

Unattended working memory items are coded by persistent activity in human medial temporal lobe neurons

Received: 21 December 2023

Accepted: 2 May 2025

Published online: 9 July 2025

 Check for updates

K. Paluch¹, M. Magnuski¹, W. Średniawa¹, D. Ivanovski², A. Rysz³, M. Stuzewska-Niedzwiedz⁴, T. Pasterski³, W. Fortuna⁵, K. Smarzewska⁵, P. C. Reinacher⁶, Sz. Kaczor³, P. Tabakow⁵, H. Babu² & J. Kamiński^{1,2}✉

The role of persistent neural activity in working memory storage is well documented, particularly in tasks where memorized items are of equal importance. However, the encoding mechanism for unattended items, often considered as ‘activity silent’, remains poorly understood. Here we recorded the activity of image-selective neurons in the medial temporal lobe while subjects ($n = 12$) shifted attention between concurrently stored memory items. Our results demonstrate that both attended and unattended memory items are encoded through persistent activity. Additionally, we observed a dynamic transformation in the neuronal subspace following cue presentation, reflecting a shift in how information was maintained. While information about the unattended item was decodable at the single-trial level from preselected image-selective cells, it was not decodable from the entire population of medial temporal lobe cells. These findings support models of persistent activity and challenge the notion that unattended items are stored via ‘activity-silent’ mechanisms.

Working memory (WM) is the ability to retain information in mind in an accessible state, manipulate it and use it to guide behaviour. This cognitive function is essential for performing complex tasks such as reasoning, reading and learning, all of which require completing multiple steps and integrating or analysing available information¹. The prevailing model of WM postulates that information is maintained through an active memory trace characterized by sustained neuronal firing^{2–4}. Evidence supporting this model has been provided by numerous studies in animals^{5–8} and, more recently, humans^{9–11}. Interestingly, in nearly all of these studies, the items to be remembered were equally relevant to the immediate task, suggesting that they were not only maintained, but also probably continuously attended¹². This leaves a gap in our understanding of how the cognitive system operates in everyday situations,

where we frequently maintain information about multiple tasks while prioritizing between them. For instance, when using navigation while driving and simultaneously engaging in conversation, we often need to fully focus on manoeuvres, keeping the conversation in the back of our mind. Once the manoeuvre is completed, we can switch our attention back to the conversation. This demonstrates the ability to prioritize the task at hand while still maintaining unattended WM content and successfully retrieving it when needed. However, it remains unclear whether the cellular mechanisms supporting the storage of prioritized and non-prioritized information are the same.

Alternative mechanisms, beyond the rate code, have been proposed to play an important role in the concurrent coding of multiple WM items¹³. Some focus on the role of spike-phase interactions^{14,15},

¹Center of Excellence for Neural Plasticity and Brain Disorders: BRAINCITY, Nencki Institute of Experimental Biology, Polish Academy of Sciences, Warsaw, Poland. ²Department of Neurosurgery, SUNY Upstate Medical University, Syracuse, NY, USA. ³Department of Neurosurgery, 1st Military Teaching Hospital in Lublin—Branch in Elk, Elk, Poland. ⁴Clinical Department of Neurology, Faculty of Medicine, Wrocław Medical University, Wrocław, Poland. ⁵Clinical Department of Neurosurgery, Faculty of Medicine, Wrocław Medical University, Wrocław, Poland. ⁶Department of Stereotactic and Functional Neurosurgery, Freiburg University Medical Center, Freiburg im Breisgau, Germany. ✉e-mail: jan.kaminski.work@gmail.com

while others suggest that WM items are stored in an activated long-term memory via mechanisms that do not rely on the electrical activity of neurons (in the form of spikes or oscillations). Instead, these mechanisms involve electrically silent, rapid changes in synaptic weights, leading to alterations in functional network connectivity^{16,17}. Recent noninvasive neuroimaging studies have provided compelling evidence supporting the hypothesis of ‘activity-silent’ coding for unprioritized WM items^{18–20}. These studies consistently reported the absence of decodable information related to the unattended memory item during the delay period, despite its successful recall later. This supposedly latent memory trace can be reactivated by presenting a retro cue^{18,19}, a magnetic impulse²⁰ or a visually silent, uninformative stimulus (a so-called ping)²¹. However, some researchers argue that active representations of unattended WM content are absent only in sensory cortices, while they persist in multimodal parietal and frontal regions²². Conversely, other studies have observed active coding of both prioritized and unprioritized items in the sensory cortices, including early visual areas²³. Notably, representations of prioritized and unprioritized items in these regions were found to be rotated relative to each other, suggesting distinct but coexisting coding strategies^{23,24}.

Given the inconsistencies in existing studies and the fact that much of the empirical evidence supporting the ‘activity-silent’ mechanism consists of negative findings, the ‘activity-silent’ hypothesis has faced criticism. For example, if the activity associated with unattended items is significantly weaker than that for attended items, it might simply fall below the detection threshold of imaging methods, which measure the average activity from hundreds of thousands of neurons^{25,26}. Indeed, a recent study demonstrated that an unattended memory trace, previously reported as ‘activity silent’, was successfully decoded in a reanalysis of the same dataset²⁷. This reanalysis focused on alpha-band oscillations instead of a broadband electroencephalogram signal, thereby increasing the signal-to-noise ratio.

To investigate the neuronal mechanisms underlying the storage of unattended WM content, we utilized a double retro-cue paradigm—a task designed to induce internal shifts of attention across multiple items concurrently held in WM²⁸. In this paradigm, a cue is presented after the items have been encoded, indicating which of them should be prioritized. Consistent with cueing effects in the perceptual domain, retro cues enhance performance (shorter reaction times and higher correctness) for the cued items and reduce performance in trials where the previously uncued items are tested²⁹. The close resemblance of the behavioural pattern observed in retro-cue tasks to the well-understood and widely studied cueing effects in the perceptual domain indicates that retro cues are an effective means of shifting attention between items concurrently stored in memory. The dual retro-cue paradigm has been used in noninvasive neuroimaging studies, which reported no delay-period activity for the unattended memory content^{18–20}. However, to our knowledge, it has not yet been used in single-neuron recording studies in either animals or human subjects.

We chose to record the activity of image-selective cells in the human medial temporal lobe (MTL). While many studies in nonhuman primates have demonstrated persistent activity during the maintenance of different types of information in WM across various brain areas—such as the dorsolateral prefrontal cortex for spatial locations^{4,5}, the entorhinal cortex for complex visual stimuli⁶ and the pre-supplementary motor area (pre-SMA) for tactile and acoustic stimuli³⁰—in humans, stimulus-specific persistent activity during WM maintenance has, to our knowledge, only been observed in recordings from image-selective cells³¹ in the MTL^{9,11}. The importance of the MTL in WM has also been reported in multiple studies using intracranial electroencephalography^{32–34}. In the original discovery of the ‘activity-silent’ mechanism for maintaining unprioritized WM content¹⁹, a supplementary analysis confined to the inferotemporal cortex, including the hippocampus and amygdala, yielded results similar to the whole-brain analysis: information about item identity was

decodable from the inferotemporal cortex when it was in the focus of attention, but not when it was outside of attention. This makes the MTL an ideal candidate for testing proposed hypotheses of ‘activity-silent’ coding of unattended WM items from both invasive and noninvasive study perspectives.

By recording the activity of image-selective cells during the dual retro-cue task, we aimed to study the impact of attention on firing rates when their preferred image was held in WM. If cell activity during the maintenance period for the unattended stimulus drops to the spontaneous level, yet the item is still correctly retrieved from memory, this would provide the strongest evidence so far for the ‘activity-silent’ coding of unattended WM items. Conversely, if persistent neuronal firing is observed for the unattended memory item, it would extend previous findings by demonstrating that the persistent activity of image-selective cells in the MTL supports the entire WM content, regardless of its attentional state.

Results

Task and behaviour

We administered a double retro-cue paradigm adapted for recording MTL image-selective cells (Fig. 1a) to epilepsy patients ($n = 13$ sessions from 12 subjects). Before the WM task, we ran a screening task to identify the four images eliciting the strongest selective responses; these images were then used as the memory set for the double retro-cue experiment (for details, see Methods; for example screening results, see Supplementary Fig. 1). In each trial of the WM task, subjects memorized two images presented concurrently above and below a fixation point for 3 s. A spatial cue (a horizontal line presented for 1 s) indicated which item should be prioritized for the following detection task. After a 5 s maintenance period, during which the representations of the attended (cued memory item) and unattended (uncued memory item) could be distinguished, subjects were presented with a rapid serial visual presentation (RSVP) stream of eight images (each displayed for 0.15 s). Subjects judged whether the cued image was present in the stream. A second spatial cue was then presented. In half of the trials, it indicated the same item as the first cue (‘same’ target trials), while in the other half, it indicated the previously uncued item (‘switch’ target trials). As with the first cue, the second cue was followed by a maintenance period, RSVP stream and response screen. The presence of switch trials ensured that the uncued memory item had to be retained in memory, as it might be used as a target for the second detection task.

Subjects performed well overall, with an average correctness of $87.82 \pm 9.31\%$ (\pm s.d.). Performance varied by target type (first/same/switch) ($F_{2,24} = 14.07, P < 0.001$, partial $\eta^2 = 0.54$, 95% confidence interval (CI) 0.10 to 0.60, permuted repeated-measures analysis of variance (ANOVA); note that ANOVA and t -test values are provided for reference, as P values are derived from permutations; Fig. 1b). Correctness was lower for the previously uncued item (‘switch’, $81.41 \pm 13.24\%$) compared with the previously cued item (‘same’, $88.94 \pm 8.73\%$; $t_{12} = 4.21, P = 0.002$, Cohen’s $d = 1.17$, CI 0.74 to 2.02) and the first target ($90.46 \pm 7.06\%$; $t_{12} = 3.88, P < 0.001$, Cohen’s $d = 1.076$, CI 0.66 to 1.78). The higher correctness in the same-target trials could not be attributed to refreshing the target representation through its brief reappearance in the first RSVP stream during target-present trials, as the same effect was observed when the analysis was limited to target-absent trials ($t_{12} = 2.62, P = 0.014$, Cohen’s $d = 0.73$, CI 0.31 to 1.28; same out: $88.14 \pm 11.89\%$ versus switch out: $81.41 \pm 15.83\%$; Supplementary Fig. 2). Notably, both same and switch targets had equal priority during encoding (the cue was presented after the images disappeared from the screen) and were maintained in WM for the same period of time (approximately 12–14 s, depending on the reaction time to the first probe). Therefore, the lower correctness for previously uncued (switch-target) items can only be attributed to their different attentional status, indicating that subjects followed task instructions and prioritized the cued item in memory.

