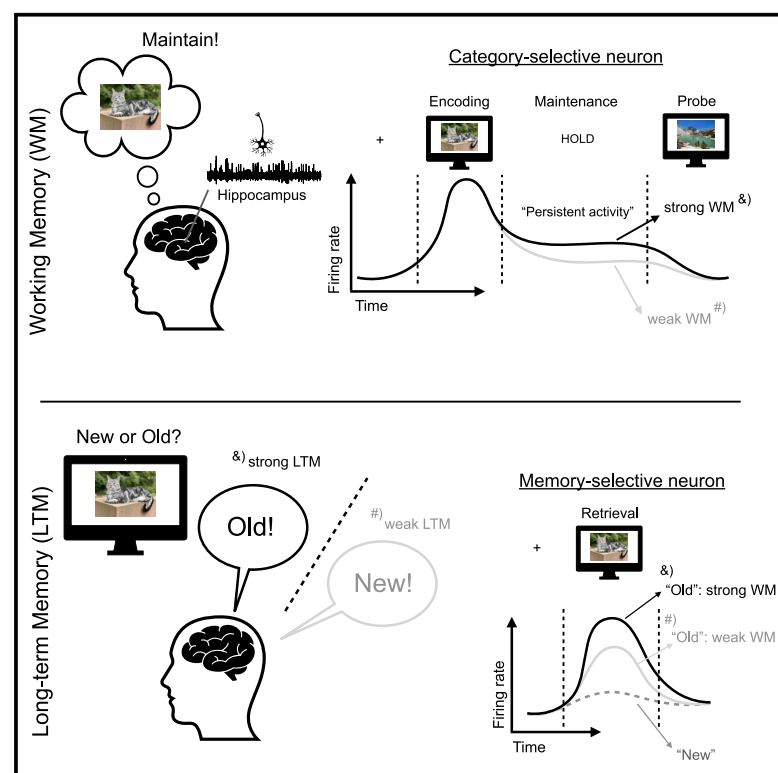


# Persistent activity during working memory maintenance predicts long-term memory formation in the human hippocampus

## Graphical abstract



## Authors

Jonathan Daume, Jan Kamiński,  
Yousef Salimpour, ...,  
Taufik A. Valiante, Adam N. Mamelak,  
Ueli Rutishauser

## Correspondence

ueli.rutishauser@cshs.org

## In brief

Daume et al. show that in the human hippocampus, working memory activity of memoranda-selective neurons predicts successful long-term memory (LTM). LTM retrieval, in turn, activates memory-selective neurons more strongly for items that were previously maintained with strong working memory activity, suggesting a shared neural mechanism supporting both memory systems.

## Highlights

- Working memory (WM) and long-term memory (LTM) interact in the hippocampus
- Persistent WM activity of category-selective neurons predicts LTM formation
- LTM-selective neurons fire more for items encoded with strong WM activity
- This suggests a single-neuron mechanism linking both WM and LTM systems in humans



## Article

# Persistent activity during working memory maintenance predicts long-term memory formation in the human hippocampus

Jonathan Daume,<sup>1,2,3</sup> Jan Kamiński,<sup>4</sup> Yousef Salimpour,<sup>5</sup> Andrea Gómez Palacio Schjetnan,<sup>6</sup> William S. Anderson,<sup>5</sup> Taufik A. Valiante,<sup>6,7,8</sup> Adam N. Mamelak,<sup>1</sup> and Ueli Rutishauser<sup>1,2,3,9,10,\*</sup>

<sup>1</sup>Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

<sup>2</sup>Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

<sup>3</sup>Center for Neural Science and Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

<sup>4</sup>Center of Excellence for Neural Plasticity and Brain Disorders: BRAINCITY, Nencki Institute of Experimental Biology, Polish Academy of Sciences, 02-093 Warsaw, Poland

<sup>5</sup>Department of Neurosurgery, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

<sup>6</sup>Krembil Research Institute, University Health Network, Toronto, ON M5T 2S8, Canada

<sup>7</sup>Division of Neurosurgery, Department of Surgery, University of Toronto, Toronto, ON M5T 1P5, Canada

<sup>8</sup>Department of Electrical and Computer Engineering, Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada

<sup>9</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

<sup>10</sup>Lead contact

\*Correspondence: [ueli.rutishauser@cshs.org](mailto:ueli.rutishauser@cshs.org)

<https://doi.org/10.1016/j.neuron.2024.09.013>

## SUMMARY

Working memory (WM) and long-term memory (LTM) are often viewed as separate cognitive systems. Little is known about how these systems interact when forming memories. We recorded single neurons in the human medial temporal lobe while patients maintained novel items in WM and completed a subsequent recognition memory test for the same items. In the hippocampus, but not in the amygdala, the level of WM content-selective persistent activity during WM maintenance was predictive of whether the item was later recognized with high confidence or forgotten. By contrast, visually evoked activity in the same cells was not predictive of LTM formation. During LTM retrieval, memory-selective neurons responded more strongly to familiar stimuli for which persistent activity was high while they were maintained in WM. Our study suggests that hippocampal persistent activity of the same cells supports both WM maintenance and LTM encoding, thereby revealing a common single-neuron component of these two memory systems.

## INTRODUCTION

Working memory (WM) is the ability to hold and manipulate a small amount of information “in mind,” an ability that is fundamental to many aspects of cognition.<sup>1</sup> Since at least the 1960s, when Atkinson and Shiffrin first proposed their model of memory,<sup>2</sup> it has been theorized that WM (then called short-term memory) and long-term memory (LTM) are two separated but connected systems. This model and later theories of WM suggest that WM acts as an intermediary between perception and LTM,<sup>3</sup> a relationship that has been studied extensively for decades. Indeed, in many instances, information held in WM is encoded better into LTM compared with information not held in WM. Behaviorally, this relationship has been shown in many studies.<sup>4</sup> For instance, words that are maintained longer in WM were later recalled better,<sup>5</sup> and items stored in WM were better remembered in a surprise recognition test compared with items

only attended or passively viewed.<sup>6</sup> A recent meta-analysis and new experiments show that the impact of holding information in WM on the quality of LTM is especially strong in the visual domain.<sup>4</sup>

Despite the ubiquity of WM-LTM interactions seen behaviorally, little is known about the mechanisms by which the two memory systems interact. Functional magnetic resonance imaging (fMRI),<sup>7–11</sup> scalp magneto-/electroencephalogram (M/EEG),<sup>6,12,13</sup> and intracranial EEG studies<sup>14,15</sup> indicate that the extent of activation of a given part of the brain as assessed by fMRI or oscillatory power during WM maintenance can be predictive of both WM maintenance success and LTM encoding success. Further, in dual-task paradigms,<sup>14,15</sup> high WM demands disrupt LTM encoding processes, arguing that they are not independent. Overall, these findings indicate that the neuronal substrate of the two processes is at least partially overlapping or shared and is located in the same areas of the brain.



However, it remains unclear what exactly is shared in terms of the neuronal substrate. One possibility, motivated by theoretical models (see below), is that WM encoding and LTM encoding engage the same cells, but this prediction has not been tested experimentally.

A hypothesis that has motivated a large body of work is that the sustained maintenance of memoranda in WM enables the gradual strengthening of LTM traces through synaptic plasticity.<sup>16</sup> A key prediction from this model is that the stronger the activation of the neurons that represent memory content during WM maintenance, the stronger the resulting LTM. Here, we test this hypothesis directly by recording from WM memory-content selective neurons (“category neurons,” see below) and assess their relationship to LTM encoding in a task with trial-unique novel stimuli for which we later test memory strength. We note that this design is different from most WM studies, in which the items held in WM are reused throughout an experiment.

Neuroimaging, intracranial EEG, and behavioral studies indicate that the medial temporal lobe (MTL), and in particular the hippocampus, are strong candidates for shared WM-LTM processes.<sup>17</sup> Indeed, recent studies indicate that the MTL, particularly the hippocampus, is critical for both WM and LTM in many circumstances as assessed by behavior and neural activity.<sup>18–27</sup> The MTL has therefore emerged as a key candidate for the brain area where the WM and LTM systems might interact.<sup>25,28</sup>

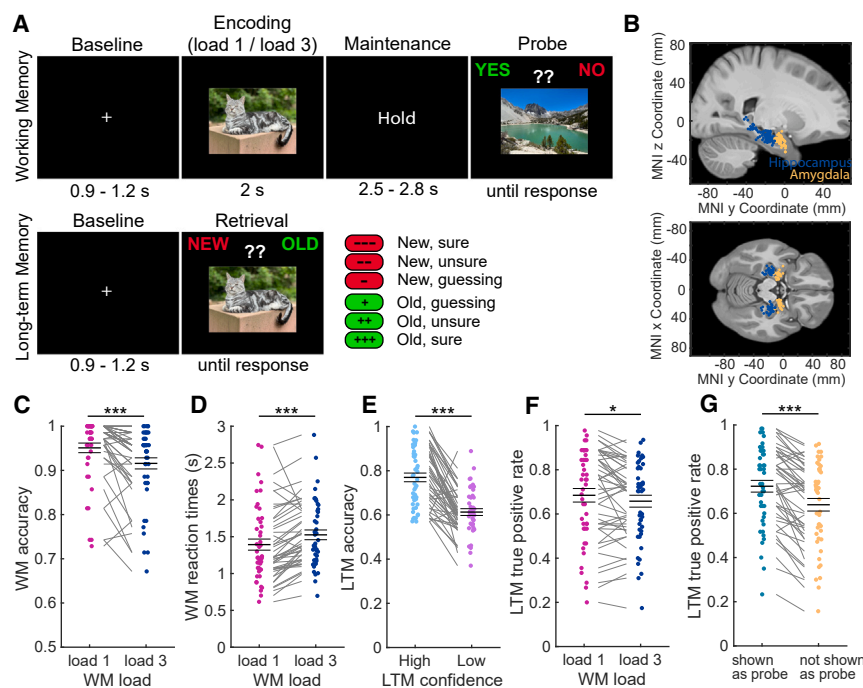
A candidate cellular substrate for a joint substrate for WM maintenance and LTM encoding is memoranda-selective persistently active neurons in the MTL. This is because such cells have persistently elevated activity throughout the maintenance period if their preferred stimulus is held in WM, which would permit synaptic plasticity as proposed by Hebb.<sup>16</sup> Persistently active cells, which constitute a relatively well-understood cellular substrate for maintaining information in WM, have been identified in both macaques and humans.<sup>29–34</sup> In humans, two types of WM-content selective persistently active neurons have been described in the MTL: concept cells<sup>32,33</sup> and category cells.<sup>34</sup> The activity of these cells is behaviorally relevant and scales with memory load during maintenance of their preferred stimuli.<sup>32–35</sup> We hypothesize that these cells might contribute to both WM maintenance and LTM encoding, thereby allowing the transient maintenance of activity to translate into structural changes through synaptic plasticity. If so, the extent of persistent activity should be indicative of later LTM strength as assessed behaviorally and/or neuronally. Here, we used the opportunity to invasively record single neurons in the human MTL in patients undergoing invasive epilepsy monitoring with depth electrodes to test this hypothesis. Patients performed both a WM and an LTM task in the same recording sessions, with a shared stimulus set between the two tasks. This design allowed us to assess whether maintaining a given trial-unique item in WM influences how well that item will later be remembered in a recognition memory test. Our study focuses on the activity of category-selective neurons—neurons that fire preferentially in response to certain categories of images, like “cars.”<sup>34,36–38</sup> We have previously found that category neurons stay persistently active when an image from their preferred category is held in WM.<sup>34</sup> Here, we test the

