011).

of these trials was to jitter the beginning of the recall phase with regard to the TR. The recall phase then began, with presentation of five question marks at the center of the screen signaling participants to start recalling. Recall was executed by overtly pronouncing as many words as possible from the last list presented, in any order, until the cue preparing them for the next list appeared on the screen. The recall phase lasted 22.5 s. Verbal responses were digitally recorded using Audacity software (http://audacity.sourceforge.net). As an incentive to enhance performance, participants were told that they would be awarded monetary prizes (comparable to $200) if they reached the highest scores in the experiment. Figure S2 illustrates the experimental design for a single list which is preceded by a pre-encoding phase.

Imaging Procedure

Participants were scanned on a GE 3T Signa Horizon LX 9.1 echo speed scanner (Milwaukee, WI). During each of the runs, whole-brain T2*-weighted EPI functional images were acquired (TR=2500 msec, 20 cm FOV, 64x64 matrix, Flip Angle=85, TE=35, 44 coronal slices perpendicular to the hippocampal axis, 3 mm thickness with 0.7 gap, sequential acquisition). In each run, 174 volumes were acquired. Four additional volumes were acquired at the beginning of each run to allow for T1 equilibration (and were excluded from the analysis).

Data Analysis

Behavioural Analysis. The behavioural recall data were transcribed manually by research assistants. Based on the transcription, the study items were classified into the following categories: (1) Words subsequently recalled at the first three output positions. These presumably reflect retrieval from working memory, or a short-term memory store [45], which was not the focus of the current study. Crucially, these items were not included in any of the analyses, including those examining temporal contiguity effects. (2) Words subsequently recalled correctly, from all output positions but the first three ("Remembered"); (3) Words which were not subsequently-recalled ("Forgotten").

The Temporal Contiguity Effect was examined using scripts from http://memory.psych.upenn.edu/Software. To reiterate, the Temporal Contiguity Effect refers to the phenomenon whereby the smaller the absolute temporal lag between two items at study, the higher the probability that these two items will be recalled consecutively. We examined conditional-response probabilities (CRPs): the probability of making transitions at a certain lag conditional on this lag being available [31]. For the analysis examining across-subject correlation with pre-activation magnitude, temporal-factor scores were calculated for each participant [46]. The temporal factor score is a measure of the tendency of a participant to successively retrieve items with short temporal lags (namely, which appeared close to each other at encoding), with a score of 0.5 indicating no effect of temporal contiguity.

Pre-processing. Imaging data were preprocessed and analyzed using SPM8 (Wellcome Department of Cognitive Neurology, London). A slice-timing correction to the first slice was performed followed by realignment of the images to correct for subject movement. Next, data

were spatially normalized to an EPI template based upon the MNI305 stereotactic space [47]. The images were resampled into 2 mm cubic voxels and spatially smoothed with an 8 mm FWHM isotropic Gaussian kernel. Finally, images were resliced to be aligned to the anatomical ROIs.

Pattern Similarity Analysis. Effects of pre-activation were investigated using a method introduced by Staresina et al. [4] to study reactivation of individual items. The pre-encoding phases were broken down to pre-encoding trials, each trial corresponding to single timepoint, or TR. Likewise, the study phase was broken down by TRs. Since study trials were not temporally aligned with the TRs, the onset of each study trial was defined as that of the closest TR. TRs corresponding to trials whose onsets overlapped those of another trial were excluded from the analyses. Since the peak BOLD response is ~5 seconds after the onset of each trial, all trials were shifted forward in time by two TRs. For each trial, a multi-voxel pattern of activity within regions of interest was extracted from the BOLD data. Data were detrended to remove linear drifts. A similarity score, indexed by Pearson correlation, was then calculated between the pattern of each pre-encoding TR and of each study trial. A Fisher transformation was applied to the correlation coefficients, aiming to make their sampling distribution approach that of the normal distribution, and the results were divided by the coefficients' standard deviations ($1/\sqrt{n-3}$) [48].

Regions of Interest Definition. Masks of 90 cortical and subcortical regions spanning the brain (excluding the cerebellum) were obtained using the AAL (Automated Anatomical Labeling) atlas [49].