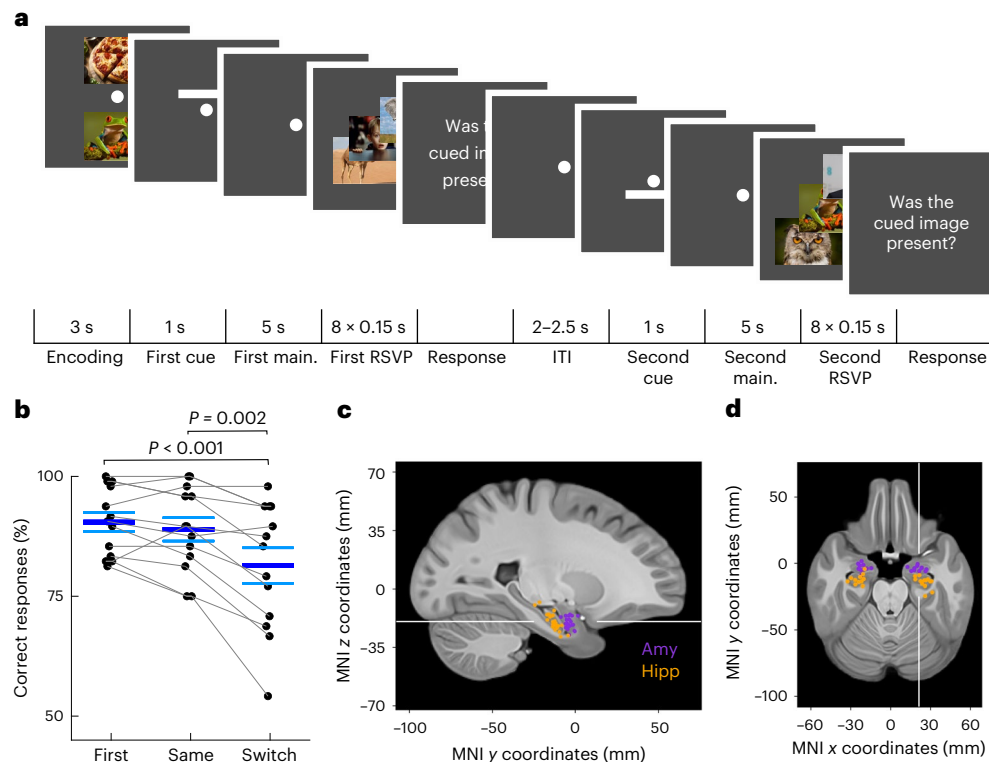


Fig. 1 | Task, recording locations and behavioural results. a, The double retro-cue task: each trial began with two images (one above and one below the fixation point), followed by a spatial cue indicating the target for the detection task. After a 5 s maintenance (main.) period, a stream of eight distractor images was presented (images are presented sequentially; only three are depicted for simplicity). Subjects indicated whether the cued image appeared in the stream. After their response, a second cue indicated either the same image or the previously uncued image. This was followed by another maintenance period, a target detection task and a response screen. **b**, The percentage of correct

responses ($n = 13$ sessions) for the first target, second previously cued target ('same') and second previously uncued target ('switch'). The dark and light blue lines represent the mean and s.e.m., respectively. Individual sessions are connected by grey lines. A permuted F test and two-sided paired t -tests were used to test differences between conditions. **c, d**, Recording site locations in MNI152 space, overlaid on the CIT168 T1 brain atlas. Sagittal slice, for plotting purposes MNI x coordinates for all sites were set to -21 (**c**). Horizontal slice, for plotting purposes MNI z coordinates were set to -20 (**d**). The white lines indicate the locations of axial and sagittal slices. Amy, amygdala; hipp, hippocampus.

We also analysed reaction times. One subject was excluded from the reaction time analysis due to unrecorded reaction times in target-absent trials caused by technical issues. Reaction times for correct responses were similar across first (1.61 ± 1.45 s), same (1.71 ± 1.76 s) and switch (1.84 ± 2.01 s) targets (permuted repeated-measures ANOVA: $F_{2,22} = 1.76$, $P = 0.159$, partial $\eta^2 = 0.13$, CI 0 to 0.25).

Our behavioural results indicate that subjects followed the spatial cues and successfully shifted attention between concurrently stored memory items. The observed correctness pattern closely aligns with previous findings from neuroimaging studies employing the double retro-cue paradigm^{18–20}.

Identification of image-selective cells

Building on previous research^{9,11,31}, which observed persistent neuronal activity during the maintenance of attended WM content, we utilized the encoding period to identify image-selective cells. To avoid any dependencies between cell selection and hypothesis testing (conducted on adjacent time windows: encoding versus maintenance), we applied a leave-one-out approach. The selectivity of each cell was tested iteratively ($n = 96$, total number of trials). In each iteration, cell selectivity was assessed on $n - 1$ trials, and if the selectivity criteria were met (Methods), maintenance effects were tested on the one trial excluded from the selectivity definition. We identified $n = 6,339$ trials as image-selective (15.68%, $P < 0.001$ compared with the surrogate data distribution). Additionally, 42 cells were consistently selective for the same image across all 96 iterations (9.99%, $P < 0.001$ compared with the surrogate data distribution). The full distribution of image selectivity across neurons

is shown in Supplementary Fig. 3a,b. Differences between conditions expected from the null distribution (obtained from data with scrambled labels) are plotted in Supplementary Fig. 3c.

Persistent neural activity for cued and uncued WM items

To account for the different number of trials entering the analysis from different neurons (Supplementary Fig. 3a), we used linear mixed-effects models (LMEs). We tested whether the activity of image selective cells depended on the image type: preferred image cued in WM, preferred image uncued in WM, or images other than the preferred one held in WM (other). The hierarchical structure of the models accounted for the nested nature of the data, where multiple trials were recorded from each neuron and multiple neurons were recorded from different subjects. Trials in which images other than the preferred image were presented and held in WM were used as the reference condition. To balance the model we required that all included cells had at least one trial in each of the compared conditions (yielding $n = 79$ neurons; from each cell, only trials defined as selective based on the leave-one-out cross-validation were included).

As per definition during encoding, we observed significant increase in firing rate when the preferred image was displayed. This increase was evident for both the later cued ($\beta = 0.359$, s.e.m. 0.034; $t_{4,936} = 10.58$, $P < 0.001$, CI 0.29 to 0.42, standardized beta $\beta' = 0.152$) and the later uncued item ($\beta = 0.390$, s.e.m. 0.034; $t_{4,936} = 11.47$, $P < 0.001$, CI 0.32 to 0.46, $\beta' = 0.165$), with no significant difference between these two conditions ($\beta = 0.031$, s.e.m. 0.039, $t_{4,936} = 0.796$, $P = 0.426$, CI -0.05 to 0.11 , $\beta' = 0.013$; Fig. 2a,c).

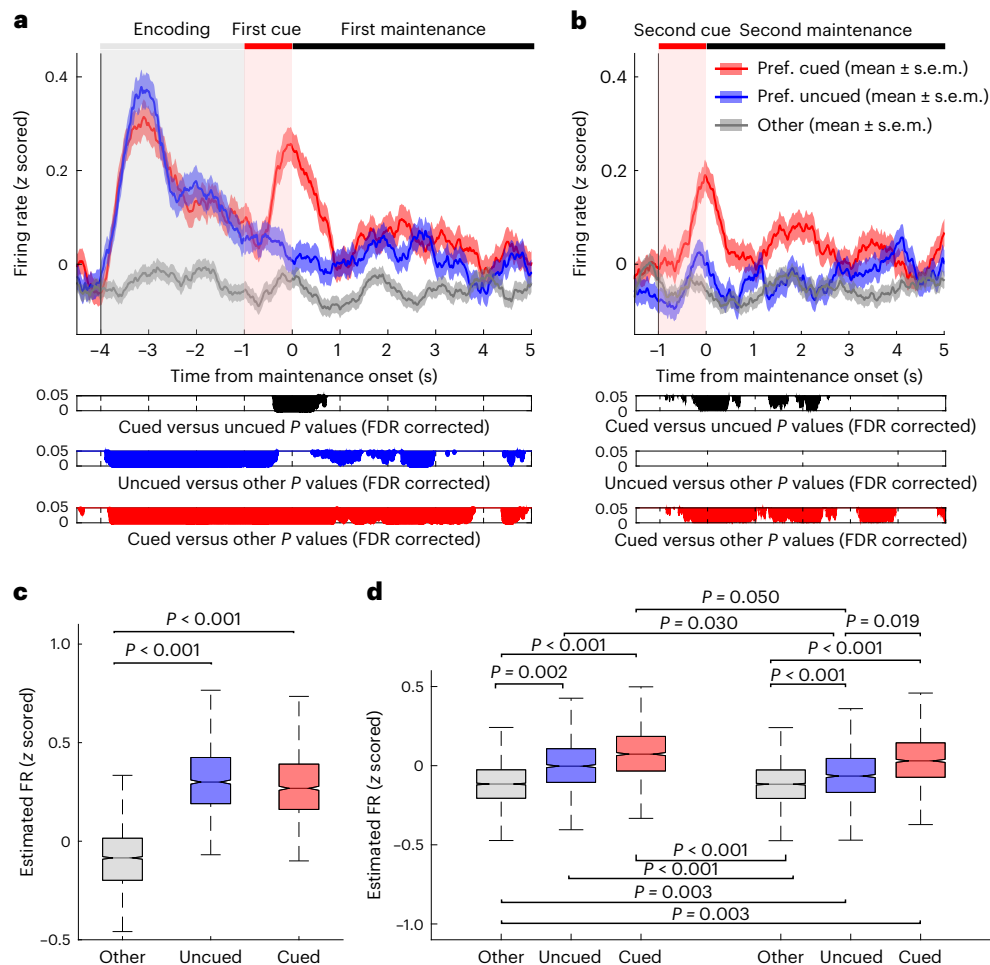


Fig. 2 | Image-selective cells in the human MTL are persistently active during the maintenance of both cued and uncued WM content. The average firing rate of image-selective neurons ($n = 79$; only trials identified as image-selective were included) during the double retro-cue task. **a**, A PSTH (bin size, 500 ms; step size, 1 ms) showing activity during encoding, the first cue presentation and the first maintenance period. **b**, A PSTH (bin size, 500 ms; step size, 1 ms) showing activity during the second cue presentation and the second maintenance period. The lines represent averages and shaded areas represent \pm s.e.m. across all included trials recorded from all neurons. Highlighted are significant differences in activity for preferred (pref.) cued, preferred uncued and other

conditions, computed using the LME at individual timepoints (bottom). P values were corrected for multiple comparisons using FDR correction. **c,d**, Box plots illustrate the average ($n = 79$ neurons), conditional firing rates (FR; which account for both fixed and random effects) during encoding, as estimated in the LME with one predictor: condition (**c**) and both maintenance periods, as estimated in the LME with two predictors: condition and time (**d**). Differences between predictor levels were evaluated using F tests on the fixed effects estimates, all tests were two sided and FDR corrected. The central line in each box plot represents the median, the box marks the interquartile range and whiskers extend to 1.5 times the interquartile range.

In the analysis of activity during maintenance periods, we added time (first maintenance and second maintenance) as an additional predictor to our model. The model revealed increased firing rates whenever the preferred item was held in WM, regardless of whether it was cued ($\beta = 0.23$, s.e.m. 0.055, $t_{9,872} = 4.19$, $P < 0.001$, CI 0.12 to 0.34, $\beta^* = 0.134$) or uncued ($\beta = 0.18$, s.e.m. 0.055, $t_{9,872} = 3.36$, $P < 0.001$, CI 0.08 to 0.29, $\beta^* = 0.105$) compared with the reference condition; Fig. 2a,d). There was no overall change in neuronal activity across the two maintenance periods (main effect of time: $\beta = -0.001$, s.e.m. 0.020, $t_{9,872} = -0.052$, $P = 0.958$, CI -0.04 to 0.04 , $\beta^* = 0$). However, the interaction between the uncued item and time was close to the significance threshold (interaction preferred uncued \times time: $\beta = -0.065$, s.e.m. 0.035, $t_{9,872} = -1.87$, $P = 0.061$, CI -0.13 to 0 , $\beta^* = -0.029$). No such effect was observed for the cued item (interaction preferred cued \times time: $\beta = -0.038$, s.e.m. 0.034, $t_{9,872} = -1.11$, $P = 0.266$, CI -0.10 to 0.03 , $\beta^* = -0.017$; Fig. 2b,d). Post hoc tests showed that activity for the cued and uncued items was similar during the first maintenance period ($F_{1,9,872} = 0.515$, $P = 0.546$; false discovery rate (FDR) corrected, Cohen's $d = 0.065$). However, during the second maintenance period, activity

for the uncued item was smaller than for the cued one ($F_{1,9,872} = 6.45$, $P = 0.019$; FDR corrected, Cohen's $d = 0.103$), as well as smaller than the activity for both the cued ($F_{1,9,872} = 4.36$, $P = 0.050$; FDR corrected, Cohen's $d = 0.158$) and the uncued items ($F_{1,9,872} = 5.39$, $P = 0.030$; FDR corrected, Cohen's $d = 0.094$) during the first maintenance period. This decrease was not due to the passage of time from encoding, as no such decrease was observed for the cued item (first preferred cued versus second preferred cued: $F_{1,9,872} = 1.24$, $P = 0.332$; FDR corrected, Cohen's $d = -0.055$). To sum up, our results indicate that both cued and uncued items evoked persistent neuronal firing during their maintenance in WM and suggest a reduction in activity for the uncued item after the second cue, when it became task irrelevant and could be dropped from WM (for pairwise comparisons of all conditions see Fig. 2d).

