Hippocampal neurons code individual episodic memories in humans

L. D. Kolibius[1,4,*], F. Roux[1,4], G. Parish[1,4], M. Ter Wal[4], M. Van Der Plas[1,4], R. Chelvarajah[2], V. Sawlani[2], D. T. Rollings[2], J. Lang[3], S. Gollwitzer[3], K. Walther[3], R. Hopfengärtner[3], G. Kreiselmeyer[3], H. Hamer[3], B. P. Staresina[4], M. Wimber[1,4], H. Bowman[4,5], S. Hanslmayr[1,4,*]

[1] Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, United Kingdom
[2]Neuroscience Department, Queen Elizabeth Hospital Birmingham, Complex Epilepsy and Surgery Service, Birmingham, United Kingdom
[3]Department of Neurology, University Hospital Erlangen, Epilepsy Centre, Erlangen, Germany
[4] School of Psychology and Centre for Human Brain Health, University of Birmingham, Birmingham, United Kingdom
[5]Centre for Cognitive Neuroscience and Cognitive Systems and the School of Computing, University of Kent, Kent, United Kingdom

* Corresponding authors:
Luca D. Kolibius
2604636K@student.gla.ac.uk

Simon Hanslmayr
Simon.Hanslmayr@glasgow.ac.uk

Abstract

The hippocampus is an essential hub for episodic memory processing. However, how human hippocampal single neurons code multi-feature associations remains unknown. Here, we provide evidence that individual neurons code discrete memory episodes, which we term Episode Specific Neurons (ESNs). These ESNs do not reflect the coding of a particular feature in the episode (i.e., concept or time). Instead, they code for the conjunction of the different elements that make up the episode.

Main text

Episodic memory refers to our ability to reinstate the what, where and when of past experiences (Tulving, 2002). It is undisputed that the hippocampus has an integral role in episodic memory processing (Lisman et al., 2017; Marr, 1971; Squire, 1992), but how it codes episodic memories remains controversial. Concept cells in the hippocampus fire in response to specific invariant features independent of the context. One prevailing idea is that the diverse elements that make up an episode are coded by the simultaneous activity of multiple of these concept cells (Quiroga, 2020). Alternatively, single units in the hippocampus might sparsely encode a specific set of features within an individual episode that act as pointers to cortical modules during memory reinstatement (Teyler & DiScenna, 1986). Unlike concept neurons, these Episode Specific Neurons (ESNs) would fire in response to a conjunctive code and not in response to individual content features (such as a specific face or a specific place). Despite simulations pointing towards their existence (Bowman & Wyble, 2007), to this day there is no evidence for such a sparse conjunctive code in humans. In search of an answer to this issue, we leveraged intracranial microwire recordings to investigate the firing patterns of single neurons in the human hippocampus. We analyzed recordings from 16 participants (7 female; age mean = 36.125 years, from 26-53 years) implanted with stereotactic Behnke-Fried depth electrodes in the hippocampus while they performed a memory association task (Figure 1a).

We hypothesized a firing rate reinstatement in a significant number of hippocampal single neurons that code for the conjunctive features present within a specific episode. As control analyses, we investigated whether this firing activity can be explained by a firing response to specific invariant features as occurs in 'concept cells' (CC; Quiroga et al., 2005), or by a time preference, as occurs in 'time cells' (TC; Reddy et al. 2020; Umbach et al., 2020).

