



# Neurons detect cognitive boundaries to structure episodic memories in humans

Jie Zheng<sup>1</sup>, Andrea G. P. Schjetnan<sup>2</sup>, Mar Yebra<sup>3</sup>, Bernard A. Gomes<sup>3</sup>, Clayton P. Mosher<sup>3</sup>, Suneil K. Kalia<sup>2</sup>, Taufik A. Valiante<sup>2,4,5,6</sup>, Adam N. Mamelak<sup>3</sup>, Gabriel Kreiman<sup>1,7</sup>✉ and Ueli Rutishauser<sup>3,8,9,10</sup>✉

**While experience is continuous, memories are organized as discrete events. Cognitive boundaries are thought to segment experience and structure memory, but how this process is implemented remains unclear. We recorded the activity of single neurons in the human medial temporal lobe (MTL) during the formation and retrieval of memories with complex narratives. Here, we show that neurons responded to abstract cognitive boundaries between different episodes. Boundary-induced neural state changes during encoding predicted subsequent recognition accuracy but impaired event order memory, mirroring a fundamental behavioral tradeoff between content and time memory. Furthermore, the neural state following boundaries was reinstated during both successful retrieval and false memories. These findings reveal a neuronal substrate for detecting cognitive boundaries that transform experience into mnemonic episodes and structure mental time travel during retrieval.**

Our lives unfold over time, weaving rich information into a continuous sequence of experiences. However, our memories are not continuous. Rather, we remember discrete episodes ('events')<sup>1</sup>, which serve as anchors to bind together the myriad different aspects (where, when and what) of a given autobiographical memory<sup>2,3</sup>, much like objects do in perception<sup>4</sup>. A fundamental unresolved question in human memory is, therefore, what marks the beginning and the end of an episode?<sup>5</sup>

The transformation from ongoing experience to distinct events is thought to rely on the identification of boundaries that separate events<sup>1,6–11</sup>. This theory is motivated by large-scale patterns of activity changes in the human brain around event boundaries<sup>5,12,13</sup>, but the underlying neural mechanisms and their relationship to memory are unknown. Neurons in the rodent hippocampus elevate their firing rates in the vicinity of investigator-imposed spatial boundaries<sup>14</sup>, and the place fields of hippocampal neurons are shaped by physical boundaries<sup>15–17</sup>. In accordance with the boundaries of subenvironments<sup>14</sup>, hippocampal place fields remap<sup>18,19</sup> in response to context shifts and are reinstated<sup>15,20</sup> when the animal is placed back into a familiar context. Additionally, rodent hippocampal neuron ensembles encode lap-specific representations in a maze irrespective of an animal's spatial location<sup>21</sup>, presumably representing cognitive boundaries between distinct events. Boundaries shape mnemonic representations of both spatial environments and the events that occur during navigation and structure the place fields and event-specific representations of cognitive maps. No such understanding at the single-cell level exists for the non-spatial episodic memories that define us as individual human beings<sup>2,22</sup>.

We investigated the neuronal mechanisms underlying the identification of event boundaries in humans under semirealistic

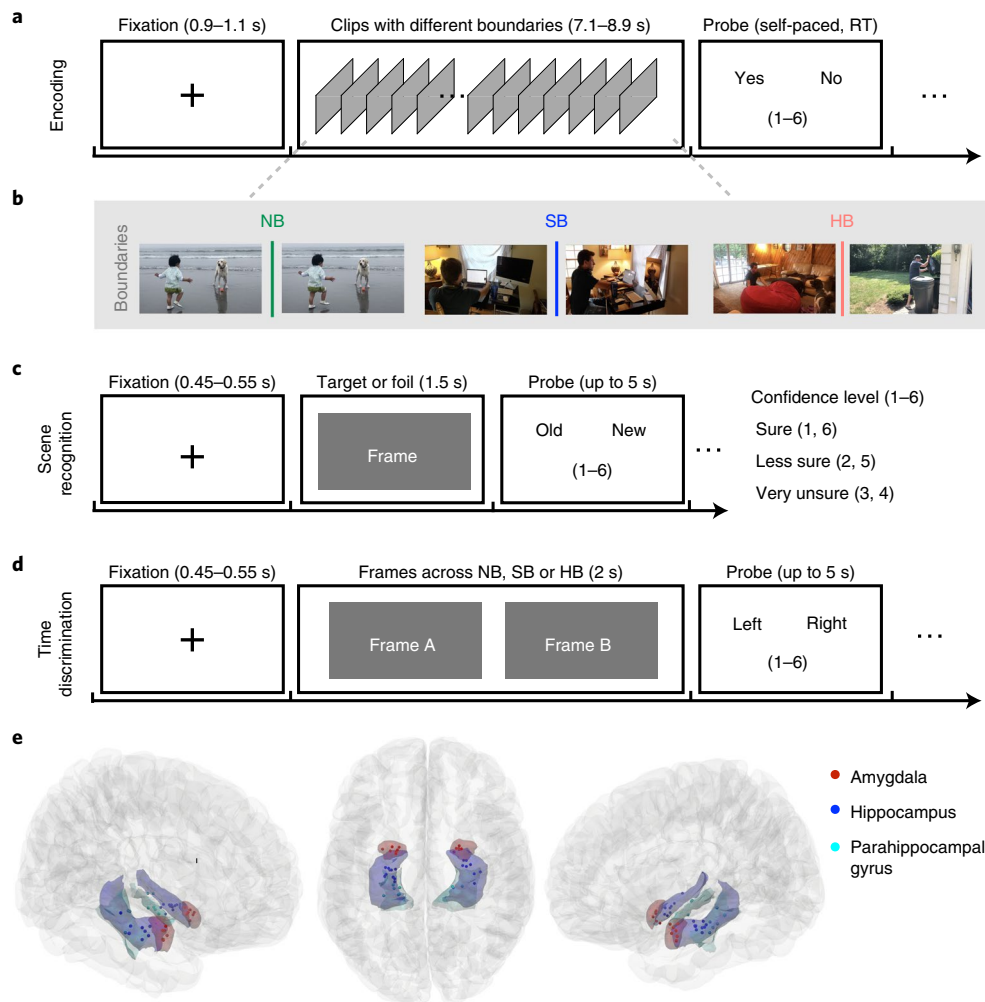
continuous experience. We recorded single-neuron activity from individuals with drug-resistant epilepsy implanted with depth electrodes<sup>23</sup> while testing their memory for the content of video clips with two kinds of embedded cognitive boundaries: soft boundaries (SBs) and hard boundaries (HBs). SBs are episodic transitions between related events within the same movie, while HBs are episodic transitions between two unrelated movies. Behaviorally, both SBs and HBs enhanced recognition of video clip content that followed a boundary, whereas HBs impaired memory of the temporal order of events. We found neurons in the MTL that signaled the timing of both types of boundaries. The activity of these boundary-responsive neurons predicted memory strength as assessed by scene recognition and temporal order discrimination accuracy.

## Results

### Boundaries boost recognition but disrupt serial order memory.

We studied how boundaries influence the formation and retrieval of memories of brief video clips. Twenty individuals performed the task while we recorded the activity of single neurons (Fig. 1e, Extended Data Fig. 1 and Supplementary Tables 2 and 3 show the participant demographics and the location of microwire bundles). The task consisted of three parts: encoding, scene recognition and time discrimination. During encoding (Fig. 1a), individuals watched 90 different and new video clips containing either no boundaries (NBs; one continuous movie shot), SBs (cuts to a new scene within the same movie) or HBs (cuts to a new scene from a different movie; Fig. 1b). A question about the prior movie appeared every four to eight clips (for example, is anyone wearing glasses?). Participants answered  $89 \pm 5\%$  of these questions accurately.

<sup>1</sup>Department of Ophthalmology, Children's Hospital, Harvard Medical School, Boston, MA, USA. <sup>2</sup>Krembil Brain Institute and Division of Neurosurgery, University Health Network (UHN), University of Toronto, Toronto, Ontario, Canada. <sup>3</sup>Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA, USA. <sup>4</sup>Department of Surgery (Neurosurgery), Institute of Biomedical Engineering, and Electrical and Computer Engineering, University of Toronto, Toronto, Ontario, Canada. <sup>5</sup>Max Planck-University of Toronto Center for Neural Science and Technology, University of Toronto, Toronto, Ontario, Canada. <sup>6</sup>Center for Advancing Neurotechnological Innovation to Application, University Health Network, University of Toronto, Toronto, Ontario, Canada. <sup>7</sup>Center for Brains, Minds and Machines, Cambridge, MA, USA. <sup>8</sup>Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, CA, USA. <sup>9</sup>Center for Neural Science and Medicine, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA. <sup>10</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. ✉e-mail: [gabriel.kreiman@childrens.harvard.edu](mailto:gabriel.kreiman@childrens.harvard.edu); [ueli.rutishauser@cshs.org](mailto:ueli.rutishauser@cshs.org)



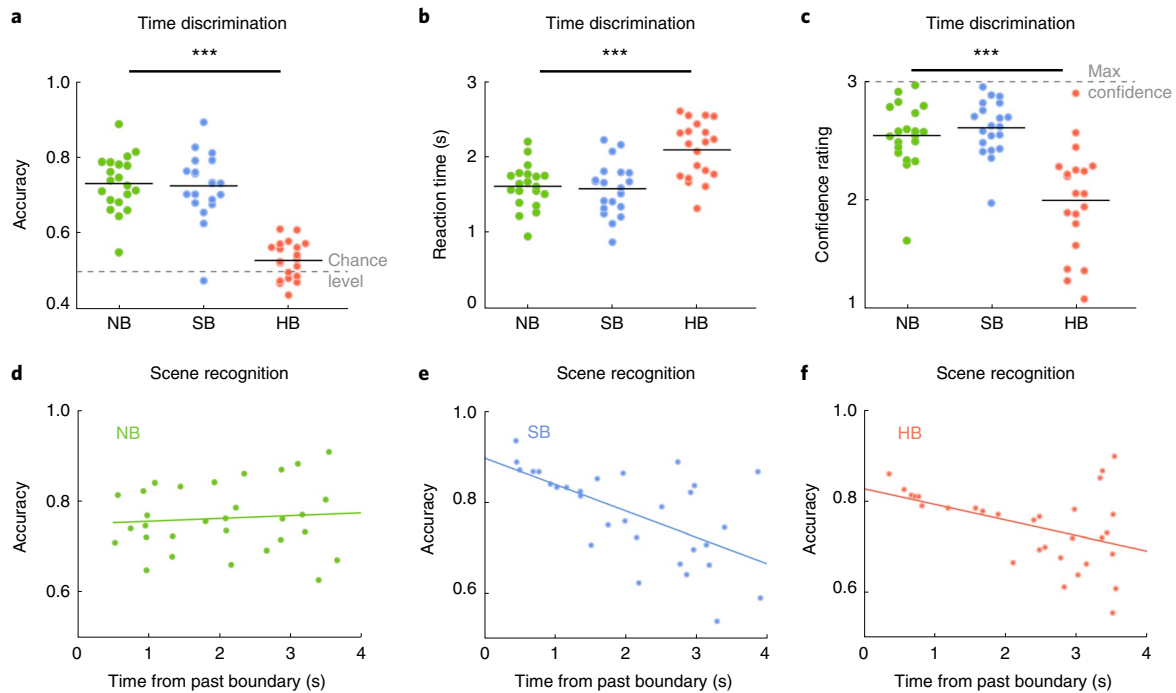
**Fig. 1 | Experiment and recording locations.** **a**, Encoding task. Individuals watched 90 video clips (~8 s each, no audio) with either NB (continuous movie shot), an SB (cut to a new scene within the same movie, one to three SBs per clip) or an HB (cut to a different movie, one HB per clip). Every four to eight clips, individuals were prompted to answer a yes or no question related to the content of the immediately preceding clip together with a confidence rating (Methods). RT, reaction time. **b**, Example boundaries (visual features of boundaries are in Supplementary Table 1). Owing to copyright restrictions, the images shown are different from those used for the experiment. **c**, Scene recognition memory task. Individuals indicated whether a static image was new or old (seen during encoding task) together with a confidence rating. **d**, Time discrimination task. Individuals indicated which of two frames they saw first during the encoding task together with a confidence rating. **e**, Recording locations of the 39 microwire bundles that contained at least one boundary/event neuron (see Montreal Neurological Institute (MNI) coordinates in Supplementary Table 3 and Extended Data Fig. 1) across all individuals (participant information is in Supplementary Table 2) in the amygdala (red), hippocampus (blue) or parahippocampal gyrus (cyan), rendered on a template brain. Each dot represents the location of a microwire bundle.

We subsequently evaluated what individuals remembered about the video clips with two memory tests: scene recognition (Fig. 1c) and time discrimination (Fig. 1d). During the scene recognition test, individuals were presented with a single static frame. These frames were chosen with equal probability from either the previously presented video clips ('targets') or from other video clips that were not shown to the participants ('foils'). Participants made an 'old' or 'new' decision together with a confidence rating (sure, less sure and very unsure) in each trial. During the time discrimination test, individuals were shown two old frames chosen from the same video clip, presented side by side, and had to indicate which frame was seen earlier in time together with a confidence rating.

In the time discrimination task, participants correctly identified which frame was shown first in  $73 \pm 7\%$  and  $73 \pm 8\%$  of trials when the two frames were separated by a NB or an SB, respectively (Fig. 2a; both above chance of 50%; one sample *t*-test, NB:  $t_{19} = 13.97$ ,

$P = 2 \times 10^{-11}$ ; SB:  $t_{19} = 11.63$ ,  $P = 4 \times 10^{-10}$ ). By contrast, participants performed significantly worse when discriminating the order of frames separated by an HB (Fig. 2a; HB:  $53 \pm 5\%$ ,  $P = 0.02$  against chance level; significantly lower than NB and SB:  $F_{2,57} = 51.33$ ,  $P = 2 \times 10^{-13}$ ). Individuals also showed longer reaction times (Fig. 2b; HB:  $2.10 \pm 0.37$  s; NB:  $1.62 \pm 0.28$  s; SB:  $1.59 \pm 0.34$  s;  $F_{2,57} = 14.25$ ,  $P = 10 \times 10^{-6}$ ) and lower confidence ratings when discriminating the order of frames separated by an HB (Fig. 2c; HB:  $1.95 \pm 0.45$ ; NB:  $2.52 \pm 0.29$ ; SB:  $2.59 \pm 0.23$ ;  $F_{2,57} = 20.41$ ,  $P = 2 \times 10^{-7}$ ). This effect on reaction times and confidence was not driven by accuracy differences, as it was observed for both correct and incorrect trials independently (Supplementary Fig. 1). Discriminating the temporal order of two frames was not possible by reasoning alone without having seen the video clips (Supplementary Fig. 2).