Additionally, we observed a phasic, cue-related increase in firing rate for the cued item during the first maintenance (until 832 ms after cue offset) and during the second maintenance period (until 438 ms after cue offset; Fig. 2a,b). As we were interested in how items were maintained within and outside the focus of attention, and to preclude any contamination by response to cue, we additionally tested firing

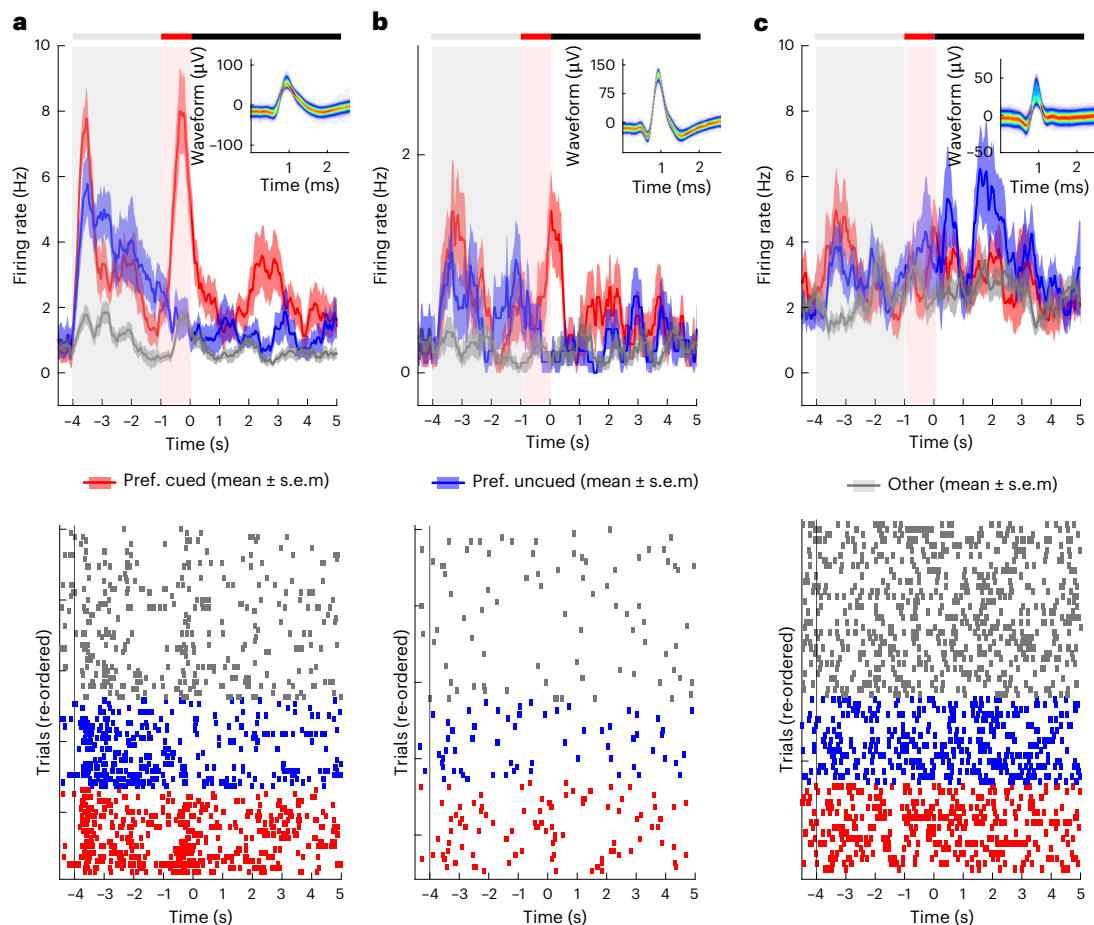


Fig. 3 | Examples of activity from image-selective cells recorded during the double retro-cue task. a–c. Three examples of image-selective cells: with greater activity for the preferred (pref.) cued item (a), similar activity for both pref. cued and pref. uncued items (b) and greater activity for pref. uncued item (c).

Top: PSTH (bin size, 500 ms; step size, 1 ms) during encoding, the first cue and the first maintenance period. The shaded areas represent \pm s.e.m. across trials. Bottom: raster plots of trials, re-ordered by condition only for plotting purposes. Insets: mean \pm s.e.m. extracellular waveform of all spikes recorded for each cell.

rate activity between 1 and 5 s in the first maintenance period and the results were qualitatively the same (Supplementary Fig. 4). For all further analyses we used an average firing rate between 1 and 5 s after maintenance onset.

Next, to test whether the observed neuronal activity was relevant to task performance, we included response accuracy (correct versus incorrect) as an additional predictor in our LME model (Methods). The model was computed separately for the first and second maintenance periods. During the first maintenance period, activity for both the cued (interaction preferred cued \times response: $\beta = -0.189$, s.e.m. 0.068, $t_{6,032} = -2.77$, $P = 0.006$, CI -0.32 to -0.06 , $\beta' = -0.095$) and uncued items (interaction uncued \times response: $\beta = -0.146$, s.e.m. 0.069, $t_{6,032} = -2.13$, $P = 0.033$, CI -0.28 to -0.01 , $\beta' = -0.073$) was significantly lower in trials that ended with incorrect compared with correct responses (Supplementary Fig. 5a,b). A similar pattern was observed during the second maintenance period. In incorrect trials, activity decreased for both the cued (interaction preferred cued \times response: $\beta = -0.164$, s.e.m. 0.066, $t_{6,031} = -2.47$, $P = 0.014$, CI -0.29 to -0.03 , $\beta' = -0.085$) and uncued items (interaction preferred uncued \times response: $\beta = -0.147$, s.e.m. 0.066, $t_{6,031} = -2.24$, $P = 0.025$, CI -0.28 to -0.02 , $\beta' = -0.077$; Supplementary Fig. 5c,d). These results indicate that image-specific persistent activity was behaviourally relevant for items maintained in both states—whether within or outside the focus of attention.

We ran two control analyses, which showed that persistent activity for the uncued item during the first maintenance period was present regardless of brain area (amygdala or hippocampus) and was

independent of the neurological condition of our subjects (that is, it was observed in both epileptogenic and healthy tissue; Supplementary Results and Supplementary Fig. 6). Furthermore, additional, control analyses confirmed that the activity for the uncued item did not change based on the selectivity of cells, indicating that it could not be attributed to the response of cells to the other image, which was concurrently in the focus of attention (for details, see Supplementary Results and Supplementary Figs. 7–9).

In summary, we observed persistent neuronal activity for both cued and uncued memory items during the first maintenance period, when both items were kept in memory. This activity was task relevant as it was greater in correct than in incorrect trials. Activity for the uncued item decreased during the second maintenance period, when the item was no longer task relevant and could be removed from memory. Thus, our data extend previous findings^{9,11}, showing that the persistent activity of MTL image-selective cells is a common mechanism underlying the storage of the entire WM content, regardless of the current focus of attention.

Cued and uncued items populate different neuronal subspaces

The presented analyses demonstrate that both cued and uncued WM content are stored using a rate code. However, inspection of single-neuron responses revealed distinct activity patterns during the maintenance period (see examples in Fig. 3a–c). Some image-selective neurons exhibited a clear response to the cue presentation, followed by greater activity for the cued items (Fig. 3a). Others responded similarly

to both cued and uncued items (Fig. 3b). Finally, some neurons showed stronger responses to the uncued WM items (Fig. 3c).

We investigated whether the variability in neuron response patterns could be attributed to variability among subjects. To control for between-subject variability, we averaged the activity of all cells recorded from each subject and re-ran the firing analysis for the encoding and first maintenance periods (Supplementary Fig. 10). This analysis confirmed the presence of persistent activity for both cued and uncued items. Thus, the observed variability appears to be intrinsic to the neural circuitry involved, rather than being attributable to differences in the behavioural strategies of subjects. This argumentation is further strengthened by the control analyses mentioned later in the text (demixed principal component analysis (dPCA) results).

This suggests that variability in these responses could serve as a mechanism for encoding item prioritization within the network. If this is the case, we would expect the cued and uncued states to occupy distinct regions of the neuronal subspace, making them differentiable even when the overall average activity appears indistinguishable. To test this hypothesis, we applied dimensionality reduction techniques at the population level, specifically dPCA³⁵. This method has previously been used to demonstrate that neuronal firing patterns create steady-state attractors during the maintenance of attended WM items⁹. In our analysis, the image type (preferred image cued/preferred image uncued/non-preferred image) was used as the marginalized variable, and only correct trials were included. We restricted the analysis to cells with at least 17 trials per image type ($n = 35$, recorded from 8 subjects). The basis functions for dPCA were computed using 6 s epochs encompassing the first cue and the maintenance period. We found that a three-dimensional space formed by the first three demixed principal components separated neural trajectories across the analysed trial types. The first component (dPC1) accounted for 38.08%, the second component (dPC2) for 27.62%, and the third component (dPC3) for 10.30% of the total data variance.

To test whether persistent activity for both cued and uncued WM items is present at the level of single trials, and whether it is possible to distinguish between cued and uncued states, we performed a decoding analysis. Decoding axes of the first three dPCs were used as linear classifiers to decode (1) the presence of the cued item in WM (cued versus other), (2) the presence of the uncued item in WM (uncued versus other) and (3) whether the preferred item was in the focus of attention or outside it (cued versus uncued). We employed leave-one-out, mean-based decoding for each component separately, as implemented in the dPCA toolbox³⁵. Decoders were trained on averaged activity during the first maintenance period (1–5 s). We performed 500 cross-validation iterations and compared the results with the accuracy of analogous decoders trained on data with scrambled labels (for details, see Methods).

The first component represented activity related to the entire memory content (attended and unattended) but was stronger for the cued item than for the uncued item. Using dPC1, we achieved an accuracy of 87.7% in decoding the presence (versus other) of the cued preferred item in WM ($P < 0.001$, estimated using scrambled data; FDR corrected, chance was 50%). This component also carried substantial information about the uncued WM item, enabling a decoding accuracy of 71.1% ($P < 0.001$, estimated using scrambled data; FDR corrected, chance was 50%). However, the signal for the cued item was significantly stronger than for the uncued item, resulting in 76.9% accuracy in decoding whether the item was inside or outside the focus of attention ($P < 0.001$, estimated using scrambled data; FDR corrected, chance was 50%; Fig. 4a,b).

The second component captured activity specific to the cued WM item. The average activity of this component during the first maintenance period was highest for the cued preferred item, intermediate for the other item and lowest for the uncued item. Using the activity of this dPC as input for our three decoders, we were able to accurately classify whether the preferred item was held in WM when it was in the focus

of attention (cued versus other: 67.7% accuracy, $P = 0.008$, estimated using scrambled data, FDR corrected, chance was 50%). However, classification was not accurate when the preferred item was outside the focus of attention (uncued versus other: 56.6% accuracy, $P = 0.293$, estimated using scrambled data, FDR corrected, chance was 50%). The difference in activity between the cued and uncued items was significant, allowing for 71.2% accuracy in decoding whether the preferred item was cued or uncued ($P < 0.001$, estimated using scrambled data, FDR corrected, chance was 50%; Fig. 4c,d).

The third component reflected activity specific to the uncued memory content. Its amplitude was higher for the uncued preferred item compared with the non-preferred images, enabling successful decoding with 66.6% accuracy ($P < 0.001$, estimated using scrambled data, FDR corrected, chance was 50%). The other two decoders, when applied to activity from dPC3, both performed at chance level (cued versus other: 57% accuracy, $P = 0.197$, and cued versus uncued, 58% accuracy, $P = 0.197$, estimated using scrambled data, FDR corrected, chance was 50%; Fig. 4e,f).

To test whether the presence of two distinct prioritization components (dPC2 and dPC3) resulted from different behavioural strategies adopted by different subjects (for example, focusing on the uncued item), we performed two control analyses. First, we performed a median split on the weights of each component and counted the proportion of subjects whose neurons had bigger-than-median weights on each of these components. If these components reflected different task strategies, neurons from different subjects should have loaded preferentially one component or the other. However, our analysis showed that six out of eight subjects included in the analysis loaded onto dPC2 and seven out of eight subjects loaded onto component dPC3 (with five subjects having higher-than-median weights on both components). Second, for each subject, we computed their average contribution to each of the two components (that is, the average weight of all neurons recorded from that subject) and correlated it with the behavioural cost of attention switching (defined as the difference in the percentage of correct responses between same and switch trials). No significant relationship was observed (dPC2: rho of 0.157, $P = 0.711$ and dPC3: rho of 0.506, $P = 0.206$). In summary, the two prioritization components did not reflect different behavioural strategies, but rather represented distinct ways in which cued and uncued information was coded in the studied population of cells.