For every single neuron, we determined the firing rate during each correctly remembered episode at encoding and retrieval. We then z-scored the firing rate across all encoding and retrieval episodes independently. We measure episode-specific firing reinstatement as the product of the standardized firing rates at encoding and retrieval (Figure 1b).
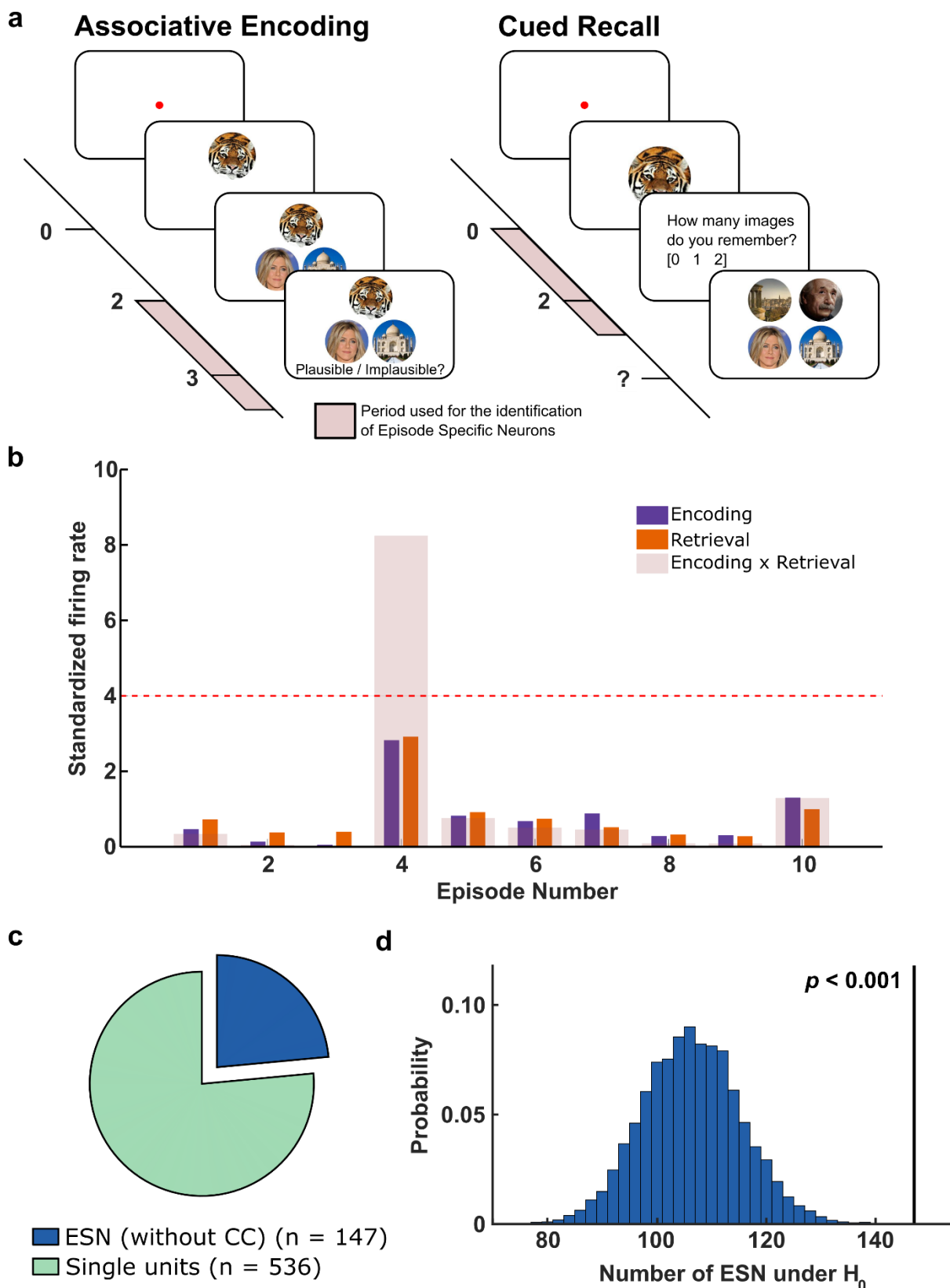
Figure 1. Experiment procedure, analysis outline, and results. (a) Outline of the experiment procedure. During encoding, participants imagined an episode involving an animal cue and two associate images (two faces, two places or a place and a face) and rated its plausibility. During recall, participants were asked to retrieve the associated images when cued with the animal cue. The experiment was self-paced and every episode was learned and tested only once. Following each encoding block of roughly 20 episodes, participants performed a short distractor task. The pink areas represent the time windows used for subsequent analyses. **(b)** A schematic

for identifying Episode Specific Neurons (ESNs) is shown. The diagram shows the z-scored firing rate on the y-axis for ten simulated episodes on the x-axis color-coded for encoding and retrieval (purple and orange, respectively). The transparent bars encompassing encoding and retrieval indicate the product of encoding and retrieval firing rates, which is used as the measure of episode-specific reinstatement. The dotted red line shows the threshold (derived from a shuffling procedure, see Methods). **(c)** The pie chart depicts the number of ESNs that show significant firing reinstatement to at least one episode when excluding putative CC activity (dark blue) and the number of single units that showed no firing reinstatement (green). **(d)** The number of ESNs as expected by chance (permutation test) and the empirical number of ESNs when excluding potential CC tuned to the animal cue (n = 147; $p < 0.001$).