hypothesis that the activity of category cells during WM maintenance trials in which images from the preferred category of a cell are maintained in WM predicts successful LTM encoding in addition to WM content and quality.

## RESULTS

41 patients (48 sessions; 1 non-binary; 20 females; 20 males) performed a modified Sternberg WM task with novel images, followed by a subsequent LTM recognition test. In the WM task, patients were asked to hold either one (load 1) or three (load 3) sequentially presented images in their minds until a probe picture appeared 2.5–2.8 s later (Figure 1A top).<sup>34</sup> The task was to indicate whether the probe picture was identical to one of the encoding pictures just presented before the delay period or not. All pictures shown during encoding were novel, i.e., never shown to the subject before. The images shown during the probe were always familiar to the subject, either from having seen them in the current trial during encoding or during a previous trial (an image is shown twice at most if used as probe; see STAR Methods). Images are drawn from five different picture categories (people, animals, cars [or tools depending on version], food, and landscapes). After a delay of 10–30 min, patients performed an LTM recognition task in which half of the images shown were the same as those shown during the WM task (familiar items) and half were novel (Figure 1A bottom). In this part of the task, subjects indicated whether a given picture was “old” (i.e., seen in the earlier performed WM task) or “new.” Patients were also asked to indicate the confidence in their response, i.e., how confident they were that a given item was old or new (“sure,” “unsure,” and “guessing”).

Patients performed well in both parts of the task, with an average accuracy of  $93.3\% \pm 7.6\%$  (mean  $\pm$  SD) in the WM task and  $70.3\% \pm 9.8\%$  ( $d' = 1.19 \pm 0.62$ ) in the LTM task, respectively (both  $p < 0.0001$  as compared with 50% chance; permutation-based  $t$  test; 2 sessions from two different patients were excluded from all analyses due to an accuracy of less than 55% in at least one of the tasks; see Table S1; see Figure S1 for behavior results of 100 healthy participants). In the WM task, patients performed with higher accuracy (Figure 1C; load 1–load 3:  $t(45) = 4.89$ ,  $p < 0.0001$ ; permutation-based paired  $t$  tests were used throughout the manuscript unless stated otherwise;  $t$  values are provided as reference only; see STAR Methods) and faster (Figure 1D;  $t(45) = -4.29$ ,  $p < 0.0001$ ) in load 1 than load 3 trials. In the LTM task, retrieval was more accurate (Figure 1E;  $t(44) = 8.66$ ,  $p < 0.0001$ ; one patient did not use confidence ratings and was therefore excluded from all confidence-related analyses) and faster (mean reaction time [RT] high:  $1.83 \pm 0.57$  s; mean RT low:  $2.62 \pm 0.66$  s;  $t(44) = -11.13$ ,  $p < 0.0001$ ) when patients indicated high confidence in their responses. Low confidence was defined as an average of unsure and guessing responses. Items that were encoded in load 1 trials in the previous WM task were remembered better than items encoded during load 3 trials (Figure 1F;  $t(45) = 2.42$ ;  $p = 0.018$ ). Moreover, items that were also used as the probe and therefore presented twice were remembered better than items not used as the probe (Figure 1G;  $t(45) = 7.25$ ,  $p < 0.0001$ ).



**Figure 1. Experimental design and behavioral analysis** (A) The study contained a WM part (top) and a subsequent LTM recognition part (bottom). In the WM task, patients had to encode either one (load 1) or three (load 3) pictures in their WM and to maintain these items until a probe picture appeared a few seconds later. Their task was to indicate whether the probe was part of the encoded items in a given trial or not. All encoding pictures were novel and drawn from five different picture categories. In the LTM task, occurring after a 10–30 min break, patients had to answer whether each presented item on the screen was old (i.e., seen in the previous WM task) or new while indicating their confidence in their response. For copyright reasons, the pictures shown are similar but not identical to those used in the study. (B) We recorded single-neuron activity from the hippocampus and the amygdala of 41 patients across 48 sessions. Each dot is a patient. (C and D) WM behavior. Patients performed (C) more accurately and (D) faster in load 1 compared with load 3 trials. (E–G) LTM behavior. (E) LTM was more accurate in high- than low-confidence trials. (F) Items previously maintained in load 1 trials were remembered more accurately than when maintained in load 3 trials. (G) Pictures that were used as probes and therefore presented twice were remembered better than items not shown as probes. In (C)–(G), we used permutation-based t tests. Center lines represent mean  $\pm$  SEM. Each dot is a session. \* $p < 0.05$ ; \*\*\* $p < 0.001$ . See also [Figures S1](#) and [S2](#) and [Table S1](#).

We recorded single-neuron activity from the hippocampus and the amygdala while patients performed the two tasks ([Figure 1B](#)). In total, 883 single units across both brain areas were included in our analyses, 351 from the hippocampus and 532 from the amygdala (see [STAR Methods](#)). The same units were recorded for both tasks. The analysis presented here is partly based on a previously published dataset,<sup>34</sup> but in that report only the WM task was analyzed, whereas here the focus is on the LTM part. Spike sorting results were assessed quantitatively ([Figure S2](#)). We use the terms neuron, unit, and cell interchangeably to refer to a putative single neuron.

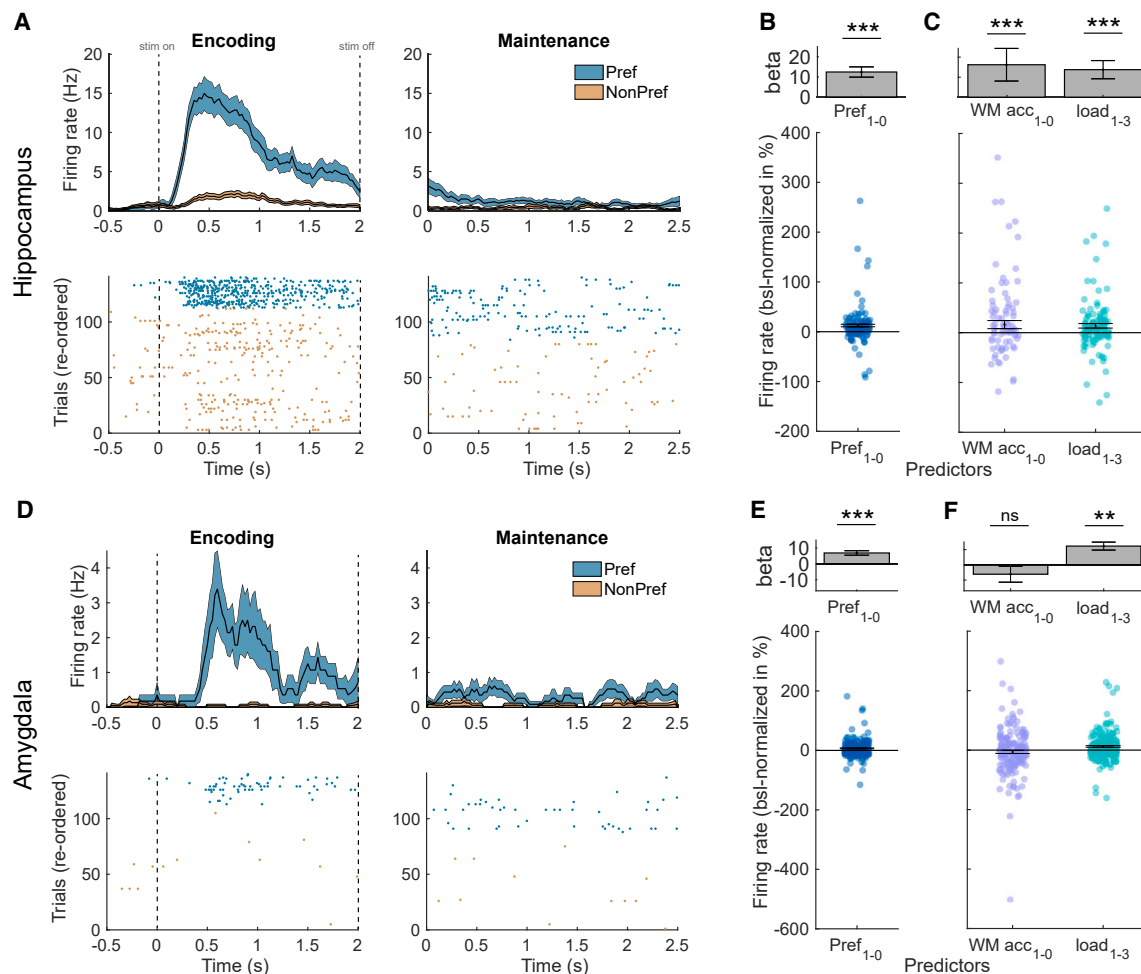
### Category neurons remain persistently active when their preferred category is held in WM

We first selected for neurons whose response following stimulus onset differed significantly between the categories of the images shown. We refer to such neurons as “category neurons” throughout (we note that this type of cell is also referred to as visually selective [VS] in other papers,<sup>36,37,39</sup> and the two terms are equivalent for the purpose of this study). To select category neurons, we assessed whether the firing rate (FR) in a 200–1,200 ms window following picture onset (encoding 1–3 and probe) was significantly correlated with the five possible picture categories (1  $\times$  5 ANOVA, followed by a right-sided permutation-based t test between the category with maximal average spike count and all other categories; if both tests were  $p < 0.05$ , we classified a neuron as a category neuron with the preferred category being the one with maximal average spike count; see [STAR Methods](#)). As shown previously, category neurons remain persistently active during the maintenance period of the WM

task when their preferred picture is held in WM.<sup>34</sup> Selecting for category neurons during picture presentation leaves their FRs during the maintenance period of the task independent for subsequent statistical analyses. In the hippocampus, 104 (29.65%) neurons qualified as category neurons, and in the amygdala, 220 (41.35%) neurons qualified as category neurons (see [Figures 2A](#) and [2D](#) for example neurons from each area and [Figure S3](#) for more examples).