The results of the dPCA analysis indicate that both attended and unattended WM content are coded via the same neuronal mechanism—the persistent firing rate of image-selective neurons. However, they also reveal plausible mechanisms of attention prioritization. First, the magnitude of the common persistent activity is modulated by the focus of attention (dPC1). Second, within MTL cell activity, there exists a subspace specifically active only for the cued WM content and another part of the network that codes exclusively for the uncued WM item (dPC2 and dPC3). Contrary to previous noninvasive studies using similar double retro-cue tasks^{18–20}, we were able to decode information about both the cued and uncued items during the first maintenance period from the activity of image-selective cells.

Prioritization of cued items during response to the probe

To understand the functional meaning of the observed dPCs, we tested their behaviour during probe presentation. Since dPC1 encoded information about both memory states (cued and uncued, representing the memory component), while dPC2 and dPC3 reflected information about a single prioritization state (attentional components), we hypothesized that these components would behave differently when a preferred item was presented as a probe, depending on whether it was prioritized or not. To test this, we projected the weights of dPC1 and dPC2 onto neuronal activity during probe presentation (dPC3 was not included in this analysis, as in trials where the preferred item was uncued, it was never presented as a probe; Methods). Trials were

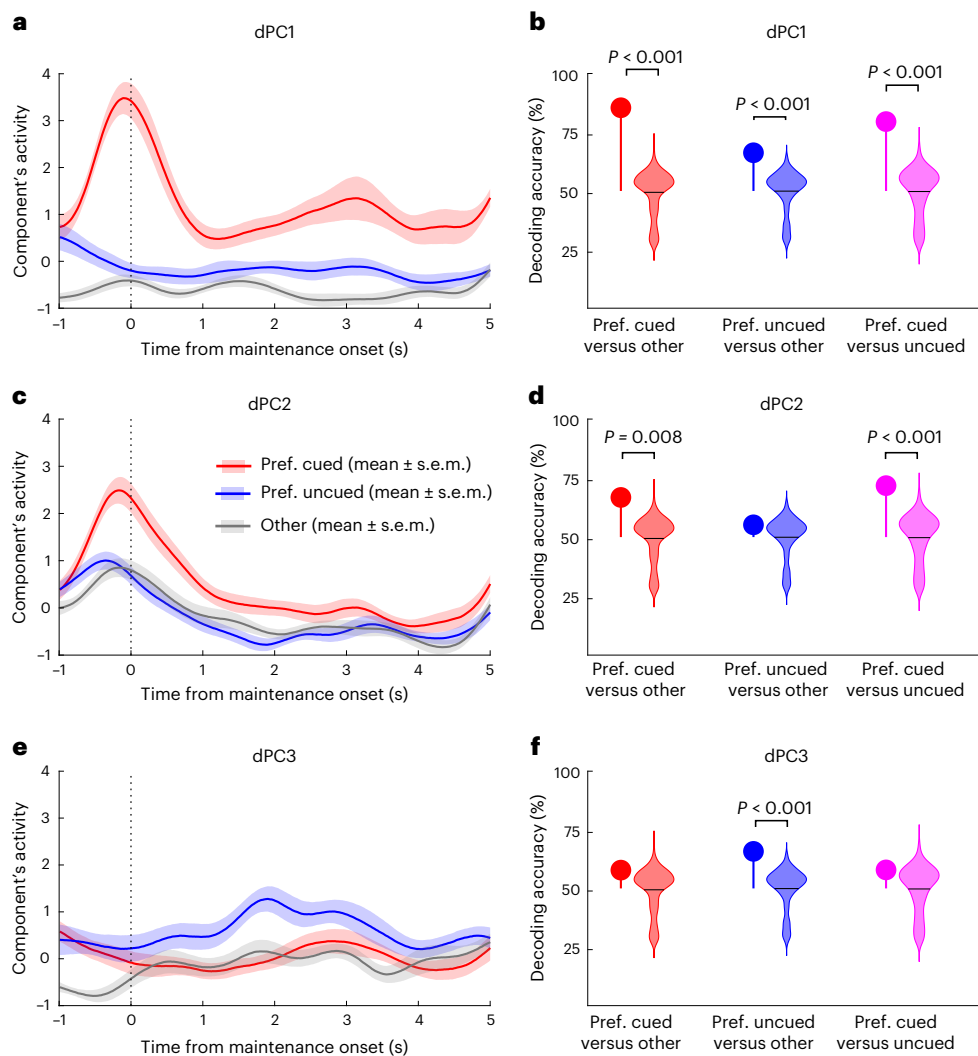


Fig. 4 | Cued and uncued items populate different neuronal subspaces. a, c, e. The activity of the first three dPCs during the first cue and maintenance period computed on the entire dataset: dPC1 reflects the entire memory content, with its amplitude modulated by the focus of attention (a); dPC2 represents the content in the focus of attention (c) and dPC3 reflects unattended memory content (e). The shaded areas represent \pm s.e.m. across trials. **b, d, f.** The accuracy of three decoders trained on the average activity of each component (dPC1 (b), dPC2 (d) and dPC3 (f)) during the maintenance period (1–5 s). Decoding accuracies obtained from image-selective cells are represented as stem plots, while violin plots show the null distributions (estimated from scrambled data; 500 permutations). The black lines indicate the mean. The chance level was 50%. All significance levels for all analyses were determined using one-sided permutation tests assessing whether decoding accuracy was greater than expected under the null hypothesis. Reported P values are FDR corrected.

categorized into four conditions based on image type (preferred item cued/non-preferred item) and two types of probe (preferred item in/preferred item out).

dPC1 exhibited an increase in activity during all presentations of the preferred item, both when it was previously cued (cued preferred in) and when it was not part of the current memory set but was presented as a lure (other preferred in; Fig. 5a). A decoder trained on the activity of this component successfully detected the presence of the preferred item in the stream of distractors in both the cued condition (accuracy: 81.8%, $P < 0.001$, estimated using scrambled data, FDR corrected, chance level 50%; Fig. 5a,b) and the other condition (accuracy: 76%, $P < 0.001$, estimated using scrambled data, FDR corrected, chance level 50%; Fig. 5d). In contrast, dPC2 showed an increase in activity during presentations of the preferred item in the stream of distractors only when it was previously cued (Fig. 5c), enabling successful classification exclusively for trials from the cued condition (accuracy 82.1%, $P < 0.001$, estimated using scrambled data, FDR corrected, chance level 50%; Fig. 5b). The activity of dPC2 when the preferred item was present in the distractor stream in the other

condition did not differ from the activity for the non-preferred item (accuracy 43.7%, $P = 0.886$, estimated using scrambled data, FDR corrected, chance was 50%; Fig. 5d).

Next, we directly compared the performance of the decoders for dPC1 and dPC2. The difference in decoding accuracy between the cued and other trial classifications for dPC1 (5.8%) was significantly lower than that for dPC2 (38.4%). To determine whether dPC2 is specifically involved in prioritizing the item currently in the focus of attention, we compared these differences with those expected from scrambled data. This analysis revealed a significant interaction between condition (cued versus other) and component (dPC1 versus dPC2), with a difference in accuracy for the cued versus other conditions of 32.6% ($P = 0.002$, estimated using scrambled data; Fig. 5e).

dPC1 exhibited a more uniform response to both cued and uncued probes, suggesting its involvement in processing less prioritized or background information. In contrast, dPC2 displayed higher activity when its preferred item was presented as a target in the subsequent visual search task, indicating its role in the dynamic prioritization of task-relevant stimuli.

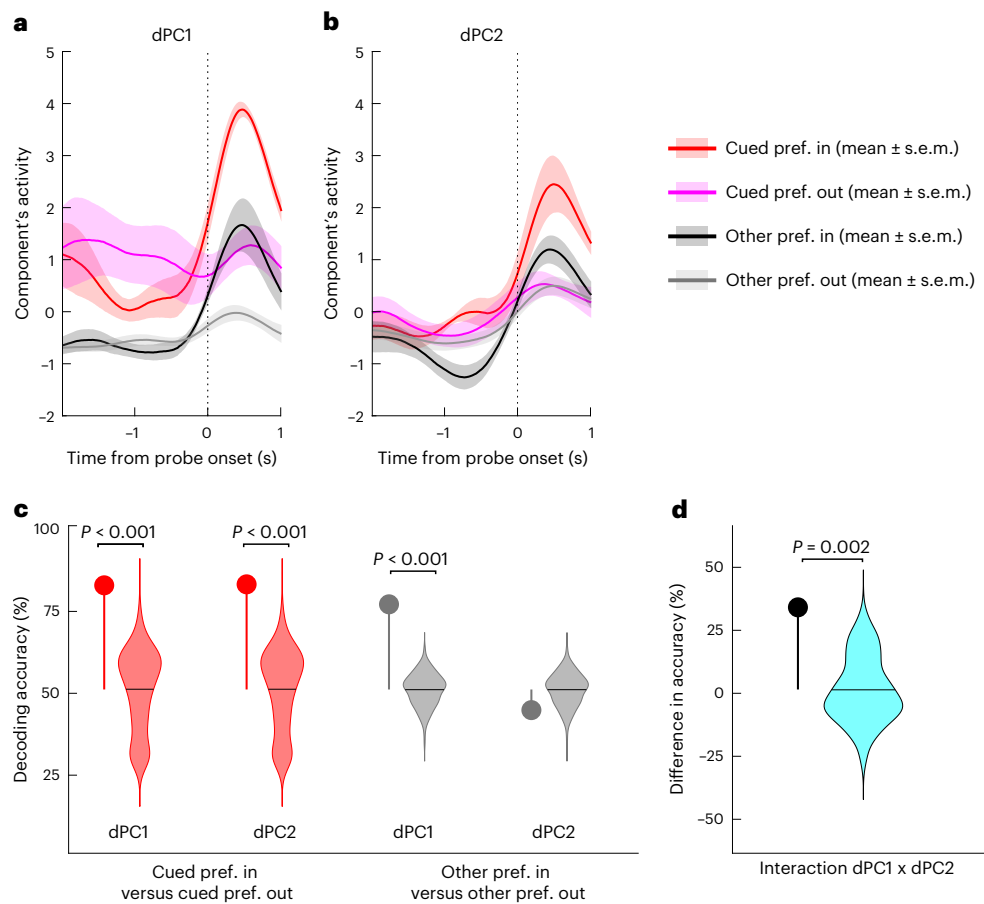


Fig. 5 | Distinct roles of dPC1 and dPC2 during probe presentation. **a, b,** The activity of the first two dPCs (dPC1 (**a**) and dPC2 (**b**)) during probe presentation (entire dataset), projected onto cued and other conditions, divided based on the presence or absence of the preferred item as a probe in the stream of distractors during the first visual search. The activity of dPC1 increases during all presentations of the preferred item, both when it was previously cued (cued preferred in) and when it was not part of the current memory set but was presented as a lure (other preferred in). dPC2 shows increased activity only when the preferred item was previously cued and presented as the current target. Shaded areas represent \pm s.e.m. across trials. **c,** The accuracy of two decoders trained to classify trials where the preferred item was present versus absent in RSVP separately for the cued and other conditions. When applied to the activity of dPC1, both decoders performed above chance. However, when applied to the

activity of dPC2, performance was above chance only for the cued condition. All significance levels for all analyses were determined using one-sided permutation tests assessing whether decoding accuracy was greater than expected under the null hypothesis. Reported P values are FDR corrected. **d,** Interaction of trial type (cued versus other) and component (dPC1 and dPC2). The difference in decoding accuracy between the cued and other conditions was greater for dPC2 than for dPC1, exceeding the expectations of the null distribution. Decoding accuracies obtained from image-selective cells are represented with stem plots. Violin plots represent null distributions (estimated using scrambled data) and black lines indicate the mean. Chance levels for decoders were at 50%, and for the interaction at 0%. Significance of the interaction was assessed similarly as in **c** but results were compared with the null distribution of differences.

