

Distinct neuronal populations in the human brain combine content and context

<https://doi.org/10.1038/s41586-025-09910-2>

Received: 10 August 2023

Accepted: 12 November 2025

Published online: 7 January 2026

Open access

 Check for updates

Marcel Bausch¹, Johannes Niediek^{1,2}, Thomas P. Reber^{1,3}, Sina Mackay¹, Jan Boström⁴, Christian E. Elger¹ & Florian Mormann¹

The medial temporal lobe, and particularly the hippocampus, has been proposed to encode items in context^{1,2}. Although hippocampal memory representations are largely context-dependent in rodents^{3,4}, concept cells in humans appear to be context-invariant⁵. However, it remains unknown how item and context information are combined to form or retrieve integrated item-in-context memories at the single-neuron level in humans. Here we show that coordinated activity of distinct neuronal populations supports item-in-context memory. In a context-dependent picture-comparison task, we recorded 3,109 neurons from 16 neurosurgical patients, identifying 597 stimulus-modulated (pre-screened) and 200 context-modulated neurons (2.95% in the amygdala, 7.68% in the parahippocampal cortex, 5.68% in the entorhinal cortex and 9.42% in the hippocampus). Their co-firing combined different comparison questions (contexts) with two subsequent pictures (stimuli) through neuronal reinstatement of question contexts. Both populations were largely separate, generalized across the preferred dimension of each other, and covaried with behavioural performance. Following experimental pairing of stimuli and context, firing of entorhinal stimulus neurons predicted that of hippocampal context neurons after tens of milliseconds. Overall, synaptic modifications and co-firing of stimulus and context neurons could contribute to item-in-context memory, specify which stimulus memories need to be retrieved, and even generalize memories through mutual reinstatement of largely separate, orthogonal representations. By contrast, only 50 stimulus–context neurons represented specific picture–question combinations, consistent with limited pattern separation in the human medial temporal lobe, favouring flexible generalization over rigid conjunctive coding.

The medial temporal lobe (MTL), and particularly the hippocampus, is essential for declarative memories of items in context^{1,2}. In humans, the hippocampus has also been implicated in generalization^{6,7}. With ease, we recall multiple memories involving a particular memory item, for example, different dinners with our friend. A uniquely human single-neuron correlate of memory items is represented by concept neurons in the MTL, which selectively and invariantly respond to stimuli containing a preferred semantic concept⁸ (for example, our friend). Concept neurons are apparently not modulated by the context of an item, and thus seem suited for generalization⁵ (but see ref. 9). As humans can also recall items based on context, concept neurons may be complemented by neurons representing context. However, remembering contexts such as time, task or location together with memory items requires the combination of item and context information¹⁰. For example, we remember when we had dinner with our friend, whether it was work related or where it took place. Furthermore, we can recall a specific restaurant visit based on item and context information. Initial evidence has suggested that separate neurons in the MTL represent contexts^{11–13}. However, it remains unclear how

neural correlates of item and context memory are combined to form or retrieve integrated item-in-context memories at the single-neuron level in humans. Recording from 16 patients undergoing surgery to treat epilepsy, we investigated how human MTL neurons combine item and context representations. We devised a task in which pairs of pictures were presented sequentially following different contexts, namely, questions (for example, ‘Bigger?’) that specified how to compare the two pictures. This required participants to remember items (that is, two pictures) in a specific context. Here ‘context’ refers to interactive or task context rather than independent or background context¹⁴. Only a small fraction of neurons encoded individual item–context combinations as reported in rodents. Instead, separate populations of neurons represented either items or contexts in an orthogonal encoding scheme, reflecting the ability of humans to generalize memories along each dimension separately. Finally, we show how item and context information is combined via co-activation, synaptic modification or reinstatement, contributing to item-in-context memory and the retrieval of contextually relevant item memories.

¹Department of Epileptology, University Hospital Bonn, Bonn, Germany. ²Machine Learning Group, Technische Universität Berlin, Berlin, Germany. ³Faculty of Psychology, UniDistance Suisse, Brig, Switzerland. ⁴Department of Neurosurgery, University Hospital Bonn, Bonn, Germany. ✉e-mail: marcel.bausch.ukb@gmail.com; florian.mormann@ukbonn.de

Distinct neurons for content and context

During 49 experimental sessions, we recorded from 3,109 neurons in the amygdala, parahippocampal cortex, entorhinal cortex and hippocampus of 16 neurosurgical patients implanted with depth electrodes for invasive seizure monitoring. Pairs of pictures were presented on a laptop screen in different contexts, namely, five questions ('Bigger?', 'Last seen in real life?', 'Older?' or 'More expensive?', 'Like better?' and 'Brighter?') that specified how the pictures should be compared (Fig. 1a). For each session, four pictures eliciting selective neuronal responses in a previous screening were selected. Each trial began with a context-providing question (such as Bigger?), followed by a sequence of two of the four pictures (stimuli) that needed to be remembered in context and compared accordingly. Patients then chose the picture that best answered the question (for example, which depicted something bigger) and indicated whether it was shown first or second via the keyboard. Most answers were highly consistent and transitive, with performance greatly exceeding chance in all but one excluded session and not differing across questions ($P = 0.083$; Kruskal–Wallis test; Extended Data Fig. 1). Furthermore, answers strongly correlated with ground truths for Bigger? and More expensive? or Older? (mean $\rho = 0.760$, $P = 6.47 \times 10^{-10}$; Extended Data Fig. 2). To quantify the representation of stimulus and context during picture presentations, repeated-measures two-way 4×5 analysis of variance (ANOVA) with factors picture, question and their interaction were computed for each neuron ($\alpha = 0.001$; effect sizes in Fig. 2e).

Although we expected to find many neurons with a significant main effect of stimulus due to our pre-selection of response-eliciting pictures, it was unclear how context would be encoded. Distinct neuronal populations were identified. Most neurons only exhibited a significant main effect of either stimulus (Fig. 1b) or context (Fig. 1c) and were termed (mere) stimulus and context neurons, respectively. Figure 1b shows an example of a (mere) stimulus neuron, selectively increasing firing to 'biscuit', irrespective of context (that is, comparison prompted by the question). Figure 1c shows a (mere) context neuron, whose firing distinguished only context during the late phase of picture presentations. Whenever the trial started with Older?, but not in any other context, this neuron responded to all pictures regardless of stimulus identity. Furthermore, we identified neurons with a significant interaction that encoded stimulus and context conjunctively (Fig. 1d). These neurons only responded when a specific picture (for example, train) was shown in a specific context (for example, Brighter image?), but not for any other picture–context combination. Finally, although mere stimulus (MS) neurons (only significant main effect of stimulus) were much more common (Fig. 1b), we also identified contextual stimulus neurons defined by significant main effects of both stimulus and context (irrespective of interaction), for example, selectively responding to the picture of a hamburger, yet particularly strongly when the context was Last seen in real life? (Fig. 1e).

Distribution of content and context

Context was represented during picture presentations in substantially more neurons than expected by chance. Out of 3,109 recorded units, 200 exhibited a main effect of question $P < 0.001$ in a repeated-measures two-way ANOVA with factors stimulus and question (including interaction; Fig. 2a), largely exceeding the expected number of roughly 3 units (two-sided binomial test, $P < 2.225 \times 10^{-308}$). This amounted to 2.95% of neurons in the amygdala (31 of 1,051), 7.68% in the parahippocampal cortex (37 of 482), 5.68% in the entorhinal cortex (25 of 440) and 9.42% in the hippocampus (107 of 1,136), all far exceeding chance (Fig. 2b; two-sided binomial tests, all $P < 2.225 \times 10^{-308}$). Although most neurons that encoded stimulus did not encode context (524 of 597), the fraction that did or that exhibited an interaction was still highly significant in two-sided binomial tests: out of the 597 neurons with a

main effect of stimulus, 73 also showed a main effect of question with $P < 0.001$ (12.23%; $P < 2.225 \times 10^{-308}$) and 31 an interaction with $P < 0.001$ (5.19%; $P < 2.225 \times 10^{-308}$). Although the populations were largely distinct, a significant main effect of context was more prevalent among neurons with than without an additional main effect of stimulus, particularly in the amygdala (7.83% versus 1.67%), the entorhinal cortex (10.29% versus 4.84%) and the hippocampus (21.26% versus 7.28%; Fig. 2c). Mere context (MC) neurons (only significant main effect of context) were most frequent in the parahippocampal cortex (6.96%, 25 of 359) and rarest in the amygdala (1.67%, 14 of 836). Moreover, significant interactions ($P < 0.001$) capturing conjunctive representations strongly exceeded chance in two-sided binomial tests (1.61%, 50 of 3,109; $P < 2.225 \times 10^{-308}$), particularly in the amygdala (1.05%, 11 of 1,051; $P = 1.588 \times 10^{-8}$), the parahippocampal cortex (2.28%, 11 of 482; $P = 4.733 \times 10^{-12}$) and the hippocampus (2.29%, 26 of 1,136; $P < 2.225 \times 10^{-308}$; Fig. 2b), especially among hippocampal stimulus neurons (10.34%, 18 of 174; Fig. 2d). To account for potential statistical dependencies between factors or neurons, ANOVA effect sizes of all neurons were compared with stratified label-shuffling controls, irrespective of significance (Fig. 2e; see Methods). Confirming previous results, patient-averaged effect sizes of stimulus ($\eta_p^2(\text{real}) = 0.114$, $\eta_p^2(\text{control}) = 0.034$, $P_{\text{patient}} = 1.313 \times 10^{-3}$ and $P_{\text{session}} = 3.330 \times 10^{-9}$), context ($\eta_p^2(\text{real}) = 0.059$, $\eta_p^2(\text{control}) = 0.034$, $P_{\text{patient}} = 1.313 \times 10^{-3}$ and $P_{\text{session}} = 4.827 \times 10^{-9}$) and their interaction ($\eta_p^2(\text{real}) = 0.036$, $\eta_p^2(\text{control}) = 0.034$, $P_{\text{patient}} = 1.840 \times 10^{-2}$ and $P_{\text{session}} = 3.140 \times 10^{-3}$) all significantly differed from controls (two-sided Wilcoxon signed-rank tests, Bonferroni corrected). Linear mixed-effects models with a fixed intercept (capturing grand population effect size differences) and random intercepts for patients, fitted across all recorded neurons, confirmed that real effect sizes significantly exceeded permutation-based baselines for all three ANOVA factors: stimulus ($\beta = 0.080$, 95% CI 0.065–0.095, $P = 2.177 \times 10^{-24}$), context ($\beta = 0.025$, 95% CI 0.019–0.032, $P = 4.186 \times 10^{-14}$) and stimulus–context interaction ($\beta = 0.002$, 95% CI 0.001–0.003, $P = 2.269 \times 10^{-6}$). Finally, we examined whether stimulus–context neurons were more prevalent in areas with higher proportions of stimulus or context neurons (Extended Data Fig. 3). Across all session–site combinations ($n = 168$), stimulus–context neuron proportions correlated significantly with both stimulus (Spearman's $\rho = 0.177$, $P = 0.022$) and context neuron proportions ($\rho = 0.487$, $P < 0.001$), particularly in the hippocampus and amygdala, where stimulus–context proportions correlated with both stimulus ($\rho = 0.274$, $P = 0.032$ for the hippocampus; and $\rho = 0.329$, $P = 0.012$ for the amygdala) and context neuron proportions ($\rho = 0.471$, $P < 0.001$ for the hippocampus; and $\rho = 0.487$, $P = 0.003$ for the amygdala). By contrast, the parahippocampal and entorhinal cortices showed significant correlations exclusively with context ($\rho = 0.569$, $P < 0.001$ for the parahippocampal cortex; and $\rho = 0.377$, $P = 0.023$ for the entorhinal cortex) but not with stimulus neuron proportions.

Abstract representations reflect task

Our next analyses sought to address whether all contexts can be decoded, how well context representations generalized across different pictures or across time, and how they were related to stimulus representations. For each session, we computed the support vector machine (linear SVM) decoding accuracy of context during picture presentations (100–1,000 ms). Decoders were trained with either all neurons or all context neurons of each session (Fig. 3a). Confirming previous ANOVA results, decoding accuracies significantly exceeded chance levels of 0.2 (one out of five questions, two-sided Wilcoxon signed-rank test, uncorrected) for all contexts when all neurons were included ($P_{\text{session}} = 1.981 \times 10^{-3}$ and $P_{\text{patient}} = 4.094 \times 10^{-3}$ for Bigger?; $P_{\text{session}} = 9.142 \times 10^{-7}$ and $P_{\text{patient}} = 6.430 \times 10^{-4}$ for Last seen?; $P_{\text{session}} = 5.187 \times 10^{-3}$ and $P_{\text{patient}} = 3.861 \times 10^{-2}$ for Older? or More expensive?; $P_{\text{session}} = 1.456 \times 10^{-3}$ and $P_{\text{patient}} = 6.624 \times 10^{-3}$ for Like better?; $P_{\text{session}} = 8.016 \times 10^{-8}$ and

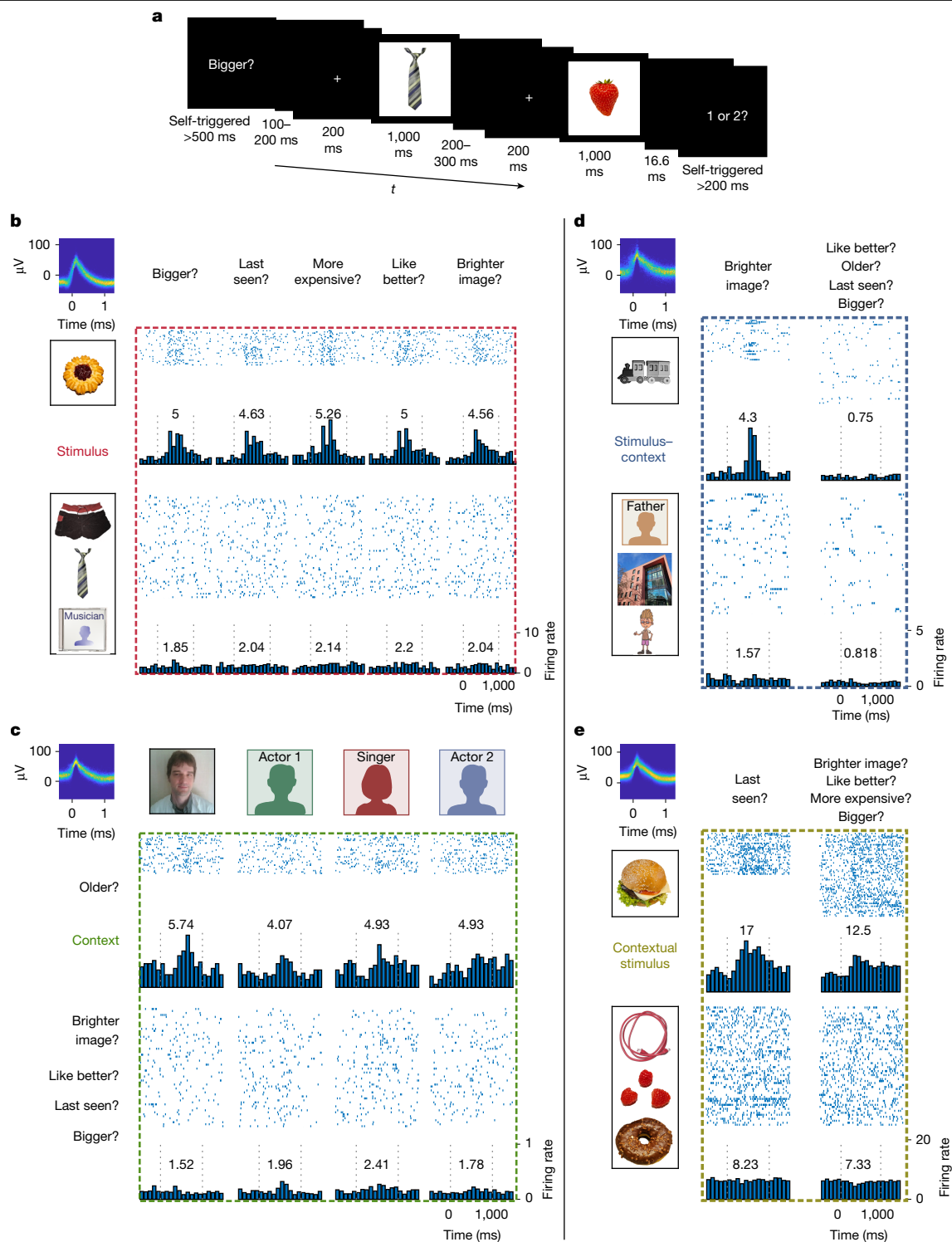


Fig. 1 | Neurons in the MTL encode stimulus and context during picture presentations. Although most stimulus neurons only encode picture identity, context is either represented alone, in conjunction with stimulus identity or as modulation of stimulus-specific firing rates. **a**, Paradigm with exemplary trial of the stimulus-comparison task (here for the context question Bigger?). During 300 trials, participants had to compare pictures according to a context question. Each trial contained one of five questions (Bigger?, Last seen (in real life)?, More expensive? or Older?, Like better? and Brighter?), a sequence of two pictures (out of four) and an answer prompt displaying '1 or 2?'. Participants indicated the sequential position of the picture that best answered the question by pressing keys 1 or 2. Question and answer screens were self-triggered. Event durations are printed underneath. **b–e**, Example neurons with spike density plots in the upper left whose firing rate (Hz) during picture presentations

(0–1,000 ms) is visualized by raster plots and histograms as a function of question and stimulus identity. **b**, A stimulus neuron that selectively increased firing whenever a biscuit was shown, irrespective of which question context was presented. The neuron did not respond to any other picture. **c**, This context neuron increased firing during picture presentations whenever the trial started with the question Older? (top left) as opposed to any other question (bottom left) regardless of stimulus identity (top row). **d**, Example of a stimulus-context interaction neuron whose firing increased when the question was Brighter image? and the stimulus depicted a train (top left) in contrast to other combinations of question and stimulus. **e**, This contextual stimulus neuron only responded to its preferred picture (hamburger, top left), but more strongly in the context of the question Last seen (in real life)?.