To confirm that category neurons remained persistently active during the maintenance period, we computed a mixed-effects generalized linear model (GLM) using *preferred category* (2 levels, true/untrue, categorical) as fixed effect and *neuron ID* nested into *patient ID* as random intercept in each area (these random intercept terms were used for all GLMs; see [STAR Methods](#)). We used baseline-normalized FRs from all correct WM trials during the maintenance period (0–2.5 s after last encoding picture offset) for this analysis. In the hippocampus, category neurons remained persistently active throughout the maintenance period (intercept:  $\beta = 17.86$ ,  $p = 1.38 \times 10^{-3}$ ; mixed-effects GLM) and had significantly higher FRs during trials in which images from the preferred category were maintained in WM ([Figure 2B](#); preferred category:  $\beta = 12.47$ ,  $p = 8.82 \times 10^{-7}$ ). Using only preferred trials and modeling WM *accuracy* (2 levels, WM correct/incorrect, categorical) as well as *load* (2 levels, load 1/load 3, categorical) as fixed effects, we further observed that category neurons in the hippocampus had higher FRs in (1) correct than incorrect trials ( $\beta = 16.22$ ,  $p = 0.046$ ) and (2) in load 1 than in load 3 trials ([Figure 2C](#);  $\beta = 13.769$ ,  $p = 0.0023$ ).

In the amygdala, category neurons also remained persistently active throughout the maintenance period across all correct



**Figure 2. Category neurons**

(A–C) Characterization of category neurons in the hippocampus. (A) Example category neuron. The preferred category of this neuron was “animals.” Top: peri-stimulus time histogram (PSTH, bin size = 200 ms, step size = 25 ms) during the first picture in each correct trial. Colored areas represent  $\pm$  SEM. Bottom: raster plot with trials reordered into preferred and non-preferred categories. Stimulus and maintenance onset is at  $t = 0$  s (left and right, respectively). (B) Firing rates of category neurons from the hippocampus remained persistently active and were higher for preferred than unpreferred trials during the WM maintenance period. Top: beta value extracted from the GLM for preferred/non-preferred regressor in units of “percent change to baseline” (–900 to –300 ms before first picture onset). Bottom: distribution of FR differences between preferred (pref = 1) and non-preferred (pref = 0) trials across all hippocampal category neurons (each dot is a neuron,  $n = 104$ ). FRs were baseline-normalized to represent percent change to baseline. (C) When the preferred category remained in WM, category neurons in the hippocampus had higher FRs in correct as compared with incorrect and in load 1 as compared with load 3 trials.

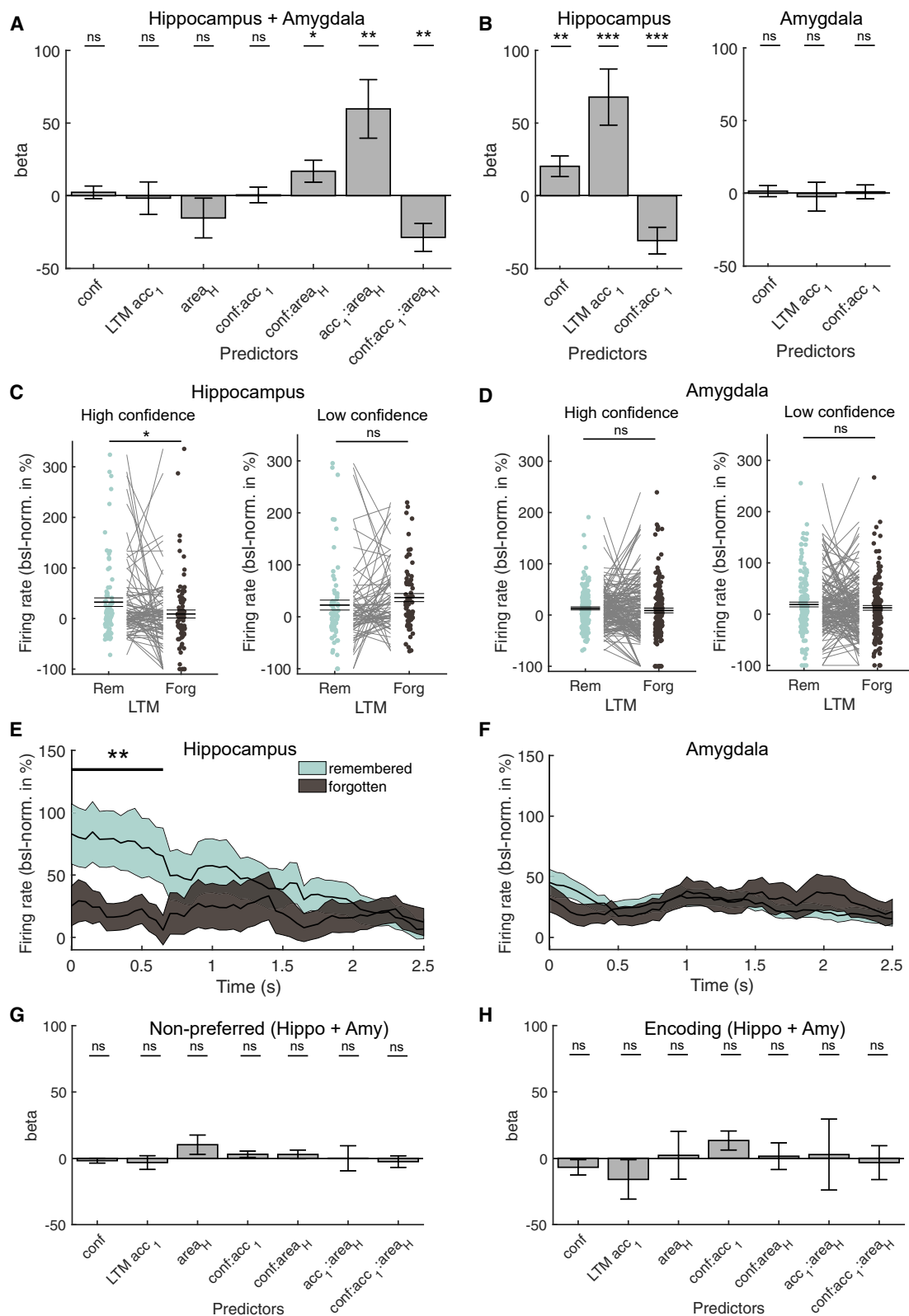
(D–F) Characterization of category neurons in the amygdala. (D) Same as in (A), but for an example category neuron from the amygdala. (E) Firing rates of category neurons from the amygdala also remained persistently active and were higher for preferred than unpreferred trials during the WM maintenance period ( $n = 220$ ). (F) Their FRs were higher in load 1 as compared with load 3 trials when their preferred category was maintained, but there was no difference between correct and incorrect trials. In (B), (C), (E), and (F), we computed mixed-effects GLMs. Error bars represent standard errors of the coefficient. In (B) and (E), each dot is a neuron. \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ; ns, not significant.

See also Figure S3.

trials (intercept:  $\beta = 9.78$ ,  $p = 6.76 \times 10^{-4}$ ) with FRs higher for preferred than unpreferred categories (Figure 2E; preferred category:  $\beta = 6.91$ ,  $p = 9.38 \times 10^{-7}$ ). However, unlike in the hippocampus, using only preferred trials category neurons in the amygdala showed a significant main effect only for load (Figure 2F;  $\beta = 12.309$ ;  $p = 2.46 \times 10^{-7}$ ), but the effect for WM accuracy ( $\beta = -6.13$ ;  $p = 0.24$ ) was not significant.

### Persistent activity of category neurons in the hippocampus, not the amygdala, predicts LTM formation

We next sought to investigate whether the activity of category neurons during the maintenance period of the WM task predicted the success of LTM formation. To that end, we computed a mixed-effects GLM with baseline-normalized FRs from the



(legend on next page)



maintenance period. For a given category neuron, we used the subset of trials for which the subject answered the WM question correctly and for which images of the preferred category were tested in the LTM recognition task (not all images seen during WM were shown during the LTM test). We modeled *LTM accuracy* of each image (2 levels, remembered/forgotten, categorical), *confidence* (3 levels, sure/unsure/guessing, continuous), *brain area* (2 levels, hippocampus/amygdala, categorical), as well as their interactions as fixed effects. We did not observe any significant main effects, nor was the interaction between *LTM accuracy* and *confidence* significant (Figure 3A; all  $p > 0.26$ ). However, all interaction terms including *area* showed significant modulations (confidence  $\times$  area: beta = 16.88,  $p = 0.025$ ; LTM accuracy  $\times$  area: beta = 59.78,  $p = 0.0030$ ; confidence  $\times$  LTM accuracy  $\times$  area: beta = -28.64,  $p = 0.0029$ ), suggesting that the relationship between FR and *LTM accuracy* and *confidence* differed between the hippocampus and the amygdala. We therefore repeated the analysis in each area separately. This revealed a significant main effect of *confidence* (Figure 3B; beta = 20.30,  $p = 4.17 \times 10^{-3}$ ) and *LTM accuracy* (beta = 67.90,  $p = 4.43 \times 10^{-4}$ ), as well as a significant interaction between the two terms (beta = -30.80,  $p = 7.74 \times 10^{-4}$ ) in the hippocampus but not in the amygdala (confidence: beta = 1.54,  $p = 0.69$ ; LTM accuracy: beta = -2.27,  $p = 0.82$ ; confidence  $\times$  LTM accuracy: beta = 1.02,  $p = 0.83$ ). These results suggest that activity during the maintenance period of the WM task of category neurons from the hippocampus was higher for later remembered than forgotten trials and therefore predictive of later LTM retrieval performance. This was not the case for persistent WM activity of category neurons from the amygdala.