Across all trials, the ability to recognize a frame as old did not differ significantly between the types of boundaries preceding the

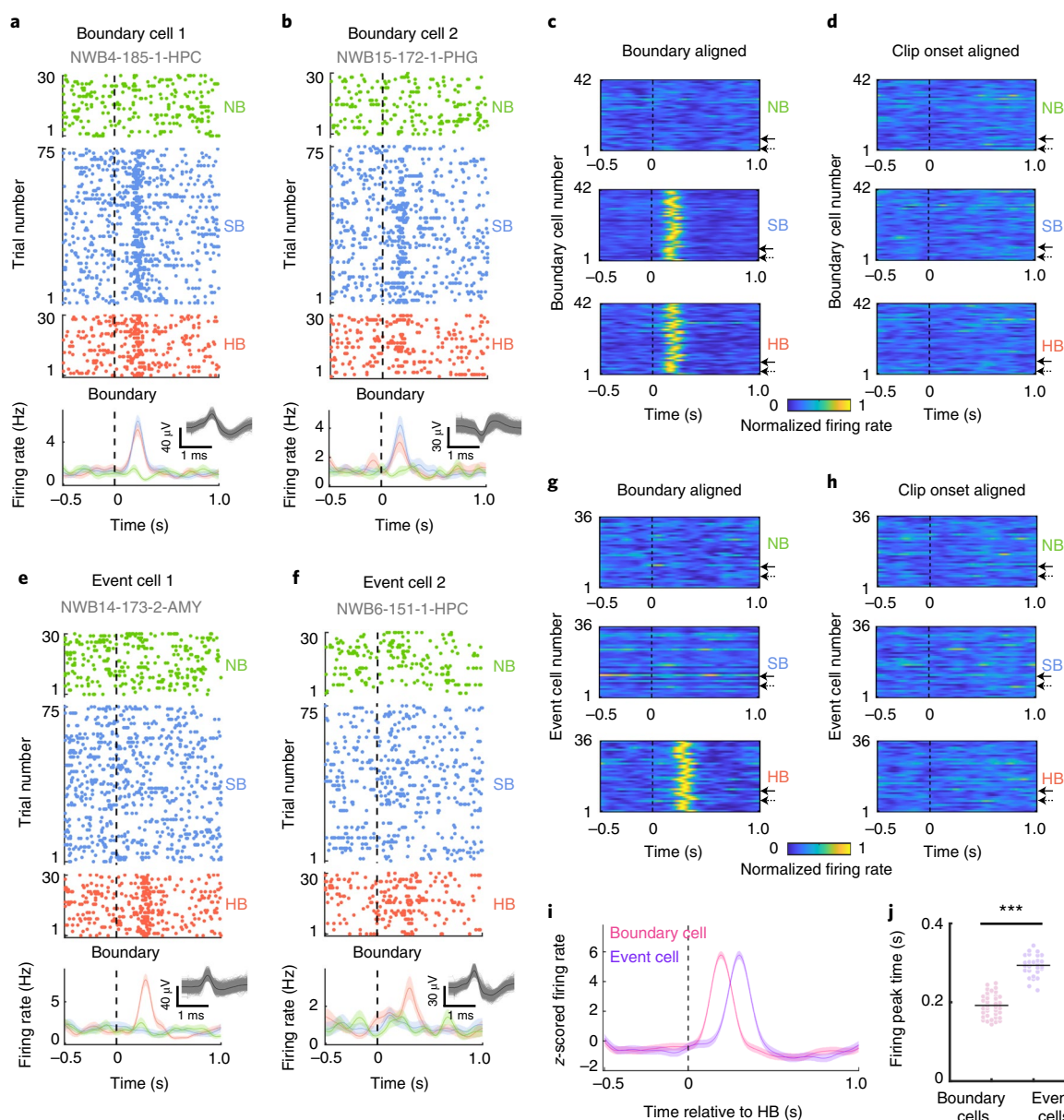


**Fig. 2 | Behavior.** HBs impaired time discrimination memory, while SBs and HBs improved scene recognition memory for frames close to them. **a–c**, Performance in time discrimination task (see also Supplementary Figs. 1 and 2) quantified by accuracy (**a**) ( $F_{2,57}=51.33$ ,  $P=2\times 10^{-13}$ , one-tailed analysis of variance (ANOVA)), reaction time (**b**) ( $F_{2,57}=14.25$ ,  $P=10\times 10^{-6}$ , one-tailed ANOVA) and mean confidence level (**c**) ( $F_{2,57}=20.41$ ,  $P=2\times 10^{-7}$ , one-tailed ANOVA) across all the trials for NBs (green), SBs (blue) and HBs (red). Behavioral data for the scene recognition task are shown in Extended Data Fig. 2. **d–f**, Scene recognition accuracy as a function of time elapsed between the target frame and its nearest past boundary (the distance effect for time discrimination accuracy and future boundaries is shown in Supplementary Fig. 3) plotted separately for NB (**d**), SB (**e**) and HB (**f**). For NB clips, time from the past boundary is measured relative to the middle of the clip. Each dot represents one recording session in **a–c** and one clip in **d–f**. Black lines in **a–c** denote the mean of the results, and colored lines in **d–f** are the fitted lines for linear regression; \*\*\* $P<0.001$ .

frame (Extended Data Fig. 2a; NB:  $76\% \pm 10\%$ ; SB:  $75\% \pm 9\%$ ; HB:  $75\% \pm 8\%$ ;  $F_{2,57}=0.07$ ,  $P=0.94$ ). The reaction times and confidence ratings during the scene recognition task were also similar across the different types of boundaries (reaction times are presented in Extended Data Fig. 2b; NB =  $1.47 \pm 0.18$  s, SB =  $1.43 \pm 0.16$  s and HB =  $1.49 \pm 0.15$  s,  $F_{2,57}=0.28$ ,  $P=0.76$ ; confidence ratings are presented in Extended Data Fig. 2c; NB =  $2.60 \pm 0.18$ , SB =  $2.60 \pm 0.20$  and HB =  $2.52 \pm 0.28$ ,  $F_{2,57}=0.54$ ,  $P=0.56$ ). Therefore, the impaired time discrimination ability across HB transitions was not due to differences in memory strength as measured by scene recognition accuracy. Even though the overall accuracy was similar among NB, SB and HB conditions, the recognition accuracy of target frames decreased as a function of the time elapsed between the target frame and its immediately preceding boundary. Target frames presented shortly after an SB and HB were remembered better than those farther away from the boundary (Fig. 2e,f; SB:  $r=-0.61$ ,  $P=4\times 10^{-4}$ ; HB:  $r=-0.44$ ,  $P=0.015$ ). By contrast, recognition accuracy did not differ significantly as a function of time relative to NBs (Fig. 2d; Pearson correlation; NB:  $r=0.085$ ,  $P=0.65$ ). The temporal distance to boundary effects was unidirectional; the temporal distance to future boundaries did not correlate with memory performance (Supplementary Fig. 3a,b). Additionally, no temporal distance effect was present in the time discrimination task accuracy (Supplementary Fig. 3c,d). Also, the scene recognition accuracy and time discrimination accuracy were not significantly related to the time at which the tested frames were shown during encoding (Supplementary Fig. 4; scene recognition:  $F_{5,114}=1.87$ ,  $P=0.11$ ; time discrimination:  $F_{5,114}=1.1$ ,  $P=0.37$ ). Together, the behavioral analyses revealed that frames that closely followed an SB or HB were more likely to be remembered. Temporal order memory, however,

was disrupted by the presence of HBs. These results reveal a tradeoff in the effect of HBs on memories, with enhanced scene recognition memory and disrupted temporal order memory.

**MTL neurons demarcate episodic transitions.** We next investigated the neuronal responses to boundaries and their relationship to memory by recording from neurons in the MTL (including the hippocampus, amygdala and parahippocampal gyrus; Fig. 1e and Extended Data Fig. 1) and other brain areas (Supplementary Tables 2 and 4). Across all areas, we recorded the activity of 985 neurons from 19 individuals (1 of the 20 individuals yielded no usable recordings; Supplementary Table 2). Of these 985 neurons, 580 were recorded from the MTL. We first tested whether neurons changed their activity following boundaries by comparing their firing rate in a 1-s-long window following boundaries relative to baseline (1 s before boundary; Methods). For video clips with NBs, we aligned responses to the middle of the clip and compared responses between before and after this virtual boundary. Fig. 3a,b shows two example neurons recorded from the hippocampus and parahippocampal gyrus, respectively. These neurons showed a transient increase in firing rates within approximately 300 ms after both SBs (blue) and HBs (red). No such change was observed in the clips without boundaries (green). We refer to this type of neuron as a 'boundary cell'; 42 of 580 MTL neurons (7.24%; expected proportion by chance for all MTL neurons = 2.11%) were classified as boundary cells (Fig. 3c). The regions with the largest proportion of boundary cells were the parahippocampal gyrus ( $n=18/68$ , 26.47%), amygdala ( $n=12/169$ , 7.10%) and hippocampus ( $n=12/343$ , 3.50%). These proportions are all significantly larger than expected by chance ( $P<0.05$ ; Supplementary Table 4).



**Fig. 3 | Boundary cells and event cells demarcate different types of episodic transitions.** **a,b**, Responses during the encoding stage from two example boundary cells located in the hippocampus (HPC) (**a**) and parahippocampal gyrus (PHG) (**b**), respectively (spike sorting quality of all detected cells is shown in Supplementary Fig. 5). Boundary cells responded to both SB (blue) and HB (red) transitions. Responses are aligned to the middle point of the clip (NB, green) or to the boundary (SB, HB); top, raster plots; bottom, poststimulus time histogram (bin size = 200 ms; step size = 2 ms; shaded areas represent  $\pm$ s.e.m. across trials); insets, all spike extracellular waveforms (gray) and mean (black). **c,d**, Firing rates of all 42 boundary cells (solid and dashed arrows denote the examples in **a** and **b**, respectively) during the encoding stage aligned to the boundaries (**c**) or clip onsets (**d**) averaged over trials within each boundary type and normalized to each neuron's maximum firing rate from the entire task recording (see color scale on bottom). **e,f**, Responses during the encoding stage from two example event cells located in the amygdala (AMY) (**e**) and hippocampus (**f**), respectively. Event cells responded to HB (red) but not SB (blue) transitions; poststimulus time histogram: bin size = 200 ms, step size = 2 ms and shaded areas represent  $\pm$ s.e.m. across trials. **g,h**, Firing rates of all 36 event cells (solid and dashed arrows denote the examples in **e** and **f**, respectively) during the encoding stage, using the same format as in **c** and **d** (aligned to boundaries (**g**) or clip onsets (**h**)). Both boundary cells and event cells in the MTL do not respond to the clip onsets (**d, h**) and clip offsets (Extended Data Fig. 3) during encoding and image onsets and offsets during scene recognition and time discrimination (Extended Data Fig. 4). No significant difference in saccades was found after clip onsets versus after boundary transitions for one individual where we could record eye movement data simultaneously with the neurophysiological data (Supplementary Fig. 7). **i**, Latency analysis. Firing rate during HB transitions (to which both boundary cells and event cells responded) reached peak response earlier for boundary cells (pink) than event cells (purple). The average z-scored firing rate normalized using the average and s.d. of the firing rates and aligned to HB is shown (bin size = 200 ms, step size = 2 ms, shaded areas represent  $\pm$ s.e.m. across all boundary cells or event cells). **j**, Peak times of average firing rate traces of all boundary cells (pink) and all event cells (purple) ( $F_{1,76} = 274.78$ ,  $P = 6 \times 10^{-27}$ , one-tailed ANOVA). Each dot represents one boundary cell (pink) or one event cell (purple). Black lines denote the mean averaged across all boundary cells or event cells; \*\*\* $P < 0.001$ , one-way ANOVA, d.f. = 1,76. The spatial distribution of boundary cells and event cells is shown in Supplementary Table 4.

Was the response of boundary cells a result of the abrupt changes in pixel-level content between the frame before and after the boundary? To answer this question, we considered the responses of the cells during other abrupt changes of visual input: video clip onsets (Fig. 3d) and offsets (Extended Data Fig. 3) during encoding and image onsets and offsets during scene recognition and time discrimination (Extended Data Fig. 4). Boundary cells did not respond significantly to either clip or image onsets or offsets ( $P > 0.05$ ; permutation  $t$ -test, Methods), showing that the response of boundary cells is likely related to higher-level cognitive discontinuities rather than pure visual changes.

We also found a second group of neurons that transiently increased their firing rate only following HBs but not SBs or NBs. Two examples of such cells, located in the amygdala and hippocampus, are shown in Fig. 3e,f. We refer to this type of neuron as an 'event cell'; 36 of 580 MTL neurons (6.20%; proportion expected by chance for all MTL neurons = 2.26%) were classified as event cells (Fig. 3g). The regions with the largest proportion of event cells were the hippocampus ( $n = 27/343$ , 7.87%), amygdala ( $n = 7/169$ , 4.27%) and parahippocampal gyrus ( $n = 2/68$ , 2.94%). These proportions are all significantly larger than expected by chance ( $P < 0.05$ ; Supplementary Table 4). Similar to boundary cells, event cells did not significantly change their firing rates following video clip onsets or offsets (Fig. 3h and Extended Data Fig. 3) during encoding nor did they respond to image onsets or offsets during scene recognition or time discrimination (Extended Data Fig. 4;  $P > 0.05$ , permutation  $t$ -test; Methods). However, boundary cells and event cells did increase their firing rates to the randomly interspersed probe questions that followed some clip offsets (randomly present every four to eight trials; Supplementary Fig. 6). Seventy six of 580 (13.1%,  $P = 0.01$ ) and 4 of 580 (0.7%,  $P = 0.43$ ) MTL cells changed their firing rate in response to clip onsets and clip offsets, respectively, but neither of these cells qualified as boundary cells or event cells (Extended Data Fig. 5).

SBs and HBs differed in terms of their high-level conceptual narrative, which is interrupted in HBs but not in SBs. To evaluate whether it is possible to determine from visual features alone whether a boundary is soft or hard, we computed the differences

between pre- and postboundary frames in terms of pixel-level characteristics (that is, luminance, contrast, complexity, entropy and color distribution), high-level visual features (that is, objects) and perceptual similarity ratings. These analyses revealed that SB and HB transitions did not differ significantly from each other in any of the attributes we tested (Supplementary Table 1). Therefore, the differential activation of event cells to HBs but not SBs was likely a result of detection of the disruption in the conceptual narrative, that is, a transition between two different episodes.

While both boundary cells and event cells responded to HB transitions, a comparison of their response dynamics indicated that boundary cells responded to HBs approximately 100 ms earlier than event cells (Fig. 3i). This latency difference was also observed when comparing the time at which the peak responses were reached; boundary cells showed a peak at  $197 \pm 49$  ms, whereas event cells showed a peak at  $301 \pm 55$  ms (Fig. 3j;  $F_{1,76} = 274.78$ ,  $P = 6 \times 10^{-27}$ ).

We also evaluated the existence of boundary and event cells in brain areas other than the MTL, such as the medial frontal cortex, insula and orbitofrontal cortex (OFC). We found 8/405 (1.96%) boundary cells and 9/405 (2.22%) event cells among the non-MTL group (Supplementary Tables 2 and 4), with only event cells in the OFC exceeding the proportions expected by chance. Thus, boundary-responsive neurons are largely within the MTL, to which we restricted the following analyses.

### Responses of boundary and event cells predict memory strength.

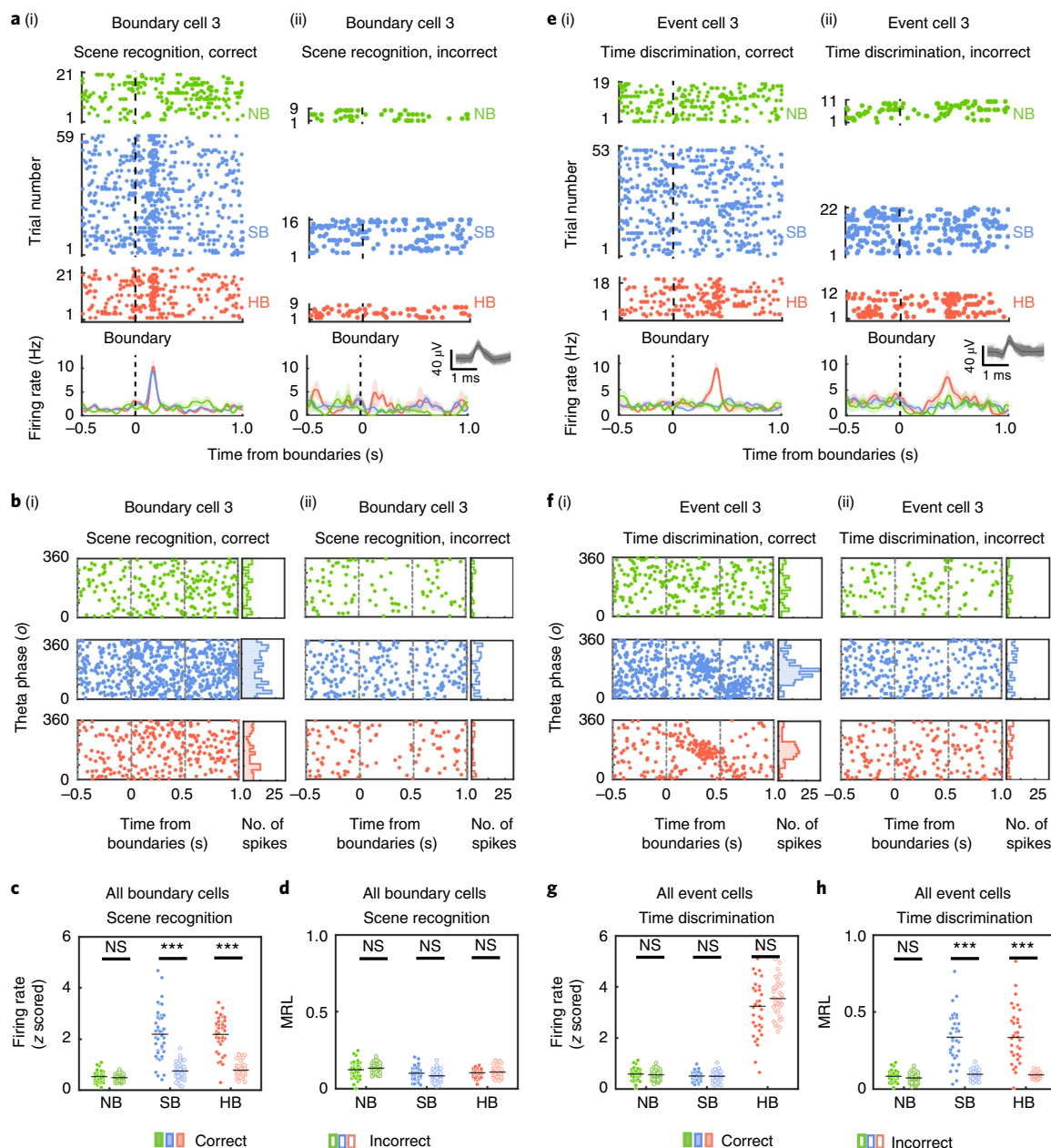
We next asked whether the responses of boundary cells and event cells during encoding correlated with later measures of memory for the content of the videos. We examined whether the strength of responses of boundary cells or event cells to boundaries varied as a function of whether the familiarity or temporal order of a stimulus was later remembered or forgotten. Fig. 4a shows an example boundary cell located in the hippocampus whose response during encoding differed between video clips, from which frames were later correctly remembered as familiar (Fig. 4a(i)) versus incorrectly identified as new (forgotten; Fig. 4a(ii)); the responses to boundaries that preceded later-remembered frames were stronger. This effect was present, on average, among boundary cells ( $n = 42$ ) for