Using a shuffling procedure, we generated a distribution of reinstatement values expected by chance. A neuron was considered an ESN if (i) the empirical reinstatement value exceeded the 99th-percentile of the shuffled distribution for at least one episode and (ii) if the standardized firing rate for encoding and retrieval of that episode each exceeded 1.645 ($\triangleq p_{right-tailed} < 0.05$; Figure 1b). This approach revealed that a significant number of hippocampal neurons were ESNs (160 out of 625 single units $\triangleq$ 25.6%; $p < 0.001$; permutation test). Two example ESNs are shown in Figure 2. Most ESNs reinstated one episode (136 out of 160 ESNs $\triangleq$ 85.00%), while the rest reinstated between two and four episodes.

To rule out that ESNs are firing in response to a stimulus category (such as faces or places) or a specific face or place, we repeated the ESN identification analysis, but only included ESNs that showed reinstatement in at least one episode with two face associate images and at least one episode with two place associate images. Such neurons would be very unlikely to code for a specific concept because the stimulus material in these episodes does not overlap. Again, we found a significant number of ESNs (6 out of 625 single units $\triangleq$ 0.96%; $p =$ 0.0097). It is important to note, that this is a very conservative analysis, as only a few ESNs reinstate two or more episodes (24 out of 160 ESNs $\triangleq$ 15%).