To complement the beta values obtained from our GLM analysis, which reflect group-level effects across all category neurons, we next directly compared FRs across category neurons during the maintenance period of the WM task using permutation-based  $t$  tests. Given the significant interaction between LTM accuracy and confidence (Figure 3B), this analysis was conducted separately for high- and low-confidence LTM retrieval trials, comparing later remembered and forgotten trials for the preferred category of each category neuron in each area. In

the hippocampus, FRs were higher for remembered than forgotten trials for high confidence (Figure 3C;  $t(82) = 2.16$ ,  $p = 0.028$ ; some neurons were removed due to insufficient data in at least one of the conditions or since they differed  $\pm 3$  SD from the mean across all neurons and conditions; see STAR Methods), but not low-confidence trials ( $t(68) = -1.38$ ,  $p = 0.17$ ). In the amygdala, we neither observed a significant effect for high (Figure 3D;  $t(158) = 0.84$ ,  $p = 0.41$ ) nor low-confidence trials ( $t(152) = 1.15$ ,  $p = 0.26$ ).

To observe when during the maintenance period of the WM task the effect between later remembered and forgotten trials was present, we tested time-resolved FRs between remembered and forgotten trials for all high-confident trials separately for category neurons from the hippocampus and amygdala. In the hippocampus, FRs during the beginning of the maintenance period (0–650 ms) differed significantly between the two conditions (Figure 3E; cluster- $p = 0.0042$ ; cluster-based permutation test; see Figure S4 for single-neuron examples). In the amygdala, we did not observe any significant cluster throughout the entire maintenance period of the WM task (Figure 3F; cluster- $p = 0.16$ ).

Lastly, we tested whether we would observe a relationship between neural activity and LTM formation also for neuronal activity during the maintenance of stimuli from the non-preferred categories of category cells (Figure 3G) or for the visually evoked response of category neurons when images were shown on the screen during encoding (Figure 3H; encoding 1; 0–2 s after picture onset; preferred trials only). None of the main effects nor interactions in either of the mixed-effects GLMs showed any significant relationship between FR and the factors tested (all  $p > 0.06$ ), indicating that the relationship between LTM accuracy and persistent activity of category cells is specific for the images from the preferred categories of each cell and temporally specific to the WM maintenance period.

### Activity during WM maintenance is linked to subsequent memory signal

During recognition memory tasks, a common observation in the MTL is that “memory-selective” (MS) cells differ in their response between novel and familiar images.<sup>37</sup> These cells are thought to represent a memory strength signal,<sup>37</sup> with stronger responses

### Figure 3. Relationship between WM maintenance activity and LTM formation

(A) Mixed-effects model using the FR obtained during the WM maintenance period in correct and preferred trials of all category cells across both regions, modeling *confidence*, subsequent *LTM accuracy* (remembered vs. forgotten), *area*, and their interactions as fixed effects and *neuron ID* nested into *patient ID* as random intercepts. We found significant modulations of FR by interactions of *confidence* and *LTM accuracy* with *area*, suggesting differences in FR modulations by LTM accuracy and confidence per area.

(B) Mixed-effects GLM results separately for the hippocampus (left) and the amygdala (right). Only in the hippocampus, we observed that persistent activity during the WM maintenance period predicted later LTM accuracy as well as confidence.

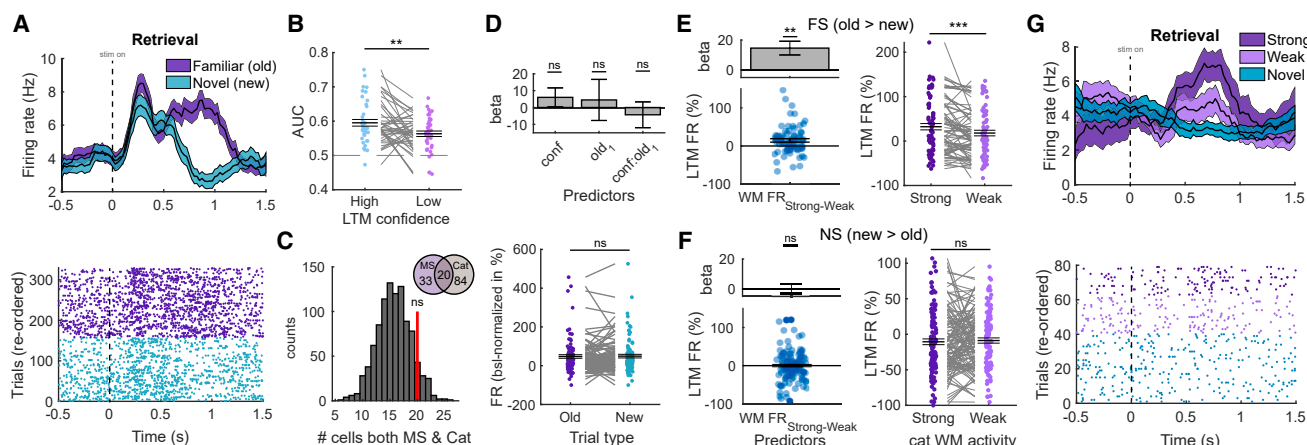
(C and D) Comparison of FRs during the maintenance period between later remembered and forgotten images, separately for high (left) and low-confidence trials (right) for category neurons from (C) the hippocampus and (D) the amygdala. Statistics are permutation-based  $t$  tests, and each dot is a neuron. FRs differed significantly between remembered and forgotten images in the hippocampus for high but not low-confidence trials.

(E and F) Time-resolved FR differences between later high-confident remembered and forgotten trials in (E) the hippocampus and (F) the amygdala for the maintenance period. FR differences between remembered and forgotten trials in the hippocampus were strongest in the first section of the WM delay period (0–650 ms). Cluster-based permutation  $t$  test. Colored areas represent  $\pm$  SEM.  $t = 0$  marks the onset of the maintenance period.

(G) Mixed-effects model results using the WM maintenance period FRs of non-preferred trials across all category neurons from both regions.

(H) Mixed-effects model results using the FRs during the first picture presentation (encoding 1; 0–2 s after picture onset; preferred images and correct trials only) across all category neurons from both regions. In (A), (B), (G), and (H), error bars represent standard errors of the coefficients. Betas are shown in units of baseline-normalized FRs (percent change to baseline, -900 to -300 ms before first picture onset). In (C) and (D), center lines represent mean  $\pm$  SEM. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ; ns, not significant.

See also Figure S4.



**Figure 4. Relationship between memory-selective and category-selective neurons in the hippocampus**

(A) Example of a memory-selective cell that had significantly higher FRs during correct old than new trials in the LTM recognition task. Colored areas in the PSTH plot (top) represent  $\pm$  SEM.  $t = 0$  is stimulus onset.

(B) ROC analysis comparing neuronal response of MS cells between new and old trials separately for low and high confidence. AUC was significantly higher for high- than low-confidence trials. Permutation-based  $t$  test. Each dot is a neuron.

(C) Null distribution of randomly selecting the same number of cells from the entire hippocampal population as category neurons for 1,000 times and determining the overlap with MS cells. The overlap between category cells and MS cells (red bar) was not significantly higher than expected by chance.

(D) Top: mixed-effects GLM model testing for a relationship between FR of category neurons and confidence or familiarity ("old<sub>1-o</sub>"). FR is estimated during picture presentation in the LTM task, and only trials of the preferred category of a given cell are used. Bottom: permutation-based  $t$  test comparing FRs of category neurons during picture presentation between old and new items (preferred category only). Category neurons neither coded for familiarity nor novelty in the LTM task. Each dot is a neuron.

(E) Response of familiarity-selective (FS) MS cells during familiar trials in the LTM task was stronger if maintenance activity for the same image was strong during WM maintenance. Left: mixed-effects GLM. Top: beta extracted from GLM result. Bottom: distribution of FR differences between strong and weak maintenance trials. Right: permutation-based  $t$  test. Each dot is an FS-category cell pair ( $n = 76$ ).

(F) Same as (E), but for novelty-selective (NS) MS cells. The response of novelty-selective MS cells to familiar images did not differ significantly between whether the WM maintenance activity was weak or strong for an image. In (E) and (F), each dot is an NS neuron-category neuron pair ( $n = 122$ ).

(G) Example of a hippocampal FS neuron during LTM retrieval for familiar (old) pictures, split into strong and weak WM activity of a previously recorded category-selective neuron from the same session, and novel items (pictures were from the preferred category of the paired category-selective neuron only). In (B, bottom), (D, bottom), (E, right), and (F, right), we used permutation-based  $t$  tests. Center lines represent mean  $\pm$  SEM. In (D, top), (E, left), and (F, left), error bars represent the standard error of the coefficients. MS, memory-selective; Cat, category-selective; FS, familiarity-selective; NS, novelty-selective; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; ns, not significant.

associated with stronger memories. We therefore next asked whether MS cells are present in our experiment and, if so, whether there is a relationship between the activity of MS cells during LTM retrieval and that of category neurons during WM maintenance. Since we only observed a relationship between persistent activity of category neurons and successful LTM formation in the hippocampus, we restricted this analysis to the hippocampus.