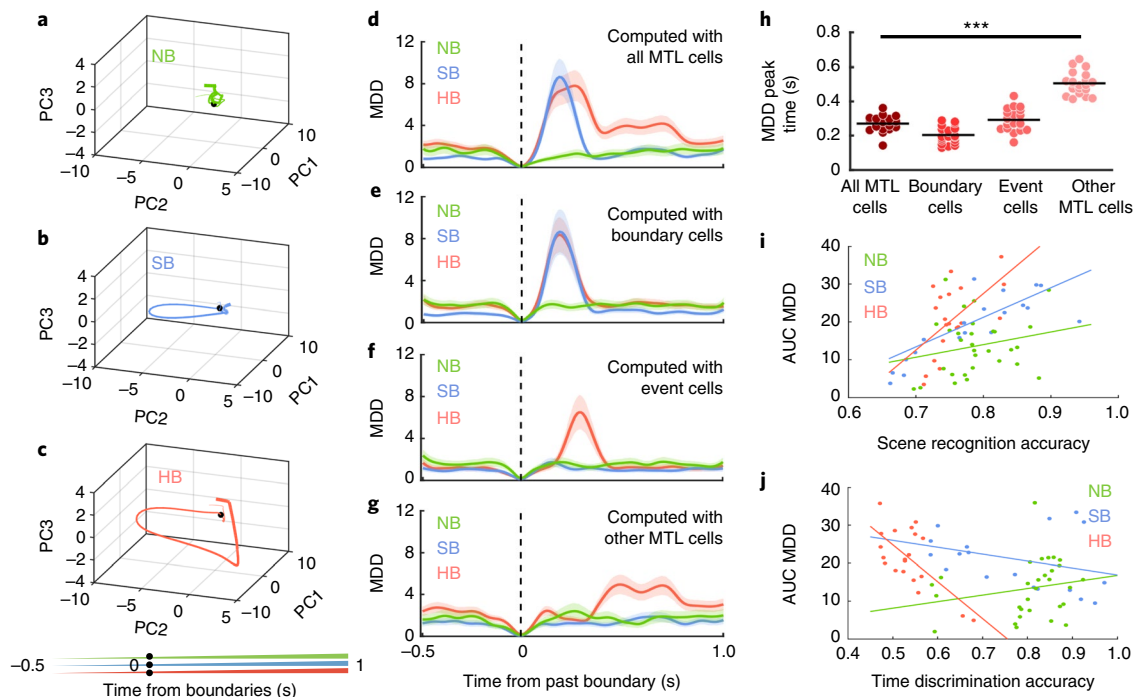
**Fig. 4 | Responses of boundary cells and event cells during encoding correlate with later retrieval success. a–d**, Response of boundary cells during encoding grouped by participants' subsequent memory performance in the scene recognition task. **a**, Boundary cell recorded in the hippocampus. During encoding, this cell responded more strongly to SB and HB transitions than NB if the frame following the boundary in that trial was correctly identified during the scene recognition task (i) compared to incorrect trials (ii). The format is as in Fig. 3. Shaded areas represent  $\pm$ s.e.m. across trials. **b**, Timing of spikes from the same boundary cell shown in **a** relative to theta phase calculated from the LFPs for clips of which frames were later remembered (i) or forgotten (ii) (left). The phase distribution of spike times in the 1-s period following the middle of the clip (NB) or boundary (SB, HB) for clips from which frames were remembered (i) and forgotten (ii) is shown on the right. **c,d**, Population summary for all 42 boundary cells. Black lines denote the mean results averaged across all 42 boundary cells. **c**, Z-scored firing rate (0–1s after boundaries during encoding) differed significantly between boundaries after which frames were remembered (color filled) versus forgotten (empty) for both SB and HB (SB:  $F_{1,82} = 82.93$ ,  $P = 4 \times 10^{-14}$ ; HB:  $F_{1,82} = 156.9$ ,  $P = 1 \times 10^{-20}$ ; NB:  $F_{1,82} = 1.18$ ,  $P = 0.28$ ; one-tailed ANOVA; NS, not significant). **d**, Mean resultant length (MRL) of spike times (that is, sum of vectors with vector lengths equal to 1 and vector angles equal to the spike timings relative to theta phases 0–1s after boundaries during encoding divided by total number of vectors; value range (0 to 1): 0, uniform distribution (i.e., neurons fire at random theta phases); 1, unimodal distribution (i.e., neurons firing at the same theta phase across all boundary cells for each boundary type did not differ significantly between correct (color filled) and incorrect (empty) clips). **e,f**, Response of boundary cells during encoding grouped by participants' subsequent memory performance in the time discrimination task. **e**, Example event cell recorded in the hippocampus that responded to HB transition regardless of whether the temporal order of the clip was later correctly (i) or incorrectly (ii) recalled in the time discrimination task. Shaded areas represent  $\pm$ s.e.m. across trials. The format is the same as in **a**, but clips were grouped based on memory outcomes in the time discrimination task. **f**, The spike timing of the same event cell shown in **e** relative to theta phase plotted for correct (i) and incorrect (ii) trials. The format is the same as in **b**, but clips were grouped based on memory outcomes in the time discrimination task. **g,h**, Population summary for all 36 event cells. Black lines denote the mean results averaged across all 36 event cells. **g**, Z-scored firing rate (0–1s after boundaries during encoding) did not differ significantly between later correctly (color filled) or incorrectly (empty) remembered temporal orders for all three boundary types. **h**, MRL of spike times (relative to theta phases, 0–1s after boundaries during encoding) was significantly larger after SB and HB transitions if the temporal order of the clip was correctly recalled (color filled) than if it was incorrectly recalled (empty) (SB:  $F_{1,70} = 81.55$ ,  $P = 2 \times 10^{-13}$ ; HB:  $F_{1,70} = 60.79$ ,  $P = 4 \times 10^{-11}$ ; NB:  $F_{1,70} = 1.53$ ,  $P = 0.22$ ; one-tailed ANOVA). Each dot represents one boundary cell (**c** and **d**) or one event cell (**g** and **h**). Black lines in **c**, **d**, **g** and **h** denote the mean of the results. Note that in **a–d**, the neural responses of boundary cells reflect whether participants remembered or forgot target frames that followed a boundary. Results computed based on trials grouped by participants' memory performance for a target frame before a boundary are shown in Supplementary Fig. 8;  $***P < 0.001$ , one-way ANOVA, d.f. = 1,82 for **c** and **d**, and d.f. = 1,70 for **g** and **h**.

frames preceded or followed by both SBs and HBs but not by NBs (Fig. 4c, SB:  $F_{1,82} = 82.93, P = 4 \times 10^{-14}$ ; HB:  $F_{1,82} = 156.9, P = 1 \times 10^{-20}$ ; NB:  $F_{1,82} = 1.18, P = 0.28$ ; Supplementary Fig. 8, SB:  $F_{1,82} = 91.67, P = 5 \times 10^{-15}$ ; HB:  $F_{1,82} = 62.78, P = 1 \times 10^{-11}$ ; NB:  $F_{1,82} = 0.05, P = 0.83$ ). This effect was specific to scene recognition and boundary cells. First, the firing rate of boundary cells did not significantly predict performance in the time discrimination task (Extended Data Fig. 6a,c; NB:  $F_{1,82} = 1.25, P = 0.27$ ; SB:  $F_{1,82} = 1.35, P = 0.25$ ; HB:  $F_{1,82} = 1.14, P = 0.29$ ). The firing rate of event cells ( $n = 36$ ) during encoding was not predictive of scene recognition memory and temporal order memory (Fig. 4e,g and Extended Data Fig. 7a,c; NB:  $F_{1,70} = 1.12, P = 0.29$ ; SB:  $F_{1,70} = 1.63, P = 0.21$ ; HB:  $F_{1,70} = 0.79, P = 0.38$ ).

Given the importance of theta frequency band spike field coherence in plasticity<sup>24</sup>, we next considered the timing of spikes with respect to the theta band in the local field potentials (LFPs; 4–8 Hz, measured on the same microwire from which the neuron was recorded; Methods and Supplementary Fig. 9). We determined the

theta phase of each spike that occurred within 1 s following boundaries and compared the resulting phase distributions among NB, SB and HB. Event cells tended to fire at a consistent phase of the theta band LFP following both HBs and SBs for clips whose temporal order was later remembered correctly (Fig. 4f). To summarize this effect across the population, we computed the MRL across all phases for all spikes fired by a given cell (Methods). If the time of spikes are randomly distributed, the MRL equals 0, whereas an identical phase for all spikes would result in an MRL of 1. The mean MRL across all event cells ( $n = 36$ ) was significantly larger following both SB and HB but not NB if temporal order was later correctly remembered (Fig. 4h; SB:  $F_{1,70} = 81.55, P = 2 \times 10^{-13}$ ; HB:  $F_{1,70} = 60.79, P = 4 \times 10^{-11}$ ; NB:  $F_{1,70} = 1.53, P = 0.22$ ). This effect was specific to event cells and temporal order memory. First, the strength of theta phase locking of event cells did not predict scene recognition memory success (Extended Data Fig. 7b,d; NB:  $F_{1,70} = 0.75, P = 0.39$ ; SB:  $F_{1,70} = 1.1, P = 0.30$ ; HB:  $F_{1,70} = 2.13, P = 0.15$ ). Second, the strength of phase locking of boundary cells neither predicted scene recognition





**Fig. 5 | Population neural state shift magnitude following episodic transitions reflects participants' subsequent memory performance.** **a–c**, Trajectories in neural state space formed by the top three PCs (with most explained variance: PC1=26.05%, PC2=10.89% and PC3=6.69%) summarizing the activity of all MTL cells during the encoding stage for clips containing NBs (**a**), SBs (**b**) and HBs (**c**). Each data point indicates the neural state at a specific time relative to boundary onset (line thickness indicates time; see scale on bottom). Black dots mark the time of the boundary (SB, HB) or the middle of the clip (NB). **d–g**, MDD (that is, Euclidean distance relative to boundaries in the PC space formed by all PCs that cover explained variance  $\geq 99\%$ ) as a function of time aligned to the middle of the clip (green, NB) and boundaries (blue, SB; red, HB). MDD is shown for all MTL cells (**d**;  $n=580$  in the top 55 PCs space), all boundary cells (**e**;  $n=42$  in the top 27 PCs space), all event cells (**f**;  $n=36$  in the top 26 PCs space) and all other MTL cells (that is, non-boundary/event cells in the MTL; **g**;  $n=502$  in the top 58 PCs space). Shaded areas represent  $\pm$ s.e.m. across trials. See neural state shifts within each individual in Extended Data Fig. 8. **h**, Latency analysis. Time when MDD shown in **d–g** reached peak value following HB (red lines) significantly differed when computed with different groups of cells ( $F_{3,76}=103.96$ ,  $P=8 \times 10^{-27}$ , one-tailed ANOVA). Black lines denote the mean results averaged across different cell populations. **i, j**, Correlation between distance traveled in state space following boundaries and behavior. **i**, Positive correlation between area under the curve (AUC) MDD (sum of Euclidean distances within a 0- to 1-s time window after boundaries in the PC space) and scene recognition accuracy. Dots mark the accuracy in the scene recognition task (x axis) and the AUC MDD during encoding (y axis) of the target frames plotted separately for frames that follow NB (green,  $r=0.214$ ,  $P=0.256$ , Pearson correlation), SB (blue,  $r=0.653$ ,  $P=0.002$ , Pearson correlation) and HB (red,  $r=0.565$ ,  $P=0.009$ , Pearson correlation). **j**, Negative correlation between the AUC MDD versus time discrimination accuracy plotted in the same format as in **i** for NB (green,  $r=0.212$ ,  $P=0.261$ , Pearson correlation), SB (blue,  $r=-0.273$ ,  $P=0.244$ , Pearson correlation) and HB (red,  $r=-0.677$ ,  $P=0.001$ , Pearson correlation); \*\*\* $P < 0.001$ , one-way ANOVA, d.f. = 3,72 in **h**.

memory success (Fig. 4b,d; NB:  $F_{1,82}=1.16$ ,  $P=0.28$ ; SB:  $F_{1,82}=1.87$ ,  $P=0.18$ ; HB:  $F_{1,82}=0.45$ ,  $P=0.5$ ) nor temporal order memory (Extended Data Fig. 6b,d;  $F_{1,82}=1.33$ ,  $P=0.25$ ; SB:  $F_{1,82}=0.14$ ,  $P=0.71$ ; HB:  $F_{1,82}=1.98$ ,  $P=0.16$ ). Third, we evaluated whether there were cells whose theta band phase locking of spikes following boundaries was predictive of the success of memory formation regardless of whether their firing rate was modulated (Methods). There were 32 of 580 MTL cells that showed an enhanced number of spikes phase locked in the theta band after a boundary compared to before a boundary and where the phase locking was correlated with correct/incorrect performance in either one of the two memory tasks. Of those 32 cells, 20 (56%) were also event cells. By contrast, there was no significant overlap between the 32 cells and boundary cells (Supplementary Table 5).

In summary, boundary cells and event cells predicted distinct aspects of memory formation; whereas the firing rate of boundary cells was predictive of later scene recognition memory performance, the phase locking of event cells was predictive of temporal order memory performance.

**Neural state shifts across boundaries reflect memory strength.** We next investigated the changes in the neural responses following

boundaries at the population level of all  $n=580$  MTL cells (pseudopopulation, Methods). We examined the dynamics of population activity around the boundaries by evaluating the change of activity using principal component analysis (PCA). During NB video clips, the neural state exhibited only slow changes as a function of time (Fig. 5a). By contrast, the neural state changed abruptly following SBs and HBs (Fig. 5b,c). These abrupt 'neural state shifts' were consistent with the changes in firing rates we reported for boundary cells and event cells, but the observations in Fig. 5 reflect the activity of all MTL cells. To quantify the size of state shifts, we computed the multidimensional Euclidean distance (MDD( $t$ )) in state space between a given time  $t$  and the boundary (Fig. 5d–g). The dimensionality of the state space we used was the number of principal components (PCs) that explained  $\geq 99\%$  of the variance. Plotting MDD as a function of time revealed an abrupt change within  $\sim 300$  ms after the boundary for SB and HB video clips (Fig. 5d–g). This abrupt change can also be seen at the level of individuals (Extended Data Fig. 8).

We evaluated what types of cells contributed most to the neural state shift. First, neural state shifts following SBs were only visible when boundary cells were included (Fig. 5d,e). Second, early neural state shifts after HBs were only visible when event cells were

included, while later HB-related shifts remained present in the absence of either event cells (Fig. 5f) or both event and boundary cells (Fig. 5g). Third, the point of time at which MDD reached its maximal value varied systematically between groups of cells; the responses carried by boundary cells appeared significantly earlier than those carried by event cells and non-boundary/event cells (Fig. 5h;  $F_{3,76} = 103.96$ ,  $P = 8 \times 10^{-27}$ ). Together, this shows that early population-level state shifts are principally due to the activity of boundary cells, whereas event cells and non-boundary/event cells in MTL contribute to slower-latency HB-related state shifts.

We next assessed whether the size of neural state shifts following boundaries during encoding was related to whether a stimulus was later remembered or not. We computed the extent of state changes in the population following a boundary by calculating the total Euclidean distance traversed in state space in the 0- to 1-s time window after boundaries (AUC MDD; Methods). The AUC MDD was positively correlated with recognition accuracy for frames following SBs and HBs but not NBs (Fig. 5i; SB:  $r = 0.653$ ,  $P = 0.002$ ; HB:  $r = 0.565$ ,  $P = 0.009$ ; NB:  $r = 0.214$ ,  $P = 0.256$ ). By contrast, the AUC MDD was negatively correlated with accuracy in the time discrimination task for HBs but not for NBs or SBs (Fig. 5j; HB:  $r = -0.677$ ,  $P = 0.001$ ; NB:  $r = 0.212$ ,  $P = 0.261$ ; SB:  $r = -0.273$ ,  $P = 0.244$ ). Together, these results reveal a neural correlate of the tradeoff between these two types of memory, with large neural state shifts beneficial for scene recognition memory but detrimental for order memory.

### Neural context after boundaries reinstates during recognition.

It is thought that reinstatement of the neural context present at encoding enables mental time travel during memory retrieval<sup>25,26</sup>. However, it remains unknown what exactly is reinstated for continuous experience and how boundaries shape the retrieval process. To address this question, for each individual, we quantified the degree and timing of reinstatement by computing the correlation between the vectors of spike counts of all recorded MTL neurons during the scene recognition task (1.5-s fixed time window) and during encoding (1.5-s sliding window, step size of 0.1 s; Methods). Correct targets (that is, frames from presented clips during encoding that were correctly remembered as old) were accompanied by a significant positive correlation between neural activity during the scene recognition and the encoding period shortly after SB/HB transitions (Fig. 6a,e;  $P < 0.01$ , permutation test; Methods). By contrast, we observed no significant correlation for forgotten targets (that is, frames from presented clips that were incorrectly marked as new; Fig. 6b,f). This effect could not be explained by individuals not attending to the scene recognition task because visually responsive cells responded equally well to both remembered and forgotten target trials (Extended Data Fig. 9).