However, the above-described ESNs could still reflect the firing of cells tuned to the image of the animal cue, which is trial-unique and presented during encoding and retrieval. To address this issue, we repeated the initial ESN identification analysis but excluded single units that showed a significant increase in firing during encoding between 300ms and 1000ms after cue onset. This approach has traditionally been used to identify putative concept neurons (Quiroga et al., 2005; see Methods). We still identified a significant number of ESNs using this restriction (147 out of 625 single units $\triangleq$ 23.52%; $p < 0.001$; permutation test; Figure 1d). The low firing rate during the first two seconds of encoding makes it unlikely that ESNs are concept cells tuned to an animal cue (see Figure 3a). There is a tendency that spike waveshapes of ESNs are wider than those of other single units (unpaired t-test; $p =$ 0.0529; Figure S3a), suggesting that ESNs might belong to a subcategory of neurons that is structurally different from other single units. However, there is no significant difference in the spike height or Fano factor between ESNs and other single units (unpaired t-tests; all p values > 0.1; Figure S3b and S4). All further analyses based on ESN classification use this ESN definition that excludes putative CC.
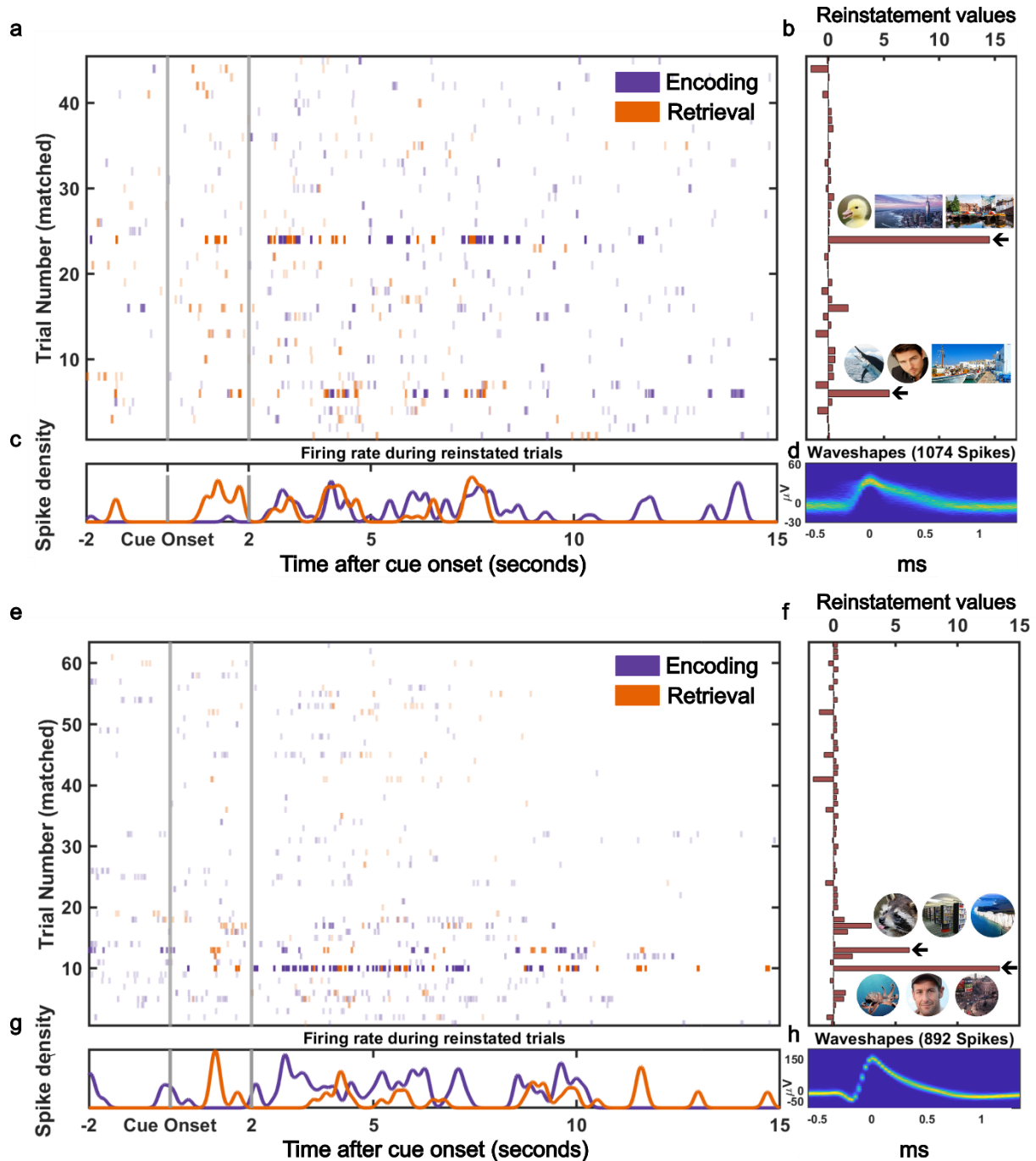
Figure 2. Firing patterns for two example Episode Specific Neurons. (a) Spike raster plot. Each line indicates a spike. On the x-axis is time and on the y-axis episodes. Color-coded in purple for encoding and orange for retrieval. The transparency is adjusted according to the reinstatement values in that specific episode. **(b)** Reinstatement values and the animal cues with the respective associate images for reinstated episodes (indicated by the black arrow) **(c)** Spike density plot for reinstated episodes. Please note that the experiment is self-paced and episode length varies. **(d)** 2D histogram of the waveshape of that particular single unit (Niediek et al., 2016). **(e-h)** same as (a-d) but for a different example ESN.
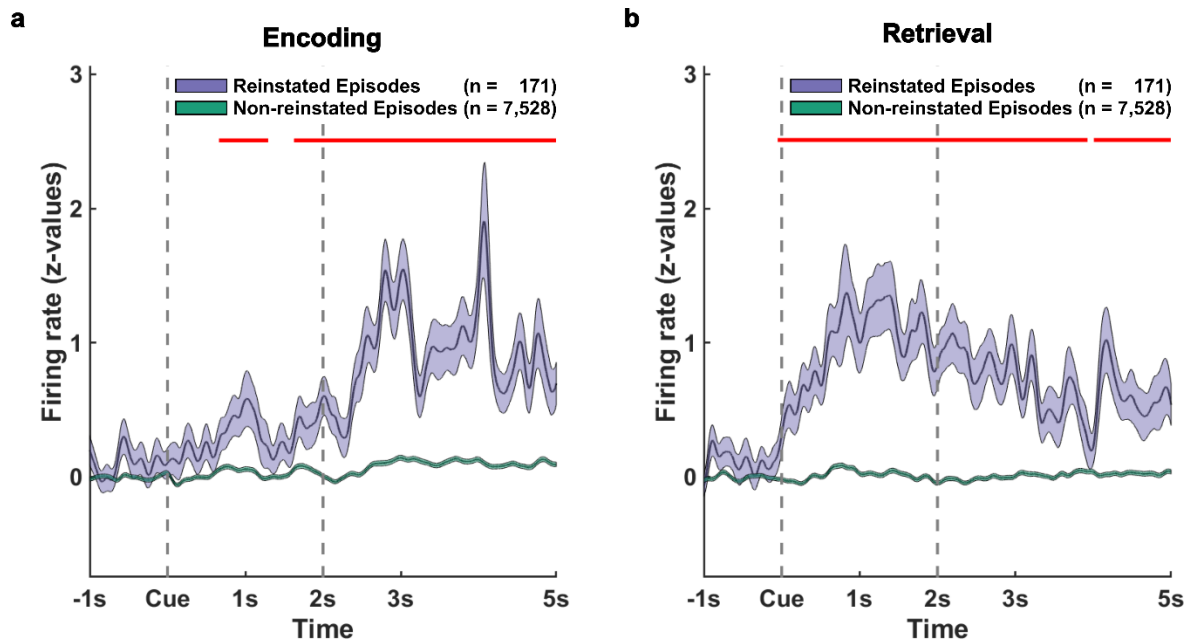
Figure 3. Firing rate of ESNs during reinstated (purple) and non-reinstated (green) episodes. (a) Firing rate of ESNs from cue onset until five seconds later during memory encoding. The red line marks time points where the firing rate during reinstated episodes (n =171) significantly exceeds the firing rate during non-reinstated episodes (n =7,528; cluster permutation test; Maris & Oostenveld, 2007). **(b)** Same as (a) but for memory retrieval. The shaded areas indicate the SEM.