We selected for MS neurons during all correct trials in the LTM recognition part (permutation-based  $t$  test comparing response between correct familiar and novel images, that is, true positives with true negatives, 200–1,200 ms after picture onset,  $p < 0.05$ ). Out of the 351 recorded cells,  $n = 53$  (15.10%) were MS cells. Of these, 25 (47%) responded more to familiar than novel items, with the remaining responding more to novel than to familiar items. Figure 4A shows an example neuron that increased its FR more for familiar than novel pictures (old > new). To assess whether MS cells, as expected, carry a memory strength signal, we next compared their response strength between trials retrieved with high and low confidence. We used receiver operating characteristics (ROCs) analysis to do so (comparing old vs. new trials; see

STAR Methods). The response of MS neurons differed significantly more for high- as compared with low-confidence trials (Figure 4B;  $t(40) = 2.89$ ,  $p = 0.0039$ ; 12 neurons excluded due to insufficient data in either one of the conditions, see STAR Methods). This data shows that MS cells are present and signal memory strength in our experiment.

We next examined whether the overlap between the group of category neurons and MS neurons is significantly higher than chance, possibly hinting toward an involvement of category neurons also in LTM retrieval processes. Of the 53 MS neurons and the 104 category neurons, 20 neurons (37.8% of MS cells; 19.2% of category neurons) were part of both groups. We determined whether this overlap is significantly higher than what would be expected by chance. We did so by randomly selecting the same number of neurons as we found for category neurons from the population of all hippocampal cells and determined how many neurons of those randomly selected neurons were also MS cells. This procedure was repeated for 1,000 times to obtain a null distribution, used to determine the statistical significance of the originally observed overlap. This revealed that the overlap between category neurons and MS cells was not significantly higher than chance (Figure 4C;  $p = 0.11$ ), suggesting that



category and MS cells were statistically independent populations of neurons (as expected<sup>37</sup>).

To further examine whether, as a group, category neurons carried a memory signal during LTM retrieval, we computed a mixed-effects GLM with FRs of category neurons during all correct trials in the LTM recognition task (0.2–1.2 s after picture onset) in which the preferred category of a neuron was shown. We modeled *confidence* (3 levels, sure/unsure/guessing, continuous), *familiarity* (2 levels, old/new, categorical), and their interaction as fixed effects. None of the effects revealed significant modulations of FRs (Figure 4D top; confidence:  $\beta = 6.05$ ,  $p = 0.28$ ; familiarity:  $\beta = 4.49$ ,  $p = 0.71$ ; confidence  $\times$  familiarity:  $\beta = -4.32$ ,  $p = 0.57$ ). To further confirm this result, we directly compared FRs of category neurons between preferred old and new items and did not observe a significant difference (Figure 4D bottom;  $t(94) = -0.28$ ,  $p = 0.78$ ).

Finally, we tested whether the response of MS cells to familiar images during LTM retrieval was correlated with the activity of simultaneously recorded category cells during WM maintenance earlier in the same session while the same picture was held in mind. To do so, we examined all possible pairs of simultaneously recorded MS and category cells ( $n = 198$ ). For each pair, we examined the trials during which images of the preferred category of the category cell were shown both in the WM and in the LTM task (i.e., preferred familiar trials; see STAR Methods). We split the familiar trials in the LTM recognition task into two groups based on the level of activity of the category cell during WM maintenance for the same images (weak vs. strong maintenance activity, median split). We used a mixed-effects GLM to model FRs of the MS cells as a function of the fixed effect *maintenance FR of category cells* (2 levels, strong/weak, categorical; based on simultaneously recorded FR of category neurons in the earlier WM task). We computed separate models for familiarity-selective (old > new) and novelty-selective (new > old) MS neurons as we hypothesized effects to be specific to neurons signaling familiarity.<sup>37</sup> This analysis revealed that FRs of familiarity-selective MS cells during LTM retrieval of items that were previously accompanied by high persistent activity in the WM task were higher than those previously accompanied by low persistent activity (Figure 4E left; maintenance FR:  $\beta = 14.63$ ,  $p = 1.39 \times 10^{-3}$ ). No significant relationship was observed for the activity of novelty-selective MS neurons (Figure 4F left; maintenance FR:  $\beta = 0.09$ ,  $p = 0.98$ ).

To again complement the beta values obtained from our GLM analysis, we next directly compared FRs of MS cells during LTM retrieval for pictures that were maintained earlier with strong vs. weak WM maintenance activity of previously recorded category neurons (from the same session) using permutation-based  $t$  tests. This analysis was performed separately for familiarity- and novelty-selective cells. We observed higher FR for images maintained with strong compared with weak WM activity for familiarity-selective (Figure 4E right;  $t(75) = 3.43$ ,  $p = 0.0003$ ; permutation-based  $t$  test) but not for novelty-selective cells (Figure 4F right;  $t(121) = -0.48$ ,  $p = 0.64$ ). For illustration, Figure 4G shows the activity of an example familiarity-selective neuron during the LTM retrieval task. Familiar (old) images are split into the two groups according to whether the activity of a

simultaneously recorded category neuron for the same images was weak or strong during WM maintenance. The response to novel items from the preferred category of the paired category neuron is shown for comparison. Together, these results suggest that the strength of category-selective WM maintenance activity is correlated with a neuronal measure of LTM strength (the activity of MS cells). These results are in addition to the correlation of WM maintenance activity and later behaviorally assessed LTM strength (previous paragraph).

## DISCUSSION

Our results reveal that the activity of category cells in the hippocampus during WM maintenance was predictive of the success of LTM encoding. This relationship was specific for activity during the maintenance period and to trials in which the preferred category of category-selective cells was maintained in WM. By contrast, there was no significant correlation between the neural activity of category cells during maintenance of stimuli from the non-preferred categories and LTM strength and no significant correlation between activity of the same category cells during the encoding period with later LTM memory strength, indicating specificity to activity during WM maintenance. Further, this effect was specific to the hippocampus as the activity of category neurons in the amygdala was not significantly predictive of successful memory formation. Together, our findings reveal that the neural code used for maintaining items in WM is at least partially overlapping with the neural code that facilitates LTM encoding.

We also observed a relationship with a neuronal measure of LTM memory strength: the stronger the level of persistent activity for a given image, the larger was the response to that same image of MS cells during the recognition memory test (Figure 4). This reveals a direct neuronal-neuronal relationship between activity related to WM maintenance and LTM retrieval. Notably, this neuronal-neuronal relationship was only the case for the MS cells that increased their FR to familiar stimuli. By contrast, the MS cells that increased their FR to novel stimuli showed no significant correlation (for familiar stimuli). This result further supports the argument that what we observed is a signature of memory because the activity of MS cells scales with memory strength (and declared confidence).<sup>37</sup>

Lesion studies indicate that the MTL is not necessary to perform simple WM tasks,<sup>23</sup> which has led to the long-standing idea of parallel memory systems, with the MTL not involved in WM. But if so, why is there persistent activity in the MTL during WM maintenance? One hypothesis is that the purpose of persistent activity is to engage the mechanism used in encoding new memories in order to be able to utilize synaptic plasticity to recover information in case it drops from the focus of attention.<sup>40</sup> Under this framework, persistent activity would enhance the strength of items in LTM.<sup>41</sup> This hypothesis is supported by more recent lesion studies, which show that subjects without a functional MTL do exhibit WM deficits in four situations: (1) in the presence of distractors, (2) when memory load is high, (3) when maintenance time is long,<sup>23</sup> or (4) when making low-confidence judgments.<sup>24,25</sup> In each of these scenarios, the probability that an item will drop out from the focus of attention is

high, and thus the network needs a mechanism for recovering this information. Here, we show that the extent of persistent activity in the hippocampus predicts whether items were encoded into LTM, thus revealing a specific example of a neural mechanism within the hippocampus that is engaged by both the WM and LTM systems. We hypothesize that the role of persistent activity in the hippocampus is to augment the encoding of new information into LTM through repetition of the firing pattern throughout the maintenance period, thereby strengthening long-lasting, long-term potentiation.<sup>41</sup> This hypothesis is supported by theoretical work that indicates that the repetition provided by prolonged activity facilitates the modification of synapses.<sup>42,43</sup>

The response properties of category cells in the amygdala and hippocampus were similar during WM processing but were remarkably different with respect to LTM encoding. In contrast to the hippocampus, activity of category cells from the amygdala did not predict LTM encoding success (Figure 3). The relatively similar tuning properties of neurons in these two areas during encoding are not surprising in the context of prior work. For example, both brain areas contain concept cells<sup>44,45</sup> as well as MS cells.<sup>37,46</sup> Here, we now find that the relationship between short-term memory maintenance and its impact on later LTM is specific to the hippocampus. This is congruent with the fact that the hippocampus is particularly crucial for encoding new memories.<sup>47</sup>

Our findings provide evidence for an interaction between WM and LTM where WM maintenance serves as a gating mechanism for LTM formation. By contrast, earlier research has shown that existing long-term memories can also be retrieved to and maintained in WM.<sup>48</sup> In this study, we used complex visual stimuli that are novel to the participants but could have pre-existing LTM associations for a specific person, animal, or object. It is thus possible that in our study such associations have been retrieved into WM to support the successful maintenance of these images. Since we did not make use of neutral images, such as fractals, without pre-existing associations or have not measured such associations for the current pictures on an individual basis, it is not possible for us to assess how LTM retrieval supported successful WM maintenance. We note, however, that even if this were the case, this would not explain the observed effect because regardless of whether WM maintenance engaged retrieval of existing images or not, encoding of a new memory was required to solve our task. We therefore interpret our findings as indicating that WM maintenance supported LTM formation. On a behavioral level, we observed that WM load had an influence on the successful formation of newly stored long-term memories since images maintained in load 3 trials were less well remembered than images maintained in load 1 trials. Moreover, the strength of persistent activity during the maintenance period predicted successful encoding into LTM, but not the neural activity observed during encoding (see Figure 3). These observations cannot be explained by LTM retrieval processes during the WM maintenance period. Nevertheless, we hypothesize that interactions between the systems likely go in both directions: persistent activity during WM maintenance predicts LTM formation, and, in turn, WM is

supported by the retrieval of pre-existing LTM associations, presumably enhancing content-selective persistent activity during WM maintenance. However, future research is needed to shed more light on these interesting questions.