The reinstated neural context during retrieval was most similar to the neural context present during encoding approximately ~1.2 s after the boundary that preceded the tested frame (Fig. 6i, filled circles; SB:  $-1.26 \pm 0.38$  s,  $t_{19} = 14.68$ ,  $P = 8 \times 10^{-12}$ ; HB:  $-1.28 \pm 0.48$  s,  $t_{19} = 11.80$ ,  $P = 3 \times 10^{-10}$ ). Notably, the point of time of maximal similarity preceded the time at which the later tested frame was shown by 1–1.5 s (Fig. 6i, empty circles; SB:  $1.53 \pm 0.61$  s,  $t_{19} = 11.18$ ,  $P = 8 \times 10^{-10}$ ; HB:  $1.72 \pm 1.03$  s,  $t_{19} = 7.44$ ,  $P = 5 \times 10^{-7}$ ). This observation remains true also for smaller window sizes used to compute the correlations (Supplementary Fig. 13), indicating that the neural state reinstated during testing is the one that was present in between the preceding boundary and the tested frame. Thus, the neural state that was reinstated is the one present well before the tested frame was shown during encoding. Together with an absence of significant reinstatement during incorrect targets (Fig. 6b), these analyses suggest that the neural correlate of reinstatement is not the result of identical sensory inputs in the clips and tested frames. Additionally, no significant correlation

for correctly identified foils was observed (that is, frames from unrepresented clips correctly marked as new; Fig. 6c,g).

Reinstatement is thought to contribute primarily to the recollection of memories that are remembered with high confidence<sup>27,28</sup>. Compatible with this view, the correlations following boundary transitions were significantly stronger in high- than in low-confidence trials (converted from three levels of confidence; Methods). This difference was observed during both correctly remembered targets and falsely recognized foils (Fig. 6j–m; SB, correct targets:  $P = 5 \times 10^{-10}$ ; HB, correct targets:  $P = 4 \times 10^{-6}$ ; NB, correct targets:  $P = 0.79$ ; SB, incorrect foils:  $P = 5 \times 10^{-7}$ ; HB, incorrect foils:  $P = 5 \times 10^{-5}$ ; NB, incorrect foils:  $P = 0.18$ ; *t*-test).

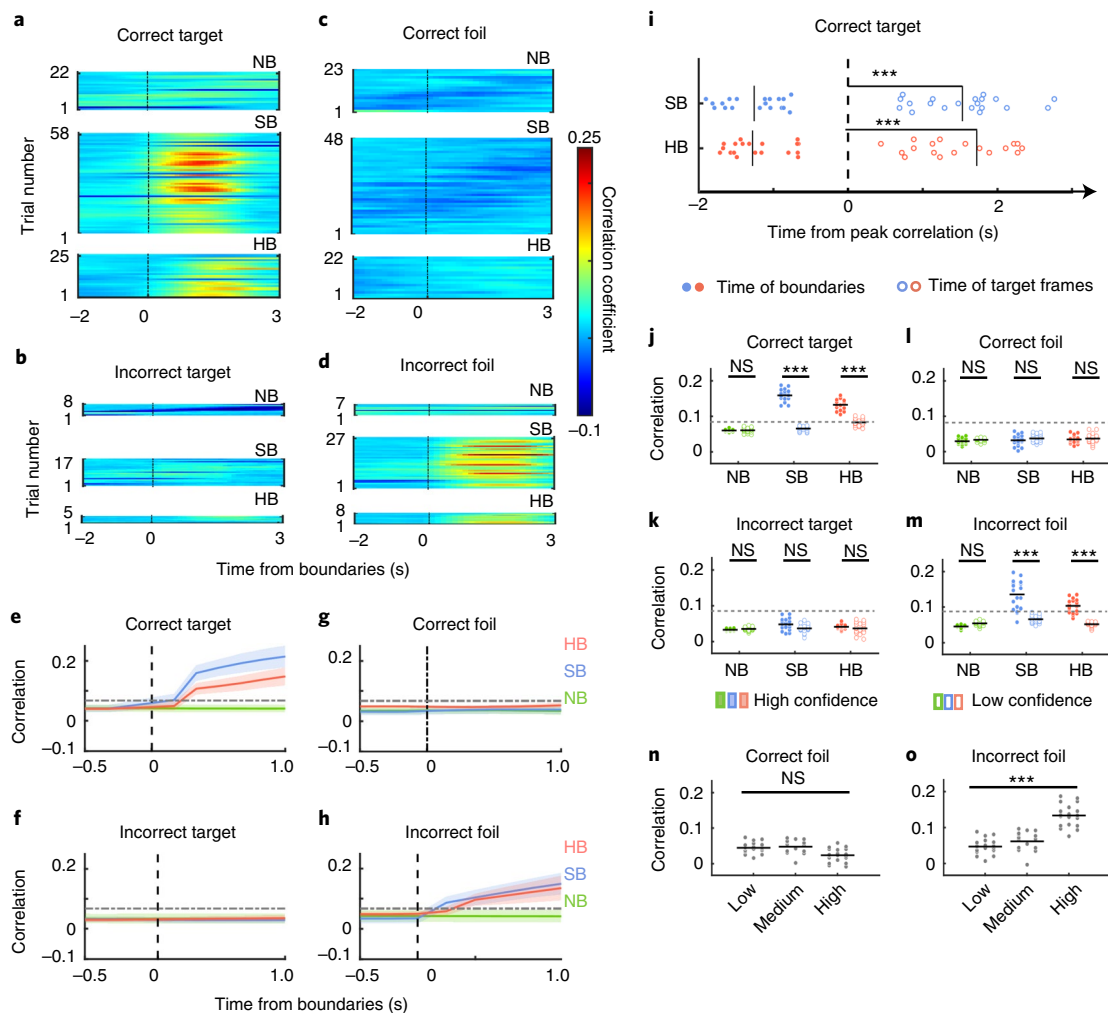
Notably, strong correlations between the neural state at retrieval and encoding also occurred when a new image was incorrectly classified as seen before (Fig. 6d,h; incorrect foil,  $P < 0.01$ , permutation test; Methods), thereby revealing a neural explanation for the false alarms. Were these false alarms, which were accompanied by neural reinstatement, caused by visual similarity between the targets and foils? To address this question, we assessed the similarity between each target and foil by acquiring similarity ratings from an independent control group of individuals ( $n = 30$ ). Similarity values were balanced across NB, SB and HB (Supplementary Fig. 2a). We split foils into low (top 66.67–100%), medium (top 33.33–66.67%) and high (top 1–33.33%) similarity groups. Correlations between encoding and scene recognition were significantly stronger for highly similar foils falsely recognized as old than for low- and medium-similarity foils (Fig. 6o; incorrect foil:  $F_{2,54} = 10.67$ ,  $P = 1 \times 10^{-4}$ ). By contrast, the extent of correlation for correctly rejected foils (true negatives) did not vary as a function of similarity (Fig. 6n; correct foil:  $F_{2,54} = 2.182$ ,  $P = 0.144$ ). Thus, the reason for false alarms is that the wrong context is reinstated due to the high similarity of the foil with a target. Together, these results support the notion that the neural state present at encoding following the boundary is reinstated during memory retrieval.

### Discussion

Memories are often conceptualized as discrete events on a narrative timeline<sup>6</sup>. However, the very definition of an event remains enigmatic. Where do events start and end, and how are multiple signals bound together over time to form a singular event? Here, we test the hypothesis that boundary detection is a mechanism that segments continuous experience into discrete events. Behaviorally<sup>29</sup>, boundaries enhance scene recognition memory for temporally proximal events while disrupting temporal order memory. We found two types of cells that responded to cognitive boundaries; one responded to both SBs and HBs, while another group responded only after HBs.

Both SBs and HBs are accompanied by salient visual changes, whereas the NB control condition does not (Fig. 1b). However, the observed responses to boundaries (Fig. 3) cannot be explained by these sharp visual input changes. First, cells differentiate between SBs and HBs even though both types encompass similar changes at the pixel level (Supplementary Table 1). Second, boundary and event cells did not respond to strong visual changes at clip onsets (Fig. 3d,h) or offsets (Extended Data Fig. 3) during encoding nor to the image onsets or offsets during the scene recognition and time discrimination tasks (Extended Data Fig. 4). In our task, trial structure is predictable, that is, the intertrial interval is always followed by a video and then the next intertrial interval (Fig. 1a–d). By contrast, whether a boundary will occur in a given video and if so what type (SB or HB) is not predictable. One hypothesis for the absence of responses to clip onsets and offsets is therefore their predictability, which would lead to an absence of prediction errors that are hypothesized to underly event segmentation<sup>30</sup>. Supporting this hypothesis, boundary and event cells also increased their responses following the unpredictable probe question during





**Fig. 6 | Reinstatement of neural context after boundaries during recognition.** **a–d**, Single subject example. The color code indicates correlation between the population response during scene recognition (0–1.5 s relative to stimulus onset) and the encoding period (sliding window of 1.5 s and 100-ms step size). Correlations are aligned to the middle of the clip (NB) or boundaries (SB, HB) and are shown separately for correctly recognized familiar targets (**a**), correctly recognized new (not seen) foils (**c**), forgotten targets (**b**) and incorrectly recognized foils (false positives; **d**) in the scene recognition task. The same plots for all subjects are shown in Supplementary Figs. 10 and 11. **e–h**, Population summary. Correlation coefficients, as shown in **a–d**, averaged across all participants for NB (green), SB (blue) and HB (red) trials. Shaded areas represent  $\pm$ s.e.m. across participants. The gray dashed horizontal lines denote the significance threshold ( $P < 0.01$ , one-tailed permutation test; Methods). Data are shown for correctly recognized familiar targets (**e**), forgotten targets (**f**), correctly recognized new foils (**g**) and incorrectly recognized foils (**h**) in the scene recognition task. See the same analyses after excluding boundary cells and event cells and only for boundary cells and event cells in Supplementary Fig. 12. **i**, The reinstated neural context was located in between the boundary and the tested frame. For trials with target frames extracted after boundaries, the time distance from when the correlation coefficient peaks to the time of SB and HB (filled circles; SB:  $-1.26 \pm 0.38$  s,  $t_{19} = 14.68$ ,  $P = 8 \times 10^{-12}$ ; HB:  $-1.28 \pm 0.48$  s,  $t_{19} = 11.80$ ,  $P = 3 \times 10^{-10}$ ; one-tailed  $t$ -test) or target frames (empty circles; SB:  $1.53 \pm 0.61$  s,  $t_{19} = 11.18$ ,  $P = 8 \times 10^{-10}$ ; HB:  $1.72 \pm 1.03$  s,  $t_{19} = 7.44$ ,  $P = 5 \times 10^{-7}$ ; one-tailed  $t$ -test) is shown. Negative/positive values denote the point of time of boundaries (negative) or target frames (positive) relative to when the correlation coefficient reaches its peak value. Asterisks indicate the significance of the peak correlation leading the time of target frames. See the same analyses with the correlation computed using different window sizes in Supplementary Fig. 13. **j–m**, Population summary (confidence). Reinstatement differed between frames remembered with high (filled circles) and low (empty circles) confidence responses for “old” decisions (correct targets (**j**) and incorrect foils (**m**)) in SB and HB conditions but not “new” decisions (correct foils (**l**) and incorrect targets (**k**)) and NB conditions, regardless of whether they were correct or incorrect (SB, correct targets:  $P = 5 \times 10^{-10}$ ; HB, correct targets:  $P = 4 \times 10^{-6}$ ; NB, correct targets:  $P = 0.79$ ; SB, incorrect foils:  $P = 5 \times 10^{-7}$ ; HB, incorrect foils:  $P = 5 \times 10^{-5}$ ; NB, incorrect foils:  $P = 0.18$ ; one-tailed  $t$ -test). Correlation coefficients are as shown in **e–h** averaged over 0–1 s after boundaries. **n, o**, Population summary (target–foil similarity). Correlation coefficients versus similarity ratings between targets and foils are plotted for correct (**n**;  $F_{2,54} = 2.182$ ,  $P = 0.144$ ; one-tailed ANOVA) and incorrectly recognized foils (**o**;  $F_{2,54} = 10.67$ ,  $P = 1 \times 10^{-4}$ ; one-tailed ANOVA). Each dot represents one recording session. Black lines in **i–m** and **o** denote the mean results averaged across all recording sessions; \*\*\* $P < 0.001$ .

encoding (present randomly every four to eight trials; Fig. 1a). In contrast to the selective single-neuron response to cognitive boundaries, prior functional magnetic resonance imaging (fMRI) work<sup>31</sup> reported a clip offset-triggered blood oxygenation level–dependent signal change in the hippocampus whose magnitude was predictive

of subsequent memory strength. This off response has been interpreted to have the same origin as the between-event responses<sup>32</sup> and was present despite the fixed clip length and trial structure.

Boundary cells respond to both SBs and HBs, whereas event cells respond only to HBs (Fig. 3). These distinct responses may reflect

the hierarchical structure of episodic memory, with event cells representing episodic transitions between distinct events, while boundary cells represent more frequent but smaller episodic transitions within the same overall narrative. These findings provide evidence for the theory that event segmentation is a hierarchical dynamic procedure, with fine-to-coarse segmentations associated with different kinds of cognitive boundaries<sup>9</sup>. The anatomical location and response latency of the cells are also compatible with this proposal; boundary cells respond first and are most common in the parahippocampal gyrus, whereas event cells respond later and are most common in the hippocampus (Supplementary Table 4). This distinction is also visible at the population level, with neural state shifts for SBs mainly driven by boundary cells, whereas HB-related state shifts occur later and are driven by a broader group of cells (Fig. 5d–h). Notably, the late HB-related state shifts are partially driven by cells that are not classified as event cells (a conclusion that holds even when using a more liberal definition of event cells at  $P < 0.01$ ). This suggests that besides the early HB-related responses, there is a secondary later response to HBs that is encoded as a distributed population response. We hypothesize that the early responses of boundary cells reflect contextual changes detected in the higher-level visual areas<sup>33,34</sup>, while event cells are the result of a late output signal from a comparator operation between predicted and received signals<sup>35,36</sup>.

Responses of boundary or event cells bring to mind border and place cells in the rodent hippocampus<sup>16,37</sup>. As rodents move between compartments, place cells cluster at boundaries (for example, doorways)<sup>14</sup>, the crossing of which is followed by remapping<sup>18,19</sup> or reinstatement<sup>15,20</sup> of a different set of hippocampal place fields. Here, boundary cells and event cells respond to transitions (boundaries) between different episodic contexts. Also, similar to place field remapping, the neural state changed abruptly following a boundary. When participants are reexposed to familiar target frames during the later recognition test, the neural state reinstates if the item is successfully recognized. Similar to place field reinstatement, the reinstated neural state is most similar to the one following boundaries even before when the tested frame is shown. This finding provides insight into the question of what neural context is reinstated during mental time travel and memory search<sup>38–42</sup>. This finding also indicates that abrupt changes in neural context are important to demarcate periods of time that can be reinstated later. We note several differences between boundary and event cells and border cells in rodents. Border cells respond to physical boundaries and are observed in tasks in which rodents are extensively trained. By contrast, boundary or event cells in humans respond to cognitive boundaries in a variety of different videos, none of which individuals have seen before. This property is an essential requirement for a process to divide experience into episodes to shape episodic memories, which, by definition, occur only once in new environments.

What roles do boundary responses play in episodic memory? At the single-cell level, the firing rate of boundary cells predicts scene recognition memory strength and the clustering of event cells' spike timing relative to theta phase predicts temporal order memory encoding success. This indicates that these two kinds of cells play distinct roles during encoding, with each strengthening only one kind of memory using a different mechanism. The response of boundary and event cells during encoding was 'content invariant' because they responded to many clips with varying content (Fig. 3). The role of boundary responses during retrieval was in guiding the points of time that would later be reinstated (Fig. 6i) but not participating in the reinstatement process itself. This is expected if boundary cells do not carry information about content. Confirming this, the results on reinstatement remain essentially unchanged after excluding boundary and event cells from the analysis (Supplementary Fig. 12). Together, these findings suggest that boundary and event cells play two roles in episodic memory;

they structure memories during encoding, and they serve as markers for periods of time that are later reinstated.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-022-01020-w>.