Recent studies in humans show that hippocampal neurons code specific time points invariant across repetitions which are referred to as time cells (Reddy et al. 2020, Umbach et al., 2020). We investigated whether our dataset contains such time cells (TC) using a similar method as employed by Umbach et al. (2020), once normalizing encoding block length and once without this normalization (see Methods). Of all recorded cells 15 (normalized) and 10 (non-normalized) fulfilled the criteria of TCs, which is below chance level (*p* values > 0.5; permutation test). Critically, there was no significant overlap between single units that behave like TCs and ESNs (*p* values > 0.7; permutation test) indicating that ESNs cannot be construed as TCs.

In conclusion, we find single neurons in the hippocampus that show firing reinstatement in response to a specific conjunctive code representing a unique episode. These Episode Specific Neurons do not fire in response to individual concepts (concept cells) or to specific, re-occurring time points (time cells). We propose that during memory formation an assembly of ESNs acts as a pointer that initially binds the features of an episode together, in line with the indexing theory (Bowman & Wyble, 2007; Teyler & DiScenna, 1986). Reactivation of this pointer allows ESNs to reinstate the episodic memory previously encoded. Importantly, because ESNs reinstate unique episodes, they contain a time and content component. However, rather than reflecting the underlying coding mechanism, this time and content aspect necessarily emerges from the conjunctive code of an episode that is unique in content and time.

Methods

Experimental procedure
During the encoding phase of the experiment the participant associated a cue with two other stimuli. For each episode, the cue was a new picture of an animal. The stimuli could be pictures of either places, faces or both. Every picture was only shown once. Two seconds after the animal cue was presented, the associate stimuli were shown, while the animal cue remained on the screen. The participant was asked to create a vivid imaginary story involving the cue and the two stimuli. This part of the experiment was self-paced. The task continued once the participant rated the plausibility of the imaginary story (plausible/implausible).

After the encoding phase, the participant performed a distractor task to rule out working memory effects. During the distractor task, participants had to indicate whether a random number (up to two digits) that appeared serially on the screen was odd or even. After each response, the participant received feedback indicating a correct or incorrect response. This task consisted of 15 trials.

During the retrieval phase, all cues from the previous encoding phase were presented sequentially in pseudorandom order. Each animal cue was presented for two seconds and subjects were tasked to retrieve the corresponding images. The participant was then asked how many associated images they remembered (none, one, two). Participants had as much time to respond as they required. If the participant indicated that they remembered one or two images, they then had to select two pictures from an array of four pictures (two targets and two distractors).

The experiment ended after the retrieval phase if the total runtime exceeded 40 minutes, or the patient asked to abort the experiment. Otherwise, the experiment continued with the next encoding block. The encoding block initially consisted of 20 episodes but could be adjusted depending on the cognitive abilities of the patient. If the hit rate fell below 66.25%, fewer episodes were shown for the next block and vice versa if the hit rate surpassed 73.75%.