Our findings further suggest that the neural mechanisms of successful LTM formation overlap with those of WM maintenance. However, we emphasize that this does not mean that the two processes share exactly the same mechanisms such that the distinction between WM and LTM could eventually be discarded. Instead, our findings should be interpreted within a Hebbian view of two distinct memory systems: a short-term memory system that depends on reverberatory activity of cell assemblies and a long-term storage system that involves strengthening of synaptic connectivity between neurons.<sup>49</sup> Our findings suggest that processes of LTM formation, which ultimately lead to successful LTM storage, become enhanced through interactions with persistent activity of WM-selective neurons. The exact mechanistic consequences of such interactions, however, remain the subject of future investigations. It also remains unclear whether other forms of WM maintenance, like activity-silent WM,<sup>50</sup> interact with LTM formation in the same way as persistent neural activity. Activity-silent WM maintenance has mainly been observed for WM content outside the focus of attention,<sup>51,52</sup> which is an important difference from our study in which the focus of attention was not manipulated per item. Earlier research, however, indicated that attention to WM items enhances successful LTM formation,<sup>4</sup> which indicates that persistent activity plays a special role in interactions with LTM formation.

In conclusion, our study reveals that the activity of hippocampal category-selective cells during WM maintenance is predictive of LTM encoding success. This relationship is unique to the hippocampus and the WM maintenance period, with no similar predictive activity observed in the amygdala or during encoding periods. Our findings suggest that persistent activity of category cells in the hippocampus contributes to the encoding of declarative memories, reinforcing the role of Hebbian-type plasticity. They further show that WM- and LTM-specific neural populations interact on a local level with stronger persistent activity predicting stronger memory-related activity during retrieval. These results provide significant insights into the neural mechanisms involved in interactions between WM and LTM with single-cellular resolution.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources should be directed to the lead contact, Ueli Rutishauser ([ueli.rutishauser@cshs.org](mailto:ueli.rutishauser@cshs.org)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

All data used in this study are publicly available in the DANDI Archive (<https://doi.org/10.48324/dandi.001187/0.240904.2347>). The published dataset contains the timestamps and waveforms of the sorted neurons, electrode coordinates, and behavioral data as well as the stimuli, triggers, experimental parameters, and anonymized patient metadata of each session.

Example code to reproduce the results is deposited at GitHub (<https://github.com/rutishauserlab/SBCAT-NO-release-NWB>) and Zenodo (<https://doi.org/10.5281/zenodo.13621888>). The DOI is also listed in the [key resources table](#).

Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## ACKNOWLEDGMENTS

We thank the staff and physicians of the Epilepsy Monitoring Unit at Cedars-Sinai Medical Center for assistance; all patients and their families for their participation; Michael Kyzar, Sophia Cheng, Nand Chandravadia, and Ian Reucroft for spike sorting, data management, and data acquisition; Matthew Chen for behavioral prototyping; and Chrystal M. Reed and Jeffrey Chung for patient care. This work was supported by the National Institute of Health (U01NS117839 to U.R.), a Postdoctoral Fellowship by the German Academy of Sciences Leopoldina (to J.D.), and a Postdoctoral Award by the Center for Neural Science and Medicine at Cedars-Sinai (to J.D.).

## AUTHOR CONTRIBUTIONS

Conceptualization, J.D., J.K., and U.R.; writing – original draft, J.D., J.K., and U.R.; writing – review and editing, all authors. Investigation: J.D., J.K., Y.S., and A.G.P.S.; formal analysis: J.D. and U.R.; methodology, J.D., J.K., and U.R.; funding acquisition, resources, and supervision, J.D., U.R., and A.N.M.; performed surgery, A.N.M., T.A.V., and W.S.A.

## DECLARATION OF INTERESTS

W.S.A. serves as a compensated consultant to Globus Medical, Turing Medical, UniQure, and Iota Biosciences. He receives royalty payments from Globus Medical.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
- [METHOD DETAILS](#)
  - Task
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)
  - Spike sorting
  - Selection of neurons
  - Relating WM maintenance activity to LTM formation
  - Single-neuron AUC analysis
  - Interaction between simultaneously recorded category- and memory-selective neurons
  - Statistics

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.neuron.2024.09.013>.

Received: September 15, 2022

Revised: July 22, 2024

Accepted: September 16, 2024

Published: October 11, 2024

## REFERENCES

1. Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annu. Rev. Psychol.* 63, 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>.
2. Atkinson, R.C., and Shiffrin, R.M. (1968). Human Memory: A Proposed System and its Control Processes. *Psychol. Learn. Motiv.* 2, 89–195. [https://doi.org/10.1016/s0079-7421\(08\)60422-3](https://doi.org/10.1016/s0079-7421(08)60422-3).
3. Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. <https://doi.org/10.1038/nrn1201>.
4. Hartshorne, J.K., and Makovski, T. (2019). The effect of working memory maintenance on long-term memory. *Mem. Cognit.* 47, 749–763. <https://doi.org/10.3758/s13421-019-00908-6>.
5. Souza, A.S., and Oberauer, K. (2017). Time to process information in working memory improves episodic memory. *J. Mem. Lang.* 96, 155–167. <https://doi.org/10.1016/j.jml.2017.07.002>.
6. Daume, J., Gruber, T., Engel, A.K., and Fries, U. (2017). Phase-Amplitude Coupling and Long-Range Phase Synchronization Reveal Frontotemporal Interactions during Visual Working Memory. *J. Neurosci.* 37, 313–322. <https://doi.org/10.1523/jneurosci.2130-16.2017>.
7. Schon, K., Hasselmo, M.E., LoPresti, M.L., Tricarico, M.D., and Stern, C.E. (2004). Persistence of Parahippocampal Representation in the Absence of Stimulus Input Enhances Long-Term Encoding: A Functional Magnetic Resonance Imaging Study of Subsequent Memory after a Delayed Match-to-Sample Task. *J. Neurosci.* 24, 11088–11097. <https://doi.org/10.1523/jneurosci.3807-04.2004>.
8. Ranganath, C., Cohen, M.X., and Brozinsky, C.J. (2005). Working Memory Maintenance Contributes to Long-term Memory Formation: Neural and Behavioral Evidence. *J. Cogn. Neurosci.* 17, 994–1010. <https://doi.org/10.1162/0899929054475118>.
9. Axmacher, N., Schmitz, D.P., Weinreich, I., Elger, C.E., and Fell, J. (2008). Interaction of Working Memory and Long-Term Memory in the Medial Temporal Lobe. *Cereb. Cortex* 18, 2868–2878. <https://doi.org/10.1093/cercor/bhn045>.
10. Blumenfeld, R.S., and Ranganath, C. (2006). Dorsolateral Prefrontal Cortex Promotes Long-Term Memory Formation through Its Role in Working Memory Organization. *J. Neurosci.* 26, 916–925. <https://doi.org/10.1523/jneurosci.2353-05.2006>.
11. Davachi, L., Maril, A., and Wagner, A.D. (2001). When Keeping in Mind Supports Later Bringing to Mind: Neural Markers of Phonological Rehearsal Predict Subsequent Remembering. *J. Cogn. Neurosci.* 13, 1059–1070. <https://doi.org/10.1162/089992901753294356>.
12. Khader, P.H., Jost, K., Ranganath, C., and Rösler, F. (2010). Theta and alpha oscillations during working-memory maintenance predict successful long-term memory encoding. *Neurosci. Lett.* 468, 339–343. <https://doi.org/10.1016/j.neulet.2009.11.028>.
13. Khader, P., Ranganath, C., Seemüller, A., and Rösler, F. (2007). Working memory maintenance contributes to long-term memory formation: Evidence from slow event-related brain potentials. *Cogn. Affect. Behav. Neurosci.* 7, 212–224. <https://doi.org/10.3758/cabn.7.3.212>.
14. Axmacher, N., Elger, C.E., and Fell, J. (2009). Working Memory-Related Hippocampal Deactivation Interferes with Long-Term Memory Formation. *J. Neurosci.* 29, 1052–1060. <https://doi.org/10.1523/jneurosci.5277-08.2009>.
15. Axmacher, N., Lenz, S., Haupt, S., Elger, C.E., and Fell, J. (2010). Electrophysiological signature of working and long-term memory interaction in the human hippocampus. *Eur. J. Neurosci.* 31, 177–188. <https://doi.org/10.1111/j.1460-9568.2009.07041.x>.
16. Hebb, D.O. (1949). *The Organization of Behavior: A Neuropsychological Theory* (Wiley).
17. Squire, L.R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiol. Learn. Mem.* 82, 171–177. <https://doi.org/10.1016/j.nlm.2004.06.005>.
18. Axmacher, N., Henseler, M.M., Jensen, O., Weinreich, I., Elger, C.E., and Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc. Natl. Acad. Sci. USA* 107, 3228–3233. <https://doi.org/10.1073/pnas.0911531107>.