Only cued items are decodable from all MTL cell activity

Previous noninvasive neuroimaging studies^{18–20} have often failed to detect signals corresponding to unattended memory content, despite participants being able to correctly recall this content when it was brought back into the focus of attention. We hypothesized that these negative results might be due to the limited sensitivity of noninvasive methods, which measure the average activity of large neuronal populations. To test this hypothesis and facilitate a more direct comparison with previous studies, we conducted a decoding analysis using all recorded neurons ($n = 421$) from the temporal lobe. As before, all significance levels were estimated using scrambled data, and the chance-level decoding was set at 50% (Methods).

We successfully identified which of the four images was presented as a probe with an average accuracy of 70.40% ($P < 0.001$, estimated using scrambled data, chance level 50%; Fig. 6a). Applying this decoder to the activity during the maintenance periods, we were able to decode the identity of the cued memory item in both the first (62.21%, $P < 0.001$, estimated using scrambled data, FDR corrected, chance level 50%) and second (57.3%, $P < 0.001$, estimated using scrambled data, FDR

corrected, chance level 50%) maintenance periods (Fig. 6b). Most importantly, decoding performance for the first uncued item did not differ from chance level (52.41%, $P = 0.174$, FDR corrected) and was significantly lower than for the first cued item ($P < 0.001$, compared with the difference expected based on scrambled data, FDR corrected). We were also unable to decode the identity of the second uncued item (50.45%, $P = 0.469$, estimated using scrambled data, FDR corrected, chance level 50%).

These findings are consistent with previous noninvasive studies^{18–20} and suggest that the signal for the unattended memory content in the overall output of the MTL is weaker than the signal for the content in the current focus of attention. As a result, in prior studies using noninvasive recordings, this signal might have remained below the detection threshold.

Discussion

We recorded the activity of image-selective neurons in the human MTL during a double retro-cue task designed to shift attention between multiple items held in memory. We observed persistent activity for

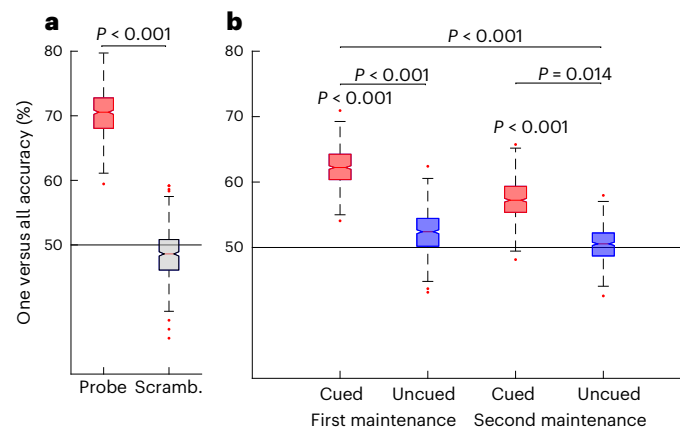


Fig. 6 | Decoding image identity from maintenance periods activity of all cells recorded from the MTL. **a**, The average decoding accuracy for identifying which image was presented as a probe ($n = 421$ neurons, 500 realizations). **b**, The average decoding accuracy when decoders trained on probe presentation were applied to decode information based on cell activity ($n = 421$ neurons, 500 realizations) during the maintenance periods (different trial types are indicated on the x axis). Consistent with previous noninvasive studies, without preselecting image-selective neurons to enhance the signal-to-noise ratio, only information about the attended memory items was decodable above chance level. ‘One versus

all’ denotes the average accuracy of decoders trained to distinguish between a given image and all others (50% chance level). P values above box plots indicate significance compared with chance (estimated using scrambled (scramb.) data with 500 permutations) obtained with one-sided permutation tests. The differences between decoding accuracy for different trial types were assessed using one-sided permutation tests. Reported P values are FDR corrected. Box plots represent quartiles (25% and 75%), horizontal lines indicate medians, whiskers extend up to 1.5 times the interquartile range and dots above whiskers show outliers.

both attended and unattended WM content, suggesting that persistent activity serves as a general mechanism for information storage rather than being solely a marker of attention focus. The attention cue modulated this activity, allowing a clear distinction between attended and unattended WM states.

The question of whether persistent activity correlates with memory storage or attention focus was first raised over a decade ago in noninvasive human studies¹⁹. However, this issue remains unexplored in single-unit recording studies. Previous human single-neuron studies on persistent activity either involved tasks with a single memorized item (automatically becoming the focus of attention) or used multiple memory items with equal probabilities of being tested, probably leading to an equal distribution of attention^{9–11}. As a result, these studies could not differentiate between attention focus and memory maintenance. Our study makes two key contributions: (1) it extends previous single-neuron research by incorporating visual cues to direct attention to one of two concurrently stored items and (2) it directly tests claims from prior noninvasive neuroimaging studies that have used null results to argue for ‘silent’ memory mechanisms for unattended WM content^{18–20,36}.

Our findings align with previous single-neuron human studies^{9–11} that reported persistent neuronal firing of image-selective cells in the human MTL during the maintenance of attended WM content, represented in our task by the cued item. Crucially, we also observed persistent activity related to the unattended WM content (uncued item) during the first maintenance period (Fig. 2a,d). Activity for the uncued WM item significantly decreased during the second maintenance period, though it remained, on average, greater than baseline (Fig. 2b,d). While it is possible that the uncued item attracted some attention, the drop in performance indicates that it had lower priority than the items in the focus of attention. It is important to note that our behavioural results (same trials: 88.94%; switch trials: 81.41% correctness; Fig. 1b and Supplementary Fig. 2) closely resemble those reported in previous noninvasive studies showing lack of decodable activity (for example, same trials: 93% and switch trials: 88% (ref. 19)). Along with clear responses to cue presentations (Fig. 2a,b) this strongly indicates that subjects followed task instructions and prioritized indicated items. Moreover, for both cued and uncued items the observed persistent activity was relevant for memory performance, as it was higher in

trials that ended with correct response (Supplementary Fig. 5). The persistent activity for the uncued item was robust, as it was consistently observed in numerous control analyses (Supplementary Fig. 6) and it could not be attributed to diffused neuron selectivity (that is ‘selectivity leakage’) (Supplementary Figs. 7–9). Our findings align with a recent simulation study³⁷, which successfully reproduced both behavioural and functional magnetic resonance imaging (MRI) results of earlier studies using a persistent firing model, showing no need for the ‘silent’ memory trace for the uncued item.

If the entire memory content is coded by the same neuronal population using the same active coding mechanism, how does neural network codes item prioritization? Although we observed similar persistent activity for both cued and uncued items at the population level, individual cell responses showed much greater complexity (Fig. 3a–c). Some recorded cells exhibited robust persistent activity when their preferred image was cued but were much less active when it was uncued. Others had similar firing rates regardless of attentional focus, while some preferentially fired when their preferred image was memorized but uncued. The dPCA analysis provided a clearer perspective on the population dynamics underlying item prioritization observed in single-neuron examples. It revealed that attention cues transformed population activity related to the cued and uncued items, creating distinct subspaces represented by the first three dPCs. The first component represented common activity associated with both cued and uncued items (memory component; Fig. 4a,b), while the two other components reflected item prioritization: one representing activity specific to the cued item (Fig. 4c,d) and the other activity specific to the uncued item (Fig. 4e,f). Differences in the activation patterns for cued and uncued items in the first two components were sufficient to decode the attentional states (cued versus uncued) of WM item at the single-trial level (Fig. 4b,d). Thus, while all information was coded via image-selective cells, different parts of the network specialized in conveying memory content and information about its prioritization. The distinct roles of the memory and prioritization components were further supported by their differential responses to the presentation of their preferred item in the RSVP stream of distractors. The memory component responded to all appearances of the preferred item in the stream, regardless of its attention status (Fig. 5a,c). In contrast, the attention component responded only when the preferred item was

cued during the maintenance period (Fig. 5b,c), demonstrating that the second component was sensitive towards item prioritization.

These results align with recent work on nonhuman primates, which demonstrated that selecting one stimulus from a two-item memory set enhances and transforms its neuronal representations. Panichello and Buschman³⁸ also reported that selected memory content had a stronger signal, as reflected by higher decoding accuracy, than unselected content. Their study showed that the same mechanism is employed whether selection occurs in the perceptual domain or among mental representations. However, their use of a single retro-cue task precluded an examination of activity corresponding to the unattended content, as uncued items could be immediately removed from WM. Our results go beyond, providing insights into the neural mechanisms supporting unattended memory representations. While the origins of this dynamic transformation remain uncertain, a likely candidate is the frontal and parietal cortex³⁹. It has been proposed that neurons in these regions, often referred to as ‘freely conjunctive’ neurons, may dynamically adjust their connection weights with lower-level stimulus-specific cells (for example, MTL image-selective cells) through rapid plasticity, enabling the formation of flexible mental representations⁴⁰.

A growing body of noninvasive studies has reported the absence of a detectable signal for the unprioritized WM item in various paradigms^{19,21,41}. However, these findings are based on null results, which are inherently prone to critique. First, the transformation of the representation of the uncued WM item may have practical implications for decoding sensitivity, showing its dependence on the data used to train the decoder. For example, the classifiers trained on separate, perceptual tasks (during which decoded items are in the focus of attention) are likely to be more sensitive to signal features characteristic of cued memory content than uncued items. Indeed, previous studies that reported no signal for the uncued memory item^{18–21} trained their decoders on separate perceptual tasks, where items were presented one at a time. In contrast, a study that trained the decoder directly on the memory task involving two concurrently presented stimuli²² successfully decoded unattended WM items from parietal and frontal areas but not from the visual cortex. Second, previous studies have demonstrated that increasing signal-to-noise ratio—either by increasing number of subjects²² or refining task-relevant signal (by limiting analysis to the alpha band instead of analysing the entire electroencephalogram spectrum)²⁷—can lead to successful decoding of both attended and unattended memory content. To test the impact of signal-to-noise ratio in our own data, we intentionally introduced noise by including non-image-selective neurons in our decoding analysis. This approach aimed to simulate the conditions faced by noninvasive studies, which measure the average activity of tens of thousands of cells. Consistent with previous noninvasive studies, we found that under these conditions, we could no longer decode the identity of the uncued WM item during the first maintenance period (Fig. 6), whereas this item was successfully decoded in the dPCA analysis that was restricted to image-selective cells (Fig. 4b,f).

How can we explain the lower behavioural performance typically reported in double retro-cue tasks and observed in our study (Fig. 1b) for the previously uncued items? We observed that the overall signal for the cued item is stronger than the signal for the uncued item. First, when we analysed output from the entire MTL, we were able to decode only cued WM items (Fig. 6b), indicating that the signal for these items is stronger than the signal for the uncued items. Second, the explained signal variance for the prioritization component specific to the cued item was larger (27.62%), than explained variance for the prioritization component specific to the uncued item (10.30%). Third, the memory component (dPC1) exhibited greater activity for the cued than uncued item, suggesting a stronger memory trace for the attended content (Fig. 4a,b).

While we are focusing here on the rate code, there is growing evidence that the information could be coded via alternative mechanisms

such as spike-phase code. The mechanism employed probably depends on the requirements of the task. For instance, it has been shown that when memorizing the order of information is required, the firing of cells encoding different items becomes locked to distinct phases of theta oscillations^{14,42,43}.

To sum up, our results support persistent activity models of unattended items, which propose that information is stored through sustained neuronal firing^{2–4}. While our findings do not rule out synaptic contributions to WM, they highlight the need for empirical evidence to substantiate ‘activity-silent’ coding. This concept remains theoretically appealing due to its potential benefits, such as lower metabolic costs and the capacity to simultaneously encode more items without interference⁴⁴. Recent single-neuron recording studies have already begun to reveal a combination of persistently active coding and ‘activity-silent’ mechanisms. For instance, a study examining serial error bias in non-human primates⁴⁵ showed that a latent representation of the previous item reactivates during the following trial and impairs the encoding of the current one. Another study in human subjects demonstrated that persistent activity can be disrupted during the presentation of other stimuli, and subsequently recovered, suggesting the involvement of additional mechanisms in information storage⁴⁶. Further research is essential to clarify the involvement of proposed ‘activity-silent’ mechanisms and to identify the specific contexts in which they operate.

Methods

Ethical statement

This study was conducted in compliance with all relevant ethical regulations and received approval from the Institutional Review Boards and Bioethics Committees of SUNY Upstate Medical University (1624521), the Medical University of Wrocław (KP-831/2021) and the Medical University of Białystok (APK.002.306.2023). All patients volunteered for the study and provided informed consent before participation. No compensation was provided in exchange for participation.