The patients performed the memory task on a laptop computer (Toshiba Tecra W50; 60 Hz refresh rate), while either seated in a chair next to their bed or their hospital bed.

Participants
Eight patients were recorded in the Queen Elizabeth Hospital Birmingham (Birmingham, UK) (4 female; mean age: 36.25 years, from 26-49 years) and eight patients in the Universitätsklinikum Erlangen (Erlangen, Germany) (3 female; mean age: 36 years, from 26-53 years) (see Table S1).

Ethical approval
Ethical approval was granted by the National Health Service Health Research Authority (15/WM/0219) and the Ethik-Kommission of the Friedrich-Alexander Universität Erlangen-Nürnberg (142_12 B).

Behavioural analysis
For our analyses, we considered an episode a hit if the participant correctly identified both stimuli. We considered an episode a miss if the participant either indicated not to remember any stimuli or did not remember both stimuli correctly. Participants correctly recalled on

average 68.15% (SE = 2.61%) episodes (see Table S1). This is substantially more than would be expected by chance (16.7%).

Statistical analysis
All statistical analyses were conducted using MATLAB R2020b on a computer running on Windows 10 Enterprise. The significance threshold for all statistical tests was set at 0.05. All permutation tests were implemented with N = 10,000 random draws.

Co-Registering
For each patient, a pre-operational T1-weighted MRI scan was co-registered with a post-operational scan and normalized in MNI space using SPM12. Each macro-electrode was localized either within the hippocampus or outside it through visual inspection. Only activity from microwires in Behnke-Fried electrodes that were assigned to the hippocampus was analyzed in the current study.

Recording System and Electrodes
Patients were implanted with one to eight (see Table S1 for an overview) depth electrodes of the Behnke Fried type with microwire bundles (Ad-Tech Medical Instrument Corporation, USA) to localize epileptic foci. The electrode location was determined by clinical need. These single-use electrodes are made from platinum, have a diameter of 1.3mm and allow for simultaneous macro- and microcontact recordings. Platinum has a high impedance for lower frequency and a low impedance for higher frequency bands. As such it is suitable to pick up local action potentials. The micro contacts extended radially past the endpoint of the macro depth electrode, and each contained eight high-impedance microwires (40-micron diameter) and one low-impedance microwire that is typically used for referencing.

The electrodes were connected to an ATLAS system (Neuralynx Inc, USA) consisting of CHET-10-A pre-amplifiers and a Digital Lynx NX amplifier and recorded with a sampling rate of either 32,000 Hz (Location: Birmingham) or 32,768 Hz (Location: Erlangen). Upon acquisition, an analogue bandpass filter from 0.1 Hz to 9,000 Hz was applied.

Spike detection and spike sorting
In the following paragraph, we will outline the process used to filter the raw data, detect spike timestamps, extract features of the waveshape and cluster spike waveshapes into putative single neurons using the wave_clus toolbox. For a more in-depth description of the wave_clus algorithm, the reader is referred to Chaure, Rey and Quiroga (2018).

The unfiltered signal included both the local field potential and the action potentials of individual neurons. Action potentials are characterized by a very steep and transient amplitude in the signal. To extract these spikes, we first applied zero-phase filtering using a second-order bandpass elliptic filter in the range of 300-3000 Hz. The resulting signal contained the information of the so-called spike-band.

Next, we segmented the continuous filtered data into epochs of five minutes. Segmenting the continuous data into smaller epochs had the advantage that noise in the signal did not increase the detection threshold for the whole recording and instead was limited to the segment in which it occurred (Chaure et al., 2018).

Spike detection was performed separately for positive and negative deflections. Once a spike was identified, 64 data points around the spike maximum were extracted. This corresponds to a 2ms window at a sampling rate of 32000 Hz. The spike peak was aligned to the 20th sampling point. To avoid misalignment of the spike, the waveshape was first upsampled to

320 data points using cubic spline-interpolated waveforms and then downsampled again (Chaure et al., 2018).

Based on the extracted spike waveform, features were computed using a four-scale multiresolution decomposition with a Haar wavelet. This results in 64-wavelet coefficients for each spike. The 10 most significant coefficients were identified using a Lilliefors test and used for the clustering procedure (Chaure et al., 2018).