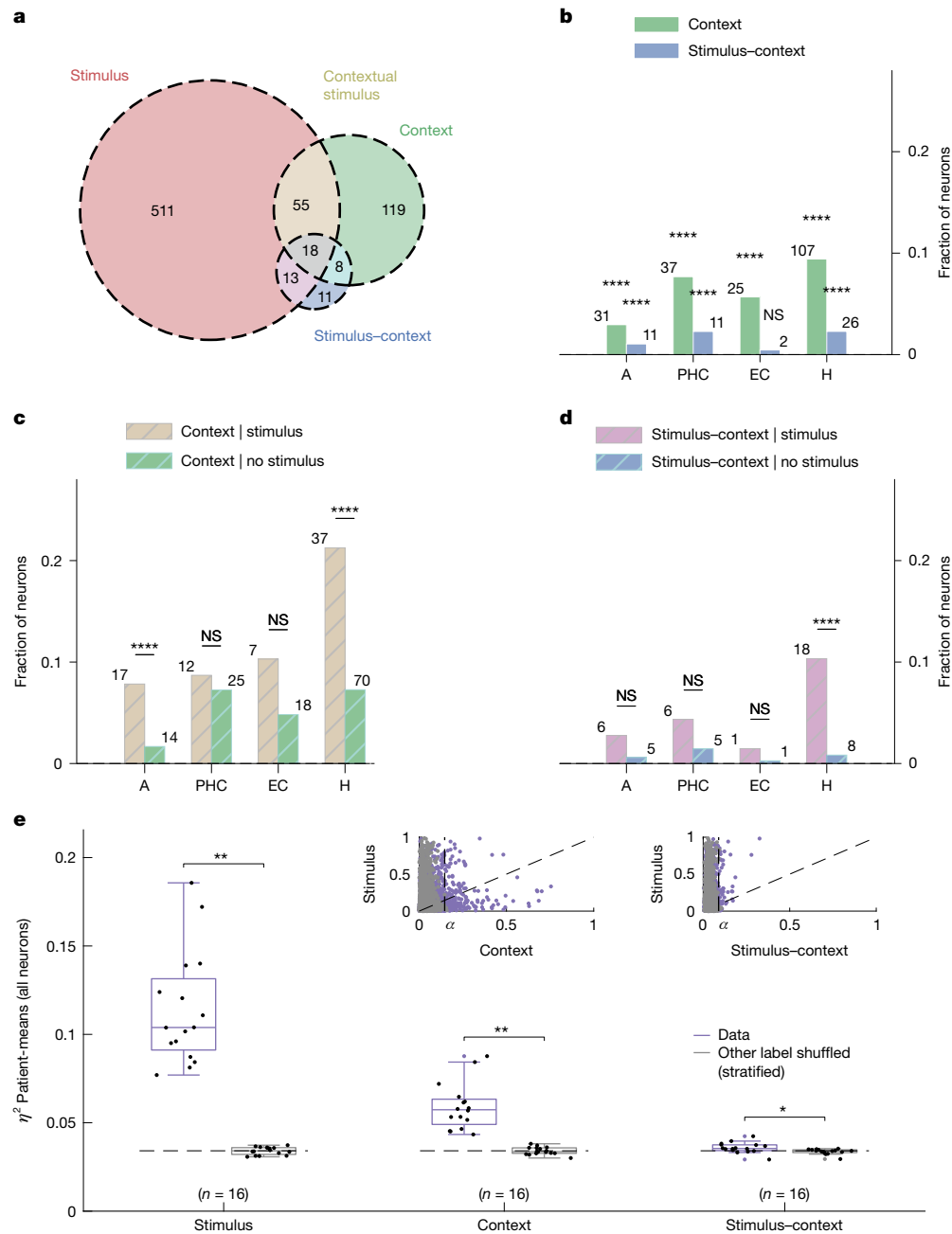


Fig. 2 | Stimulus and context are mainly encoded separately and rarely conjunctively. **a**, Venn diagram of a significant number of neuron types defined by repeated-measures two-way ANOVA with factors stimulus (stimulus neurons in red, ochre, grey or pink), question (context neurons in green, ochre, grey or light blue) or their interaction (stimulus–context interaction neurons in grey, pink, blue or light blue) during picture presentations ($\alpha = 0.001$) from 3,109 neurons. Intersections denote multiple effects. Most stimulus neurons were not modulated by context but overlapped with context neurons (contextual stimulus neurons) or stimulus–context interaction neurons. **b**, Probabilities of context or stimulus–context interaction neurons across brain regions (absolute numbers above the bars). The significance asterisks denote binomial tests versus chance (dotted line). Neurons in all MTL regions were strongly modulated by context and stimulus–context. A, amygdala; EC, entorhinal cortex; H, hippocampus; PHC, parahippocampal cortex. **c**, Analogous to panel **b** with probabilities of context conditioned on stimulus modulation (stimulus | no stimulus). The significance asterisks denote Fisher's exact test results from regional contingency tables. Stimulus neurons were more likely to be modulated

by context (contextual stimulus neurons) than non-stimulus neurons in all MTL regions (significant in amygdala and hippocampus), especially the hippocampus. **d**, Analogous to panel **c**, with probabilities of significant interactions (stimulus–context). Stimulus–context interaction neurons with conjunctive representations were most frequent in the hippocampus. NS, not significant. **e**, Boxplots of patient-averaged effect sizes from all 3,109 neurons (η_p^2) for factors stimulus, context and interaction (stimulus–context). Data (purple) are compared with stratified-shuffled controls (grey; Wilcoxon signed-rank test). Effect sizes markedly exceeded chance for stimulus ($\eta_p^2(\text{real}) = 0.114$, $\eta_p^2(\text{control}) = 0.034$, $P_{\text{patient}} = 1.313 \times 10^{-3}$, $P_{\text{session}} = 3.330 \times 10^{-9}$); and context ($\eta_p^2(\text{real}) = 0.059$, $\eta_p^2(\text{control}) = 0.034$, $P_{\text{patient}} = 1.313 \times 10^{-3}$ and $P_{\text{session}} = 4.827 \times 10^{-9}$). The difference for stimulus–context was smaller but significant ($\eta_p^2(\text{real}) = 0.036$, $\eta_p^2(\text{control}) = 0.034$, $P_{\text{patient}} = 1.840 \times 10^{-2}$ and $P_{\text{session}} = 3.140 \times 10^{-3}$). The insets depict scatter plots of η_p^2 of context (left) or stimulus–context (right) versus stimulus. The dashed lines indicate the smallest significant effect size ($\alpha = 0.001$). **** $P < 0.0001$, ** $P < 0.01$ and * $P < 0.05$ (Bonferroni corrected).

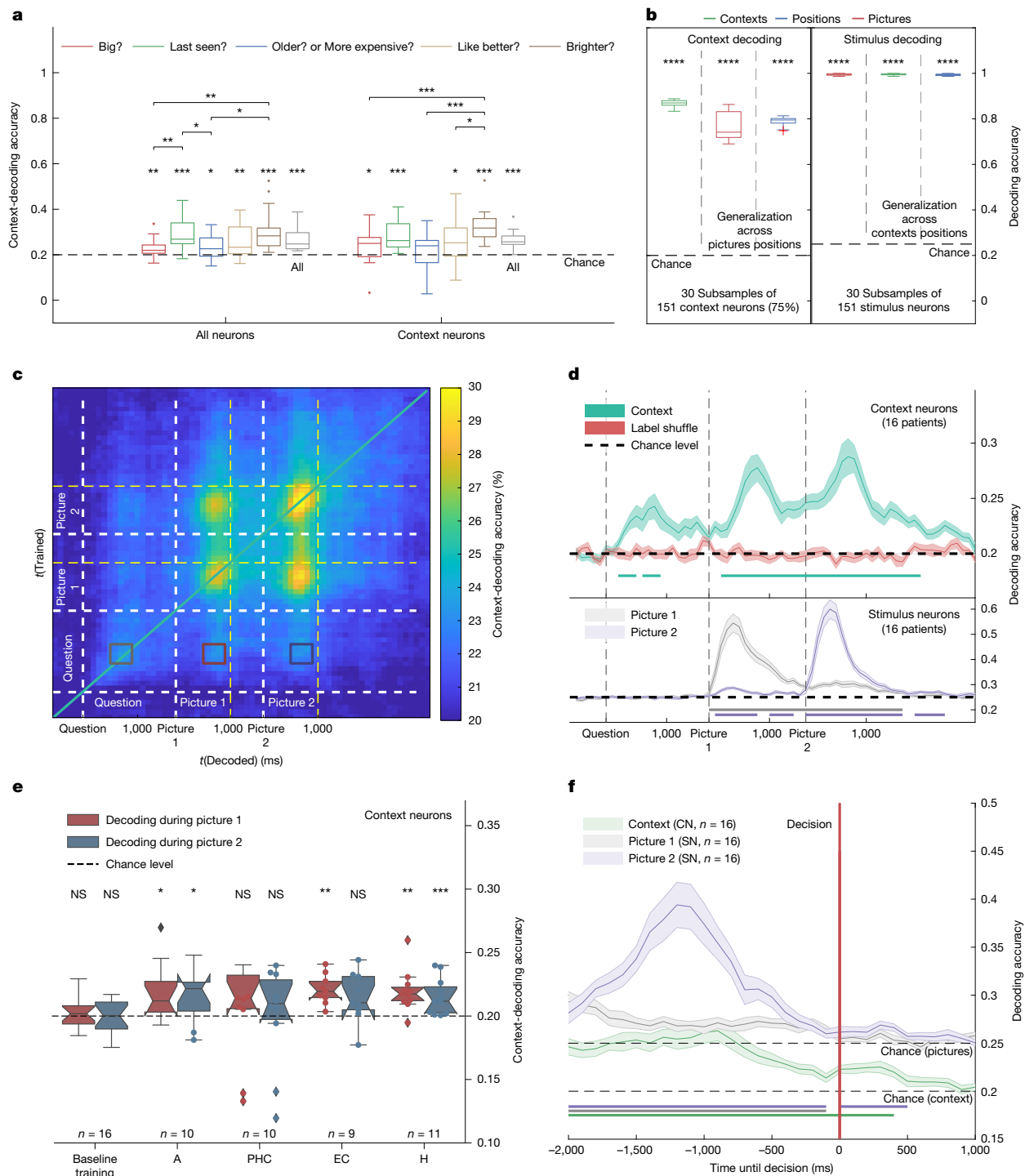


Fig. 3 | Generality and temporal dynamics of context representations imply relevance for stimulus processing. Abstract context representations invariant to picture identity or position combine with content representations across temporal gaps during picture presentations via context reactivation until context-dependent decisions. **a**, Population SVM-decoding accuracies of context (16 patients) during picture presentations comparing different questions (colours) using all versus only context neurons. **b**, Pooled context (green) or stimulus (red) decoding accuracies from context (left) and stimulus (right) neurons, for cross-validation or generalization across pictures (red), contexts (green) and serial picture positions (blue). **c**, A heat map of context decoding across time. The green diagonal line denotes identical training or testing times (see panel **d**). The dashed lines show event onsets (white) and offsets (yellow). The boxes indicate times analysed in panel **e**. **d**, Patient-wise context (top; green) or stimulus decoding (bottom; picture 1 or picture 2 in grey or lavender, respectively) and label-shuffled controls (red). Data are mean $\pm$ s.e.m. (solid lines

and shaded areas, respectively), with dashed lines indicating chance. Significant differences (cluster permutation test, $P < 0.01$) from controls (top) or chance (below) are shown as solid lines below. Context decoding exceeded chance throughout, peaking after stimulus representations and remaining elevated. **e**, Patient-wise context-decoding accuracies trained with context neuron activity during questions (400 neurons per region) or baseline, and decoded during late first (red) or second (blue) picture presentations (boxes in panel **c**). The dashed line indicates chance. Question activity was reactivated during pictures, especially in the hippocampus. **f**, As in panel **d**, but before context-dependent decisions. Both context and content representations exceeded chance (dashed lines) until decisions (red line). CN, context neuron; SN, stimulus neuron. For the boxplots (**a**, **b**, **e**), quantile 1, median, quantile 3 and whiskers (points within $\pm 1.5 \times$ the interquartile range) are shown, with the asterisks denoting differences from zero (Wilcoxon signed-rank test, uncorrected) or each other (Mann-Whitney U -test, uncorrected). **** $P < 0.0001$, *** $P < 0.001$, ** $P < 0.01$ and * $P < 0.05$.

$P_{\text{patient}} = 4.378 \times 10^{-4}$ for Brighter?; and $P_{\text{session}} = 4.378 \times 10^{-4}$ and $P_{\text{patient}} = 4.378 \times 10^{-4}$ for all contexts). Similar results were obtained when restricting these analyses to context neurons ($P_{\text{session}} = 1.713 \times 10^{-2}$ and $P_{\text{patient}} = 4.938 \times 10^{-2}$ for Bigger?; $P_{\text{session}} = 5.888 \times 10^{-5}$ and $P_{\text{patient}} = 4.378 \times 10^{-4}$ for Last seen?; $P_{\text{session}} = 3.875 \times 10^{-2}$ and $P_{\text{patient}} = 4.08 \times 10^{-1}$ for Older? or More expensive?; $P_{\text{session}} = 8.126 \times 10^{-3}$ and $P_{\text{patient}} = 4.373 \times 10^{-2}$ for Like better?; $P_{\text{session}} = 9.826 \times 10^{-7}$ and $P_{\text{patient}} = 4.368 \times 10^{-4}$ for Brighter?; and $P_{\text{session}} = 4.368 \times 10^{-4}$ and $P_{\text{patient}} = 4.368 \times 10^{-4}$ for all contexts). Although session-wise decoding accuracies were highest for Last seen? (0.292) and Brighter image? (0.303), they significantly exceeded chance for all contexts (0.229 for Bigger?, 0.234 for Older? or More expensive? and 0.257 for Like better?).

Next, we assessed decoding performances from neurons pooled across sessions. To estimate their variance, random subsamples were drawn (see Methods). Pooled analyses yielded highly significant context (0.866, $P = 1.708 \times 10^{-6}$; green) and stimulus (0.995, $P = 1.627 \times 10^{-6}$; red) decoding accuracies (two-sided Wilcoxon signed-rank test against 0.2 or 0.25, uncorrected; Fig. 3b). In addition, we quantified generalization across stimuli, contexts or serial picture positions. When training and decoding contexts across different stimuli (0.764, $P = 1.727 \times 10^{-6}$) or picture positions (0.791, $P = 1.716 \times 10^{-6}$), accuracies remained highly significant. The same was true when stimuli were decoded across different contexts (0.995, $P = 1.538 \times 10^{-6}$) or picture positions (0.994, $P = 1.638 \times 10^{-6}$).