Double retro-cue task

We used a modified double retro-cue task with four customized images for memorization (Fig. 1a). Each trial began with a 3 s presentation of two images, one positioned above and the other below the fixation point (‘encoding’). Subjects were instructed to memorize both images along with their respective positions on the screen. Next, a spatial cue (a horizontal line) appeared for 1 s, either above or below the fixation point, indicating whether the upper or lower image would serve as the target in the upcoming detection task. This cue was followed by a 5 s maintenance period, which was critical for distinguishing between the representations of attended (cued) and unattended (uncued) memory items.

After the first maintenance period, an RSVP stream of eight images was displayed, with each image shown for 0.15 s. Distractor images were randomly selected for each trial from a pool of 12 images that did not elicit selective neuronal responses during the screening session. The RSVP stream increased trial difficulty and encouraged subjects to follow the cues, focusing their attention on only one memorized item at a time. A probe stimulus was inserted at positions three to six within the RSVP stream. In half of the trials, the probe matched the cued image, in the other half, it was one of the two images from the memory set not presented during encoding, serving as a lure.

After the RSVP, a response screen was displayed and subjects were asked whether the cued image had appeared in the stream. The response screen remained visible until the subject provided an answer. Following a 2–2.5 s intertrial interval (ITI), a second spatial cue was presented for 1 s, indicating the target for the second detection task. In 50% of the trials, the target was the same item as in the first detection task (‘same trials’); in the remaining 50% of the trials, it was the previously uncued item (‘switch trials’). The switch trials ensured that the item not prioritized during the first task was retained in memory,

as it could become the target in the second detection task. The second cue was followed by another 5 s maintenance period, an RSVP stream (eight images, each presented for 0.15 s), a response screen (which remained visible until a response was made) and a short ITI (2–2.5 s). Participants responded by pressing either a green or red button on a response pad.

In each session, subjects performed 96 trials, which were balanced in terms of image position and pairing with other images. Each image appeared in 48 trials, equally often at the top and bottom of the screen, and was paired with each of the other images 16 times. Additionally, each image was cued in 24 trials during the first maintenance period and 24 times during the second maintenance period (12 in the ‘same’ and 12 in the ‘switch’ condition). Data collection and analysis were not performed blind to the conditions of the experiments.

Screening session

Following the procedure described by Kamiński and coworkers, we customized the set of images to be memorized for each subject (for a detailed description, see the original work⁹). Each subject completed a screening task 2–3 h before the main experiment. During this task, subjects viewed between 63 and 126 images representing people, animals, landscapes, everyday objects, food and so on. The image sets were tailored to each participant based on their interests and background, maximizing the likelihood of eliciting selective neuronal responses. Each image was displayed six times in a randomized order, with each presentation lasting 1 s. To maintain engagement, subjects were occasionally asked a control question about the image shown immediately before (for example, ‘Did the last image depict a person/landscape/animal?’), which appeared randomly every few trials.

After the screening task, we immediately analysed the data to select four images for the dual retro-cue task. Selection was based on the mean response during the screening session in a 200–1,000 ms window relative to stimulus onset. To identify cells with selective responses to specific images, we performed a permuted one-way ANOVA with image ID as the factor. The selection criteria included the *F* statistic computed from a one-way ANOVA with the image as a factor, the mean and standard deviation ($M \pm$ s.d.) of the firing rate during repeated image presentations, and visual inspection of the putative unit’s waveform. The four images eliciting the strongest selective responses were chosen as the memory set for the double retro-cue experiment (see Supplementary Fig. 1 for example screening results).

We tested whether correctness in the double retro-cue task was consistent across all four selected images. Since the images were tailored separately for each subject, we first performed a chi-square analysis with the label of the cued image (1, 2, 3 and 4) as a factor separately for each subject. Two out of thirteen sessions showed a significant effect of image label. Next, we compared this result with the null distribution of the number of sessions with significant effects obtained from 500 iterations of this analysis run on data with permuted image labels. The result was not significant ($P = 0.12$), indicating that correctness was, at the group level, similar across different images.

Both the screening and double retro-cue tasks were implemented in MATLAB using Psychtoolbox. It is important to note that the statistical tests conducted during the screening and retro-cue tasks were statistically independent, as they were performed at different times.

Subjects

Twelve epilepsy patients (six men and six women; mean (M)_{age} = 33.67 \pm 13.33 years (\pm s.d.)) participated in the study, with one patient performing the task twice on different days. All patients were implanted with Behnke-Fried electrodes for localization of the epilepsy focus before potential surgical treatment. The localization and number of implanted electrodes were determined solely based on clinical considerations (for detailed information about individual subjects’ age, sex and electrodes localization see Supplementary Table 1).

Surgeries were performed at three locations: SUNY Upstate Medical University Hospital in Syracuse, New York; University Clinical Hospital in Wrocław, Poland; and the 1st Military Clinical Hospital in Lublin, Branch in Elk, Poland. Microelectrode positions were verified using preoperative T1-weighted MRI scans and postoperative computed tomography (CT) images. Individual T1 scans were registered to the Anterior Commissure–Posterior Commissure coordinate frame (ACPC), and CT scans were registered first to the head-based coordinate frame (CTF) and then transformed to the ACPC frame using FieldTrip. CT and MRI images were automatically coregistered, and each coregistration was visually inspected. Microelectrode positions were manually identified in the CT images using the FieldTrip graphical interface and transformed from individual to MNI152 standard space using Statistical Parametric Mapping volume-based normalization. Anatomical labels were assigned to the microelectrode positions using the Automated Anatomical Labeling atlas. For visualization, the microelectrode positions were projected onto the two-dimensional sagittal plane of the CIT168 template brain (Fig. 1c,d shows the recording site locations in Montreal Neurological Institute (MNI) space).

The average IQ score for patients from SUNY Upstate Medical University (9 out of 12 patients) measured with Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV) full scale, was 87.11 \pm 7.90% (\pm s.d.). Owing to the specifics of the consent, we cannot share the IQ score of other subjects. All patients understood and followed the task instructions.

Spike sorting

Each depth electrode contained eight microwires with a diameter of 40 μ m. We recorded broadband signals (0.1–9,000 Hz filter) sampled at 32 kHz using a Neuralynx Atlas system. Signals were locally referenced to one of the eight microwires in each brain area. The raw signal was filtered with a zero-phase lag filter in the 300–3,000 Hz band, and spikes were detected and sorted using a semiautomated template-matching algorithm OSort⁴⁷. Owing to local referencing, units recorded on the reference channel could appear on all other channels from the same depth electrode. To identify such cases, we computed the time coincidence of spike firing for all pairs of units recorded from microwires localized at the same depth electrode. Units with spike coincidence rates $\geq 30\%$ were considered duplicates, and only one of these units was retained for further analyses⁴⁸. Across all subjects, we isolated an average of 32 \pm 15.3 putative single units per session from the MTL, including 221 from the amygdala and 200 from the hippocampus. Throughout this paper, the terms ‘neuron’ and ‘cell’ are used interchangeably to refer to a putative single unit.

Statistics

To analyse time-resolved neuronal activity during the double retro-cue task, we employed an LME with the following formula: activity ~ condition + (1|subject) + (1|subject:neuron) + (1|subject:neuron:trial_id_in_condition). In this model, activity refers to subsequent points in the peri-stimulus time histogram (PSTH), and condition was treated as a fixed effect to account for the type of image held in WM (preferred cued, preferred uncued or other than preferred). Random intercepts were included for three nested levels of the data hierarchy (subject, neuron and trial). This approach appropriately accounts for the repeated-measures design and the hierarchical structure of the data, ensuring robust estimation of fixed effects while addressing individual differences and within-subject dependencies. We used trials in which images other than preferred were presented and held in WM as a reference condition. To directly compare the preferred cued and uncued conditions and assess interactions of additional predictors (see below) with the other condition, the reference group was switched to the preferred cued condition. Obtained *P* values were corrected for multiple comparisons using the FDR, as implemented in the MATLAB function `fdr_bh`. The same model was used to analyse

average neuronal activity during predefined time periods (encoding and the first maintenance). For analysing neuronal activity across both maintenance periods and their potential interaction with condition we added time (first or second) as an additional predictor. The resulting model had the formula $\text{activity} \sim \text{condition} \times \text{time} + (1|\text{subject}) + (1|\text{subject:neuron}) + (1|\text{subject:neuron:trial_id_in_condition})$. Post hoc tests comparing different combinations of predictor levels were performed using the MATLAB function `coefTest` with appropriate contrasts. To balance the LME, we required that the included cells have at least one observation for each predictor level. For comparing neuronal activity between correct and incorrect trials, we used model with the formula: $\text{activity} \sim \text{condition} \times \text{response} + (1|\text{subject}) + (1|\text{subject:neuron}) + (1|\text{subject:neuron:trial_id_in_condition})$. To examine the effects of brain area, we applied the following model: $\text{activity} \sim \text{condition} \times \text{area} + (1|\text{subject}) + (1|\text{subject:neuron}) + (1|\text{subject:neuron:trial_id_in_condition})$. Similarly, for the analysis by recording site, we used the model $\text{activity} \sim \text{condition} \times \text{site} + (1|\text{subject}) + (1|\text{subject:neuron}) + (1|\text{subject:neuron:trial_id_in_condition})$. In all plots representing average firing rates during different task phases, we visualized the predicted conditional means, which account for both fixed and random effects.

The remaining analyses were conducted using permutation tests, with null distributions estimated from 2,000 iterations. Surrogate distributions were generated by permuting condition labels within each cell for paired data or between cells for unpaired data. For two-group comparisons, we used permuted paired or unpaired t -tests. For analyses involving more than two groups, F statistics were calculated. It is important to note that while we report ANOVA and t -test values for reference, P values are derived from permutations. Permutation statistics were computed using the `statcond` function from the EEGLAB toolbox⁴⁹ (parameters: 'method' = perm; 'naccu' = 2000; 'paired' = on or off, depending on the type of test).

All analyses, except from the comparison of correct and incorrect trials, were performed exclusively on correct trials. This ensured that subjects had attended to the images during encoding and successfully memorized them, thereby minimizing noise. Sample sizes were not predetermined by statistical methods but were similar with prior studies investigating persistent neuronal activity during WM in humans⁹. Effect sizes for behavioural data were computed using Cohen's d for paired comparisons and partial η^2 for F tests. Effect sizes for predictors in the LME model were estimated using standardized beta coefficients, while effect sizes for paired comparisons of predictor levels were calculated by dividing the estimated contrast value by the residual standard deviation from the linear mixed-effects model (equivalent to using Cohen's d).

Selection of image-selective neurons

One of our goals was to identify the same population of cells that was previously reported^{9,11} to be persistently active during maintenance of attended WM content. Following their methodology, we used neuronal activity during the encoding period to identify image-selective neurons in our data. For each recorded neuron, we conducted a statistical test to determine whether its firing rate varied based on the identity of the presented images. We counted spikes within a window from 0.2 to 3 s following stimuli onset, and divided all trials into those containing or not containing a particular image and used permuted paired t -tests to compare their average activity. Definition of selectivity was met when at least one test yielded $P < 0.05$ and a positive t value, indicating that the firing rate was significantly bigger in response to pairs containing a specific image. If multiple comparisons met this threshold, the image with the highest t value was designated as the 'preferred image,' while all others were categorized as 'not preferred'.

To avoid any dependencies between cell selection and hypothesis testing, which were conducted on adjacent time windows (encoding versus maintenance), we employed a leave-one approach. Each cell was

tested multiple times ($n = 96$, total number of trials). In each iteration, cell selectivity was tested on 95 trials, and if the selectivity criteria were met, maintenance effects were tested on the one trial excluded from the selectivity definition. Additionally, we required an increase in firing rate during the encoding period compared to the baseline (0–0.5 s before encoding onset).