References and Notes:

1 Sekeres, M. J., Moscovitch, M. & Winocur, G. in *Cognitive Neuroscience of Memory Consolidation* (eds Nikolai Axmacher N. & Rasch B.) (2017).

2 Kitamura, T. *et al.* Engrams and circuits crucial for systems consolidation of a memory. *Science* **356**, 73-78 (2017).

3 Gruber, M. J., Ritchey, M., Wang, S.-F., Doss, M. K. & Ranganath, C. Post-learning hippocampal dynamics promote preferential retention of rewarding events. *Neuron* **89**, 1110-1120 (2016).

4 Staresina, B. P., Alink, A., Kriegeskorte, N. & Henson, R. N. Awake reactivation predicts memory in humans. *Proceedings of the National Academy of Sciences* **110**, 21159-21164 (2013).

5 Sadeh, T., Ozubko, J. D., Winocur, G. & Moscovitch, M. How we forget may depend on how we remember. *Trends in cognitive sciences* **18**, 26-36 (2014).

6 Yonelinas, A. P. & Ritchey, M. The slow forgetting of emotional episodic memories: an emotional binding account. *Trends in cognitive sciences* **19**, 259-267 (2015).

7 Dragoi, G. & Tonegawa, S. Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature* **469**, 397-401 (2011).

8 Dragoi, G. & Tonegawa, S. Distinct preplay of multiple novel spatial experiences in the rat. *Proceedings of the National Academy of Sciences* **110**, 9100-9105 (2013).

9 Buhry, L., Azizi, A. H. & Cheng, S. Reactivation, replay, and preplay: how it might all fit together. *Neural plasticity* **2011** (2011).

10 Johnson, A. & Redish, A. D. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *The Journal of neuroscience* **27**, 12176-12189 (2007).

11 Pfeiffer, B. E. & Foster, D. J. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* **497**, 74-79 (2013).

12 Diba, K. & Buzsaki, G. Forward and reverse hippocampal place-cell sequences during ripples. *Nat Neurosci* **10**, 1241-1242, doi:http://www.nature.com/neuro/journal/v10/n10/suppinfo/nn1961_S1.html (2007).

13 Grosmark, A. D. & Buzsáki, G. Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. *Science* **351**, 1440-1443 (2016).

14 Silva, D., Feng, T. & Foster, D. J. Trajectory events across hippocampal place cells require previous experience. *Nature neuroscience* **18**, 1772-1779 (2015).

15 Eichenbaum, H. Does the hippocampus preplay memories? *Nat Neurosci* **18**, 1701-1702, doi:10.1038/nn.4180 (2015).

16 Sweeney-Reed, C. M. *et al.* Pre-stimulus thalamic theta power predicts human memory formation. *NeuroImage* (2016).

17 Cohen, N. *et al.* Peri-encoding predictors of memory encoding and consolidation. *Neuroscience & Biobehavioral Reviews* **50**, 128-142 (2015).

18 Brown, T. I. *et al.* Prospective representation of navigational goals in the human hippocampus. *Science* **352**, 1323-1326 (2016).

19 Otten, L. J., Quayle, A. H., Akram, S., Ditewig, T. A. & Rugg, M. D. Brain activity before an event predicts later recollection. *Nat Neurosci* **9**, 489-491, doi:http://www.nature.com/neuro/journal/v9/n4/suppinfo/nn1663_S1.html (2006).

20 Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D. & Daw, N. D. Model-based choices involve prospective neural activity. *Nature neuroscience* **18**, 767-772 (2015).