Received: 1 March 2021; Accepted: 19 January 2022;  
Published online: 7 March 2022

### References

1. Ezzayat, Y. & Davachi, L. What constitutes an episode in episodic memory? *Psychol. Sci.* **22**, 243–252 (2011).
2. Tulving, E. Episodic memory: from mind to brain. *Annu. Rev. Psychol.* **53**, 1–25 (2002).
3. Radvansky, G. A. & Zacks, J. M. *Event Cognition* (Oxford University Press, 2014).
4. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
5. Zacks, J. M. et al. Human brain activity time-locked to perceptual event boundaries. *Nat. Neurosci.* **4**, 651–655 (2001).
6. Avrahami, J. & Kareev, Y. The emergence of events. *Cognition* **53**, 239–261 (1994).
7. DuBrow, S. & Davachi, L. The influence of context boundaries on memory for the sequential order of events. *J. Exp. Psychol. Gen.* **142**, 1277–1286 (2013).
8. Chen, J. et al. Shared memories reveal shared structure in neural activity across individuals. *Nat. Neurosci.* **20**, 115–125 (2017).
9. Kurby, C. A. & Zacks, J. M. Segmentation in the perception and memory of events. *Trends Cogn. Sci.* **12**, 72–79 (2008).
10. Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S. & Reynolds, J. R. Event perception: a mind–brain perspective. *Psychol. Bull.* **133**, 273–293 (2007).
11. Jafarpour, A., Griffin, S., Lin, J. J. & Knight, R. T. Medial orbitofrontal cortex, dorsolateral prefrontal cortex, and hippocampus differentially represent the event saliency. *J. Cogn. Neurosci.* **31**, 874–884 (2019).
12. Ben-Yakov, A. & Henson, R. N. The hippocampal film editor: sensitivity and specificity to event boundaries in continuous experience. *J. Neurosci.* **38**, 10057–10068 (2018).
13. Baldassano, C. et al. Discovering event structure in continuous narrative perception and memory. *Neuron* **95**, 709–721 (2017).
14. Spiers, H. J., Hayman, R. M., Jovalekic, A., Marozzi, E. & Jeffery, K. J. Place field repetition and purely local remapping in a multicompartment environment. *Cereb. Cortex* **25**, 10–25 (2015).
15. Derdikman, D. et al. Fragmentation of grid cell maps in a multicompartment environment. *Nat. Neurosci.* **12**, 1325–1332 (2009).
16. Lever, C., Burton, S., Jeewajee, A., O'Keefe, J. & Burgess, N. Boundary vector cells in the subiculum of the hippocampal formation. *J. Neurosci.* **29**, 9771–9777 (2009).
17. O'Keefe, J. & Burgess, N. Geometric determinants of the place fields of hippocampal neurons. *Nature* **381**, 425–428 (1996).
18. Alme, C. B. et al. Place cells in the hippocampus: eleven maps for eleven rooms. *Proc. Natl Acad. Sci. USA* **111**, 18428–18435 (2014).
19. Colgin, L. L., Moser, E. I. & Moser, M. B. Understanding memory through hippocampal remapping. *Trends Neurosci.* **31**, 469–477 (2008).
20. Grieves, R. M., Jenkins, B. W., Harland, B. C., Wood, E. R. & Dudchenko, P. A. Place field repetition and spatial learning in a multicompartment environment. *Hippocampus* **26**, 118–134 (2016).
21. Sun, C., Yang, W., Martin, J. & Tonegawa, S. Hippocampal neurons represent events as transferable units of experience. *Nat. Neurosci.* **23**, 651–663 (2020).
22. Levine, B. et al. Episodic memory and the self in a case of isolated retrograde amnesia. *Brain* **121**, 1951–1973 (1998).
23. Rutishauser, U. Testing models of human declarative memory at the single-neuron level. *Trends Cogn. Sci.* **23**, 510–524 (2019).
24. Rutishauser, U., Ross, I. B., Mamelak, A. N. & Schuman, E. M. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature* **464**, 903–907 (2010).
25. Pacheco Estefan, D. et al. Coordinated representational reinstatement in the human hippocampus and lateral temporal cortex during episodic memory retrieval. *Nat. Commun.* **10**, 2255 (2019).

26. Manning, J. R., Polyn, S. M., Baltuch, G. H., Litt, B. & Kahana, M. J. Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proc. Natl Acad. Sci. USA* **108**, 12893–12897 (2011).
  27. Folkerts, S., Rutishauser, U. & Howard, M. W. Human episodic memory retrieval is accompanied by a neural contiguity effect. *J. Neurosci.* **38**, 4200–4211 (2018).
  28. Howard, M. W., Fotedar, M. S., Datey, A. V. & Hasselmo, M. E. The temporal context model in spatial navigation and relational learning: toward a common explanation of medial temporal lobe function across domains. *Psychol. Rev.* **112**, 75–116 (2005).
  29. Swallow, K. M., Zacks, J. M. & Abrams, R. A. Event boundaries in perception affect memory encoding and updating. *J. Exp. Psychol. Gen.* **138**, 236–257 (2009).
  30. Richmond, L. L., Gold, D. A. & Zacks, J. M. Event perception: translations and applications. *J. Appl. Res. Mem. Cogn.* **6**, 111–120 (2017).
  31. Ben-Yakov, A. & Dudai, Y. Constructing realistic engrams: poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. *J. Neurosci.* **31**, 9032–9042 (2011).
  32. Ben-Yakov, A., Eshel, N. & Dudai, Y. Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. *J. Exp. Psychol. Gen.* **142**, 1255–1263 (2013).
  33. Isik, L., Singer, J., Madsen, J. R., Kanwisher, N. & Kreiman, G. What is changing when: decoding visual information in movies from human intracranial recordings. *Neuroimage* **180**, 147–159 (2018).
  34. Aminoff, E. M., Kveraga, K. & Bar, M. The role of the parahippocampal cortex in cognition. *Trends Cogn. Sci.* **17**, 379–390 (2013).
  35. Lisman, J. E. & Grace, A. A. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* **46**, 703–713 (2005).
  36. Vinogradova, O. S. Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus* **11**, 578–598 (2001).
  37. Solstad, T., Boccara, C. N., Kropff, E., Moser, M. B. & Moser, E. I. Representation of geometric borders in the entorhinal cortex. *Science* **322**, 1865–1868 (2008).
  38. Xiao, X. et al. Transformed neural pattern reinstatement during episodic memory retrieval. *J. Neurosci.* **37**, 2986–2998 (2017).
  39. Favila, S. E., Samide, R., Sweigart, S. C. & Kuhl, B. A. Parietal representations of stimulus features are amplified during memory retrieval and flexibly aligned with top-down goals. *J. Neurosci.* **38**, 7809–7821 (2018).
  40. Jang, A. I., Wittig, J. H. Jr, Inati, S. K. & Zaghoul, K. A. Human cortical neurons in the anterior temporal lobe reinstate spiking activity during verbal memory retrieval. *Curr. Biol.* **27**, 1700–1705 (2017).
  41. Howard, M. W. & Natu, V. S. Place from time: reconstructing position from a distributed representation of temporal context. *Neural Netw.* **18**, 1150–1162 (2005).
  42. Polyn, S. M. & Kahana, M. J. Memory search and the neural representation of context. *Trends Cogn. Sci.* **12**, 24–30 (2008).
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- © The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

## Methods

**Participants.** *Individuals with drug-resistant epilepsy.* Twenty individuals with drug-resistant epilepsy volunteered for this study and gave their informed consents. The institutional review boards of Toronto Western Hospital and Cedars-Sinai Medical Center approved all protocols. The task was conducted while the individuals stayed in the hospital after implantation of depth electrodes for monitoring seizures. The location of the implanted electrodes was solely determined by clinical needs. The behavioral analyses included results from all 20 individuals, and the neural results were analyzed across 19 individuals (participant 20 had no usable recordings; Supplementary Table 2).

*Amazon Mechanical Turk (MTurk) workers.* MTurk workers were recruited for similarity ratings (Similarity ratings), including 30 individuals (age  $23.25 \pm 3.42$  years, nine female) for rating the visual properties of different boundaries (Supplementary Table 1), 30 individuals (age  $22.79 \pm 5.73$  years, 11 female) for rating the similarity between target and foil frames (Supplementary Fig. 2a) and 30 individuals (age  $25.06 \pm 6.11$  years, 7 female) for performing the time discrimination task without an encoding session (Supplementary Fig. 2b). All control tasks conducted on Amazon MTurk workers were under the approval of the institutional review board of Boston Children's Hospital, and informed consents were obtained with digital signatures for each individual.

**Task.** The task consisted of three parts: encoding, scene recognition and time discrimination (Fig. 1a,c,d).

*Encoding.* Participants watched a series of 90 unique clips with no sound and were instructed to remember as much of the clips as possible. Each trial started with a baseline period, a fixation cross reminding individuals to fixate at the center of the screen throughout the task. The duration of the baseline period ranged from 0.9 to 1.1 s (randomized, sampled from uniform distribution). The fixation period was followed by the presentation of a video clip that contained either NBs (continuous movie shots; virtual NBs for analysis purposes were always located in the middle of the clip), SBs (cuts to a new scene within the same movie, one to three SBs per clip randomly distributed in the clips) or an HB (cuts to a new scene from a different movie, one HB per clip located at 4 s after the start of the clip). Examples of SBs and HBs are shown in Fig. 1b. A yes or no question related to the content of the clip (for example, is anyone in the clip wearing glasses?) appeared randomly after every four to eight clips.

*Scene recognition.* After watching all 90 clips, participants' memory for the content of the videos was evaluated in a scene recognition test. During scene recognition, frames extracted from encoded clips (target frames) and frames from new, never shown clips (foil frames) were presented to the participants. Individuals were instructed to identify whether these frames were old or new (that is, whether they had seen the frame during the encoding session). To generate the testing frames for scene recognition, we first extracted two frames from each clip, one randomly pulled out from the first half of the clip and the other one from the second half. We then kept half of these frames extracted from the first/second half of the clip (in total,  $n = 90$ ) as target frames and used the other half as templates to search for foil frames (in total,  $n = 90$ ) from a different movie played by different actors/actresses ( $n = 30$ ), a different movie played by the same actors/actresses ( $n = 30$ ) or the unrepresented portion from the same movie played by the same actors/actress ( $n = 30$ ) to introduce different levels of similarity between the target frames and foil frames. The total number of target and foil frames (30 target frames and 30 foil frames for each boundary type) and the average similarity level of foil frames were counterbalanced across different boundary types ( $F_{2,87} = 1.72$ ,  $P = 0.19$ ; rated by Amazon MTurk workers; Similarity ratings).

*Time discrimination.* After the scene recognition test, we evaluated participants' memory about the temporal structure of the clip with a time discrimination test. In each trial, two frames (half of them picked at 1 s and 7 s and the other half picked at 3 s and 5 s of the clip) separated by different kinds of boundaries (NB, SB or HB) were extracted from the same video clip and were presented side by side. Participants were instructed to indicate which of the two frames (that is, 'left' or 'right') appeared first (earlier in time) in the videos they watched during encoding. In both the time discrimination and recognition memory test, the duration of the baseline period ranged from 0.45 to 0.55 s (randomized, sampled from uniform distribution).

*Confidence measurement.* All binary choices through the encoding session, scene recognition and time discrimination were made together with a subjective confidence judgment (that is, sure, less sure and very unsure). Thus, there were always six possible responses for each question. Given that there were fewer 'less sure' and 'very unsure' responses than 'sure' responses, we grouped 'very unsure' and 'less sure' responses together as 'low confidence' and labeled 'sure' responses as 'high confidence' in Fig. 6j–m.

**Similarity ratings.** *Visual properties of SBs and HBs.* Both SB and HB transitions were accompanied by transient visual changes (cuts to a new scene), whereas there were no such visual changes for the NB condition. We quantified the visual changes

of each boundary type by computing metrics that relate to pixel-level differences, luminance, contrast, complexity, entropy and color distribution between pre- and postboundary frames. In addition, to quantify visual differences not directly captured at the pixel level, we used pre- and postboundary frames as inputs for the AlexNet network (pretrained on ImageNet dataset)<sup>43</sup>, extracted the activation matrices from the layer 'fc7' for both images and computed the Euclidean distance between their activation matrices. Moreover, we collected perceptual ratings (that is, similarity ratings between pre- and postboundary frames) from Amazon MTurk workers. During similarity ratings, pre- and postboundary frames were presented side by side, and MTurk workers were instructed to rate the similarity of the image pairs by clicking on a scaling bar (0–1: 0, different; 1, identical). See results in Supplementary Table 1.

*Similarity ratings between target and foil frames.* When selecting foil frames, we used target frames as templates to search for foil frames with different similarity levels (Task). We presented the target frame with its corresponding foil frame side by side and instructed MTurk workers to rate the similarity between them (Supplementary Fig. 2a).

*Time discrimination without encoding.* To ensure that the time discrimination task could not be solved by pure reasoning, we recruited MTurk workers to perform the time discrimination test without watching clips (Supplementary Fig. 2b).

**Electrophysiology.** We recorded bilaterally from the amygdala, hippocampus and parahippocampal gyrus and other regions outside the MTL using hybrid depth electrodes (Ad-Tech Company), which contained eight microwires (40  $\mu$ m in diameter) at the tip of each electrode shank. For each microwire, broadband signals (0.1–9,000 Hz filtered) were recorded at 32 kHz using the ATLAS system (Neuralynx).

**Spike sorting and quality metrics of single units.** The recorded signals were filtered offline in the 300- to 3,000-Hz band with a zero-phase lag filter. Spikes were detected and sorted using the semiautomated template-matching algorithm Osort<sup>44,45</sup> v4. We computed spike sorting quality metrics for all putative single units (Supplementary Fig. 5) to quantify our recording and sorting quality<sup>46–48</sup>.

**Electrode localization.** Electrode localization was based on postoperative MRI/computed tomography scans. We co-registered postoperative and preoperative MRIs using Freesurfer's `mri_robust_register`<sup>49</sup>. To summarize electrode positions and to provide across-study comparability, we aligned each participant's preoperative scan to the CIT168 template brain in MNI152 coordinates<sup>50</sup> using a concatenation of an affine transformation followed by a symmetric image normalization (SyN) diffeomorphic transform<sup>51</sup>. The MNI coordinates of the right microwires from the same electrode shank were marked as one location. MNI coordinates of microwires with putative neurons detected across all participants were plotted on a template brain for illustration (Fig. 1e).

**Data analyses.** *Boundary cell.* For each recorded neuron, we counted spikes within the 0- to 1-s (postboundary) and –1- to 0-s time interval (baseline) relative to boundaries during encoding. A cell was considered a boundary cell if it met the following criteria: (1) its spike counts within postboundary time windows were significantly different from its spike count within baseline time windows for SB and HB but not for NB ( $P < 0.05$ , permutation *t*-test), and (2) its spike counts within postboundary time windows were significantly greater in SB and HB than NB ( $P < 0.05$ , permutation *t*-test).

*Event cell.* A cell was considered an event cell if it met the following criteria: (1) its spike counts within postboundary time windows were significantly different from its spike count within baseline time windows for HB but not for NB and SB ( $P < 0.05$ , permutation *t*-test), and (2) its spike counts within postboundary time windows were significantly greater in HB than NB and SB ( $P < 0.05$ , permutation *t*-test).

*Boundary and event cell responses to clip onsets and offsets.* For each selected boundary cell and event cells, we counted spikes within the 0- to 1-s (post) and –1- to 0-s (pre) time interval relative to clip onsets/offsets during encoding for clips with NB, SB and HB separately. The boundary cell or event cell was considered as not responding to clip onsets/offsets if their spike counts within each boundary condition did not differ between post- and prewindow ( $P > 0.05$ , permutation *t*-test).

*Phase-tuning cells.* We computed the MRL for the theta band phases of all spikes fired within a 0- to 1-s window following a boundary and –1 to 0 s preceding a boundary (baseline) during encoding. We also computed the MRLs in the 0- to 1-s postboundary time window separately for spikes fired in trials that were later remembered (correct) versus forgotten (incorrect) during the scene recognition and time discrimination task. We defined a cell as a 'phase-tuning neuron' if it met the following criteria: (1) its MRL within the postboundary time window was significantly different from its MRL within the baseline time window for SB and

HB but not for NB trials ( $P < 0.05$ , permutation  $t$ -test), and (2) its MRL within the postboundary time window was significantly different between correct and incorrect trials in either the scene recognition and/or the time discrimination tasks ( $P < 0.05$ , permutation  $t$ -test).

**Chance level for cell response analyses.** To estimate the number of neurons that would be considered boundary cells or event cells by chance, we repeated the same procedures for boundary cell and event cell analyses after randomly shuffling the trial labels (NB, SB and HB) 1,000 times. For each iteration, we obtained the proportion of selected boundary cells and event cells relative to the total number of neurons within each region. These 1,000 values formed the empirically estimated null distribution for the proportion of boundary cells and event cells expected by chance. A region was considered to have a significant amount of boundary cells or event cells if its actual fraction of significant cells exceeded 95% of the null distribution (Supplementary Table 4;  $P < 0.05$ ).

**Association between spiking activity during encoding and later memory retrieval accuracy.** *Firing rate modulation.* For each boundary cell and event cell, we grouped its spike activity within 0 to 1 s after boundaries during encoding based on participants' subsequent memory performance either in the scene recognition task (correct versus incorrect recognition) or the time discrimination task (correct versus incorrect discrimination). We then computed the firing rate as a function of time (bin size = 200 ms, step size = 2 ms) for each trial, which was further  $z$  score-normalized using the mean and s.d. of the firing rate across the whole trial. For each boundary cell and event cell, we then computed the mean  $z$ -scored firing rate within 0- to 1-s time intervals relative to boundaries for each trial and averaged this value across trials within each boundary type. The resulting values across all boundary cells and event cells were used for comparisons across NB, SB and HB conditions (Fig. 4c,g and Extended Data Figs. 6c and 7c).