Subsequently, we computed context-decoding accuracies from binned population activity across different training and decoding periods during the trial (Fig. 3c; see Methods). Starting approximately 400 ms after question onset, context-decoding accuracies markedly exceeded chance (more than 20%; light blue or warmer colours) for most of the trial, peaking during question and late picture presentations. Consistent with previous analyses (Fig. 3b), training during first-picture presentations resulted in high decoding accuracies during second-picture presentations and vice versa, confirming generalized context representations. Figure 3d depicts patient-wise context-decoding accuracies of context neurons (top; green) for identical training and decoding times (green diagonal line in Fig. 3c) together with stimulus-decoding accuracies of pictures 1 and 2 by stimulus neurons (bottom; grey and lavender). Context-decoding accuracies (Fig. 3d, top) exceeded chance for the entire duration of the trial (solid green line; two-sided cluster permutation test, $P < 0.01$), even compared with label-shuffled controls (red line; two-sided cluster permutation test, $P < 0.01$). In parallel, stimulus neurons significantly encoded currently depicted pictures with high accuracies (Fig. 3d, bottom), as well as previous or even upcoming pictures (solid lines in grey and lavender). The latter was possible as second pictures could be inferred stochastically by design (probability of one-third instead of one-quarter due to no repetition). Of note, neural representations of first pictures were reactivated during second-picture presentations in all contexts (Extended Data Fig. 4), confirming that stimulus neurons encoded contents rather than question-related features. Furthermore, visual stimuli and contexts could be decoded simultaneously during picture presentations (first stimuli, then contexts) and with high accuracy even within single sessions (Extended Data Fig. 5). Context decoding across brain regions is shown in Extended Data Fig. 6. During question presentations, context decoding significantly exceeded chance (two-sided cluster permutation test, $P < 0.01$) in the parahippocampal cortex (turquoise), the entorhinal cortex (violet), the hippocampus (light blue) and even in the amygdala (mustard; $P < 0.05$). Context decoding was most accurate in the hippocampus (blue), particularly during picture presentations, exceeding chance in all four MTL regions.

Next, we determined whether activity patterns during question presentations were recapitulated during picture presentations (Fig. 3e). Session-wise SVM decoders were trained with regional context neuron activity either during question presentations (600–1,000 ms) or preceding baseline intervals (–500 to 100 ms), whereas contexts were

decoded in the late phases of first (red; 600–1,000 ms) or second (blue; 600–1,000 ms) picture presentations. Training and decoding intervals correspond to the coloured boxes in Fig. 3c. To account for regional differences in the number of recorded neurons (1,051 in the amygdala, 482 in the parahippocampal cortex, 440 in the entorhinal cortex and 1,136 in the hippocampus), 400 neurons were randomly sampled per region. Decoders trained during questions, but not during preceding baselines (Fig. 3e, far left), significantly distinguished contexts during both subsequent picture presentations (Fig. 3e, right side; Wilcoxon signed-rank test, uncorrected) in the amygdala ($A_{\text{picture1}}: P = 1.953 \times 10^{-2}$ and $A_{\text{picture2}}: P = 4.883 \times 10^{-2}$), the entorhinal cortex ($EC_{\text{picture1}}: P = 3.906 \times 10^{-3}$ and $EC_{\text{picture2}}: P = 5.469 \times 10^{-2}$) and the hippocampus ($H_{\text{picture1}}: P = 1.953 \times 10^{-3}$ and $H_{\text{picture2}}: P = 6.601 \times 10^{-4}$), but not in the baseline control condition or in the parahippocampal cortex ($BL_{\text{picture1}}: P = 7.057 \times 10^{-1}$ and $BL_{\text{picture2}}: P = 1$; and $PHC_{\text{picture1}}: P = 4.316 \times 10^{-1}$ and $PHC_{\text{picture2}}: P = 4.922 \times 10^{-1}$).

Finally, we tested whether content and context representations of stimulus and context neurons persisted until the end of the trial when patients made context-dependent picture choices (Fig. 3f). Both context and picture contents could be decoded above chance until a decision was indicated (two-sided cluster permutation test, $P < 0.01$), with significantly enhanced decoding when choices matched ground truth ($P = 0.0128$ for context and $P = 0.0351$ for stimulus, one-sided Wilcoxon signed-rank test; Extended Data Fig. 7).

Sequential activation arises during task

To assess whether pairing stimuli and contexts induced neuronal interactions or synaptic modifications reflecting stimulus–context associations, we computed shift-corrected, trial-by-trial cross-correlograms (see Methods) from pairs of mere (single main factor $P < 0.001$ in repeated-measures two-way ANOVA) stimulus and MC neurons during picture presentations from different MTL regions (Extended Data Figs. 8 and 9).

Only for pairs of neurons from the entorhinal cortex and the hippocampus of the same hemisphere (Fig. 4a and Extended Data Fig. 8) did firing of MS neurons (entorhinal MSs) predict firing of MC neurons (hippocampal MCs) after approximately 40 ms, as evidenced by asymmetric shift-corrected cross-correlograms (blue, peak on the left), but not that of other neurons (hippocampal ON; black, no prominent peaks). Cross-correlograms differed significantly between groups (that is, MS–MC versus MS–ON) on short timescales (–131 to –17 ms; two-sided cluster permutation test, $P < 0.01$ for neural pairs and $P < 0.05$ for sessions; Fig. 4a and Extended Data Fig. 8), but not for pairs from different regions (Extended Data Fig. 8) or across hemispheres (Extended Data Fig. 9). This is consistent with unidirectional synaptic changes reflecting stimulus–context associations. Of six sessions from four participants with simultaneous recordings of entorhinal stimulus and hippocampal context neurons, three (50%) from two participants (50%) exhibited significant cross-correlation peaks around –40 ms (cluster permutation test, $P < 0.05$).

Next, we analysed whether asymmetric cross-correlations between entorhinal MS and hippocampal MC neurons arise from pairings of stimuli with contexts and persist post-experiment. We computed cross-correlations during time windows before (Fig. 4b, grey) and after the experiment (Fig. 4b, red; see Methods). Strong asymmetric cross-correlation peaks were absent before the experiment, but present after its termination, differing significantly at –87 to –19 ms ($P < 0.01$, firing of MS predicts that of MC). Sequential firing therefore appeared to be associated with the experiment and to persist after its termination. We then quantified mean cross-correlation peaks from –100 to –10 ms throughout recordings (Fig. 4c, left). Mean cross-correlations did not differ significantly from zero (two-sided Wilcoxon signed-rank test) before ($P = 2.889 \times 10^{-1}$, pre in grey), but did both during the first ($P = 2.036 \times 10^{-5}$, exp. in violet) and second ($P = 1.106 \times 10^{-5}$, exp. in blue) half of the experiment and after its termination ($P = 1.834 \times 10^{-5}$,

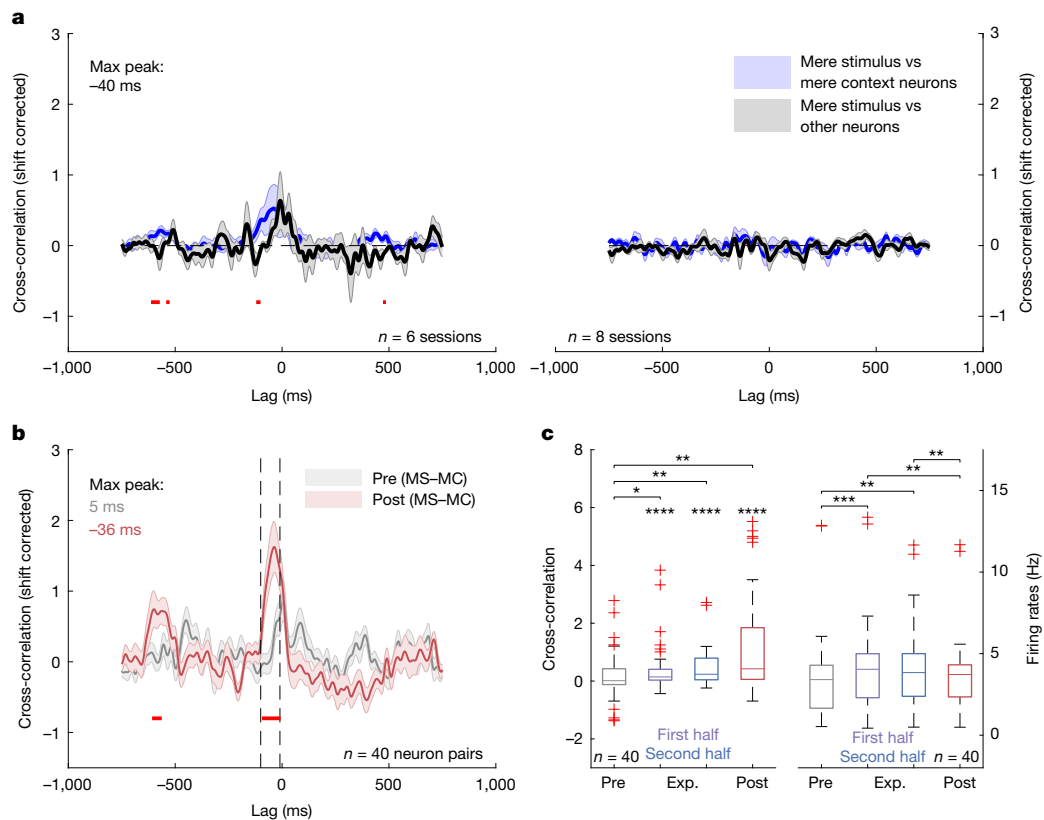


Fig. 4 | Emergence of sequential stimulus and context neuron activity suggests a storage mechanism. a, Mean cross-correlograms (six sessions) during picture presentations between entorhinal MS and either hippocampal MC neurons (each with only one significant main effect, blue) or a matched number of other hippocampal neurons (non-significant, black; left). Data are mean $\pm$ s.e.m. (solid lines and shaded areas, respectively). The peak lag time is shown in the top left. The red horizontal lines indicate significant differences ($P < 0.05$, cluster permutation test). Stimulus–neuron firing in the entorhinal cortex predicted hippocampal context–neuron, but not other neuron firing after approximately 40 ms. The right panel is the same as the left panel, but with the region order reversed. The hippocampal stimulus neuron firing did not predict entorhinal context–neuron firing. **b**, Cross-correlations of all 40 entorhinal cortex stimulus and hippocampus context neuron pairs (MS–MC) before (pre, grey) and after (post, red) the experiment. Asymmetric correlations around -40 ms

post in red). Cross-correlations in all subsequent time periods significantly exceeded pre-experiment levels ($P_{(\text{pre}, \text{exp1})} = 3.258 \times 10^{-2}$, $P_{(\text{pre}, \text{exp2})} = 6.625 \times 10^{-3}$ and $P_{(\text{pre}, \text{post})} = 3.106 \times 10^{-3}$, two-sided Wilcoxon signed-rank test). By contrast, mean firing rates (Fig. 4c, right) did not differ before and after the experiment ($P_{(\text{pre}, \text{post})} = 4.103 \times 10^{-1}$), but were significantly higher in its first and second half than before ($P_{(\text{pre}, \text{exp1})} = 1.424 \times 10^{-4}$ and $P_{(\text{pre}, \text{exp2})} = 7.183 \times 10^{-3}$) and after ($P_{(\text{post}, \text{exp1})} = 3.244 \times 10^{-3}$ and $P_{(\text{post}, \text{exp2})} = 7.782 \times 10^{-3}$).

Pre-activation predicts reactivation

We tested whether pre-activation of context neurons by preferred contexts (that is, questions) modulates their activity during subsequent picture presentations and their response likelihood following stimulus neuron activation.

First, we assessed whether normalized activity (baseline normalization; see Methods) of MC neurons during question presentations predicts activity during subsequent picture presentations (Fig. 5b and baseline control in Fig. 5c). Participant means of question ($r = 0.46$, $P = 2.07 \times 10^{-5}$), but not baseline activity ($r = 0.03$, $P = 7.79 \times 10^{-1}$) significantly ($P < 0.01$, Pearson correlation) predicted subsequent responses

were absent before but emerged during the experiment and persisted afterwards ($P < 0.01$, cluster permutation test). **c**, Boxplots (quantile 1, median, quantile 3 and whiskers (points within $\pm 1.5 \times$ the interquartile range) are shown) of mean cross-correlations (-10 to 100 ms; see dashed lines in **b**) between the same neuron pairs as in panel **b** (left). Correlations did not differ from zero before the experiment (exp.; $P = 0.29$, pre in grey; Wilcoxon signed-rank test), but exceeded zero during its first (exp., violet) and second (exp., blue) halves and after its termination (post, red; all $P < 0.001$). Correlations of all later periods exceeded pre-experiment levels ($P_{(\text{pre}, \text{exp1})} = 3.258 \times 10^{-2}$; $P_{(\text{pre}, \text{exp2})} = 6.625 \times 10^{-3}$; and $P_{(\text{pre}, \text{post})} = 3.106 \times 10^{-3}$). Firing rates did not differ significantly before and after the experiment but were significantly lower ($P < 0.01$) than during each experimental half (right). **** $P < 0.0001$, *** $P < 0.001$, ** $P < 0.01$ and * $P < 0.05$. All tests are two-sided and uncorrected.

of MC neurons to picture presentations (Fig. 5b,c). The same results were obtained for all MC neurons (Extended Data Fig. 10a,b). Similarly, question activity of MC neurons predicted subsequent picture activity significantly more strongly than baseline activity trial by trial (Extended Data Fig. 10c). Specifically, participant means of Pearson correlation effect sizes were significantly higher for correlations between question and picture activity ($r = 0.12$, left in green) than between baseline and picture activity ($r = 0.03$, right in grey; $P = 6.13 \times 10^{-3}$, two-sided Wilcoxon signed-rank test). Next, we determined whether cross-correlations between stimulus and context neurons were affected by whether a preferred context (question with the strongest context neuron response) or a preferred stimulus (picture with the strongest stimulus neuron response) was presented (permutation test; see Methods). Cross-correlations between entorhinal MS and hippocampal MC neurons were significantly stronger on short (-66 to 28 ms) timescales when a preferred versus non-preferred context of context neurons was displayed (Fig. 5d; $P < 0.01$, two-sided cluster permutations test), but not when a preferred versus non-preferred stimulus of stimulus neurons (Extended Data Fig. 10d,f) was depicted. A schematic model relating these findings to stimulus–context association storage and retrieval is depicted in Fig. 5a.