Nonparametric clustering in the feature space was performed using superparamagnetic clustering (SPC). SPC grouped spike waves into clusters based on nearest-neighbour interactions (Blatt et al., 1996). Template-matching in Euclidian space was performed to assign unclassified waveforms to one of the identified clusters. The resulting clustering solution was then manually inspected and further optimized by rejecting artefact clusters, splitting clusters that represent multi-unit activity and merging clusters that likely stem from the same neural source. See Figure S3 to S5 for an overview of the spike width, spike height, the Fano factor and the firing rate separately for ESNs and all other single units.

Identification of Episode Specific Neurons (ESNs)

For every single unit, we determined the number of spikes within each correctly remembered episode. During encoding, spikes from the onset of the associate images (two seconds after the cue onset i.e., when the whole information of the episode was present) until the end of the episode were considered. During the retrieval phase, spikes from cue onset until the time point at which participants indicated how many images they remembered were considered. We chose this time window because an episode could be reinstated following cue presentation, while after the response patients were presented with an array of images that could have potentially induced single unit firing. Because the experiment was self-paced and longer episodes trivially contained more spikes, the firing rate (in hertz) was computed for each episode and single unit. In the next step, we z-scored this firing rate per single unit within all encoding episodes and retrieval episodes independently. We then multiplied this standardized firing rate for encoding and retrieval episodes elementwise to gain an indicator for the reinstatement of firing (Figure 1b).

To estimate a threshold at which episode-specific firing reinstatement occurs on a single unit level, we permuted the order of the encoding episodes and recomputed the elementwise product of the shuffled episode series. We repeated this permutation step 10,000 times and stored all output values. The 99th percentile of these pooled values was then used as a threshold for firing reinstatement. As an additional constraint, z-scored firing during encoding and retrieval each had to exceed 1.645 ($\triangleq$ $p_{right-tailed}$ < 0.05) to make sure the elementwise product was not predominantly driven by a high firing rate in one of the two phases alone (i.e., either encoding or retrieval). This procedure is allowing us to threshold, but we do not have family-wise error corrected statistical significance at the single unit level. Furthermore, we assume that single units fire independently.

In a second step, we calculated whether the number of ESNs (as identified in the above procedure) was above chance level. We did this by randomly choosing one of the permutations calculated in the first step for every single unit and checked whether it would be classified as an ESN under the same criteria outlined above. This approach is similar to a set-level effect in SPM (Penny et al., 2011). Note that for the shuffled data as well only correctly remembered episodes were used, hence lower values for shuffled data cannot be due to memory failure. This process was repeated 10,000 times and the total number of

single units which would be classified as an ESN in every single iteration of this process was used to build a distribution against which we compared our empirically discovered number of ESNs.

Identification of putative Concept Cells

We repeated the analysis under *Identification of Episode Specific Neurons (ESNs)* using a more conservative definition of ESN to exclude units coding for a specific face/place or faces and places in general. Here, firing had to exceed the threshold for at least one episode in which two places were used as associate stimuli and at least one episode in which two faces were used as associate stimuli. This definition seeks to ensure that ESN firing did not merely reflect firing to a specific concept (e.g., a certain category of faces) or specific stimulus, as these concepts and stimuli were orthogonal between episodes in which two places were used as associate stimuli and the episodes in which two faces were used as associate stimuli. To ensure concept neurons tuned to the animal cue were not falsely interpreted as ESN activity, we repeated the analysis outlined in *Identification of Episode Specific Neurons* again but excluded ESNs that showed a significant firing increase in response to the animal cue. Putative concept cells tuned to the animal cue were defined by the number of spikes between 300 and 1,000ms after cue onset during reinstated encoding episodes. Baseline activity was derived from the average number of spikes between 1,000 and 300ms before all cue onsets. We considered an ESN a putative concept neuron if the median number of spikes during the activity period of all reinstated trials was (1) at least two and (2) exceeded the mean spike number during the baseline plus five standard deviations (see Quiroga et al., 2005).