19. Nichols, E.A., Kao, Y.C., Verfaellie, M., and Gabrieli, J.D.E. (2006). Working memory and long-term memory for faces: Evidence from fMRI and global amnesia for involvement of the medial temporal lobes. *Hippocampus* 16, 604–616. <https://doi.org/10.1002/hipo.20190>.
20. Leszczynski, M., Fell, J., and Axmacher, N. (2015). Rhythmic Working Memory Activation in the Human Hippocampus. *Cell Rep.* 13, 1272–1282. <https://doi.org/10.1016/j.celrep.2015.09.081>.
21. Piekema, C., Kessels, R.P.C., Mars, R.B., Petersson, K.M., and Fernández, G. (2006). The right hippocampus participates in short-term memory maintenance of object–location associations. *NeuroImage* 33, 374–382. <https://doi.org/10.1016/j.neuroimage.2006.06.035>.
22. Libby, L.A., Hannula, D.E., and Ranganath, C. (2014). Medial Temporal Lobe Coding of Item and Spatial Information during Relational Binding in Working Memory. *J. Neurosci.* 34, 14233–14242. <https://doi.org/10.1523/jneurosci.0655-14.2014>.
23. Jenson, A., and Squire, L.R. (2012). Working memory, long-term memory, and medial temporal lobe function. *Learn. Mem.* 19, 15–25. <https://doi.org/10.1101/lm.024018.111>.
24. Goodrich, R.I., Baer, T.L., Quent, J.A., and Yonelinas, A.P. (2019). Visual working memory impairments for single items following medial temporal lobe damage. *Neuropsychologia* 134, 107227. <https://doi.org/10.1016/j.neuropsychologia.2019.107227>.
25. Yonelinas, A., Hawkins, C., Abovian, A., and Aly, M. (2024). The role of recollection, familiarity, and the hippocampus in episodic and working memory. *Neuropsychologia* 193, 108777. <https://doi.org/10.1016/j.neuropsychologia.2023.108777>.
26. Borders, A.A., Ranganath, C., and Yonelinas, A.P. (2022). The hippocampus supports high-precision binding in visual working memory. *Hippocampus* 32, 217–230. <https://doi.org/10.1002/hipo.23401>.
27. Xie, W., Chapeton, J.I., Bhasin, S., Zawora, C., Wittig, J.H., Inati, S.K., Zhang, W., and Zaghloul, K.A. (2023). The medial temporal lobe supports the quality of visual short-term memory representation. *Nat. Hum. Behav.* 7, 627–641. <https://doi.org/10.1038/s41562-023-01529-5>.
28. Husain, M. (2024). The human hippocampus contributes to short-term memory. *Brain* 147, 2593–2594. <https://doi.org/10.1093/brain/awae194>.
29. Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349. <https://doi.org/10.1152/jn.1989.61.2.331>.
30. Chafee, M.V., and Goldman-Rakic, P.S. (1998). Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. *J. Neurophysiol.* 79, 2919–2940. <https://doi.org/10.1152/jn.1998.79.6.2919>.
31. Rainer, G., Asaad, W.F., and Miller, E.K. (1998). Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* 393, 577–579. <https://doi.org/10.1038/31235>.
32. Kamiński, J., Sullivan, S., Chung, J.M., Ross, I.B., Mamelak, A.N., and Rutishauser, U. (2017). Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nat. Neurosci.* 20, 590–601. <https://doi.org/10.1038/nn.4509>.
33. Kornblith, S., Quiroga, R.Q., Koch, C., Fried, I., and Mormann, F. (2017). Persistent Single-Neuron Activity during Working Memory in the Human Medial Temporal Lobe. *Curr. Biol.* 27, 1026–1032. <https://doi.org/10.1016/j.cub.2017.02.013>.
34. Daume, J., Kamiński, J., Schjettan, A.G.P., Salimpour, Y., Khan, U., Kyzar, M., Reed, C.M., Anderson, W.S., Valiente, T.A., Mamelak, A.N., and Rutishauser, U. (2024). Control of working memory by phase–amplitude coupling of human hippocampal neurons. *Nature* 629, 393–401. <https://doi.org/10.1038/s41586-024-07309-z>.
35. Boran, E., Fedele, T., Klaver, P., Hilfiker, P., Stieglitz, L., Grunwald, T., and Sarnthein, J. (2019). Persistent hippocampal neural firing and hippocampal–cortical coupling predict verbal working memory load. *Sci. Adv.* 5, eaav3687. <https://doi.org/10.1126/sciadv.aav3687>.
36. Rutishauser, U., Reddy, L., Mormann, F., and Sarnthein, J. (2021). The Architecture of Human Memory: Insights from Human Single-Neuron Recordings. *J. Neurosci.* 41, 883–890. <https://doi.org/10.1523/jneurosci.1648-20.2020>.
37. Rutishauser, U., Ye, S., Koroma, M., Tudusciuc, O., Ross, I.B., Chung, J.M., and Mamelak, A.N. (2015). Representation of retrieval confidence by single neurons in the human medial temporal lobe. *Nat. Neurosci.* 18, 1041–1050. <https://doi.org/10.1038/nn.4041>.
38. Kreiman, G., Koch, C., and Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat. Neurosci.* 3, 946–953. <https://doi.org/10.1038/78868>.
39. Bausch, M., Niediek, J., Reber, T.P., Mackay, S., Boström, J., Elger, C.E., and Mormann, F. (2021). Concept neurons in the human medial temporal lobe flexibly represent abstract relations between concepts. *Nat. Commun.* 12, 6164. <https://doi.org/10.1038/s41467-021-26327-3>.
40. Kamiński, J., and Rutishauser, U. (2020). Between persistently active and activity-silent frameworks: novel vistas on the cellular basis of working memory. *Ann. N. Y. Acad. Sci.* 1464, 64–75. <https://doi.org/10.1111/nyas.14213>.
41. Huang, Y.Y., and Kandel, E.R. (1994). Recruitment of long-lasting and protein kinase A-dependent long-term potentiation in the CA1 region of hippocampus requires repeated tetanization. *Learn. Mem.* 1, 74–82. <https://doi.org/10.1101/lm.1.1.74>.
42. Jensen, O., Idiart, M.A., and Lisman, J.E. (1996). Physiologically realistic formation of autoassociative memory in networks with theta/gamma oscillations: role of fast NMDA channels. *Learn. Mem.* 3, 243–256. <https://doi.org/10.1101/lm.3.2-3.243>.
43. Jensen, O., and Lisman, J.E. (1996). Theta/gamma networks with slow NMDA channels learn sequences and encode episodic memory: role of NMDA channels in recall. *Learn. Mem.* 3, 264–278. <https://doi.org/10.1101/lm.3.2-3.264>.
44. Quiroga, R.Q., Kraskov, A., Koch, C., and Fried, I. (2009). Explicit Encoding of Multimodal Percepts by Single Neurons in the Human Brain. *Curr. Biol.* 19, 1308–1313. <https://doi.org/10.1016/j.cub.2009.06.060>.
45. Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature* 435, 1102–1107. <https://doi.org/10.1038/nature03687>.
46. Rutishauser, U., Schuman, E.M., and Mamelak, A.N. (2008). Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. *Proc. Natl. Acad. Sci. USA* 105, 329–334. <https://doi.org/10.1073/pnas.0706015105>.
47. Squire, L.R., Stark, C.E.L., and Clark, R.E. (2004). The Medial Temporal Lobe. *Annu. Rev. Neurosci.* 27, 279–306. <https://doi.org/10.1146/annurev.neuro.27.070203.144130>.
48. Fukuda, K., and Woodman, G.F. (2017). Visual working memory buffers information retrieved from visual long-term memory. *Proc. Natl. Acad. Sci. USA* 114, 5306–5311. <https://doi.org/10.1073/pnas.1617874114>.
49. Nobre, A.C. (2022). Opening Questions in Visual Working Memory. *J. Cogn. Neurosci.* 35, 49–59. [https://doi.org/10.1162/jocn\\_a.01920](https://doi.org/10.1162/jocn_a.01920).
50. Stokes, M.G. (2015). “Activity-silent” working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405. <https://doi.org/10.1016/j.tics.2015.05.004>.
51. Rose, N.S., LaRocque, J.J., Riggall, A.C., Gosseries, O., Starrett, M.J., Meyering, E.E., and Postle, B.R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science* 354, 1136–1139. <https://doi.org/10.1126/science.aah7011>.
52. Wolff, M.J., Jochim, J., Akyürek, E.G., and Stokes, M.G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nat. Neurosci.* 20, 864–871. <https://doi.org/10.1038/nn.4546>.
53. Minxha, J., Adolphs, R., Fusi, S., Mamelak, A.N., and Rutishauser, U. (2020). Flexible recruitment of memory-based choice representations by

- the human medial frontal cortex. *Science* 368, eaba3313. <https://doi.org/10.1126/science.aba3313>.
54. Tyszka, J.M., and Pauli, W.M. (2016). In vivo delineation of subdivisions of the human amygdaloid complex in a high-resolution group template. *Hum. Brain Mapp.* 37, 3979–3998. <https://doi.org/10.1002/hbm.23289>.
55. Rutishauser, U., Schuman, E.M., and Mamelak, A.N. (2006). Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. *J. Neurosci. Meth.* 154, 204–224. <https://doi.org/10.1016/j.jneumeth.2005.12.033>.
56. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Meth.* 164, 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.



## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Data generated in this study	<a href="http://www.dandiarchive.org">www.dandiarchive.org</a>	<a href="https://doi.org/10.48324/dandi.001187/0.240904.2347">https://doi.org/10.48324/dandi.001187/0.240904.2347</a>
Software and algorithms		
OSORT 4.1	<a href="https://rutishauserlab.org/osort">https://rutishauserlab.org/osort</a>	RRID:SCR_015869
Fieldtrip (ft_freqstatistics.m)	<a href="https://www.fieldtriptoolbox.org">https://www.fieldtriptoolbox.org</a>	Version: 20200409; RRID:SCR_004849
EEGLAB (statcond.m)	<a href="https://sccn.ucsd.edu/eeglab/index.php">https://sccn.ucsd.edu/eeglab/index.php</a>	Version: 13_6_5b; RRID:SCR_007292
MATLAB	The MathWorks, Inc.	RRID:SCR_001622
Code to reproduce results from current study	<a href="http://www.zenodo.org">www.zenodo.org</a>	<a href="https://doi.org/10.5281/zenodo.13621888">https://doi.org/10.5281/zenodo.13621888</a>
Code to reproduce results from current study	<a href="http://www.github.com">www.github.com</a>	<a href="https://github.com/rutishauserlab/SBCAT-NO-release-NWB">https://github.com/rutishauserlab/SBCAT-NO-release-NWB</a>

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

41 patients (48 sessions; 1 binary; 20 females; 20 males; age:  $39.9 \pm 12.9$  years; Table S1; of these, 38 sessions are also analyzed in Daume et al.<sup>34</sup> and 10 are newly added), undergoing invasive monitoring to assess treatment options for drug-resistant epilepsy, participated in the study. Two patients with task performance lower than 55% correct in either the WM or the LTM task were excluded from further analyses (see Table S1). All patients had Behnke-Fried hybrid electrodes (AdTech Inc.) implanted for intracranial seizure monitoring, gave their informed consent, and participated voluntarily. This study was part of an NIH Brain consortium between three institutions (Cedars-Sinai Medical Center, Toronto Western Hospital, and Johns Hopkins Hospital) and approved by the Institutional Review Board of the institution at which the patient was enrolled. Electrode localization was performed using a pre-operative MRI together with either MRI or CT post-operative images and Freesurfer as previously described.<sup>53</sup> Electrode positions are plotted on the CIT168 Atlas Brain<sup>54</sup> in MNI152 coordinates for the sole purpose of visualization (Figure 1B). Coordinates appearing in white matter or outside of the target area is due to usage of a template brain. Electrodes that were localized outside of the target area in native space were excluded from analysis (4 out of a total of 149 recording sites).