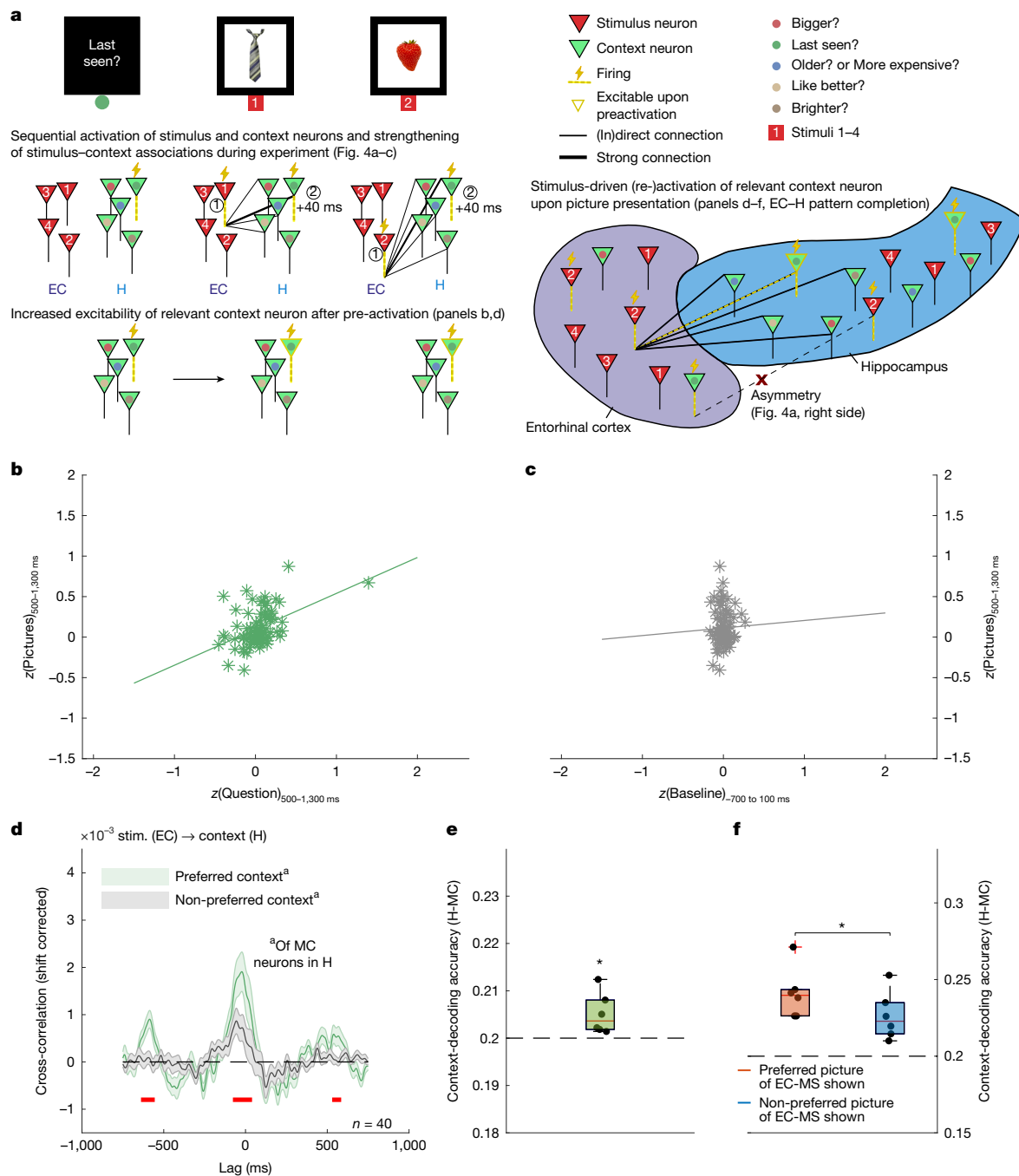


Fig. 5 | Sequential activation after pre-activation could guide stimulus-driven pattern completion. **a**, Schematic model of stimulus-context storage. Activation of entorhinal cortical (EC) MS (red) and hippocampal (H) MC neurons (green) during trial events. Specific context neurons in the hippocampus respond to particular questions (coloured circles) and are reactivated during picture presentations (top left). Picture presentations elicit responses in selected stimulus neurons (numbers), whose activity predicts context neuron firing after approximately 40 ms (Fig. 4), strengthening connections encoding stimulus-context associations. Question response strengths predict context neuron activity and excitability (yellow triangles) during subsequent picture presentations (panels b,d; Extended Data Fig. 10a–c; lower left). Pre-activation by preferred contexts guides stimulus-driven pattern completion in hippocampal context neurons (right; dashed yellow line; panels d–f). **b**, Scatter plot of mean z-values for the five contextual questions during question versus picture presentations, computed per MC neuron and patient (5 questions $\times$ 16 patients, green stars in panel b and grey stars in panel c). Pearson correlations are visualized by regression lines. Question-evoked responses during picture responses

($r = 0.46$, $P = 2.07 \times 10^{-5}$). **c**, As in panel b, but baseline versus picture presentations showing no correlation ($r = 0.03$, $P = 0.78$). **d**, Shift-corrected cross-correlations of EC-MS and H-MC during picture presentations after a preferred (maximum response, green) versus a non-preferred context (grey). The red horizontal lines indicate significant differences ($P < 0.01$, two-sided cluster permutation test). Data are mean $\pm$ s.e.m. (solid lines and shaded areas, respectively) with the dashed line indicating zero. MS firing predicts MC firing more strongly in preferred contexts. $n = 40$ neurons. **e**, Context SVM-decoding accuracy of H-MC trained with question and decoded with picture activity time locked (0–100 ms) to EC-MS firing (6 sessions: $P = 0.031$; 40 neuron pairs: $P = 5.794 \times 10^{-4}$; two-sided Wilcoxon signed-rank test). **f**, Context SVM decoding of H-MC distinguishing whether preferred (red) versus non-preferred (blue) pictures of the corresponding EC-MS were presented (6 sessions: $P = 0.0156$; 40 neuron pairs: $P = 0.0338$; one-sided Wilcoxon signed-rank test). For the boxplots (e,f), quantile 1, median, quantile 3 and whiskers (points within $\pm 1.5 \times$ the interquartile range) are shown; dashed lines indicate chance. * $P < 0.05$.

For further validation, we performed SVM-decoding analyses regarding the role of entorhinal stimulus and hippocampal context neuron interactions in stimulus-driven reinstatement of context. First, we computed session-wise context SVM-decoding accuracies of hippocampal MC neurons trained with question activity and decoded with subsequent picture activity (100–1,500 ms) time locked to entorhinal stimulus neurons firing (100-ms windows). Decoding accuracies slightly, but significantly, exceeded chance (0.21; Fig. 5e; 6 sessions: $P = 0.03$; 40 neuron pairs: $P = 5.79 \times 10^{-4}$; two-sided Wilcoxon signed-rank test). Second, we computed session-wise context SVM-decoding accuracies of hippocampal MC neurons for trials with preferred versus non-preferred pictures of the corresponding entorhinal stimulus neuron. Hippocampal MC neurons represented context more strongly when preferred (0.241) versus non-preferred (0.226) pictures of entorhinal stimulus neurons were presented (Fig. 5f; 6 sessions: $P = 0.016$; 40 neuron pairs: $P = 0.034$; one-sided Wilcoxon signed-rank test).

Discussion

Contexts constrain which memories are relevant for future decisions¹⁵. Single neurons in the MTL represent both semantic content and context, that is, aspects of an episode¹, such as tasks¹¹, episodes¹³ or rules¹². It remains unclear, however, how content and context are combined to form or retrieve integrated item-in-context memories in humans. In a context-dependent stimulus-comparison task, pairs of pictures were presented following a question that specified how the pictures were to be compared (Fig. 1a). Sixteen of seventeen participants remembered picture contents in their respective question context, showing consistent and transitive choices (Extended Data Fig. 1) closely matching objective ground truths (Extended Data Fig. 2). Analysing 3,109 neurons recorded from 16 neurosurgical patients, we found that co-activation of mostly separate populations of 597 stimulus-modulated neurons and 200 context-modulated neurons combined question and picture representations across temporal gaps. During picture presentations, when both representations became task-relevant, context neurons encoded question context, and stimulus neurons encoded picture contents (Fig. 3).

Context and content representations persisted until the trial ended, when participants selected one picture according to context (Fig. 3f), and were enhanced during behaviourally correct trials (Extended Data Fig. 7). During and after, but not before the experiment, firing of entorhinal stimulus neurons predicted hippocampal context neuron activity (Fig. 4a–c), consistent with storage of stimulus–context associations via spike-timing-dependent plasticity. Increased excitability of context neurons after pre-activation by their preferred context suggested a gating mechanism of stimulus-driven pattern completion in the hippocampus (Fig. 5a,d) that enhanced context representations (Fig. 5f). This implies that stimulus and context neurons contribute to context-dependent stimulus processing via reinstatement of context (Figs. 3e and 5e,f). Their co-activation could support memory of stimuli within their context and refine content processing according to context. The generality of the representations of both populations (Fig. 3b) appears well suited to support flexible decision-making by broadening or constraining memories through reinstatement or co-activation. Conjunctive representations of stimulus–context, conversely, reflected limited pattern separation in humans and may specify item-in-context or their attributes.

Operationalization of context

Contexts include independent elements (encoded alongside contents without altering processing) and interactive elements that influence content processing¹⁴. For example, evaluating the price of a melon in a supermarket versus its taste at home illustrates interactive context effects. Our task operationalizes interactive contexts, as questions define comparison rules that shape stimulus processing¹⁶.

Behaviourally, we successfully captured these effects as pictures were ranked differently across questions (median Kendall's $W = 0.232$) yet consistently and transitively within questions, tracking objective ground truths (mean $\rho = 0.760$; Extended Data Fig. 2), which is not expected for fixed stimulus-value judgements. Neural data reflected both interactive contexts and the contents that they acted on. First, stimulus neurons represented picture contents rather than isolated features, as evidenced by consistent encoding across contexts (Fig. 3b) and reactivation of first-picture representations during second-picture presentations for all questions (Extended Data Fig. 4), not only for feature-relevant questions. Second, context neurons encoded interactive context rather than transient attention, distinguishing questions before stimulus onset (Figs. 3d and 5b and Extended Data Fig. 10a,c), reinstating representations during picture viewing (Figs. 3e and 5e), and maintaining them until the decision (Fig. 3f). Supporting their functional relevance, both context and stimulus decoding were higher on correct trials (Extended Data Fig. 7). Although the task captures how interactive context shapes memory processing, future studies should examine these mechanisms under independent context manipulations and in naturalistic settings.

Content and context across the MTL

The MTL is thought to contain parallel streams of content and context converging in the hippocampus, enabling mutual cuing of contents and context^{17,18}. In this view, neocortical context representations are processed via the parahippocampal cortex or medial entorhinal cortex, and content representations via the lateral entorhinal cortex or perirhinal cortex before convergence in the hippocampus¹⁷. Approximately 6% of MTL neurons strongly encoded context during picture presentations. Consistent with the literature, context representations were more frequent in the parahippocampal cortex, entorhinal cortex and hippocampus than in the amygdala, and were highest among hippocampal stimulus neurons (21.26%). Although the large majority of stimulus neurons (88%) were invariant to context, and most context neurons (63.5%) were invariant to stimulus, a small but significant fraction of stimulus neurons represented either both or specific conjunctions of context and stimulus, particularly in the hippocampus (Fig. 2b–d). Conjunctive representations were most abundant in the hippocampus (2.29%), particularly among stimulus neurons (10.34%). Consistent with this, stimulus–context neurons were more prevalent in sessions with higher proportions of both stimulus and context neurons in the hippocampus and amygdala, but only with higher context neuron proportions in the parahippocampal and entorhinal cortices (Extended Data Fig. 3). These regional patterns support the convergence of parallel content and context streams in the hippocampus, where they may create conjunctive representations. Nevertheless, distinct representations remained most prevalent. The context invariance of most stimulus neurons (Fig. 2a,e) is notable given potential attentional fluctuations across questions. Although stimulus neurons were overrepresented due to pre-selection of response-eliciting pictures, this did not bias (but rather supported) context invariance, as stimulus representations persisted from pre-screenings to main experiments (Extended Data Fig. 11).

In our paradigm, questions provided contextual constraints for relevant information (for example, value information from the amygdala¹⁹ for Like better?) about stimuli that were to be related to each other and to serial picture positions. The MTL, and particularly the hippocampus, has been associated with such relational processing. Lesions cause relational memory deficits, in both long-term²⁰ and working^{21,22} memory, impair the ability to imagine details in hypothetical contexts^{23–25} and disrupt episodic memory²⁶. Through pattern completion, inputs from the entorhinal cortex (representing content) are thought to be complemented in the hippocampal formation (thereby retrieving associated context)²⁷. Supporting this view, during and after, but not before experimental pairing of stimuli and context, firing of

entorhinal stimulus neurons predicted hippocampal context neuron activity after tens of milliseconds. Picture presentations served as cues for the recall of both content (from the current or the preceding picture) and context (the relevant question). Correspondingly, representations of current visual stimuli were followed by reinstatement of previous stimuli or contexts. Synaptic modifications of direct or indirect connections offer one potential explanation and could be tested by pairing only subsets of contents and contexts, which should produce cross-correlation changes only in corresponding neuron pairs. Furthermore, electrostimulation of hippocampal context neurons during context reactivation peaks following stimulus representations (Fig. 3d) should disrupt cross-correlations, reinstatement patterns and behaviour. In addition, increased excitability of context neurons suggests a gating mechanism of stimulus-driven pattern completion. After repeated pairings of contents and contexts strengthen connections between entorhinal stimulus neurons and multiple hippocampal context neurons, stimulus neuron firing could preferentially activate context neurons pre-activated by their preferred question. This could be tested by modulating context neuron excitability during question presentations via electrostimulation or attention. However, as the exact placement of microwires within a target region cannot be determined reliably with the noninvasive imaging methods available in humans, it remains uncertain whether hippocampal microwires were recorded from the CA1/CA3 region, the dentate gyrus or the subiculum.

Overall, experiment-induced neural interactions between distinct neuronal populations may support memory of stimuli in context. Their resulting co-activations could guide contextual memory processing of stimuli, for example, via indexing of brain areas that contain context-specific stimulus information.

Contextual memory search

Neural representations of content^{28–30} and context^{31–34} are reinstated during memory search and potentially serve as memory indices. In humans, stimulus neurons in the MTL encode semantic content^{29,35,36} and are reactivated during recall of their preferred stimulus, that is, concept^{28–30}. Other neurons encode contextual features such as rules¹², or episodic boundaries¹³ or task instructions¹¹. Moreover, ensemble activity during recognition memory reinstates contextual neural states¹³. However, it remains unclear whether distinct neuronal populations simultaneously reinstate content and context during context-dependent relational memory processing. Context neurons encoded question context throughout trials, with decoding peaking at question onset and during late picture phases when content needed to be processed according to context. Context neuron representations also generalized across stimulus identities and serial picture positions, with decoding accuracies exceeding chance even when SVM decoders were trained on activity during specific pictures or picture positions and tested on the remaining ones. In addition, response patterns to question context were reinstated during the late phase of picture presentations, particularly in the entorhinal cortex and hippocampus. Overall, the simultaneous (re-)instatement of both context and content at retrieval-related latencies³⁷, together with the generalization of their representations, suggests that they contribute to memory search.

Pattern separation and human cognition

The balance between memory specificity and generalization is fundamental to adaptive cognition. Besides declarative memories, the MTL is essential for decision-making^{15,38} and abstract human cognition such as generalization⁷, deliberation³⁸ or inference³⁹. In rodents, hippocampal neural representations often indicate conjunctions of stimuli and are context dependent^{3,4}. This ‘pattern separation’, that is, the creation of distinct neural codes for similar inputs, renders memory representations distinguishable. By contrast, in humans, selective concept neurons reliably encode stimulus content across contexts^{11,40} and

may constitute building blocks of memory⁴¹. Their context invariance and lack of pattern separation⁵ (but see ref. 9) may enable knowledge generalization across contexts (at the expense of ‘semantic deviations’ leading to false memories)³⁶.

Abstract content representations³⁶ appear well suited for a broad generalization across contexts. However, contexts also constrain which memories are relevant^{42,43}. Recalling items from studied word lists enhances the recall of items in temporal or semantic proximity⁴³, an effect that correlates with memory performance and general cognitive ability⁴⁴. Moreover, items are more readily recalled when sharing encoding contexts, such as task instructions or video clips¹³.

Our MTL neurons showed a low but detectable degree of pattern separation (stimulus–context interaction neurons in the hippocampus), supporting both generalization and differentiation of content and context. Consistent with previous reports^{11,40}, stimulus neuron activity reliably distinguished picture identity irrespective of context, and reflected picture-specific memories, as evidenced by consistent first-picture reactivations during second-picture presentations across all contexts (Extended Data Fig. 4). In addition, we identified neurons that represented context independent of picture identity or position. The invariance of both populations allows us to generalize memories, whereas their co-activation could differentiate otherwise overlapping memories. Although co-activation may narrow memory searches, reinstatement of associated contexts via pattern completion could render additional memories associated with this context more accessible and thereby promote generalization¹. A combinatorial indexing scheme of this kind would support flexible decision-making and potentially reduces the need for explicit pattern-separated, conjunctive representations.

Online content

Any methods, additional references, Nature Portfolio 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/s41586-025-09910-2>.