Firing-rate spike convolution

To produce the visualisations in Figure 3, we extracted spikes from one second before the cue onset until five seconds after cue onset for each episode. Binary spike times were convolved with a 251ms gaussian kernel (width factor: 2.5) to create a time-resolved signal of spike activity. We computed the average firing rate over time for all episodes (*ep*) during the baseline (*BL*) period 1,000ms preceding the animal cue ($\overline{x}_{BL}$). We then z-scored the spike activity during the episode ($x_{ep,t}$) using the standard deviation ($std(\overline{x}_{BL})$) and mean ($\overline{\overline{x}}_{BL}$) across all pre-cue baseline periods (see equation (1)). To account for instances where no spiking activity occurred during the baseline period, 0.1 (see Ison et al., 2015) was added to the standard deviation ($std(\overline{x}_{BL})$). Episodes were then split into reinstated and non-reinstated episodes. Firing rates for each episode type (reinstated/non-reinstated) were then averaged over ESNs.

$$z_{ep,t} = \frac{x_{ep,t}-\overline{\overline{x}}_{BL}}{std(\overline{x}_{BL})+0.1} \quad (1)$$

Identification of Time Cells

We defined the beginning of an encoding block as the most salient event. Based on Umbach and colleagues (Umbach et al., 2020), we then extracted all spikes within each block and convolved them with a 251ms gaussian kernel (width factor: 2.5). This created a block

number x time points matrix. For our first analysis, we cut each encoding block into 40 equally sized bins, thereby normalizing block duration. We then used a Kruskal-Wallis test to determine whether any of the 40 bins significantly differed from each other.

We then performed a circular shifting permutation test to calculate whether we found a significant number of Time Cells. This is done by shifting a random number of values from the beginning of the vector to the end. This shifting was imposed on each block separately and repeated N = 10,000 times for every single unit.

In a second test, the block length was determined by the longest block and shorter blocks were filled up with NaN values. This resulted in no normalization of time between blocks. The rest of the procedure is the same as described in the above paragraph.

Resource Availability

Data and analysis code will be made publicly available upon publication.

Author contributions

LDK and SH wrote the initial version of the manuscript. LDK and SH conceptualized the data analyses. LDK conducted the analyses and created the figures. FR and SH conceptualized and created the experiment. LDK, FR, GP, MTW, MVDP and KW were responsible for data acquisition. RC and VS were responsible for electrode implantation. RC, VS, DR, JL, SG and HH were responsible for patient care. RH and GK provided technical assistance with the data acquisition. BS, MW and HB contributed to the analyses. All authors were involved in finalizing the manuscript.

Competing interest statement

The authors declare no conflict of interest.

References

Blatt, M., Wiseman S., Domany E (1996). Superparamagnetic clustering of data. *Physical Review Letters, 76*, 3251.

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review, 114*(1), 38.

Chaure, F. J., Rey, H. G., & Quiroga, R. Q. (2018). A novel and fully automatic spike-sorting implementation with variable number of features. *Journal of Neurophysiology, 120*(4), 1859-1871.

Ison, M. J., Quiroga, R. Q., & Fried, I. (2015). Rapid encoding of new memories by individual neurons in the human brain. *Neuron, 87*(1), 220-230.

Lisman, J., Buzsáki, G., Eichenbaum, H., Nadel, L., Ranganath, C., & Redish, A. D. (2017). Viewpoints: how the hippocampus contributes to memory, navigation and cognition. *Nature Neuroscience, 20*(11), 1434-1447.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190.

Marr, D. (1971): Simple memory: theory for archicortex. *Philosophical Transactions of the Royal Society B, 262*, 23- 81.

Niediek, J., Boström, J., Elger, C. E., & Mormann, F. (2016). Reliable analysis of single-unit recordings from the human brain under noisy conditions: tracking neurons over hours. *PloS One, 11*(12), e0166598.

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (Eds.). (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.

Quiroga, R. Q. (2020). No pattern separation in the human hippocampus. *Trends in Cognitive Sciences, 24*, 994-1007.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435*(7045), 1102-1107.

Reddy, L., Zoefel, B., Possel, J., Peters, J., Dijksterhuis, D., Poncet, M., ... & Self, M. W. (2020). Human hippocampal neurons track moments in a sequence of events. *bioRxiv*.

Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99*(2), 195.

Teyler, T. J., & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral Neuroscience, 100*(2), 147.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*(1), 1-25.

Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B. E., Sperling, M., & Lega, B. (2020). Time cells in the human hippocampus and entorhinal cortex support episodic memory. *Proceedings of the National Academy of Sciences, 117*(45), 28463-28474.