*Phase modulation.* For each spike of each boundary cell and event cell, we computed the phase of the spike relative to the theta frequency band-filtered LFP signals recorded from the same microwire. To eliminate potential contamination by the spike waveform, we removed the LFP signal within the 3 ms around each spike and replaced it with a linear interpolation. The cleaned LFP signals were then bandpass filtered between 1 and 300 Hz (a zero-phase delay finite impulse response filter with Hamming window) and downsampled from 32 kHz to 500 Hz. We performed automatic artifact rejection<sup>52</sup> and manual visual inspection (using `ft_databrowser.m` from Fieldtrip toolbox<sup>53</sup> version 20190527) to remove large transient signal changes from the downsampled LFPs. Next, we extracted neural activity within the theta band by bandpass filtering in the 4- to 8-Hz band (`eegfilt.m` function in EEGLAB toolbox<sup>54</sup>, a two-way, zero-phase lag, least-squares finite impulse response filter to prevent phase distortion), followed by the Hilbert transform to obtain theta phase as a function of time. We then extracted the phase for each spike (that is, spike phases) by boundary cells and event cells. The phase-locking strength of each boundary cell or event cell was quantified as the MRL of all spike phases of all spikes that occurred within a 0- to 1-s window after boundaries (0, no phase locking; 1, strongest phase locking). The resulting MRL values were then compared between trials with correct or incorrect subsequent memory performance for NB, SB and HB trials separately (Fig. 4d,h and Extended Data Figs. 6d and 7d). The computation of MRL is sensitive to the number of spikes. Therefore, the comparison of MRL between correct and incorrect trials was conducted with balanced spike counts.

**State space analyses.** *Neural state trajectories.* For each trial, we binned each neuron's spike counts during encoding into non-overlapping 10-ms-wide bins, followed by smoothing with a 200-ms s.d. Gaussian kernel and  $z$  score normalization (mean and s.d. were calculated across the entire trial). We used these  $z$ -scored smoothed spike density estimates from all recorded MTL cells across all participants to form a pseudopopulation. We applied PCA to reduce the dimensionality of the pseudopopulation (MATLAB R2019b function `svd.m`). We then rank ordered the resulting PCs by their explained variance (function `dpca_explainedVariance.m` from `dpca` toolbox<sup>55</sup>) and plotted the average neural state trajectories for each boundary type in a three-dimensional space constructed by the top three PC components (Fig. 5a–c).

**MDD.** MDD was defined as the Euclidean distance between two points in the PC space (with all the first  $n$  PCs that accounted for >99% explained variance). Note that while this space captured 99% of explained variance, it was nevertheless substantially lower dimensional than the original space.

**AAUC MDD.** AAUC MDD was defined as the cumulative sum of all Euclidean distance values within the 0- to 1-s time window after boundaries (in the PC space).

**Reinstatement of neural context.** This analysis was done separately for each session of simultaneously recorded neurons and did not rely on the pseudopopulations defined in the previous section.

*Correlation between encoding and retrieval.* Neural activity was quantified for each neuron in 1.5-s-wide bins and a step size of 100 ms. We computed the Pearson correlation coefficients (`corrcoef.m` from MATLAB R2019b) between the neural population activity during scene recognition (1 time bin  $\times$  number of neurons) and encoding (80 time bins  $\times$  number of neurons) at each time step.

*Significant correlation threshold.* We computed the same correlation values after randomly shuffling the trial labels (that is, disrupting the correspondence between encoding and scene recognition trials) to obtain the average correlation strength across trials and neurons expected by chance. This procedure was repeated 1,000 times to form a null distribution, in which the 2.5th and 97.5th percentile values were used as the threshold to determine significance of the actual correlation values (dashed horizontal lines in Fig. 6).

**Statistics and reproducibility.** No statistical method was used to predetermine sample size, but our sample sizes are similar to those reported in previous publications<sup>46,56</sup>. Data collection and analysis were not performed blind to the conditions of the experiments. The experiments were not randomized. No data were excluded from the analyses. For comparisons between two groups, we used the permutation  $t$ -test statistic, and for comparisons between more than two groups, we used a parametric one-way ANOVA. For statistical thresholding, permutation tests were conducted to generate a null distribution estimated from 1,000 runs on data with scrambled labels, which avoids the assumption of normality when evaluating significance.

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

## Data availability

The data (in NWB format) that supports the key findings of this study are publicly available on the DANDI archive (<https://doi.org/10.48324/dandi.000207/0.220216.0323>).

## Code availability

Codes that support the key findings of this study are publicly available on GitHub (<https://github.com/rutishauserlab/cogboundary-zheng>).

## References

- Krizhevsky, A., Sutskever, I. & Hinton, G. E. in *NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems*, Vol. 1 (eds Bartlett, P., Pereira, F. C. N., Burges, C. J. C., Bottou, L., & Weinberger K. Q.) 1097–1105 (Morgan Kaufmann, 2012).
- 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).
- Fried, I., Rutishauser, U., Cerf, M. & Kreiman, G. *Single Neuron Studies of the Human Brain: Probing Cognition* (The MIT Press, 2014).
- Kaminski, J. et al. Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nat. Neurosci.* **20**, 590–601 (2017).
- Pouzat, C., Mazor, O. & Laurent, G. Using noise signature to optimize spike-sorting and to assess neuronal classification quality. *J. Neurosci. Methods* **122**, 43–57 (2002).
- Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H. & Buzsáki, G. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J. Neurophysiol.* **84**, 401–414 (2000).
- Reuter, M., Rosas, H. D. & Fischl, B. Highly accurate inverse consistent registration: a robust approach. *Neuroimage* **53**, 1181–1196 (2010).
- Pauli, W. M., Nili, A. N. & Tyszka, J. M. A high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. *Sci. Data* **5**, 180063 (2018).
- Avants, B. et al. Multivariate analysis of structural and diffusion imaging in traumatic brain injury. *Acad. Radiol.* **15**, 1360–1375 (2008).
- Banaie Boroujeni, K., Tiesinga, P. & Womelsdorf, T. Adaptive spike-artifact removal from local field potentials uncovers prominent beta and gamma band neuronal synchronization. *J. Neurosci. Methods* **330**, 108485 (2020).
- Oostenfeld, R., Fries, P., Maris, E. & Schoffelen, J. M. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).
- 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).
- Kobak, D. et al. Demixed principal component analysis of neural population data. *eLife* **5**, e10989 (2016).
- Bausch, M. et al. Concept neurons in the human medial temporal lobe flexibly represent abstract relations between concepts. *Nat. Commun.* **12**, 6164 (2021).

## Acknowledgements

We thank C. Katz and K. Patel for helping set up the recording system for single-unit recordings at Toronto Western Hospital, N. Chandravadia and V. Barkely for data transferring and organization, C. Reed, J. Chung and the clinical teams at both Cedars-Sinai Medical Center and Toronto Western Hospital and M. Zhang, J. Kaminski and other members of the Rutishauser and Kreiman labs for discussion. We are especially indebted to the volunteers who participated in this study. This work was supported by NIH U01NS103792 and U01NS117839 (to U.R.), NSF 1231216 (G.K.) and Brain Canada (to T.A.V.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions

J.Z. conceived the project. J.Z., G.K. and U.R. contributed ideas for experiments and analysis. S.K.K., T.A.V. and A.N.M. managed participants and surgeries. J.Z., A.G.P.S., M. Y. and C.P.M. collected data. J.Z. performed the analyses. B.A.G. performed electrode localization. B.A.G., J.Z. and U.R. and produced the Neurodata Without Borders (NWB)

formatted dataset for public release. J.Z., G.K. and U.R. wrote the manuscript with input from all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

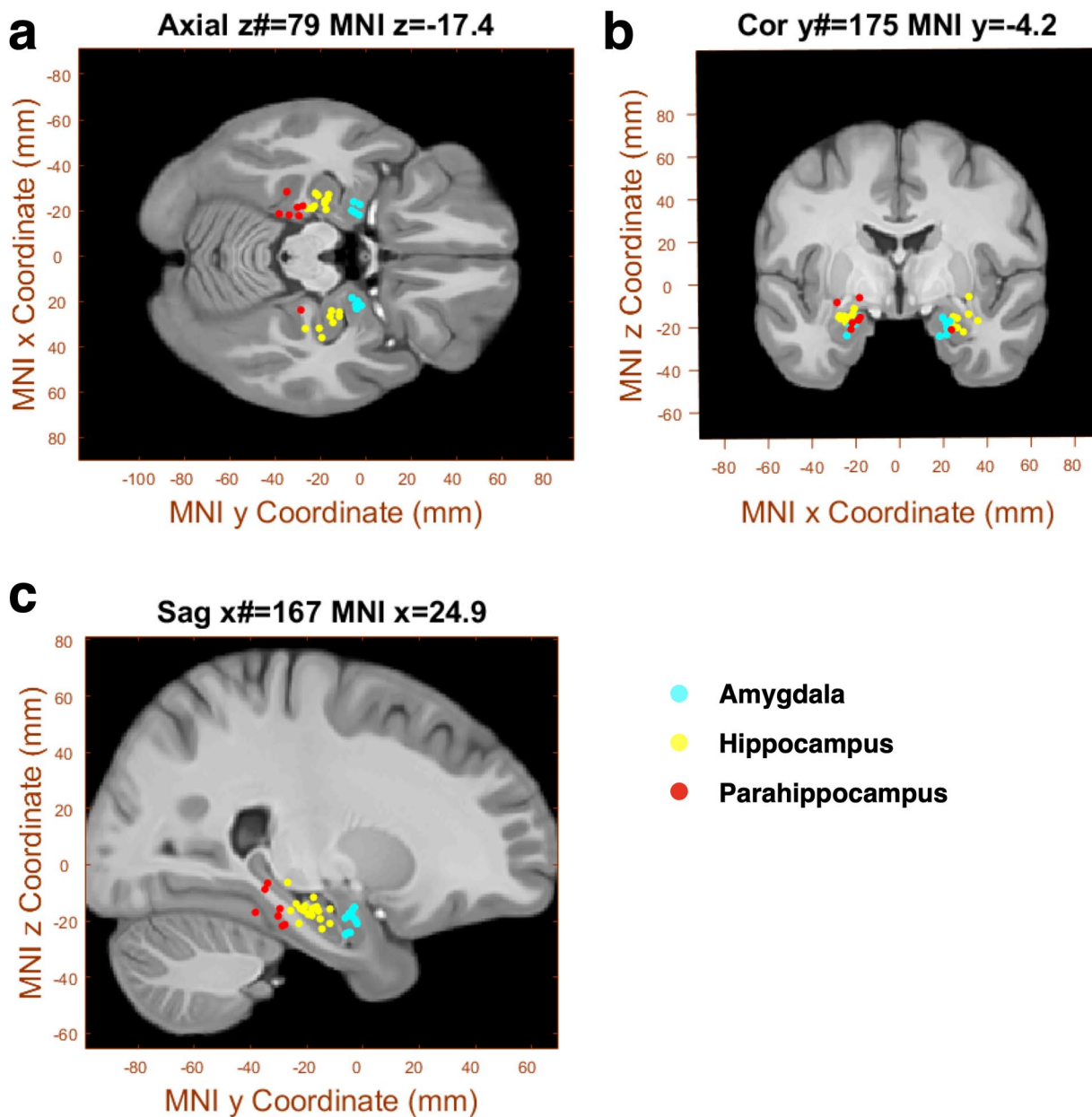
**Extended data** is available for this paper at <https://doi.org/10.1038/s41593-022-01020-w>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41593-022-01020-w>.

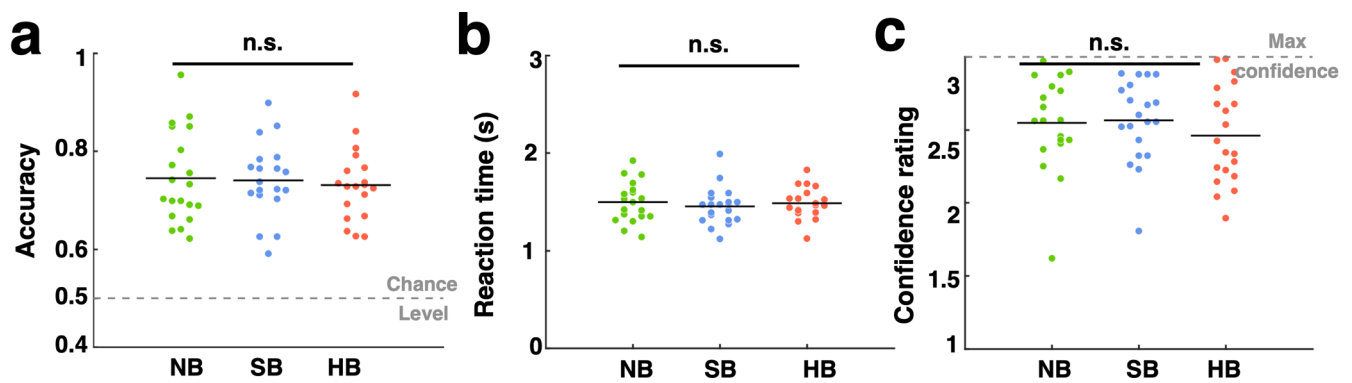
**Correspondence and requests for materials** should be addressed to Gabriel Kreiman or Ueli Rutishauser.

**Peer review information** *Nature Neuroscience* thanks Christopher Baldassano 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](http://www.nature.com/reprints).

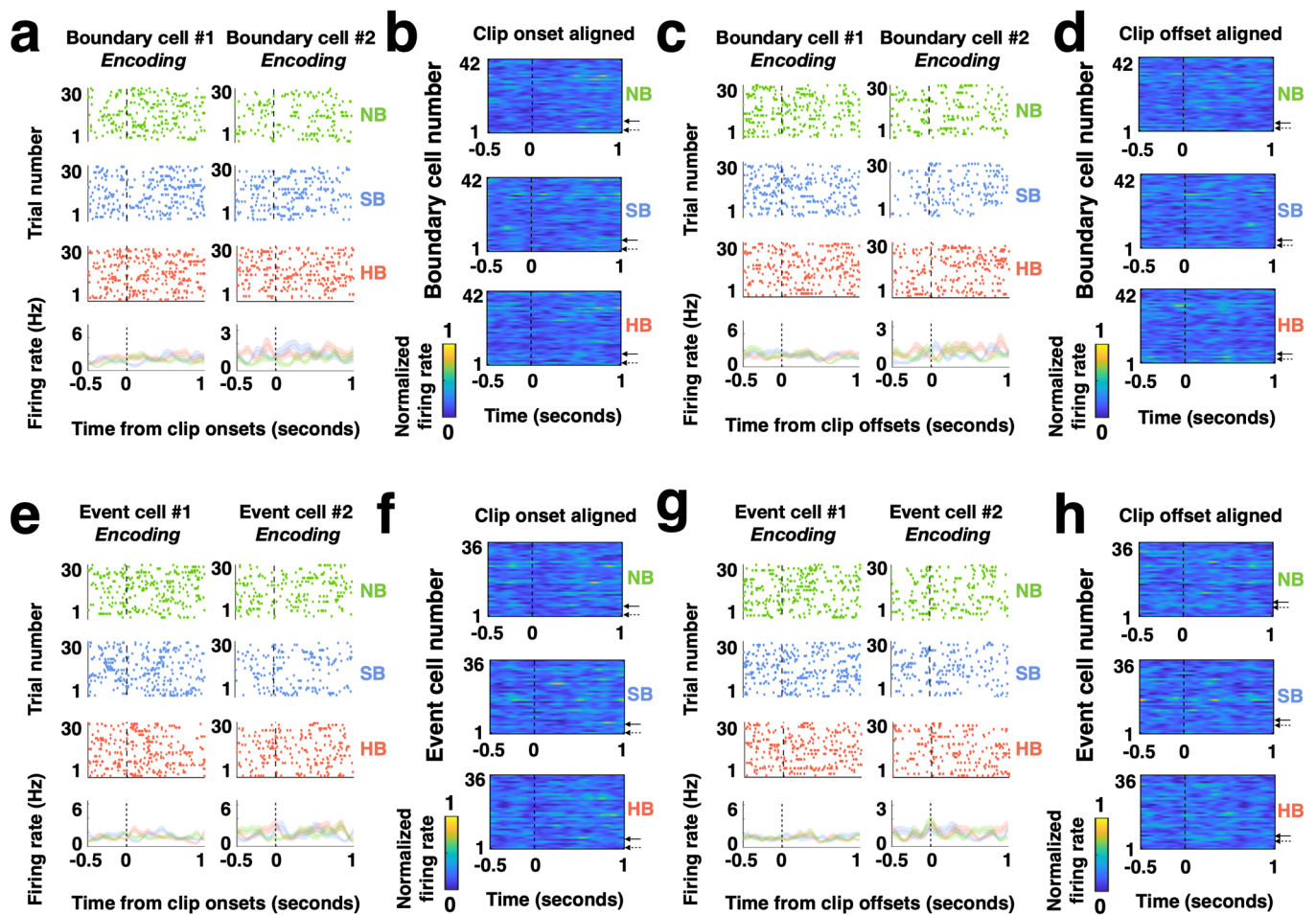


**Extended Data Fig. 1 | Electrode locations in MNI coordinates, Related to Fig. 1. a-c,** Each dot is the location of a microwire bundle in either the amygdala (cyan), hippocampus (yellow) or parahippocampus (red) on which at least one event or boundary cell was recorded, also presented in a template brain in Fig. 1e. Coordinates are in Montreal Neurological Institute (MNI) 152 space, here plotted on top of the CIT168 brain template for axial (**a**), coronal (**b**), and sagittal (**c**) view (see Methods).