- Yonelinas, A. P., Ranganath, C., Ekstrom, A. D. & Wiltgen, B. J. A contextual binding theory of episodic memory: systems consolidation reconsidered. *Nat. Rev. Neurosci.* **20**, 364–375 (2019).
- Eichenbaum, H., Yonelinas, A. P. & Ranganath, C. The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* **30**, 123–152 (2007).
- O’Keefe, J. & Krupic, J. Do hippocampal pyramidal cells respond to nonspatial stimuli? *Physiol. Rev.* **101**, 1427–1456 (2021).
- Kubie, J. L., Levy, E. R. J. & Fenton, A. A. Is hippocampal remapping the physiological basis for context? *Hippocampus* <https://doi.org/10.1002/hipo.23160> (2019).
- Quiroga, R. Q. No pattern separation in the human hippocampus. *Trends Cogn. Sci.* **24**, 994–1007 (2020).
- Zeithamova, D. & Bowman, C. R. Generalization and the hippocampus: more than one story? *Neurobiol. Learn. Mem.* **175**, 107317 (2020).
- Kumaran, D., Summerfield, J. J., Hassabis, D. & Maguire, E. A. Tracking the emergence of conceptual knowledge during human decision making. *Neuron* **63**, 889–901 (2009).
- Quiroga, R. Q. Concept cells: the building blocks of declarative memory functions. *Nat. Rev. Neurosci.* **13**, 587–597 (2012).
- Suthana, N., Ekstrom, A. D., Yassa, M. A. & Stark, C. Pattern separation in the human hippocampus: response to Quiroga. *Trends Cogn. Sci.* **25**, 423–424 (2021).
- Polyn, S. M. & Kahana, M. J. Memory search and the neural representation of context. *Trends Cogn. Sci.* **12**, 24–30 (2008).
- Minxha, J., Adolphs, R., Fusi, S., Mamelak, A. N. & Rutishauser, U. Flexible recruitment of memory-based choice representations by the human medial frontal cortex. *Science* **368**, eaba3313 (2020).
- Kutter, E. F., Boström, J., Elger, C. E., Nieder, A. & Mormann, F. Neuronal codes for arithmetic rule processing in the human brain. *Curr. Biol.* **32**, 1275–1284.e4 (2022).
- Zheng, J. et al. Neurons detect cognitive boundaries to structure episodic memories in humans. *Nat. Neurosci.* **25**, 358–368 (2022).
- Baddeley, A. D. Domains of recollection. *Psychol. Rev.* **89**, 708–729 (1982).
- Bornstein, A. M. & Norman, K. A. Reinstated episodic context guides sampling-based decisions for reward. *Nat. Neurosci.* **20**, 997–1003 (2017).
- Polyn, S. M., Norman, K. A. & Kahana, M. J. Task context and organization in free recall. *Neuropsychologia* **47**, 2158–2163 (2009).
- Eichenbaum, H., Savage, M., Fortin, N., Komorowski, R. & Lipton, P. Towards a functional organization of episodic memory in the medial temporal lobe. *Neurosci. Biobehav. Rev.* **36**, 1597–1608 (2012).

18. Davachi, L. Item, context and relational episodic encoding in humans. *Curr. Opin. Neurobiol.* **16**, 693–700 (2006).
19. Mormann, F., Bausch, M., Knieling, S. & Fried, I. Neurons in the human left amygdala automatically encode subjective value irrespective of task. *Cereb. Cortex* **29**, 265–272 (2019).
20. Squire, L. R. Memory for relations in the short term and the long term after medial temporal lobe damage. *Hippocampus* **27**, 608–612 (2017).
21. Konkel, A. & Cohen, N. J. Relational memory and the hippocampus: representations and methods. *Front. Neurosci.* **3**, 166–174 (2009).
22. Hannula, D. E., Tranel, D. & Cohen, N. J. The long and the short of it: relational memory impairments in amnesia, even at short lags. *J. Neurosci.* **26**, 8352–8359 (2006).
23. Hassabis, D., Kumaran, D., Vann, S. D. & Maguire, E. A. Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl Acad. Sci. USA* **104**, 1726–1731 (2007).
24. Rosenbaum, R. S., Gilboa, A., Levine, B., Winocur, G. & Moscovitch, M. Amnesia as an impairment of detail generation and binding: evidence from personal, fictional, and semantic narratives in K.C. *Neuropsychologia* **47**, 2181–2187 (2009).
25. Maguire, E. A. & Hassabis, D. Role of the hippocampus in imagination and future thinking. *Proc. Natl Acad. Sci. USA* **108**, E39–E39 (2011).
26. Dickerson, B. C. & Eichenbaum, H. The episodic memory system: neurocircuitry and disorders. *Neuropsychopharmacology* **35**, 86–104 (2010).
27. Teyler, T. J. & Rudy, J. W. The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus* **17**, 1158–1169 (2007).
28. Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R. & Fried, I. Internally generated reactivation of single neurons in human hippocampus during free recall. *Science* **322**, 96–101 (2008).
29. Bausch, M. et al. Concept neurons in the human medial temporal lobe flexibly represent abstract relations between concepts. *Nat. Commun.* **12**, 6164 (2021).
30. Ison, M. J., Quiroga, R. Q. & Fried, I. Rapid encoding of new memories by individual neurons in the human brain. *Neuron* **87**, 220–230 (2015).
31. Kragel, J. E. et al. Distinct cortical systems reinstate the content and context of episodic memories. *Nat. Commun.* **12**, 4444 (2021).
32. 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).
33. Staresina, B. P. et al. Recollection in the human hippocampal–entorhinal cell circuitry. *Nat. Commun.* **10**, 1503 (2019).
34. Folkerts, S., Rutishauser, U. & Howard, M. W. Human episodic memory retrieval is accompanied by a neural contiguity effect. *J. Neurosci.* **38**, 4200–4211 (2018).
35. Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107 (2005).
36. Reber, T. P. et al. Representation of abstract semantic knowledge in populations of human single neurons in the medial temporal lobe. *PLoS Biol.* **17**, e3000290 (2019).
37. Staresina, B. P. & Wimber, M. A neural chronometry of memory recall. *Trends Cogn. Sci.* **23**, 1071–1085 (2019).
38. Shadlen, M. N. & Shohamy, D. Decision making and sequential sampling from memory. *Neuron* **90**, 927–939 (2016).
39. Shohamy, D. & Daw, N. D. Integrating memories to guide decisions. *Curr. Opin. Behav. Sci.* **5**, 85–90 (2015).
40. Rey, H. G. et al. Lack of context modulation in human single neuron responses in the medial temporal lobe. *Cell Rep.* **44**, 115218 (2025).
41. Mackay, S. et al. Concept and location neurons in the human brain provide the ‘what’ and ‘where’ in memory formation. *Nat. Commun.* **15**, 7926 (2024).
42. Healey, M. K. & Uitylgt, M. G. The role of control processes in temporal and semantic contiguity. *Mem. Cognit.* **47**, 719–737 (2019).
43. Healey, M. K., Long, N. M. & Kahana, M. J. Contiguity in episodic memory. *Psychon. Bull. Rev.* **26**, 699–720 (2019).
44. Healey, M. K., Crutchley, P. & Kahana, M. J. Individual differences in memory search and their relation to intelligence. *J. Exp. Psychol. Gen.* **143**, 1553–1569 (2014).

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



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

© The Author(s) 2026

Methods

Patients and recording

All studies were approved by the Medical Institutional Review Board at the University of Bonn (accession number 095/10 for single-unit recordings in humans in general and 248/11 for the current study in particular). Each patient gave informed written consent both for the implantation of microwires and for participating in the experiment. We recorded from 17 patients with pharmacologically intractable epilepsy (9 male patients; 22–65 years of age), implanted with intracranial electrodes to localize the seizure onset zone for surgical resection⁴⁵. Each depth electrode contained a microwire bundle consisting of nine microwires, consisting of a reference electrode with low impedance and eight high-impedance recording electrodes (AdTech), which protruded from the tip of the electrode by approximately 4 mm. All bundles were localized using a post-implantation CT scan co-registered with a pre-implantation MRI scan normalized to Montreal Neurological Institute space. The differential signal from the microwires was amplified using a Neuralynx ATLAS system, filtered between 0.1 Hz and 9,000 Hz, and sampled at 32 kHz. These recordings were stored digitally for further analysis. Our original dataset consisted of 50 experimental sessions from 17 patients with recordings from the amygdala, parahippocampal cortex, entorhinal cortex and hippocampus. One session from one patient was excluded due to low behavioural performance, resulting in a final sample of 49 sessions from 16 patients. The excluded session stemmed from the only patient for which no context neurons could be detected (others exhibited an average of 12.5). For most patients, ten microwire bundles (median of 80 channels) were placed with an average of 69.34 recorded channels (range of 32–80) and a neuron yield of 0.92 per channel. Each experimental session was recorded after a screening for visually selective responses in the morning of the same day. The mean ($\pm$ s.d.) time interval between screenings and subsequent experiments was 4.71 ± 2.02 h (range of 1.40–9.15 h). In 13 of the 16 patients, more than one session was recorded. The mean ($\pm$ s.d.) time interval between multiple sessions was 46.94 ± 24.19 h (range of 16.57–118.06 h). Neural dependence across 33 consecutive pairs of these sessions was assessed using intraclass correlation coefficients (ICCs(2,1))^{46,47}. ICCs were computed across channels (that is, microwires), separately for context and stimulus neurons. ICCs quantify cross-session stability, with lower values indicating weaker dependence. Observed ICCs were compared with null distributions obtained by permuting channels. ICCs did not strongly exceed these null distributions (Hedges' g : stimulus = 0.3415; context = 0.1568) and only reached significance ($\alpha = 0.01$) in 7 of 33 consecutive stimulus neuron session pairs, 3 of 31 consecutive context neuron session pairs, and in no session pair for both, all with low overall effect sizes (Extended Data Fig. 12). We used the spike-sorting software Combinato⁴⁸ with default parameters for the exclusion of noisy recording channels, artefact removal, spike detection and spike sorting. Afterwards, we manually removed remaining artefacts, merged potentially over-clustered units from the same channel and distinguished single from multi-units using the graphical user interface in Combinato. Highly similar spike shapes, inter-spike interval distributions, neural responses to visual stimuli, asymmetric cross-correlations and the absence of neural activity during refractory periods guided this procedure, which predated all further analyses.

Experimental design

The paradigm was performed on a laptop computer with the psychtoolbox3 (www.psychtoolbox.org) and octave (www.octave.org/) running on a Debian 8 operating system (www.debian.org) for stimulus delivery. Before the experiment, approximately 100 pictures of people, animals, scenes and objects were presented on a laptop screen in pseudo-random order (presentation for 1 s; 6 or 10 trials). Then, after automatic spike extraction and sorting with Combinato, neural responses to these pictures were evaluated based on raster plots

and histograms. The aim of this procedure was to identify a subset of four pictures for the following experiment while maximizing the number of neurons that were expected to respond selectively to only one of the pictures. These four pictures were then presented in the experiment. Each self-triggered trial contained one out of five questions, a sequence of two of four pictures with jittered onsets, and an answer prompt displaying 1 or 2?. Patients indicated the sequential position of the picture that best answered the question by pressing key 1 or 2. We resolved ambiguous meanings of each depicted picture before the experiment and elaborated on the meaning of each question in a short test run of the paradigm. The questions were Bigger? (volume), Last seen in real life?, More expensive? or Older? (if the pictures set included a person), Like better? and Brighter?. Patients were instructed to try to stick to one answer for a given picture pair, but to keep mentally computing the answer to each question in the course of the experiment. In total, the experiment consisted of 300 trials in which each of the 5 questions and all 12 possible ordered picture pairs out of 4 pictures were presented equally often and in an unpredictable pseudo-random order. This resulted in 60 trials per question, 25 trials per picture pair and 5 trials per specific combination of question and picture pair.

Definition of neural populations

For each unit, repeated two-way repeated-measures ANOVAs with factors stimulus (picture), context (question) and interactions were computed from the activity during picture presentations (100–1,000 ms) at an α -level of 0.001. Units with (at least) a significant main effect of stimulus or context are referred to as stimulus or context neurons, respectively. If a unit only (at most) showed a significant main effect of either stimulus or context, it is termed a MS or a MC neuron, and if both main effects were significant, it is called a contextual stimulus neuron. Conversely, stimulus–context interaction neurons were defined by the presence of a significant interaction. We determined whether the observed number of neurons pertaining to each of these populations or all effect sizes exceeded chance in the following manner. First, the probability of obtaining observed numbers of significant neurons was assessed for each population separately by binomial tests with an expected false-positive probability of 0.001 corresponding to the α -level above. Second, to account for potential statistical dependencies between factors or neurons, repeated-measures ANOVA effect sizes of all neurons were compared with stratified label-shuffling controls, irrespective of statistical significance (Fig. 2e, see insets). Specifically, null distributions of partial η^2 were obtained either by permuting picture labels for each question separately (factor stimulus) or question labels for each picture separately (factor context or stimulus–context interaction). Afterwards, effect sizes of data and controls were averaged for each patient (see Fig. 2e in lavender and grey) and their distributions compared (two-sided Wilcoxon signed-rank test, Bonferroni corrected). To additionally account for patient-level variability, we fit separate linear mixed-effects models for each ANOVA factor across all recorded neurons. Each model included a fixed intercept (capturing grand population effect size differences) and patient-specific random intercepts. Neural populations remained largely consistent, regardless of whether they were identified via label-shuffled ANOVA effect sizes or analytical ANOVA P values.

To assess the stability of stimulus representations across pre-screenings and main experiments, SVM decoders were trained on main experiment activity and tested either on pre-screening (across) or main experiment activity (within, leave-out cross-validation with matched trial counts). For 45 recording sessions, all four pictures of the main experiment were part of the pre-screening before the experiment. Picture-decoding accuracies highly exceeded chance, particularly for microwire channels that did versus did not contain stimulus neurons (Extended Data Fig. 11). There were no significant differences in decoding accuracy when testing on pre-screening data (across) or on

main experiment activity (within), both for stimulus neuron channels and for non-stimulus neuron channels (two-sided Wilcoxon signed-rank test). Finally, Spearman correlations were computed between stimulus–context neuron proportions and stimulus or context neuron proportions across all session-site combinations ($n = 168$; Extended Data Fig. 3). Site-specific correlations were assessed using permutation tests (1,000 permutations of stimulus–context proportions, one-sided test for positive associations).

SVM population decoding

Linear SVM-decoding accuracies of context were computed from session-wise population activity of context neurons (Fig. 3a,c,d,f; see Fig. 3a for decoding from all neurons) and averaged for each patient, except in Fig. 3b in which a pooled decoding scheme was used (30 random subsamples of 151 stimulus or context neurons, corresponding to 75% of context neurons, were drawn to estimate variance). Figure 3a depicts these accuracies for each context separately. In general, decoders were implemented with functions from the LIBSVM library (v3.24) using standard parameters (unless stated otherwise) in custom scripts written in MATLAB R2021a or Python (scikit-learn 0.24.1), and a fivefold cross-validation scheme without overlap between training and testing data was used with five repetitions. Generalization of context decoding across picture identities (Fig. 3b, left in red), of stimulus decoding across contexts (Fig. 3b, right in green) or of either decoding across serial picture positions (in blue) was assessed. This was achieved by training with activity in response to each picture, each picture–context or each picture–position and decoding during the respective remaining ones. Training was performed either during picture presentations (Fig. 3a,b; 100–1,000 ms), question presentations (Fig. 3e) or during the course of the experiment (Fig. 3c,d,f; see below). Thus, obtained decoding accuracies were statistically compared with chance (one-fifth for context and one-quarter for visual stimuli; two-sided Wilcoxon signed-rank test in Fig. 3a,e; two-sided cluster permutation tests $P < 0.01$ in Fig. 3d,f) or to each other (Mann–Whitney U -test in Fig. 3a; and two-sided cluster permutation tests in Fig. 3d).