To quantify the selectivity of identified neurons, we computed the depth of selectivity index (DOS)⁷

$$s = \frac{n - \left(\frac{\sum R_i}{R_{\max}} \right)}{n - 1},$$

where n is the number of images presented, R_i is the firing rate of the neuron during the presentation of the i th picture and R_{\max} is the highest firing rate across all presented images. DOS takes values from 0 (indicating equal response to all stimuli) to 1 (indicating response to only one of the presented stimuli). Since during the encoding period two images were presented concurrently on the screen, trials containing the preferred image were used to assess the firing rate for the preferred image. Activity recorded in trials without the preferred image was used twice, once for each of the two presented images. As a reliable estimation of responses to both preferred and non-preferred items is not feasible when different neurons have varying numbers of trials included in the analysis, we calculated DOS using a subset of neurons that were identified as selective across all leave-one-out iterations ($n = 42$; Supplementary Fig. 3). The average DOS during encoding (0.2–3 s) was 0.41 ± 0.21 , which was significantly higher ($t_{42,72} = 6.9$, $P < 0.001$, permuted paired t -test) than the DOS for images used as distractors in RSVP stream (0.18 ± 0.1). This indicates that preferred images elicited a selective increase in firing rate, which was significantly greater than the response to other items from the memory set and also greater than the response to any of the distractor images used in the RSVP stream.

Simulation of the response of potential neurons to multiple images

We used probe presentation—the only moment in the task when images were presented one at a time, to estimate neuron response to individual images (Supplementary Fig. 9a). To this end, we counted the number of spikes generated when each of the images from the memory set was presented as a probe (within the time window 0.2–0.8 s after probe onset, note that probes were presented in a stream of rapidly displayed distractors, hence the short time window). We will refer to these responses as 'selectivity profiles'. Next, we used these profiles in a simple simulation. We assumed that neurons during maintenance respond exclusively to the image that is currently in the focus of attention and they do not respond to image outside of focus of attention. This was to check whether any activity for cued condition is spilling to uncued condition when uncued items are coded in an activity-silent mechanism. Analogous to our main analysis, using the design matrix we grouped trials into three conditions: preferred item (item with maximal firing rate) cued, uncued or outside the trial. Next, we simply averaged activity in response to a given image that was currently cued. The main question was to test whether the second image maintained in the trial—here the cued one—can affect observed activity in the preferred uncued condition due to not perfect image selectivity. Additionally, we tested a second, more extreme scenario, where we took the maximal cell response and assigned it to one of the three non-preferred images (chosen at random from the images that did not have maximal firing rate), thus creating profile where cell responds to two images with the same, maximal activity. The rest of the steps were identical as in the first scenario. Note that here we used the probe response as an approximation of image selectivity during maintenance, which may not necessarily be accurate. A different selectivity pattern during maintenance could alter specific results but, as shown in the second simulation, it probably does not affect the main conclusion. We also

need to mention that the simulation does not account for all random factors that could be present in the experiment, such as fluctuations in patient focus, drift in neuronal firing rates and so on, which could impact the observed results. However, since these factors are random, they are unlikely to systematically produce significant activity during the uncued condition.

Dimensionality reduction and single-trial decoding

To analyse differences in neuronal activity dynamics for cued and uncued WM content, we applied dPCA, which considers both condition-independent components and those related to experimental manipulations. We used dPCA on a pseudo-population of image-selective neurons, using one marginalized variable with three conditions: (1) preferred image cued in WM, (2) preferred image uncued in WM and (3) non-preferred image. Time was treated as an experiment-independent marginalized variable. We computed the dPCA basis functions using 6-s-long data epochs, encompassing the first cue (1 s) and the first maintenance period (5 s). Only neurons with at least 17 correct trials per condition were included in the analysis, resulting in $n = 35$ neurons. Raster representation of spike times was smoothed with a Gaussian window (1.5 s size) and z scored before the analysis. To assess stability of the components and avoid overfitting, we employed stratified Monte Carlo leave-group-out cross-validation (repeated 500 times). In each iteration, we held out one trial per neuron for each condition and averaged the remaining trials to form the training set³⁵. Component weights were computed from the training set and then projected onto the ‘unseen’ trials. Further analysis focused on the first three components, which were consistently replicated during ‘leave-one-out’, as assessed by correlations of the weights of these components with the corresponding ones obtained in the analysis of the full dataset. Specifically, the correlations were as follows: dPC1 $M = 0.99 \pm 0.001$ (mean $\rho \pm$ s.d.); dPC2 $M = 0.99 \pm 0.001$ and dPC3 $M = 0.98 \pm 0.006$.

We used the activity of these three dPCs components computed on ($n - 1$ trials) to train a decoder to readout the presence of the preferred cued item in WM (cued versus other), the presence of the preferred uncued item in WM (uncued versus other) and whether the preferred item was within or outside the focus of attention (cued versus uncued). First, we computed the average activity of each dPC for each class (preferred cued, preferred uncued or other) during the first maintenance period (1–5 s). Then, we projected previously unseen test trials onto the corresponding decoding axis, averaged component activity during the same time period and classified it according to the closest class mean. Classification accuracy was measured as the proportion of correctly classified trials, averaged over 500 iterations. Each decoder, during training, always used an equal number of trials for each condition and was tested on previously unseen trials of the same type. Significance levels for all analyses were determined using data with scrambled condition labels (chance-level accuracy at 50%). Following the dPCA toolbox pipeline, null distributions were obtained by shuffling condition labels before dPCA computation. The dPCA was computed and the decoders were trained on $n - 1$ trials and tested on the unseen trials. The first dPC was used to generate null distributions³⁵. Data labels were permuted 500 times.

For the analysis of activity during the probe presentation, we computed dPCA components during cue and maintenance on training data ($n - 1$) and projected them to the activity during probe time. The trials were categorized into four categories: (1) preferred item was cued and presented as a probe (cued preferred in), (2) preferred item was cued but not presented as a probe (cued preferred out), (3) other than preferred item was cued but preferred item was presented as a probe (other preferred in) and (4) other than preferred item was cued and presented as a probe. For each category, we calculated the average activity of each component during the probe presentation (0–1 s) for the training dataset as well as for the test trial. The test trial

was classified according to the closest class mean. The accuracy of the decoders was assessed the same way as for decoding of the maintenance period activity.

Population decoding

We conducted decoding analysis on all neurons recorded from the MTL ($n = 421$), aiming to replicate the conditions encountered in noninvasive neuroimaging studies^{18–20}, which analyse outcomes from large neuronal populations without preselection. Trials in which neurons showed image-selectivity ($n = 6,339$) constituted only a small fraction of the analysed data (15.68%). By including all cells in the decoding analysis, we intentionally introduced noise to our data, to simulate the challenges faced by noninvasive studies. Similar to the stimulus categories used in previous studies (for example, words versus lines versus faces), we aimed to extract information about the identity of the stimuli held in memory (image 1, 2, 3 or 4). To achieve this, we aggregated all recorded neurons across sessions and subjects to create a pseudo-population, resulting in an $N \times T$ matrix. Here, N represents the total number of neurons, and T is the minimum required number of trials observed across all subjects. To ensure sufficient resampling we set $T = 15$ (equivalent to 60% of correct trials). The more frequent categories were subsampled to match the least frequent one, ensuring balance across all decoding analyses. We employed a support vector machine as implemented in the Neural Decoding Toolbox⁵⁰ and the LIBSVM library⁵¹.

Since previous noninvasive studies trained their decoders on separate perceptual tasks, we trained our decoder on the probe presentation—the phase of the task most resembling a simple perceptual task, in which images were shown one at a time. For each trial, we counted the number of spikes in time 0–1 s from probe onset. To increase analysis power, we pooled data from both the first and second probe presentations. Images from different subjects were arbitrarily labelled as 1–4 (note that each subject saw a different set of pictures). Images with the same label were pooled across all subjects into single categories. For the learning algorithm, all that matters is that each subject saw four distinct images. The decoder assigned weights to each neuron independently, based on its ability to discriminate between specific categories. Subsequently, we trained the decoder to discern patterns within this high-dimensional space. To estimate performance, we used leave-one-out cross-validation. For each iteration, one trial (randomly selected for each neuron) was designated as the test trial, while the remaining trials were used for training. We employed an ‘all versus one’ decoder designed for multiple categories, with a chance accuracy at 50%. This process was repeated 500 times to estimate the cross-validated testing error. To determine whether decoding performance exceeded chance levels, we generated null distributions by applying the same decoding method to data with scrambled category labels. To compare the accuracy of different decoders, we computed the accuracy difference across all 500 realizations for both decoders and compared it with the difference between the corresponding null distributions.

Since two images were held in memory simultaneously during the maintenance period, we attempted to limit the influence of the other image in the pair on the decoding results. To achieve this, we trained four decoders on the probe presentation, each time leaving one image out of the analysis to make the decoders ‘blind’ to that particular image. For example, one decoder was trained to distinguish between images 1, 2 and 3, excluding image 4; another was trained to distinguish images 2, 3 and 4, excluding image 1; and so on for all possible combinations. These decoders were then tested on the average activity (1–5 s) during the first and second maintenance periods. Trials were grouped based on the identity of the ‘to-be-ignored’ image (the image not being decoded), and the decoding process was run four times, each time using the decoder that was blind to the ‘to-be-ignored’ image. For instance, to decode the identity of the first cued item, the decoder

trained to distinguish images 1, 2 and 3 was applied to all trials where the uncued item was image 4. Similarly, the decoder trained on images 2, 3 and 4 was used for trials where the uncued item was image 1. The reported decoding accuracies are the averages of all four runs. The same procedure was used to decode the identity of the first uncued, second cued and second uncued items.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available under the link: <https://repositorio.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/YC4TXY>

Code availability

Analysis was performed in MATLAB using the publicly available software packages Osort, ndt, EEGLAB, SMP12 and dPCA together with custom-developed analysis routines.

References

- Baddeley, A. Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* **63**, 1–29 (2012).
- Compte, A., Brunel, N., Goldman-Rakic, P. S. & Wang, X. J. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* **10**, 910–923 (2000).
- Goldman-Rakic, P. S. Cellular basis of working memory. *Neuron* **14**, 477–485 (1995).
- Constantinidis, C. et al. Persistent spiking activity underlies working memory. *J. Neurosci.* **38**, 7020–7028 (2018).
- Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* **61**, 331–349 (1989).
- Suzuki, W. A., Miller, E. K. & Desimone, R. Object and place memory in the macaque entorhinal cortex. *J. Neurophysiol.* **78**, 1062–1081 (1997).
- Rainer, G., Asaad, W. F. & Miller, E. K. Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* **393**, 577–579 (1998).
- Fuster, J. M. & Jervey, J. P. Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* **212**, 952–955 (1981).
- Kamiński, J. et al. Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nat. Neurosci.* **20**, 590–601 (2017).
- Boran, E., Hilfiker, P., Stieglitz, L., Sarnthein, J. & Klaver, P. Persistent neuronal firing in the medial temporal lobe supports performance and workload of visual working memory in humans. *Neuroimage* **254**, 119123 (2022).
- Kornblith, S., Quiroga, R., Koch, C., Fried, I. & Mormann, F. Persistent single-neuron activity during working memory in the human medial temporal lobe. *Curr. Biol.* **27**, 1026–1032 (2017).
- Larocque, J. J., Lewis-Peacock, J. A. & Postle, B. R. Multiple neural states of representation in short-term memory? It's a matter of attention. *Front. Hum. Neurosci.* **8**, 5 (2014).
- Kamiński, J. & Rutishauser, U. Between persistently active and activity-silent frameworks: novel vistas on the cellular basis of working memory. *Ann. N. Y. Acad. Sci.* **1464**, 64–75 (2020).
- Kamiński, J., Brzezicka, A., Mamelak, A. N. & Rutishauser, U. Combined phase-rate coding by persistently active neurons as a mechanism for maintaining multiple items in working memory in humans. *Neuron* **106**, 256–264.e3 (2020).
- Siegel, M., Warden, M. R. & Miller, E. K. Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl Acad. Sci. USA* **106**, 21341–21346 (2009).
- Mongillo, G., Barak, O. & Tsodyks, M. Synaptic theory of working memory. *Science* **319**, 1543–1546 (2008).
- Stokes, M. G. 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* **19**, 394–405 (2015).
- LaRocque, J. J., Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K. & Postle, B. R. Decoding attended information in short-term memory: an EEG study. *J. Cogn. Neurosci.* **25**, 127–142 (2013).
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K. & Postle, B. R. Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* **24**, 61–79 (2012).
- Rose, N. S. et al. Reactivation of latent working memories with transcranial magnetic stimulation. *Science* **354**, 1136–1139 (2016).
- Wolff, M. J., Jochim, J., Akyürek, E. G. & Stokes, M. G. Dynamic hidden states underlying working-memory-guided behavior. *Nat. Neurosci.* **20**, 864–871 (2017).
- Christophel, T. B., Jamshchinnina, P., Yan, C., Allefeld, C. & Haynes, J.-D. Cortical specialization for attended versus unattended working memory. *Nat. Neurosci.* **21**, 494–496 (2018).
- Yu, Q., Teng, C. & Postle, B. R. Different states of priority recruit different neural representations in visual working memory. *PLoS Biol.* **18**, e3000769 (2020).
- van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J. & Olivers, C. N. Current and future goals are represented in opposite patterns in object-selective cortex. *eLife* **7**, e38677 (2018).
- Nunez, P. L. & Srinivasan, R. *Electric Fields of the Brain: the Neurophysics of EEG* (Oxford Univ. Press, 2006).
- Ip, B. & Bridge, H. Mapping the visual world to the human brain. *eLife* **10**, e75171 (2021).
- Barbosa, J., Lozano-Soldevilla, D. & Compte, A. Pinging the brain with visual impulses reveals electrically active, not activity-silent, working memories. *PLoS Biol.* **19**, e3001436 (2021).
- Souza, A. S. & Oberauer, K. In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Atten. Percept. Psychophys.* **78**, 1839–1860 (2016).
- Griffin, I. C. & Nobre, A. C. Orienting attention to locations in internal representations. *J. Cogn. Neurosci.* **15**, 1176–1194 (2003).
- Vergara, J., Rivera, N., Rossi-Pool, R. & Romo, R. A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron* **89**, 54–62 (2016).
- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107 (2005).
- Cocina, F., Vitalis, A. & Caflisch, A. Spiking burstiness and working memory in the human medial temporal lobe. *Cereb. Cortex Commun.* **3**, tgac039 (2022).
- Li, J. et al. Successful working memory linked to theta connectivity patterns in the hippocampal-entorhinal circuit. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.09.08.507081> (2022).
- Watanabe, Y., Ikegaya, Y. & Yanagisawa, T. Hippocampal neural fluctuation between memory encoding and retrieval states during a working memory task in humans. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.04.28.538785> (2023).
- Kobak, D. et al. Demixed principal component analysis of neural population data. *eLife* **5**, e10989 (2016).
- Mallett, R. & Lewis-Peacock, J. A. Working memory prioritization impacts neural recovery from distraction. *Cortex* **121**, 225–238 (2019).
- Schneegans, S. & Bays, P. M. Restoration of fMRI decodability does not imply latent working memory states. *J. Cogn. Neurosci.* **29**, 1977–1994 (2017).