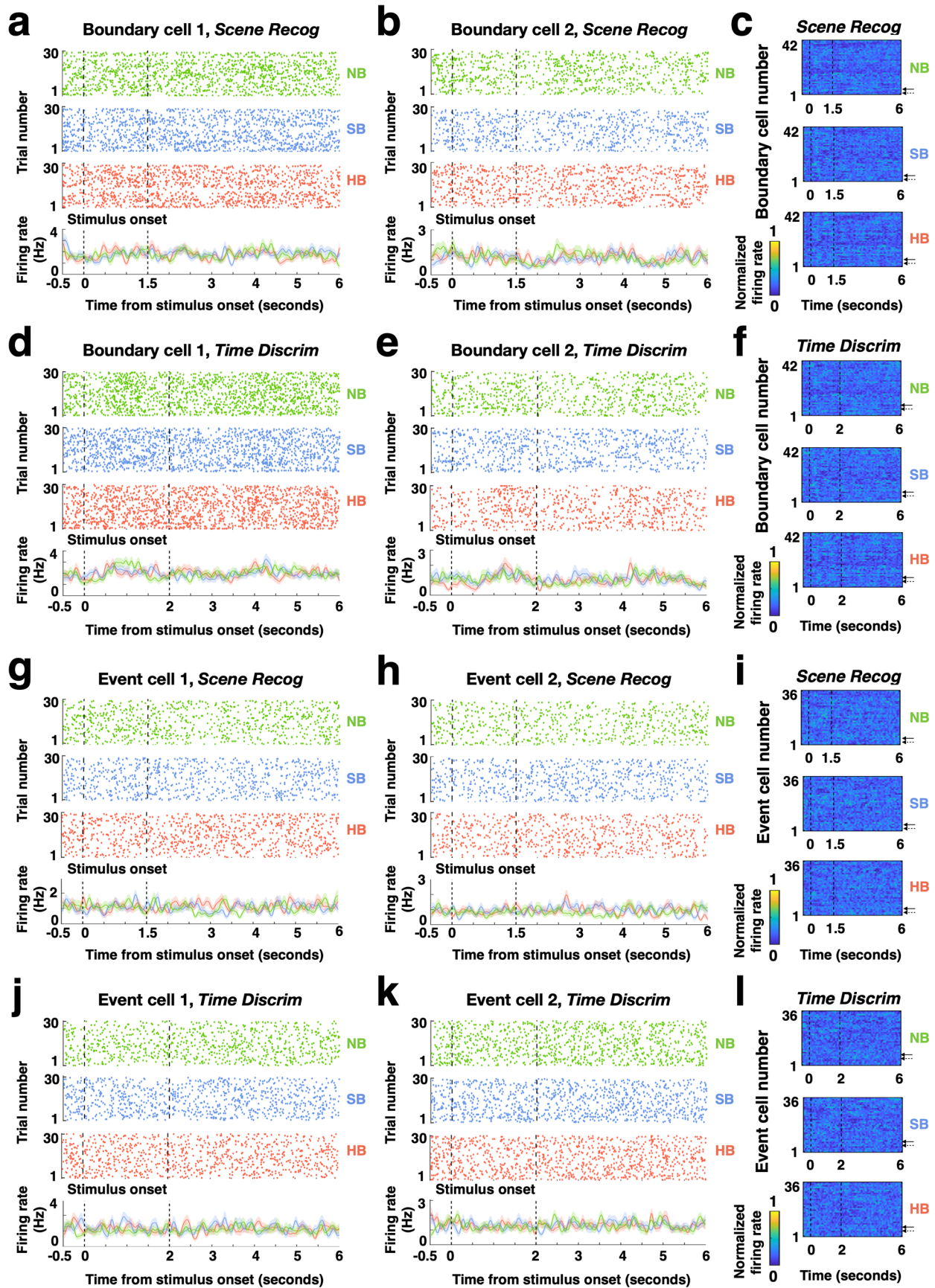


**Extended Data Fig. 2 | Participants' performance in the scene recognition task did not differ significantly across different boundary types, Related to Fig. 2. a-c.** Behavior quantified by accuracy (a), reaction time (b), and confidence level (c) across all trials. Results are shown for boundary type NB (green), SB (blue), and HB (red) during the scene recognition task. The horizontal dashed lines in (a) show chance levels (0.5) and in (c) show the maximum possible confidence value (3 = high confidence). Each dot represents one recording session. Black lines in (a-c) denote the mean results averaged across all recording sessions. One-way ANOVA between NB/SB/HB, degrees of freedom = (2, 57).



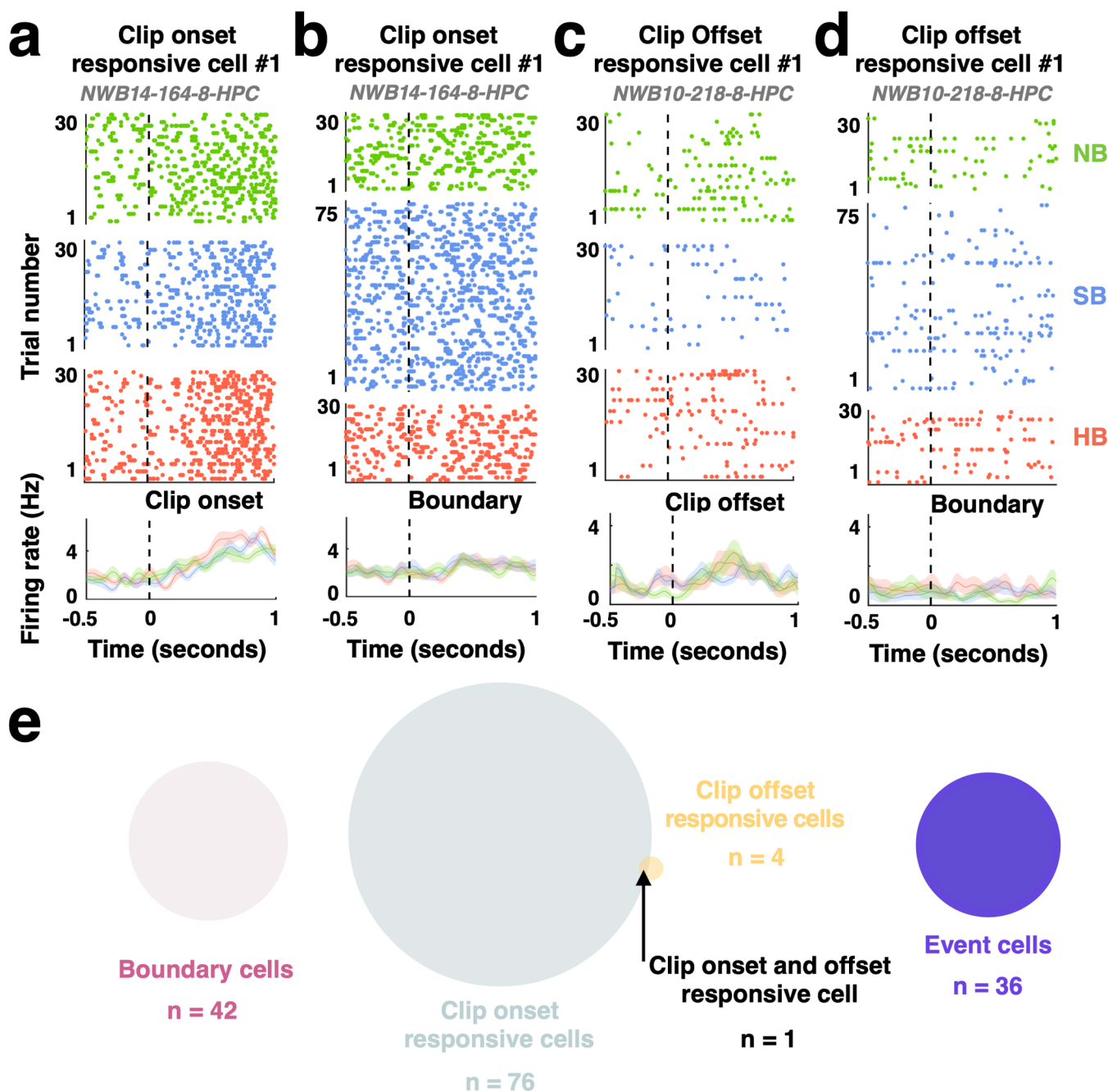


**Extended Data Fig. 3 | Boundary cells and event cells do not respond to clip onsets and clip offsets during encoding, Related to Fig. 3.** **a**, Responses during the encoding stage from the same example boundary cells shown in Fig. 3a,b aligned to the clip onsets. **b**, Firing rates of all 42 boundary cells (solid and dashed arrows denote the examples in **a**) during the encoding stage aligned to the clip onsets, averaged over trials within each boundary type and normalized to each neuron's maximum firing rate throughout the entire task (see color scale on bottom). **c**, Responses during the encoding stage from the same example boundary cells shown in (**a**) aligned to the clip offsets. **d**, Firing rates of all 42 boundary cells during the encoding stage aligned to the clip offsets using the same format as (**b**). **e**, Responses during the encoding stage from the same example event cells shown in Fig. 3e,f aligned to the clip onsets. **f**, Firing rates of all 36 event cells (solid and dashed arrows denote the examples in **e**) during the encoding stage aligned to clip onsets, using the same format as (**b**). **g**, Responses during the encoding stage from the same example event cells shown in **e** aligned to the clip offsets. **h**, Firing rates of all 36 event cells during the encoding stage aligned to the clip offsets using the same format as (**b**). For (**a**), (**c**), (**e**), (**f**), Top: raster plot color coded for different boundary types (green: NB; blue: SB; red: HB). Bottom: Post-stimulus time histogram (bin size = 200 ms, step size = 2 ms, shaded areas represented  $\pm$  s.e.m. across trials). (**b** and **f**) are copied from Fig. 3d,h for comparison purposes.



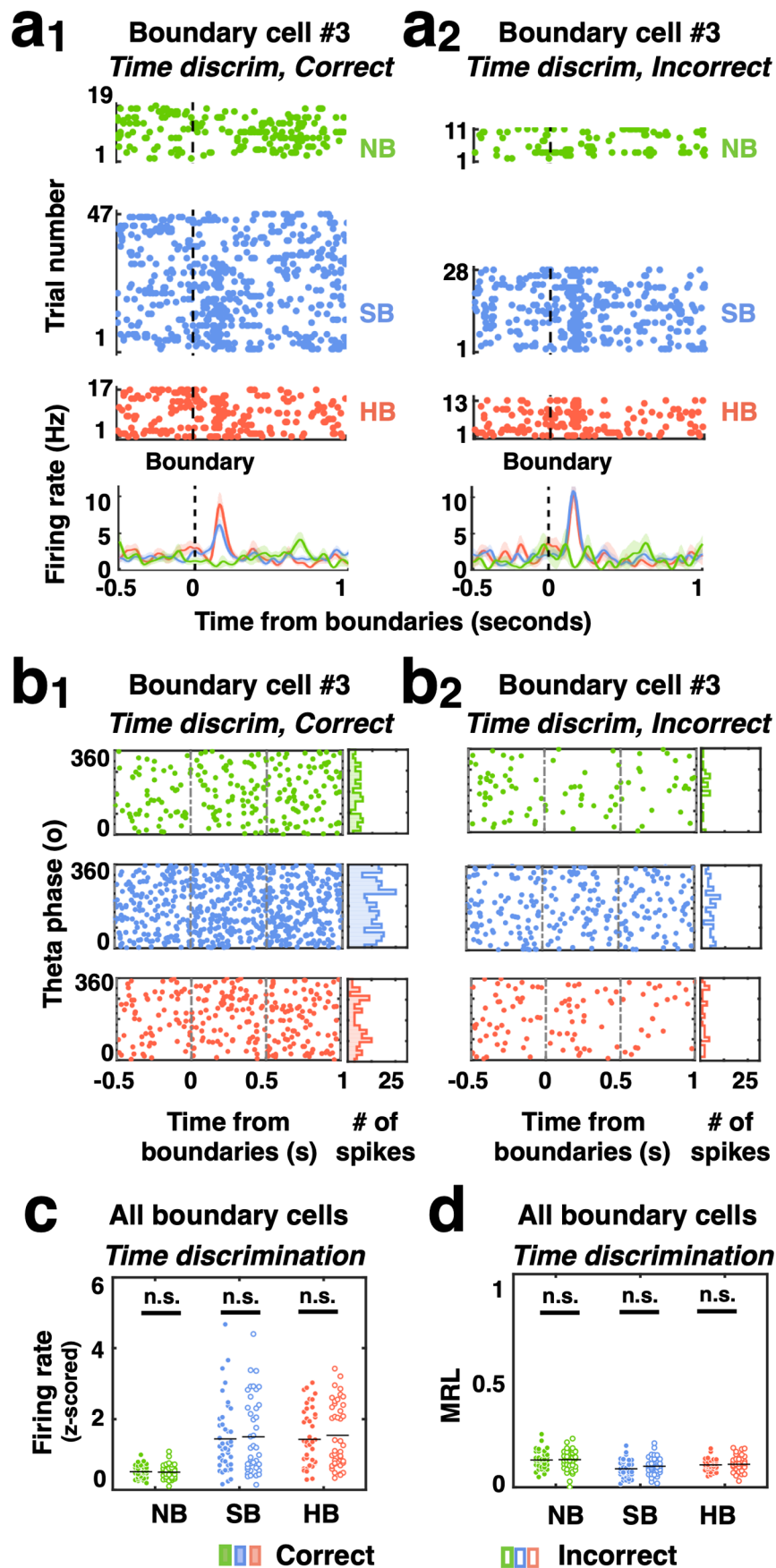
Extended Data Fig. 4 | See next page for caption.

**Extended Data Fig. 4 | Boundary cells and event cells do not respond to image onsets and offsets during scene recognition and time discrimination, Related to Fig. 3.** **a-b**, Responses during scene recognition from the same example boundary cells shown in Fig. 3a,b aligned to stimulus onset. **c**, Firing rates of all 42 boundary cells (solid and dashed arrows denote the examples in **a** and **b**) during scene recognition aligned to the stimulus onsets, averaged over trials within each boundary type and normalized to each neuron's maximum firing rate throughout the entire task (see color scale on bottom). **d-e**, Responses during time discrimination from the same example boundary cells shown in (**a** and **b**) aligned to stimulus onset. **f**, Firing rates of all 42 boundary cells during time discrimination aligned to the stimulus onset using the same format as in **c**. **g-h**, Responses during scene recognition from the same example event cells shown in Fig. 3e,f aligned to stimulus onsets. **i**, Firing rates of all 36 event cells (solid and dashed arrows denote the examples in **g** and **h**) during scene recognition aligned to the stimulus onset, using the same format as in **a** and **b**. **j**, Responses during time discrimination from the same example event cells shown in **g** and **h** aligned to stimulus onset. **k**, Firing rates of all 36 event cells during time discrimination aligned to the stimulus onsets using the same format as in **f**. For (**a**), (**b**), (**d**), (**e**), (**g**), (**h**), (**j**), (**k**), Top: raster plot color coded for different boundary types (green: NB; blue: SB; red: HB). Bottom: Post-stimulus time histogram (bin size = 200 ms, step size = 2 ms, shaded areas represented  $\pm$  s.e.m. across trials).