Session-wise context-decoding accuracies computed from binned population activity (400 ms, stride = 100 ms) at differing training and decoding times throughout the trial of all context neurons were visualized as a heat map (Fig. 3c). Context-decoding accuracies from identical decoding and training times (diagonals) were then plotted with standard errors for all 16 patients with context neurons (Fig. 3d, top in green) and compared with a label-shuffling control condition (Fig. 3d, top red). The latter was obtained by randomly permuting question context labels for each session before training and decoding. Time periods of significant differences of decoding accuracies from real data and controls were evaluated with a cluster permutation test (Fig. 3d, solid line below). Specifically, paired-samples Student's t -tests (two-sided) were computed for each time bin ($\alpha = 0.01$) to determine clusters of contiguous decoding accuracy time bin differences. Sums of t -values within each cluster were assessed relative to the null distribution of t -value sums obtained from 1,000 permutations of data and control labels (or chance levels, see below). If t -value sums fell into the top percentile of the permutation distribution, differences during the time period of the respective cluster were considered significant. Decoding accuracies of either first (Fig. 3d, bottom in grey) or second pictures (Fig. 3d, bottom in lavender) obtained by patient-wise training and decoding with the activity of stimulus neurons in the course of the experiment were plotted and compared with chance (one-quarter) in an analogous manner (Fig. 3d, bottom). In addition, we examined whether neural activity patterns of stimulus neurons reflected content rather than feature representations (Extended Data Fig. 4). Session-wise linear SVMs were trained on stimulus neuron activity (400–800 ms after first-picture onset) to decode first picture identities from neural activity throughout second picture presentations (400-ms bins, stride = 100 ms) for each of the five question contexts. Statistical significance across sessions

was assessed using cluster-based, two-sided permutation tests against chance ($\alpha = 0.05$, chance = one-quarter).

Afterwards, decoding performances of context were further analysed for each brain region separately (Extended Data Fig. 6). Furthermore, we determined whether activity patterns associated with the presentation of question instructions were recapitulated during picture presentations (Fig. 3e). For this purpose, patient-wise SVM decoders were trained with regional context neuron activity either during question presentations (600–1,000 ms) or preceding baseline intervals (–500 to 100 ms), whereas contexts were decoded in the late phases of first (red; 600–1,000 ms) or second (blue; 600–1,000 ms) picture presentations. Specifically, a random subset of neurons ($n = 400$) was repeatedly drawn ($n = 30$) from each brain region. Afterwards, decoding accuracies were averaged for each patient if context neurons were part of the drawn subsets before averaging across repetitions. The deviation of the resulting patient-wise distributions of decoding accuracies from chance (one-fifth) was assessed by means of two-sided Wilcoxon signed-rank tests (uncorrected).

Finally, we assessed decoding accuracies of stimulus and context neurons preceding the end of the trial when one picture was chosen over the other according to context. Again, patient-wise SVM-decoding accuracies of contents and context were computed from the activity of stimulus and context neurons, respectively, and compared with chance (Fig. 3f; two-sided cluster permutation tests $P < 0.01$).

Interaction analyses between populations

Session-wise cross-correlations were computed for pairs of MS–MC pairs (each with only one main effect $P < 0.001$ in repeated-measures two-way ANOVA) for different brain regions on a trial-by-trial basis during picture presentations (Fig. 4a and Extended Data Figs. 8 and 9 with neuron pairs; 0–1,500 ms after picture onsets) or during different contiguous time windows (Fig. 4b,c with neuron pairs; before, after or in the first and second half of the experiment). Each cross-correlogram was normalized according to its window length and the mean firing rate of its neural pair, and then convolved with a Gaussian kernel (sigma = 10 ms). To capture stimulus-independent interactions, normalized cross-correlations from non-simultaneous activity of the same pairs (shift predictors) were subtracted from the former. Shift predictors were obtained either by randomly permuting activity patterns of one of the neurons across relevant trials (Figs. 4a and 5d,e and Extended Data Figs. 8 and 9) or by randomly shifting spike trains in time (up to 10,000 ms). For the evaluation of effect sizes, analogous shift-corrected cross-correlations (noise correlations) were computed for a matched number of MS and other neurons (MS–ON pairs) defined by their lack of stimulus and/or context modulation ($P_{\text{stimulus}} > 0.01$, $P_{\text{context}} > 0.01$ and $P_{\text{stimulus-context}} > 0.01$ in repeated-measures two-way ANOVA). For three sessions, in which matching failed due to a sparsity of other neurons, randomly chosen other neurons from the same brain region were used instead. Cross-correlograms between MS and MC neurons or their session averages were then compared with these controls by means of two-sided cluster permutations tests ($P < 0.01$; Fig. 4a,b and Extended Data Figs. 8 and 9). For analyses in Fig. 4b,c, the largest possible symmetric time windows immediately before and after the experiment were used. The duration of these time windows (range of 16.57–106.05 s, mean of 48.79 s) either amounted to the time between the first recorded spike and the beginning of the experiment, or between the end of the experiment and the last recorded spike for each neuron pair (entorhinal MS–hippocampal MC pairs) depending on which was smallest. Afterwards, mean cross-correlations at small negative lags (MS followed by MC firing; averaged between –10 to –100 ms) as well as all mean firing rates of each neuron pair were compared against zero or against each other in different phases of the experiment (two-sided Wilcoxon signed-rank tests; Fig. 4c). In addition, cross-correlations during picture presentations (0–1,500 ms) were re-computed depending on whether preferred contexts of hippocampal MC neurons or preferred stimuli

of entorhinal MS neurons were presented ($P < 0.01$, two-sided cluster permutations tests; Fig. 5d and Extended Data Fig. 10d–f). Preferred versus non-preferred contexts or stimuli were defined as the question or stimulus that elicited the strongest neural response in each context or stimulus neuron during picture presentations (0–1,000 ms) versus a randomly drawn different question or stimulus. As a further validation of this model, two additional SVM interaction-decoding analyses were performed. First, decoders were trained with hippocampal MC neuron activity per session during question presentations and decoded with their activity during subsequent picture presentations (100–1,500 ms) time locked to entorhinal MS neurons firing (in 100-ms windows after each entorhinal MS action potential). Second, decoders were trained with hippocampal MC neuron activity per session for trials during which the preferred versus non-preferred pictures of the corresponding entorhinal stimulus neuron were presented.

Next, Pearson correlations between (patient means of) the mean question-wise activity of MC neurons during question presentations (500–1,300 ms; Fig. 5b and Extended Data Fig. 10a) or its baseline (–700 to 100 ms; Fig. 5c and Extended Data Fig. 10b) versus subsequent picture presentations (500–1,300 ms) were computed ($P < 0.01$), normalized to baseline and visualized as scatter plots with regression lines (Fig. 5b,c and Extended Data Fig. 10a,b). Finally, Pearson correlation effect sizes were computed for each MC neuron from activity between question and subsequent picture presentations or between baseline and subsequent picture presentations across all trial conditions (60 combinations of 5 questions and 12 picture pairs) and averaged per patient. Afterwards, effect sizes of both conditions were compared and visualized as boxplots (two-sided Wilcoxon signed-rank test; Extended Data Fig. 10c).

Behavioural performance

Memory of contents (that is, pictures) in context (that is, after a particular question) was evaluated by determining the consistency and transitivity of the answers that were provided at the end of each trial. Consistency (γ) captures the degree to which preference for one picture over another was maintained when the presentation order was reversed. A picture was considered to be preferred if it was chosen in at least three out of the five trials that it was paired with a specific other picture in a given temporal order and context. If this picture was preferred both when it was shown first or second (chosen at least 3 out of 5 times in each order), all 10 trials were assigned a consistency value of 1 instead of 0. The overall consistency of a session was computed by averaging the consistency values of zeros and ones across all 300 trials. For random choices, a consistency of 0.5 is expected (Extended Data Fig. 1a).

Transitivity (τ) captures the logical consistency of preferences across three pictures (x, y, z) when considered in pairs. For each of the four unordered picture triples, we computed preference probabilities $P(x > y)$, $P(y > z)$ and $P(x > z)$ for all three unordered picture pairs based on how many times out of the 10 trials one picture was chosen over the other in a given context. A triple was considered transitive if preference probabilities satisfied the condition of weak stochastic transitivity: $P(x > y) \geq 0.5$ and $P(y > z) \geq 0.5$, then $P(x > z) \geq 0.5$. This condition was tested for all permutations of x, y and z and a transitivity value of 1 instead of 0 was assigned to the triple if the condition was satisfied for any permutation. For random choices, a transitivity of 0.75 is expected (Extended Data Fig. 1a). Finally, context dependence of performance was assessed by deriving an overall behavioural performance measure from mean of consistency (γ) and transitivity (τ) values per question of each session and comparing them across contexts (Kruskal–Wallis test, $\alpha = 0.05$; Extended Data Fig. 1b). Moreover, for questions with objective ground truths (Bigger?, and More expensive? or Older?), we validated behavioural performance by computing Spearman correlations between

behaviourally determined picture ranks and objectively determined picture ranks, averaged across available questions per session. Session-wise correlations were tested against zero using a Wilcoxon signed-rank test, and the observed mean across sessions was compared with a null distribution generated from 1,000 permutations of picture ranks within each session–question pair (Extended Data Fig. 2). In addition, decoding accuracies were related to behavioural performance by comparing correct trials (in which behavioural choices matched ground-truth picture rankings) versus incorrect trials (in which choices deviated from ground truth). Session-wise linear SVMs (fivefold cross-validation, five repeats) were trained on context, stimulus or stimulus–context neuron activity during picture presentations (100–1,000 ms) to decode questions, pictures and picture–question combinations, respectively. Decoding accuracies for correct versus incorrect trials were compared using one-sided Wilcoxon signed-rank tests (correct > incorrect; $\alpha = 0.05$). Chance levels were one-quarter for stimulus, one-fifth for context and one-twentieth for stimulus–context decoding (Extended Data Fig. 7).

Statistics and reproducibility

Forty-nine sessions from 16 participants were analysed (one session of one participant was excluded based on a low behavioural performance). A total of 200 context neurons were detected in 38 sessions from the 16 participants. Twenty-five sessions contained at least 3 stimulus and context neurons each, 12 sessions more than 5.

Reporting summary

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

Data availability

All data supporting the findings of this study are available on GitHub (<https://github.com/mabausch/ContentContextNeurons.git>).

Code availability

All code related to the analyses of the article is available on GitHub (<https://github.com/mabausch/ContentContextNeurons.git>).

45. Fried, I., MacDonald, K. A. & Wilson, C. L. Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* **18**, 753–765 (1997).
46. Shrout, P. E. & Fleiss, J. L. Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* **86**, 420–428 (1979).
47. McGraw, K. O. & Wong, S. P. Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* **1**, 30–46 (1996).
48. Niediek, J., Boström, J., Elger, C. E. & Mormann, F. Reliable analysis of single-unit recordings from the human brain under noisy conditions: tracking neurons over hours. *PLoS ONE* **11**, e0166598 (2016).

Acknowledgements We thank all patients for their participation. This research was supported by the Volkswagen Foundation, the German Research Council (DFG MO 930/4-2, MO 930/15-1, SPP 2411 and SFB 1089) and a NRW Network Grant (iBehave).

Author contributions M.B., J.N. and F.M. designed the research. F.M. and C.E.E. recruited and managed patients. J.B. and F.M. performed the surgeries. M.B., T.P.R., S.M. and J.N. collected the data. M.B. analysed the data. F.M. supervised the project. M.B. and F.M. wrote the draft of the manuscript. All authors edited the manuscript.

Competing interests The authors declare no competing interests.

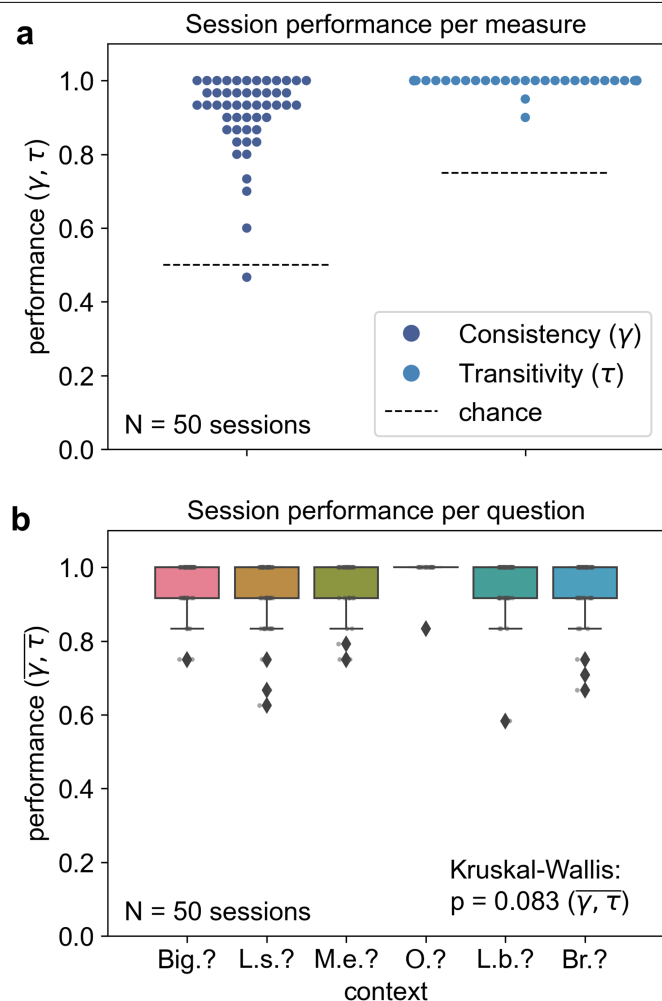
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-025-09910-2>.

Correspondence and requests for materials should be addressed to Marcel Bausch or Florian Mormann.

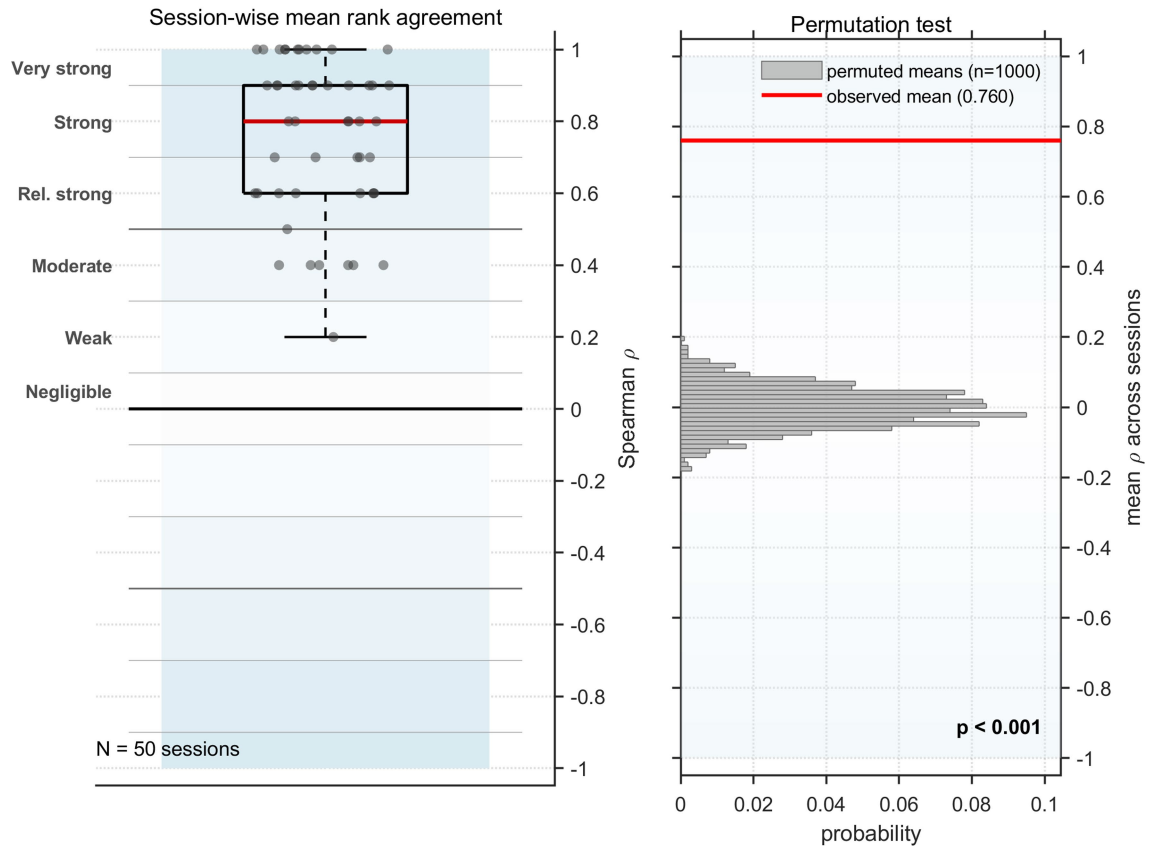
Peer review information Nature thanks Sameer Sheth and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Consistency and transitivity of answers strongly exceeded chance and approached ceiling levels. Pairwise comparisons of contents were both consistent and transitive (**a**) across different contexts (**b**) despite unique and distinct orders among pictures, indicating memory of content-in-context. Consistency measures how well context-specific preferences for one picture over another are maintained across presentation orders, while transitivity checks if a preference for A over B and B over C also results in a preference for A over C. **a.** Swarmplots depict consistency scores γ (dark blue, left) and transitivity scores τ (light blue, right) for each participant's answers in each recording session. Dotted lines represent chance performance. Only one recording session did not exceed chance levels and was excluded from further analyses. **b.** Boxplots (Q1, median, Q3; whiskers: points within ± 1.5 IQR) of mean session-wise behavioral performances (averages of consistency and transitivity) across contexts. Performance did not differ significantly between contexts ($p = 0.083$; Kruskal-Wallis-Test).