## METHOD DETAILS

## Task

The study consisted of two separate tasks: a modified Sternberg WM task followed by a subsequent LTM recognition task. The WM task has been described elsewhere.<sup>34</sup> It consisted of 140 trials and 280 novel pictures. In each trial, the onset of a fixation cross presented for 0.9 to 1.2 s (see Figure 1A top) indicated the start of the trial. The fixation cross was followed by either one (load 1; 70 trials) or three (load 3; 70 trials) consecutively presented pictures, each presented for 2 s. A maintenance period of 2.55 to 2.85 s length followed the picture presentation, during which only the word “HOLD” was shown on the screen. Then, a probe picture was presented, which was either one of the pictures shown earlier in the same trial (match) or a picture already presented in one of the previous trials (non-match). The task was to indicate whether the probe picture matched one of the pictures shown earlier in the same trial or not. Note that the probe image shown was always one that had been shown before and thus familiar, with the answer ‘Yes’ if the image was shown in this particular trial and ‘No’ if it was shown in a previous trial. For trials where the correct answer was ‘No’ (i.e. the probe image was not shown during encoding in this trial), we used images that were presented in a previous trial to assure that all probe images were equally familiar, thereby preventing the use of novelty as a signal to answer the probe question. Probe images in the ‘No’ category were chosen from one of the categories for which no images were shown during encoding in a given trial. The probe picture was shown until patients provided their response via button press. All pictures shown during encoding were novel (i.e., the patient had never seen this particular image) and were drawn from five different visual categories: faces, animals, cars (or tools depending on the version), fruits, and landscapes. In load 3 trials, each image shown during encoding was from a different category.

After a brief delay (lasting 10 - 30min), patients completed a LTM recognition task. During this task, 400 images were shown one at a time. 200 of these images were new (not used in the WM task), whereas 200 were old (previously shown in the WM task). Each trial started with a fixation cross (Figure 1A, bottom), followed by a single image for which the subject was asked to decide whether they had seen this image before (during the WM task) and to indicate the confidence in their response (sure, unsure, guessing). The image

stayed on the screen until a response was given (no timeout). The ‘new’ images (foils) were chosen from the same five visual categories as the ‘old’ images. Note that due to this design, solving the recognition memory task required remembering the specific stimuli seen because the new images used were similar to the old images.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Spike sorting

For each hybrid depth electrode, we recorded the broadband LFP signal between 0.1 and 8,000 Hz at a sampling rate of 32 kHz (ATLAS system, Neuralynx Inc.; Cedars-Sinai Medical Center and Toronto Western Hospital) or 30 kHz (Blackrock Neurotech Inc.; Johns Hopkins Hospital) from a total of eight microwires. All recordings were locally referenced within each recording site by using either one of the eight available micro channels or a dedicated reference channel with lower impedance provided in the bundle, especially when all channels contained recordings of neuronal spiking. We used the semiautomated template-matching algorithm OSort (version: 4.1)<sup>55</sup> to detect and sort spikes from putative single neurons in each wire. Spikes were detected after bandpass filtering the raw signal in the 300 – 3,000 Hz band (see Figure S2 for single cell quality metrics). The two tasks (WM and LTM) were acquired in a single recording and all neurons were jointly sorted for both tasks. In total, we isolated 950 neurons across both areas of the MTL. Neurons with a firing rate lower than 0.1 Hz in either the WM or the LTM tasks were excluded from analysis (67 neurons (7.1%)). Analyses are based on 351 isolated neurons in the hippocampus and 532 in the amygdala (a total of 883 neurons across both areas).

### Selection of neurons

To select for category neurons whose firing rate differed systematically between the picture categories during image presentation (encoding) in the WM task, we counted the number of spikes a given neuron fired in a window between 200 to 1,200 ms after picture onset across all trials in each category (all encoding periods and the probe period). We then computed a 1x5 permutation-based ANOVA with visual category as the grouping variable. In addition, we computed a post-hoc right-sided permutation-based t-test between the category with maximum spike count and all other categories combined. We classified a given neuron as a category neuron if both tests were significant (both  $p < 0.05$ ).<sup>34</sup> We refer to the category with the maximum average spike count as the preferred category of the cell. We note that category cells are selected only using spiking activity from the encoding period, leaving the firing rates during the maintenance period independent for later analyses.

In the recognition task, we selected for neurons that were memory-selective by comparing the number of spikes fired following image onset (window of 0.2–1.2 seconds after image onset) between correct familiar and novel trials using a permutation-based t-test ( $p < 0.05$ , two-sided). Neurons with higher firing rates for familiar than novel items (old > new) were classified as familiarity-selective, the other way round (new > old) as novelty-selective.

### Relating WM maintenance activity to LTM formation

We tested whether category-selective WM maintenance activity predicted LTM formation using a mixed-effects GLM across all category-selective cells. The analysis was performed on baseline-normalized firing rates during the maintenance period (0–2.5 s) of the WM task in trials in which the preferred category of a given neuron was held in WM. We only considered WM trials for which the probe question was answered correctly, and which contained an image that appeared in the subsequent LTM recognition test. We used LTM accuracy (2 levels, correct (remembered)/forgotten (wrong), categorical), LTM confidence (3 levels, high/medium/low, continuous), and area (2 levels, hippocampus / amygdala, categorical) as fixed effects and neuronID nested into patientID as random intercepts.

$$FR \sim 1 + acc * conf * area + (1 | patientID) + (1 | patientID : neuronID)$$

We hypothesized that firing rates during the WM maintenance period should be lower for pictures that were later forgotten (i.e., rated by mistake as “novel”) with high confidence (that is, “high-confidence wrong” trials) than those forgotten with low confidence (“low-confidence wrong” trials). The reason for this hypothesis is that for items that were forgotten with high confidence, there should be a weaker memory trace than for items for which patients were unsure whether they have seen the image before. We therefore labeled the confidence ratings for forgotten trials as high = 1, medium = 2, low = 3. This way the confidence labeling was consistent with our hypothesis. For remembered trials, in turn, we hypothesized that firing rates should be higher for high compared to low confidence trials, so we used the confidence labels high = 3, medium = 2, low = 1. To test whether category-selective activity predicted LTM formation during the encoding window, we based our GLM analysis on firing rates determined during the picture 1 window of preferred images (0–2 s).

### Single-neuron AUC analysis

For MS neurons, we performed ROC analysis to assess how well the firing rate of individual cells distinguished between novel and familiar trials.<sup>37</sup> Spike counts between 200 and 1,200 ms after stimulus onset in the LTM recognition task were used for each neuron’s ROC analysis. We varied the detection threshold between the minimal and maximal spike count observed, linearly spaced in 25 steps. Only neurons with at least ten correct novel and familiar trials each were included. A separate ROC analysis was performed

for high and low confidence trials. Only one of the two groups used for the ROC analysis was modified according to confidence while the other was kept constant. For familiarity-selective neurons, the fixed group was all true-negative trials (regardless of confidence) which was compared with high-confident true-positive and low-confident true-positive trials separately. For novelty-selective neurons, the fixed group was all true-positive trials which were compared with high-confident true-negative and low-confident true-negative trials separately.

### Interaction between simultaneously recorded category- and memory-selective neurons

We determined whether MS neurons were more active during LTM retrieval of a familiar picture when the persistent activity of a simultaneously recorded category-selective neuron for that same picture was also high during the earlier WM maintenance period in the same session. To do so, we median-split the FRs of each category neuron during the maintenance period of correct trials that contained a preferred picture later tested in the LTM task into weak and strong FR trials (separately for load 1 and 3 trials to avoid a bias in FRs across loads). We then tested FRs of simultaneously recorded MS cells during LTM retrieval (determined during 0.2–1.2s after picture onset) between items that have been previously maintained with strong vs weak persistent activity. For that we used a mixed-effects GLM with *Maintenance FR* (2 levels, strong vs weak, categorical) as fixed effect and *neuronID* nested into *patientID* as random intercept.

$$FR_{Retrieval} \sim 1 + Maint. FR + (1 | patientID) + (1 | patientID : neuronID)$$

We performed this analysis separately for familiarity-selective (old > new) and novelty-selective (new > old) cells.

### Statistics

For all statistical tests, we use (cluster-based) non-parametric permutation tests (statcond.m as implemented in EEGLab, using option 'perm', or ft\_freqstatistics.m in FieldTrip), i.e., tests that do not make assumptions about the underlying distributions, or mixed-effects GLMs (fitglm.m in MATLAB) to assess statistical differences between conditions. Before each test, we removed neurons that had no spikes in at least one of the tested conditions (including the baseline period) or differed  $\pm 3$  SD from the mean across all neurons and all tested conditions. This explains differences in degrees of freedom between comparisons. In the permutation-based tests, random permutations of condition labels were performed to estimate an underlying null distribution, which was then used to assess the statistical significance of the effect. All permutations statistics used 10,000 permutations, and t-tests were tested two-sided unless stated otherwise. The corresponding t estimates, which are computed based on a normal distribution, are provided as reference only. Cluster-based permutation statistics were performed as implemented in FieldTrip<sup>56</sup> with 10,000 permutations and an alpha level of 0.025 for each one-sided cluster. Lastly, error bars shown in figures reflect standard errors of the mean for permutation-based t-tests or standard errors of the coefficient for mixed-effect GLM results.