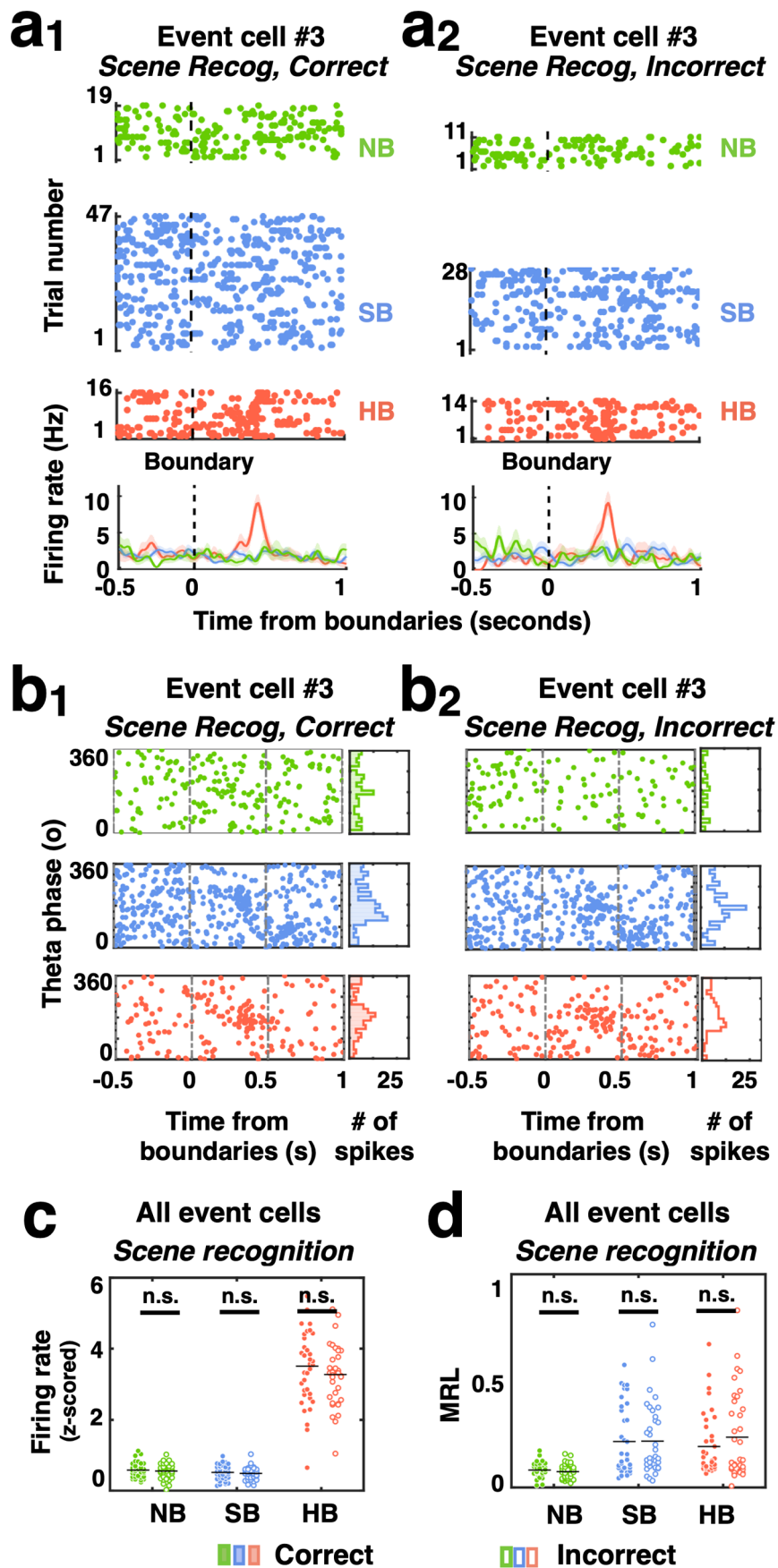
**Extended Data Fig. 5 | Neurons that respond to clip onsets and clip offsets do not overlap with boundary and event cells, Related to Fig. 3.**

**a-b**, Responses during the encoding stage from an example clip onset-responsive cell located in the amygdala aligned to clip onsets (**a**), and boundaries (**b**). Top: raster plots. Bottom: Post-stimulus time histogram (bin size = 200 ms, step size = 2 ms, shaded areas represented  $\pm$  s.e.m. across trials). A cell was considered as a clip onset cell if its firing rate differed significantly between a 1 s window immediate before and after clip onset ( $p < 0.05$ , one-tailed permutation t-test). **c-d**, Responses during the encoding stage from an example clip offset-responsive cell located in the hippocampus aligned to clip offsets (**c**), and boundaries (**d**). A cell was considered as a clip offset cell if its firing rate differed significantly between a 1 s window immediate before and after clip offsets ( $p < 0.05$ , one-tailed permutation t-test). Same format as (**a** and **b**). **e**, Seventy six out of 580 cells in the MTL qualified as clip onset-responsive cells and four out of 580 cells in the MTL qualified as clip offset-responsive cells. None of these were also selected as either boundary or event cells.



Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 6 | Responses of boundary cells during encoding grouped by memory outcomes from the time discrimination task, Related to Fig. 4.** **a<sub>1</sub>-a<sub>2</sub>**, Response of the same example boundary cell in Fig. 4a and Fig. 4b. During encoding, this cell responded to SB and HB transitions regardless of whether the temporal order of the clip was later correctly (**a<sub>1</sub>**) or incorrectly (**a<sub>2</sub>**) retrieved in the time discrimination test. Shaded areas represented  $\pm$  s.e.m. across trials. **b<sub>1</sub>-b<sub>2</sub>**, Left: timing of spikes from the same boundary cell shown in (**a<sub>1</sub>** and **a<sub>2</sub>**) relative to theta phase calculated from the local field potentials, for clips whose temporal order were later correctly (**b<sub>1</sub>**) or incorrectly (**b<sub>2</sub>**) retrieved. Right: phase distribution of spike times within [0, 1] seconds time windows following the middle of the clip (NB) or boundary (SB, HB) for clips whose temporal order were later correctly (**b<sub>1</sub>**) or incorrectly (**b<sub>2</sub>**) retrieved. **c-d**, Population summary for all 42 boundary cells. **c**, Z-scored firing rate (0-1s after boundaries during encoding) for each boundary type did not differ between clips whose temporal orders were later correctly (color filled) vs. incorrectly (empty) retrieved. **d**, Mean resultant length (MRL) of spike times (relative to theta phases, 0-1s after boundaries during encoding) across all boundary cells for each boundary type did not differ between clips whose temporal orders were later correctly (color filled) vs. incorrectly (empty) retrieved. Each dot represents one boundary cell. Black lines in **c** and **d** denote the mean results averaged across all boundary cells. One-tailed permutation t-test, degrees of freedom = (1, 82).



Extended Data Fig. 7 | See next page for caption.

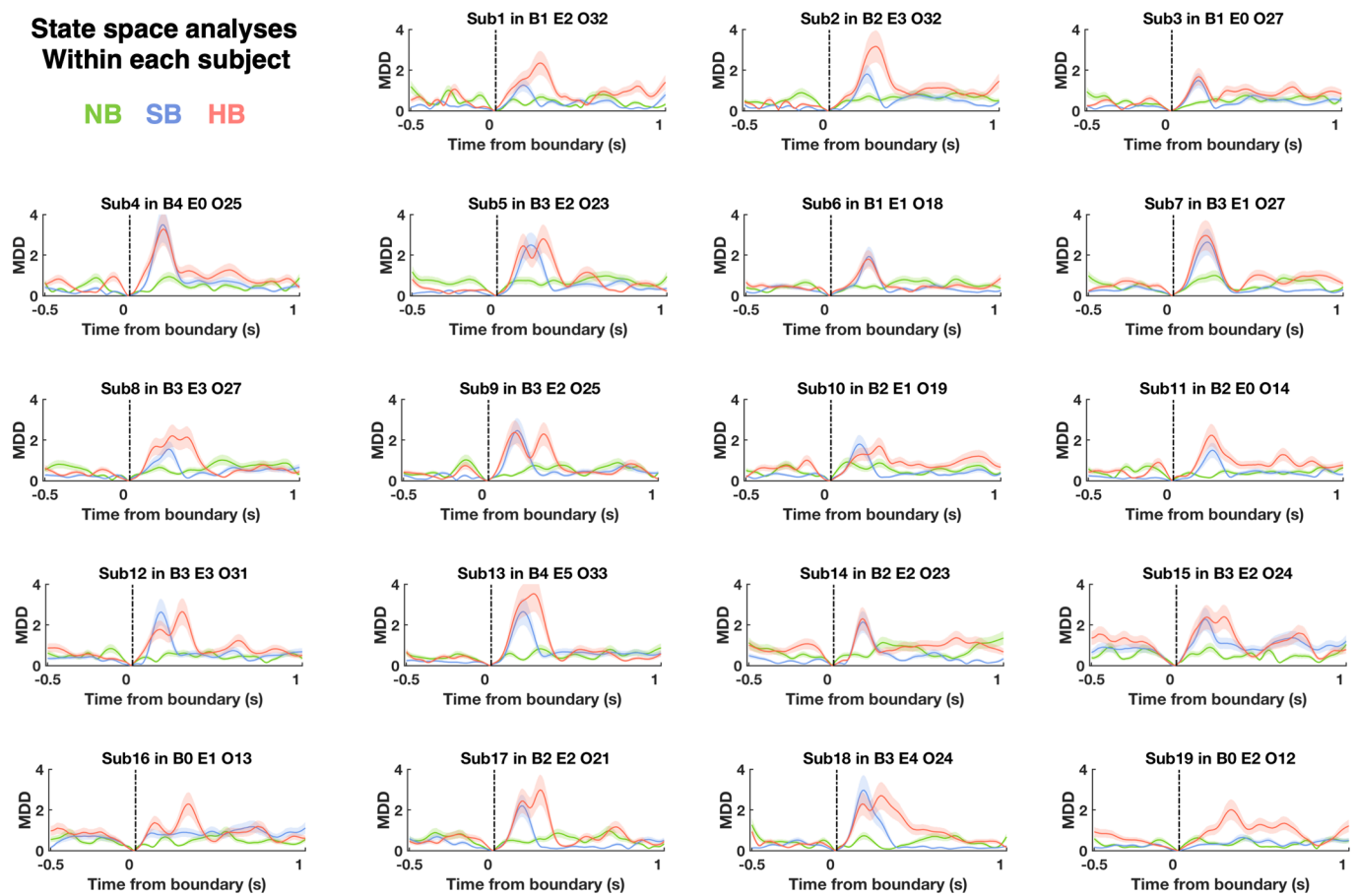
**Extended Data Fig. 7 | Responses of event cells during encoding grouped by memory outcomes from the scene recognition stage, Related to Fig. 4.**

**a<sub>1</sub>-a<sub>2</sub>**, Response of the same example event cell in Fig. 4e,f. During encoding, this cell responded to HB transitions regardless of whether frames were later correctly (**a<sub>1</sub>**) or incorrectly (**a<sub>2</sub>**) recognized in the scene recognition task. Shaded areas represented  $\pm$  s.e.m. across trials. **b<sub>1</sub>-b<sub>2</sub>**, Left: timing of spikes from the same event cell shown in **a<sub>1</sub>-a<sub>2</sub>** relative to theta phase calculated from the local field potentials, for frames that were later correctly (**b<sub>1</sub>**) or incorrectly (**b<sub>2</sub>**) recognized. Right: phase distribution of spike times within [0, 1] seconds time windows following the middle of the clip (NB) or boundary (SB, HB) for frames that were later correctly (**b<sub>1</sub>**) or incorrectly (**b<sub>2</sub>**) recognized. **c-d**, Population summary for all 36 event cells. **c**, Z-scored firing rate (0-1 s after boundaries during encoding) for each boundary type did not differ between frames that were later correctly (color filled) vs. incorrectly (empty) recognized. **d**, Mean resultant length (MRL) of spike times (relative to theta phases, 0-1 s after boundaries during encoding) across all event cells for each boundary type did not differ between frames that were later correctly (color filled) vs. incorrectly (empty) recognized. Each dot represents one event cell. Black lines in **c** and **d** denote the mean results averaged across all event cells (**c**, **d**). One-tailed permutation t-test, degree of freedom = (1, 70).

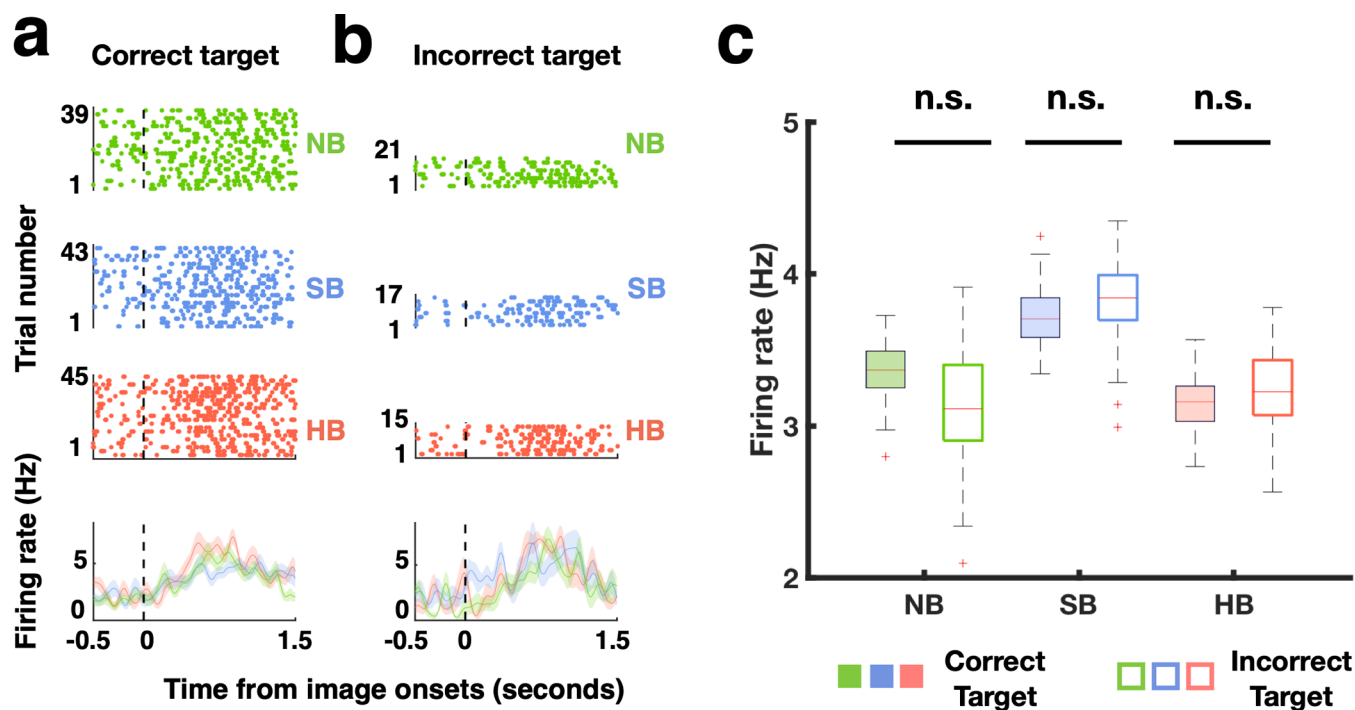


### State space analyses Within each subject

NB SB HB



**Extended Data Fig. 8 | Neural state changes following soft and hard boundaries shown for individual participants, Related to Fig. 5.** Multidimensional distance (MDD, see Fig. 5d–g for definition) as a function of time aligned to the middle of the clip (green: NB) and boundaries (blue: SB, red: HB). MDD is shown for all MTL cells within each participant (for example, ‘Sub1 in B1 E2 O32’ denotes MDD computed by 1 boundary cell, 2 event cells and 32 other MTL cells in participant 1). Shaded areas represent  $\pm$  s.e.m. across trials.



**Extended Data Fig. 9 | Clip-onsets responsive neurons respond to both correct and incorrect targets during scene recognition, Related to Fig. 6.**

**a-b**, Responses during scene recognition from an example clip onset-responsive cell (see definition in Extended Data Fig. 12) located in the amygdala aligned to image onsets in correctly recognized target (**a**) and forgotten target (**b**) trials. Top: raster plots. Bottom: Post-stimulus time histogram (bin size = 200 ms, step size = 2 ms, shaded areas represented  $\pm$  s.e.m. across trials). **c**, Comparison (across all 76 identified clip-onsets responsive neurons) between mean firing rates averaged within [0 1.5]s after image onsets for remembered vs forgotten targets. On each box, the central mark indicates the mean results averaged across all clip-onsets responsive neurons, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' marker symbol. One-way ANOVA, degrees of freedom = (1, 150).

## 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

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- 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)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- 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** Neurophysiological data were collected using the ATLAS system (Neuralynx Inc., Bozeman, Montana, USA) and hybrid depth electrodes (Ad-Tech company, Oak Creek, Wisconsin, USA). See detailed description in the Methods section.

**Data analysis** Data analyses were performed using MATLAB R2019b and Fieldtrip Toolbox 20190527. Spike sorting was done using the open source toolbox OSort v4 (<https://rutishauserlab.org/osort>). Anatomical data analysis was performed using FreeSurfer version 6 ([surfer.nmr.mgh.harvard.edu](http://surfer.nmr.mgh.harvard.edu)) and dcm2niix ([github.com/rordenlab/dcm2niix](https://github.com/rordenlab/dcm2niix)). Specific functions used in this study was described in the Methods section.

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](#)

We have made the data and code that supports the findings of this study publicly available on the NIH DANDI archive and Github. The access links are listed in the "Data availability" and "Code availability" section.

## 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	Our analysis is based on 985 neurons recorded from 20 subjects. This sample size is very large compared to similar studies. No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications.
Data exclusions	No data were excluded
Replication	The analyses were performed at single neuron level. The effect reported in the study were consistent and replicated across 20 subjects.
Randomization	Our design is a within-subject analysis: all the subjects were in the same analysis set and had all types of trials. We performed permutation testing where appropriate to ensure statistical validity of our results.
Blinding	Subjects were not aware of the goals of the study. There was no subjective measurement or decision that the investigator needed to make during the experiment. All the data are collected and analyzed off-line. Data collection and analysis were not performed blind to the conditions of the experiments as conditional information is required for further analyses.

## 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

### Methods

n/a	Involvement in the study	n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern		

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	We studied a group of 20 subjects where 12 were female, and the average age was 40.65 years old (range 20-68), and all were diagnosed with pharmacologically-intractable epilepsy. Supplementary table 3 provides information about each subject. 90 Amazon Mechanical Turk Workers (27 females, age range 21 - 43) have participated in our behavioral only tasks.
Recruitment	Subjects undergoing invasive electrophysiological recording for clinical purposes were recruited and consented to participate in this research study. Subjects who was capable of and was willing to participate the task were recruited. All the Amazon Mechanical Turk Workers recruited in this study were under 50 years old. Their behavioral results might not be generalizable to the older age group.
Ethics oversight	The study was approved by the institutional review boards of Toronto Western Hospital, Cedars-Sinai Medical Center and Boston Children's Hospital. Subjects provided informed consent.

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