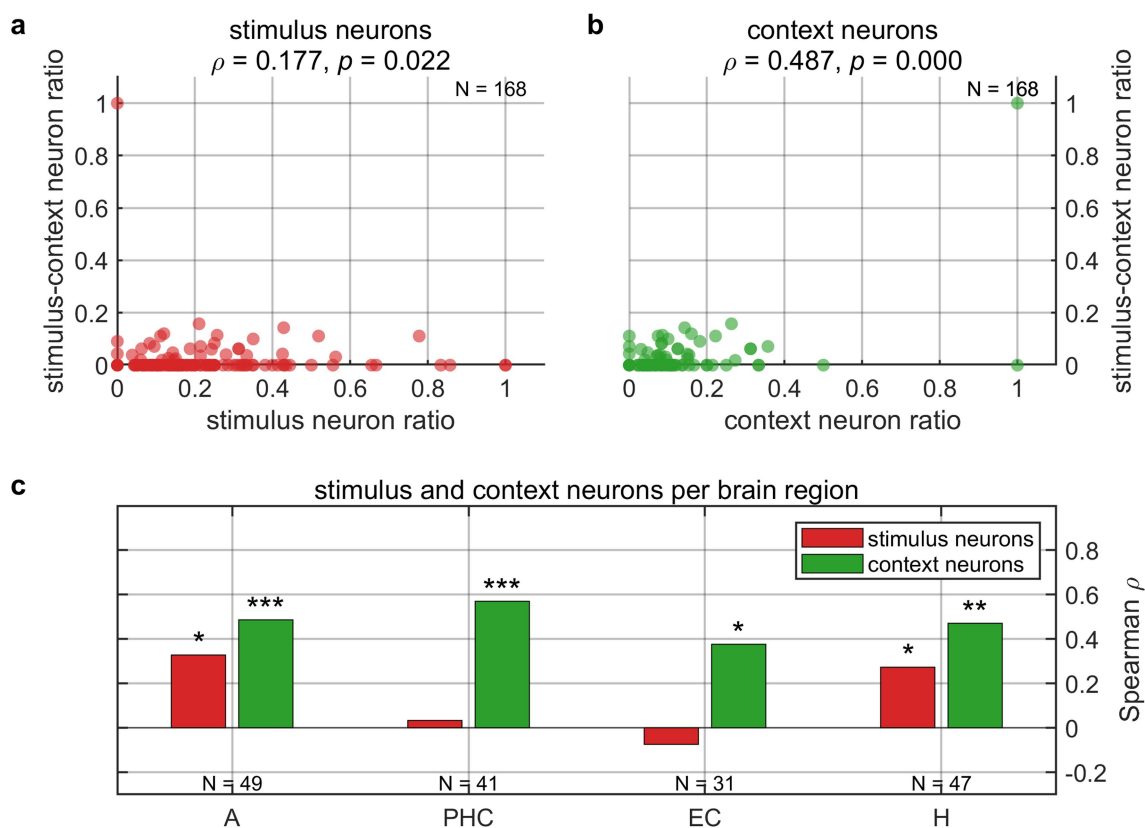
Question Rank Agreement (with "Bigger?", "More expensive?"/"Older?" Ground Truth)



Extended Data Fig. 2 | Session-wise rank agreement with objective ground truth validates successful task performance. Left: Boxplot (Q1, median, Q3; whiskers: points within ± 1.5 IQR) of session-wise mean Spearman rank correlations between behaviorally determined picture ranks and objectively determined picture ranks for the questions "Bigger?" and "More expensive?"/"Older?". Grey dots represent individual sessions (N = 50). All session means were significantly greater than zero ($p = 6.47 \times 10^{-10}$, Wilcoxon signed-rank test).

Shaded background areas indicate effect size categories. **Right:** Null distribution of mean correlations across sessions from 1,000 permutations (grey histogram), where picture ranks were randomly shuffled within each session-question pair before averaging. Observed mean shown in red ($\rho = 0.760$). The observed mean significantly exceeded the permutation null distribution ($p < 0.001$).

Correlations of stimulus-context neuron proportions with proportions of stimulus and context neurons

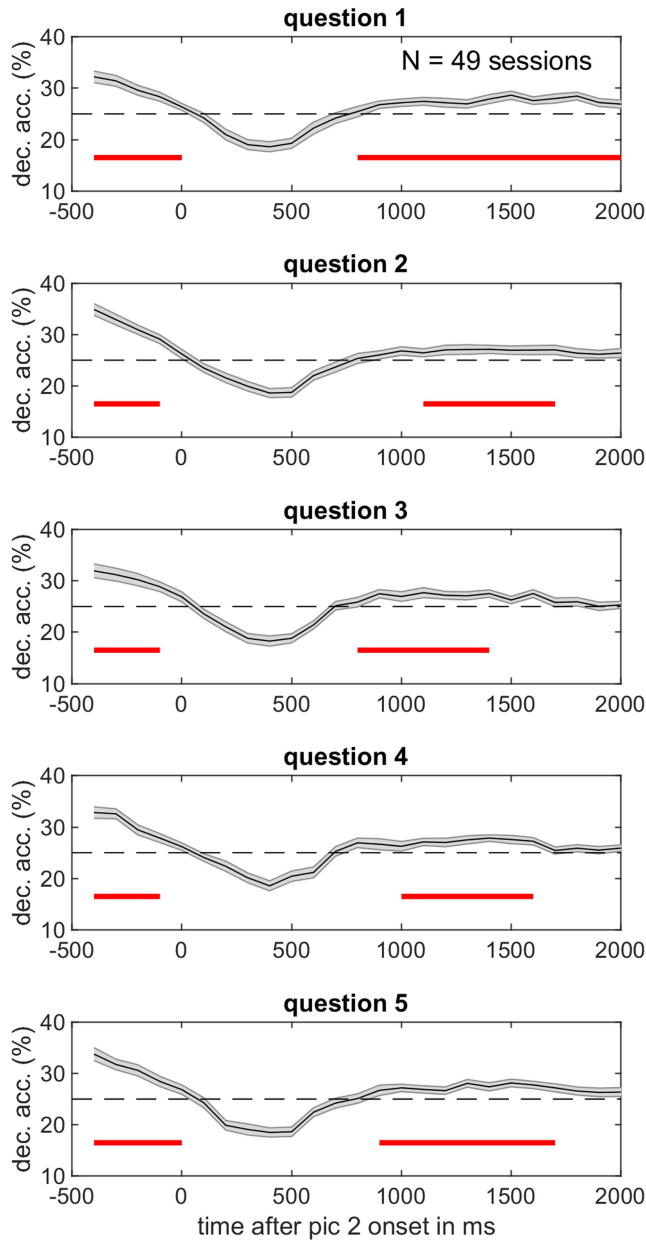


Extended Data Fig. 3 | Stimulus-context neurons are more prevalent in brain regions with higher proportions of context neurons and, to a lesser extent, stimulus neurons. a-b. Scatter plots of stimulus-context neuron ratios vs. stimulus neuron ratios (a, red) or context neuron ratios (b, green) across all session-site combinations (N = 168). Spearman correlations revealed weak positive associations with stimulus neurons ($\rho = 0.177, p = 0.022$) and strong positive associations with context neurons ($\rho = 0.487, p < 0.001$). **c.** Spearman correlation coefficients between stimulus-context neuron proportions

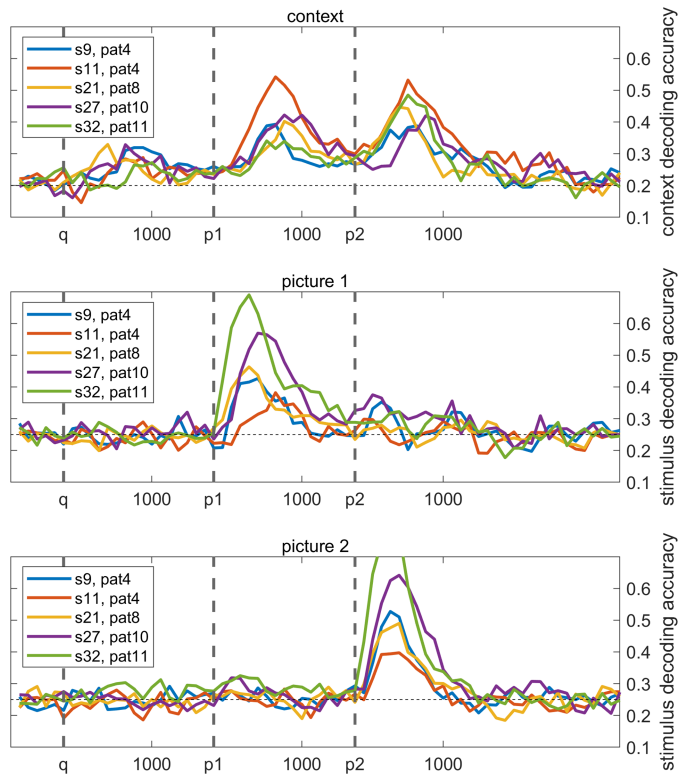
and stimulus (red) or context (green) neuron proportions for each brain region. Statistical significance was assessed using permutation tests (1000 permutations, one-sided). Amygdala (A, N = 49): stimulus $\rho = 0.329, p = 0.012$; context $\rho = 0.487, p = 0.003$. Parahippocampal cortex (PHC, N = 41): stimulus $\rho = 0.033, p = 0.387$; context $\rho = 0.569, p < 0.001$. Entorhinal cortex (EC, N = 31): stimulus $\rho = -0.075, p = 0.324$; context $\rho = 0.377, p = 0.023$. Hippocampus (H, N = 47): stimulus $\rho = 0.274, p = 0.028$; context $\rho = 0.471, p < 0.001$. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$ (uncorrected).

Picture 1 decoding with stimulus neuron activity

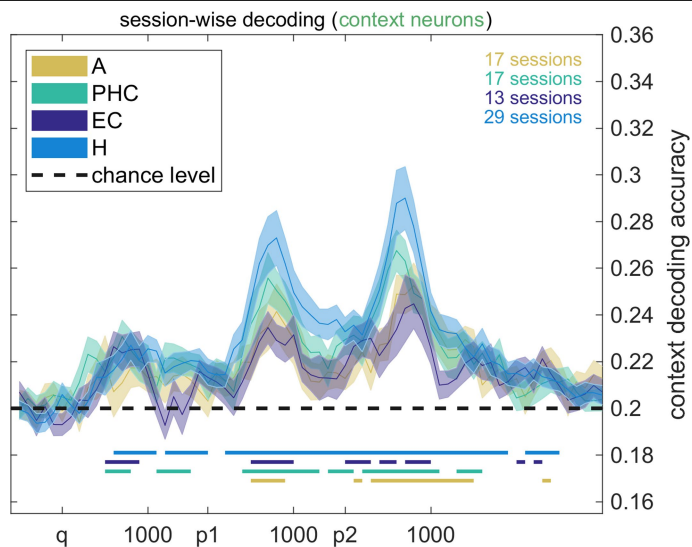
t(train): pic 1 (600 ms), t(decode): pic 2 (-500 to 2000 ms)



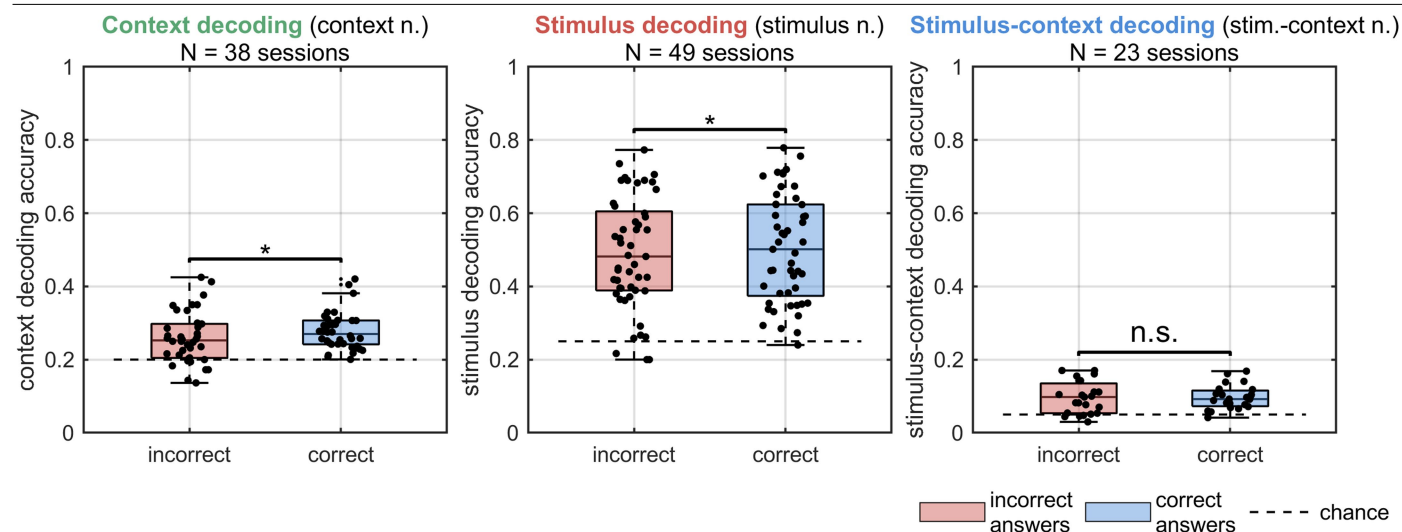
Extended Data Fig. 4 | Stimulus neurons recapitulate first-picture activity during second-picture presentations in all question contexts, indicating content-based rather than feature-specific encoding. Session-wise decoding accuracies of first stimuli during second-picture presentations in all five question contexts using SVM decoders trained with stimulus neuron activity - 600 ms after first-picture onset (400–800 ms bin). Solid lines and shaded areas depict means and standard errors, respectively. Data are compared to chance (25%, dashed horizontal black line). Solid lines in red below indicate significant differences (cluster permutation test, two-sided, $p < 0.05$). During second-picture presentations, first-picture decoding accuracies exceeded chance beginning - 600–1000 ms after second-picture onsets in all question contexts, arguing against a mere recollection of stimulus features.



Extended Data Fig. 5 | During picture presentations visual stimuli and contexts could be decoded simultaneously and with high accuracy in single recording sessions. Support vector machine decoding accuracies (linear SVM) of contexts (top, context neurons), first (middle, stimulus neurons) or second pictures (bottom, stimulus neurons) for different example sessions (s) distinguished by color from different patients (pat). Data comprise a subset of that presented in Fig. 3d. Vertical dashed lines in gray denote onsets of different trial events (q: question, p1 or p2: picture 1 or 2) and horizontal dashed lines indicate chance levels (1/5 for context, 1/4 for pictures). Despite highly sparse sampling of neurons by micro-wires, even within single recording sessions decoding accuracies were remarkably high (e.g. see sessions 11, 27 and 32 in red, violet and green with accuracies up to 50 or 70 percent for contexts and pictures, respectively). Importantly, within single recording sessions both pictures and contexts could be decoded simultaneously above chance during picture presentations (e.g. see sessions 9, 21, 27 and 32 in blue, yellow, violet and green, contrasting with pre-question baseline periods).