38. Panichello, M. F. & Buschman, T. J. Shared mechanisms underlie the control of working memory and attention. *Nature* **592**, 601–605 (2021).
39. Yeo, B. T. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
40. Manohar, S. G., Zokaei, N., Fallon, S. J., Vogels, T. P. & Husain, M. Neural mechanisms of attending to items in working memory. *Neurosci. Biobehav. Rev.* **101**, 1–12 (2019).
41. Trübtschek, D. et al. A theory of working memory without consciousness or sustained activity. *eLife* **6**, e23871 (2017).
42. Liebe, S. et al. Phase of firing does not reflect temporal order in sequence memory of humans and recurrent neural networks. *Nat. Neurosci.* **28**, 873–882 (2025).
43. Reddy, L. et al. Theta-phase dependent neuronal coding during sequence learning in human single neurons. *Nat. Commun.* **12**, 4839 (2021).
44. Lundqvist, M., Herman, P. & Miller, E. K. Working memory: delay activity, yes! Persistent activity? Maybe not. *J. Neurosci.* **38**, 7013–7019 (2018).
45. Barbosa, J. et al. Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nat. Neurosci.* **23**, 1016–1024 (2020).
46. Bausch, M. et al. Concept neurons in the human medial temporal lobe flexibly represent abstract relations between concepts. *Nat. Commun.* **12**, 6164 (2021).
47. Rutishauser, U., Schuman, E. M. & Mamelak, A. N. Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. *J. Neurosci. Methods* **154**, 204–224 (2006).
48. Magnuski, M. et al. Pylabianca: comprehensive and user-friendly Python package for single-neuron data analysis. *Acta Neurobiol. Exp.* **83**, 432–446 (2023).
49. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
50. Meyers, E. M. The neural decoding toolbox. *Front. Neuroinform.* **7**, 8 (2013).
51. Chang, C.-C. & Lin, C.-J. LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**, 27 (2011).

Acknowledgements

We are indebted to all patients who volunteered their time to participate in our study. We thank J. Barbosa for discussion of our results and sharing his valuable insights. Intern students M. Padarz, K. Rak and technician O. Matysiak for their help in additional visual inspection of sorted units as well as intern student K. Szalata for her help in marking localization of microelectrodes on individual CT scans. The staff and physicians of all hospitals in which we did our recordings for their invaluable assistance. This work was supported by Polish National Science Centre NCN grant nos 2019/34/E/HS6/00257 and 2022/47/B/HS6/02748 to J.K.; BRAINCITY project

carried out within the International Research Agenda Programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development; The Kościuszko Foundation Exchange Program to USA to K.P. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

K.P. and J.K. conceived of the study idea. K.P. and J.K. performed the formal analysis. K.P., J.K., M.M., W.Ś. and D.I. collected the data, A.R., M.S., T.P., W.F., K.S., P.C.R., Sz. K., P.T. and H.B. qualified patients, performed surgeries and/or took care for the patients during their in-ward stay. K.P., M.M. and J.K. contributed to the interpretation. K.P. wrote the paper with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41562-025-02235-0>.

Correspondence and requests for materials should be addressed to J. Kamiński.

Peer review information *Nature Human Behaviour* thanks Thomas Christophel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

- | | |
|-----------------|--|
| Data collection | Data was recorded using Pegasus Software provided by Neuralynx (https://neuralynx.fh-co.com/clinical-overview/clinical-products/pegasus-software/). Experimental procedures were written in Matlab using the Psychtoolbox (Brainard, D. H. The Psychophysics Toolbox. <i>Spat. Vis.</i> 10, 433–436 (1997)). |
| Data analysis | Analysis was performed in Matlab using the publicly available software packages: OSort (for single unit sorting); EEGLAB and FieldTrip (for statistical tests); FieldTrip and SPM12 (for electrodes localization); ndt toolbox and LIBSVM library (for decoding) and demixed PCA was computed using Kobak, D. et al. Demixed principal component analysis of neural population data. <i>Elife</i> 5, (2016). Custom MATLAB scripts are available from corresponding author upon request. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data that support the findings of this study are available under the link: <https://repor.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/YC4TX>

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

This study recorded participants' self-reported sex and age (at the time of recording). These data are provided in the Supplementary Information file in Table 1, the sex and average age are also reported in the Methods section. Out of 12 participants, 6 reported to be females. Since our research question regarded cellular mechanisms of maintenance of working memory content depending on its attentional status based on existing literature we expected it to be generalizable across sex groups. Post-hoc sex-based analyses were not conducted given insufficient sample size in the split-groups analysis.

Reporting on race, ethnicity, or other socially relevant groupings

Information about race, ethnicity or any other socially relevant groupings (except self-reported sex and age) were not recorded.

Population characteristics

Participants were subjects with drug-resistant epilepsy who for epilepsy treatment required implantation of iEEG electrodes. Location of all implanted electrodes were determined based solely on clinical premises. To control for effects of epilepsy on our results we performed control analysis comparing effect between recordings within healthy tissue and recordings within tissue later diagnosed as epileptogenic (results of these analyses are presented in Supplementary Information Figures 1 and 5.). Localization of putative epileptic focus was determined after completion of the inpatient seizure monitoring from the implanted electrodes at the end of hospital stay is reported in Supplementary Information Table 1. The main effects presented in this manuscript were similar in both locations showing that they are not contaminated by pathological activity of epileptic tissue.

Recruitment

Subjects were recruited from the population of drug-resistant epilepsy subjects who were undergoing invasive epilepsy monitoring. Only patients who gave informed consent participated in the study. Participation in the study was fully voluntary.

Ethics oversight

This study was conducted in compliance with all of the relevant ethical regulations and received approval from the Institutional Review Boards and Bioethics Committees of SUNY Upstate Medical University (1624521), the Medical University of Wrocław (KP-831/2021), and the Medical University of Białystok (APK.002.306.2023). All patients volunteered for the study and provided informed consent prior to participation. No compensation was provided in exchange for participation.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

13 sessions recorded from 12 subjects. No statistical methods were used to predetermine sample size, but our sample size is similar to those reported in previous publications in the field (Kamiński, J. et al. Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nat. Neurosci.* 20, 590–601 (2017) and Kornblith, S., Quian Quiroga, R., Koch, C., Fried, I. & Mormann, F. Persistent Single-Neuron Activity during Working Memory in the Human Medial Temporal Lobe. *Curr. Biol.* 27, 1026–1032 (2017).); Boran, E., Hilfiker, P., Stieglitz, L., Sarnthein, J. & Klaver, P. Persistent neuronal firing in the medial temporal lobe supports performance and workload of visual working memory in humans. *Neuroimage* 254, 119123 (2022)).

Data exclusions

We included participants who had: (1) at least one electrode in amygdala or hippocampus; (2) performed double retro-cue task at above chance level and (3) during screening showed selective response of one or more putative neurons in response to at least one presented image. We excluded one subject from reaction times analysis because reaction times in target absent trials were not recorded due to technical issues.

Replication	<p>We used behavioral paradigm previously only used during noninvasive recordings such as fMRI and EEG during recordings of single neuron activity in patients with drug resistant epilepsy, which gave us an excellent sensitivity necessary for rigorous testing of "activity silent" coding hypothesis. We also replicated results of previous single neuron human studies, showing once again that attended working memory content is maintained by persistent neuronal activity and extended them showing for the first time that the same mechanism underlies maintenance of the unattended content. In this work we used different analytical approaches such as firing rate analysis, machine learning and demixed principal components analysis which all have replicated the main findings.</p> <p>Particularly, we replicated behavioral results from previous studies using retro-cue paradigm combined with fMRI (e.g. Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K. & Postle, B. R. Neural evidence for a distinction between short-term memory and the focus of attention. <i>J. Cogn. Neurosci.</i> 24, 61–79 (2012)), EEG (LaRocque, J. J., Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K. & Postle, B. R. Decoding attended information in short-term memory: an EEG study. <i>J. Cogn. Neurosci.</i> 25, 127–142 (2013).) and TMS (Rose, N. S. et al. Reactivation of latent working memories with transcranial magnetic stimulation. <i>Science</i> 354, 1136–1139 (2016)).</p> <p>As well as, results of previous human single neuron studies showing persistent neuronal activity during maintenance of attended working memory item: Kamiński, J. et al. Persistently active neurons in human medial frontal and medial temporal lobe support working memory. <i>Nat. Neurosci.</i> 20, 590–601 (2017) and Kornblith, S., Quian Quiroga, R., Koch, C., Fried, I. & Mormann, F. Persistent Single-Neuron Activity during Working Memory in the Human Medial Temporal Lobe. <i>Curr. Biol.</i> 27, 1026–1032 (2017).</p>
Randomization	<p>Randomization of participants was not relevant to these study as participants were not randomly assigned to any groups a priori. For every participant order of trials was randomized. Each subject performed this task on individually set of stimuli selected based on preceding screening procedure, such that each stimuli evoked selective response of single neuron in medial temporal cortex. For similar approach see Kamiński, J. et al. Persistently active neurons in human medial frontal and medial temporal lobe support working memory. <i>Nat. Neurosci.</i> 20, 590–601 (2017). Main analyses were repeated-measures in which the same putative neuron was tested under different experimental conditions. Obtained results were then tested against random distribution assessed on scrambled data.</p>
Blinding	<p>Participants were blinded to the major hypothesis of this study. Experimentators who collected the data were not blinded and were present in patients' room during task performance since patients required close supervision due to potential seizure onset. Blinding was not possible during data analysis, as participants were not assigned to any groups. Whenever shuffling or re-sampling were used, investigators were blinded to random assignment of task condition labels.</p>

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks	<input type="text" value="n/a"/>
Novel plant genotypes	<input type="text" value="n/a"/>
Authentication	<input type="text" value="n/a"/>