21 Rogerson, T. *et al.* Synaptic tagging during memory allocation. *Nature Reviews Neuroscience* **15**, 157-169 (2014).
22 Zhou, Y. *et al.* CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. *Nature neuroscience* **12**, 1438-1443 (2009).
23 Sano, Y. *et al.* CREB regulates memory allocation in the insular cortex. *Current Biology* **24**, 2833-2837 (2014).
24 Cai, D. J. *et al.* A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* **534**, 115-118, doi:10.1038/nature17955 (2016).
25 Healey, M. K. & Kahana, M. J. Memory search is governed by universal principles—not idiosyncratic strategies. *Manuscript submitted for publication* (2013).
26 Kahana, M. J. Associative retrieval processes in free recall. *Memory & Cognition* **24**, 103-109 (1996).
27 Sadeh, T., Moran, R. & Goshen-Gottstein, Y. When items 'pop into mind': variability in temporal-context reinstatement in free-recall. *Psychonomic Bulletin & Review*, 1-12, doi:10.3758/s13423-014-0746-7 (2014).
28 Kahana, M. J. *Foundations of human memory*. (Oxford University Press, 2012).
29 Sederberg, P. B., Howard, M. W. & Kahana, M. J. A context-based theory of recency and contiguity in free recall. *Psychological Review* **115**, 893 (2008).
30 Polyn, S. M. & Kahana, M. J. Memory search and the neural representation of context. *Trends in Cognitive Sciences* **12**, 24 (2008).
31 Howard, M. & Kahana, M. J. Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **25**, 923-941 (1999).
32 Cowan, N. Metatheory of storage capacity limits. *Behavioral and brain sciences* **24**, 154-176 (2001).
33 Luck, S. J. & Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature* **390**, 279-281 (1997).
34 D'esposito, M. & Postle, B. R. The cognitive neuroscience of working memory. *Annual review of psychology* **66** (2015).
35 Jonides, J. *et al.* The mind and brain of short-term memory. *Annu. Rev. Psychol.* **59**, 193-224 (2008).
36 Loftus, G. & Masson, M. Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review* **1**, 476-490 (1994).
37 Nadel, L. & Moscovitch, M. Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology* **7**, 217-227 (1997).
38 Teyler, T. J. & DiScenna, P. The hippocampal memory indexing theory. *Behavioral neuroscience* **100**, 147 (1986).
39 Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289-300 (1995).
40 Aly, M. & Turk-Browne, N. B. Attention promotes episodic encoding by stabilizing hippocampal representations. *Proceedings of the National Academy of Sciences* **113**, E420-E429 (2016).
41 Ezzyat, Y. *et al.* Direct brain stimulation modulates encoding states and memory performance in humans. *Current Biology* **27**, 1251-1258 (2017).

42 Bendor, D. & Spiers, H. J. Does the Hippocampus Map Out the Future? *Trends in Cognitive Sciences* **20**, 167-169, doi:https://doi.org/10.1016/j.tics.2016.01.003 (2016).
43 Spreng, R. N., Mar, R. A. & Kim, A. S. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of cognitive neuroscience* **21**, 489-510 (2009).
44 Gusnard, D. A., Akbudak, E., Shulman, G. L. & Raichle, M. E. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences* **98**, 4259-4264 (2001).
45 Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J. & Usher, M. The Demise of Short-Term Memory Revisited: Empirical and Computational Investigations of Recency Effects. *Psychological Review* **112**, 3-42 (2005).
46 Polyn, S. M., Norman, K. A. & Kahana, M. J. A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review* **116**, 129 (2009).
47 Cocosco, C. A., Kollokian, V., Kwan, R. K.-S. & Evans, A. C. BrainWeb: Online Interface to a 3DMRI Simulated *NeuroImage* **5**, S425 (1997).
48 Rissman, J., Gazzaley, A. & D'Esposito, M. Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage* **23**, 752-763 (2004).
49 Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**, 273–289 (2002).

Acknowledgments: T.S. is grateful to the Azrieli Foundation for the award of an Azrieli Fellowship. We thank Richard Henson and Bradley Buchsbaum for their thoughtful and insightful comments on this manuscript.

Author Contributions: Conceptualization, T.S. and M.M..; Methodology, T.S., J.C. Y.G.G. and M.M.; Software, J.C; Formal Analysis, T.S. and J.C..; Resources, Y.G.G. and M.M.; Writing Original Draft, T.S..; Writing – Review & Editing, T.S., J.C., Y.G.G. and M.M.; Visualization, T.S. and J.C.; Supervision, M.M.; Funding Acquisition, Y.G.G. and M.M.

[i] It has recently been suggested that this finding may only be obtained if the animal is already familiar with the spatial environment or the goal, but the path to the goal is novel.