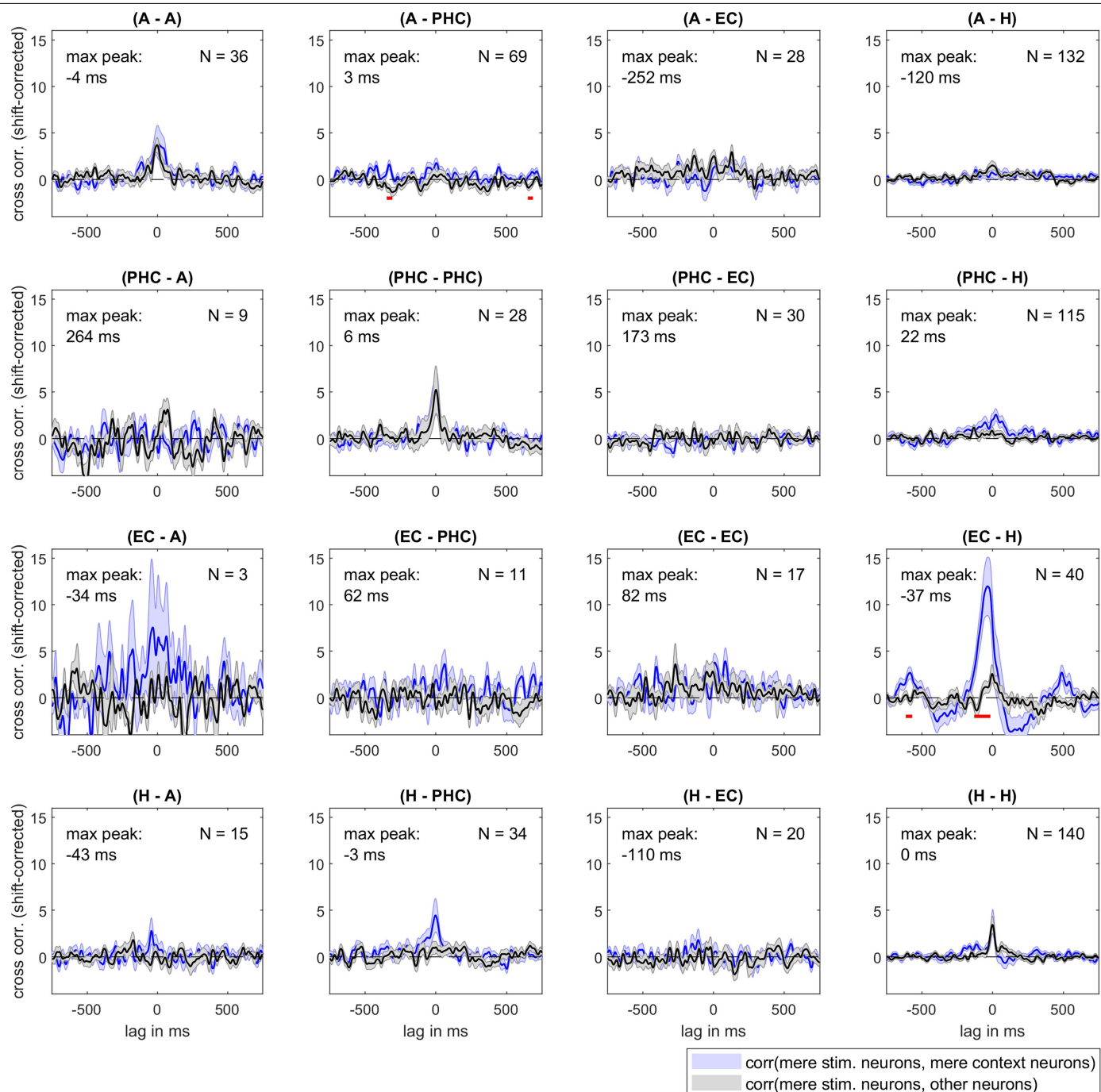


Extended Data Fig. 6 | Representations of context across different MTL brain regions. Session-wise decoding accuracies of context with context neurons from different brain regions (PHC: parahippocampal cortex, A: amygdala, EC: entorhinal cortex, H: hippocampus) distinguished by color. Contexts were encoded most strongly in hippocampus during late stimulus presentations. Solid lines and shaded areas depict means and standard errors, respectively. Horizontal lines below indicate significant differences from chance (cluster permutation test, $p < 0.01$).



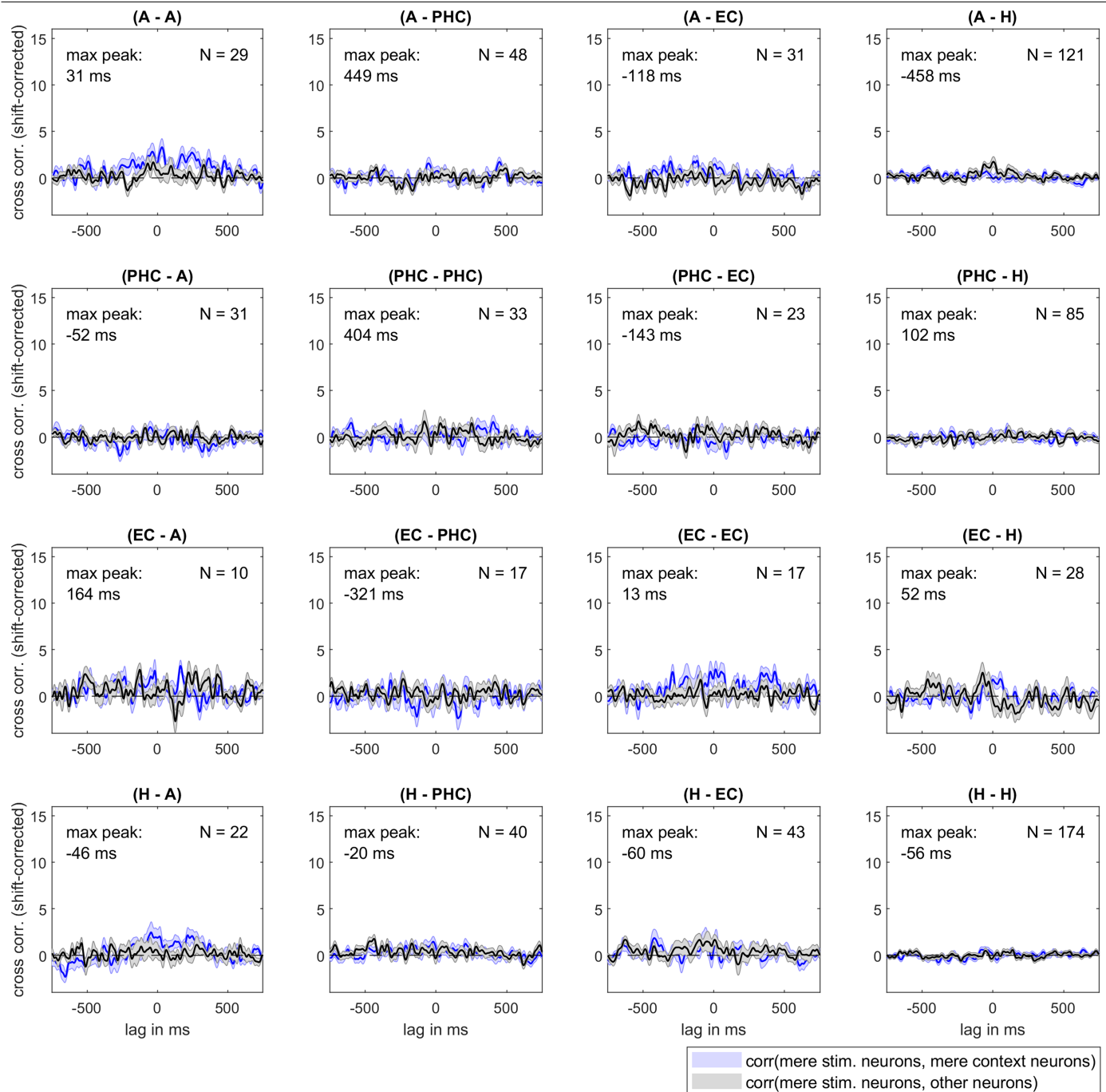
Extended Data Fig. 7 | Decoding of stimulus and context is enhanced when behavioral answers correspond to ground truth. Boxplot (Q1, median, Q3; whiskers: points within ± 1.5 IQR) of session-wise mean decoding accuracies during picture presentations for trials with incorrect (red) vs. correct (blue) answers relative to objective ground truth. **Left:** context decoding from context neurons (N = 38; chance = 1/5). **Middle:** stimulus (picture) decoding from stimulus neurons (N = 49; chance = 1/4). **Right:** stimulus-context (picture

and question) decoding from stimulus-context neurons (N = 23; chance = 1/20). Dashed lines indicate chance levels. Decoding accuracies were significantly higher for correct vs. incorrect trials for context ($p = 0.0128$) and stimulus neurons ($p = 0.0351$), while stimulus-context neurons showed no reliable difference ($p = 0.4098$, n.s.). * $p < 0.05$ (uncorrected, Wilcoxon signed-rank test, one-sided, correct > incorrect).

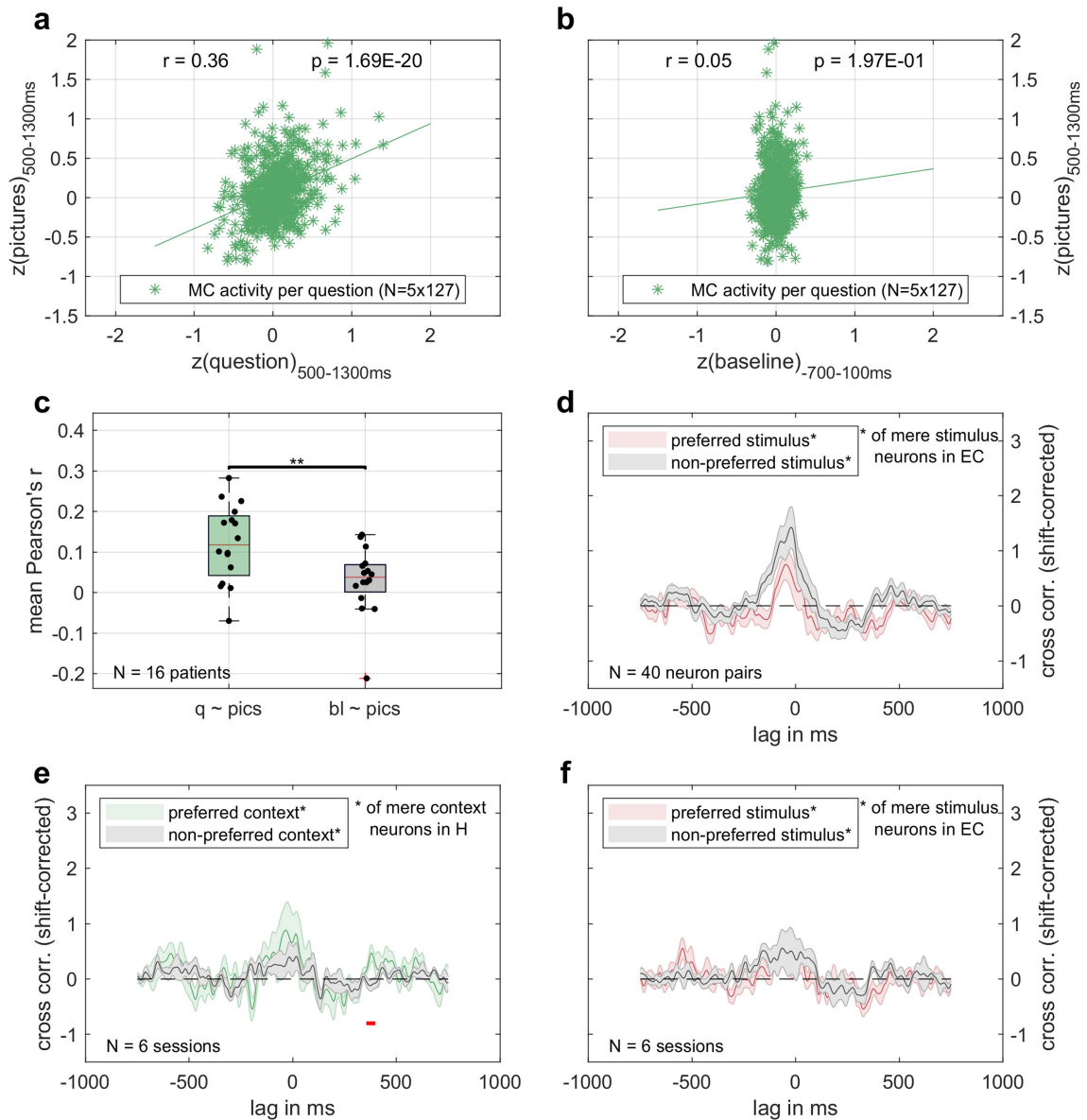


Extended Data Fig. 8 | Cross-correlograms between mere stimulus and mere context neurons of the same hemisphere during picture processing. Trial-by-trial cross-correlograms during picture presentations (shift-corrected, see Methods) either between pairs of mere stimulus and mere context neurons (each with only one of two factors $p < 0.001$ in ANOVA, blue) or a matched number of mere stimulus and other neurons without significant main effects (all factors $p > 0.1$ in ANOVA, black) from different MTL regions of the same hemisphere. Data are presented as mean values $\pm$ SEM (solid lines and shaded areas). Red horizontal lines indicate significant differences ($p < 0.01$) between pair types (blue vs black) as quantified by a two-sided cluster permutation test.

Neurons were drawn from brain region pairs written on top (PHC: parahippocampal cortex, A: amygdala, EC: entorhinal cortex, H: hippocampus) and maximum peak lag times as well as number of neuron pairs are depicted in the upper corners. Only between EC and H, firing of stimulus neurons predicted firing of context neurons, but not of other neurons, after approximately 40 (or 600) milliseconds. Asymmetric peaks on the order of tens of milliseconds are indicative of interactions between stimulus and context neurons and suggest stimulus-context associations might be stored via spike-timing-dependent plasticity (STDP) for later reinstatement of context upon stimulus presentations or contextual encoding of stimuli.

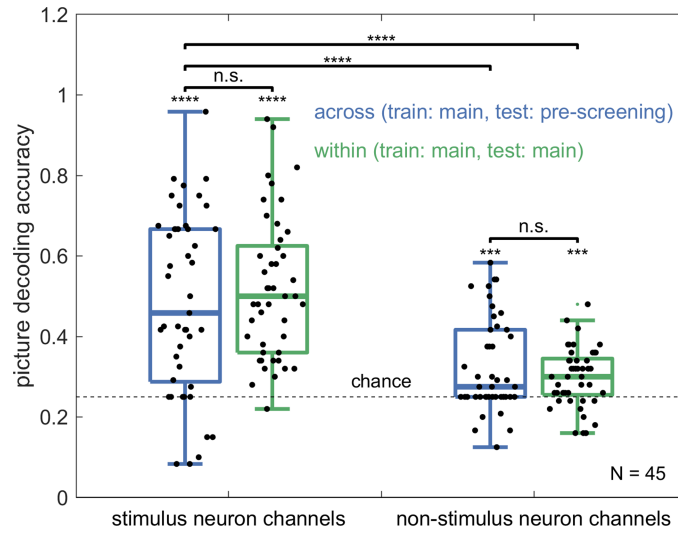


Extended Data Fig. 9 | Cross-correlograms between mere stimulus and mere context neurons of different hemispheres during picture processing. As in Extended Data Fig. 8, but for different hemispheres. No significant differences were detected.



Extended Data Fig. 10 | Sequential activation after pre-activation of context neurons by their preferred context could guide stimulus-driven pattern completion. **a.** Scatter plot of mean z-values for each of the 5 contextual questions (baseline normalization) during question vs. subsequent picture presentations computed for each of the 127 mere context neurons ($N = 5 \times 127$). Pearson correlation strengths and p values (uncorrected) are shown at the top and visualized by regression lines. Question response strengths significantly predict responses to subsequent pictures ($r = 0.36$, $p = 1.68 \times 10^{-20}$). **b.** As in **a**, but comparing baseline to subsequent picture presentations. Baseline response strengths do not predict subsequent picture responses ($r = 0.05$, $p = 0.20$). **c.** Patient-means of Pearson correlation effect sizes computed from mere context neuron activity between question and subsequent picture presentations across trial conditions. Activity between question and subsequent picture presentations ($r = 0.12$, left in green) was correlated significantly more strongly

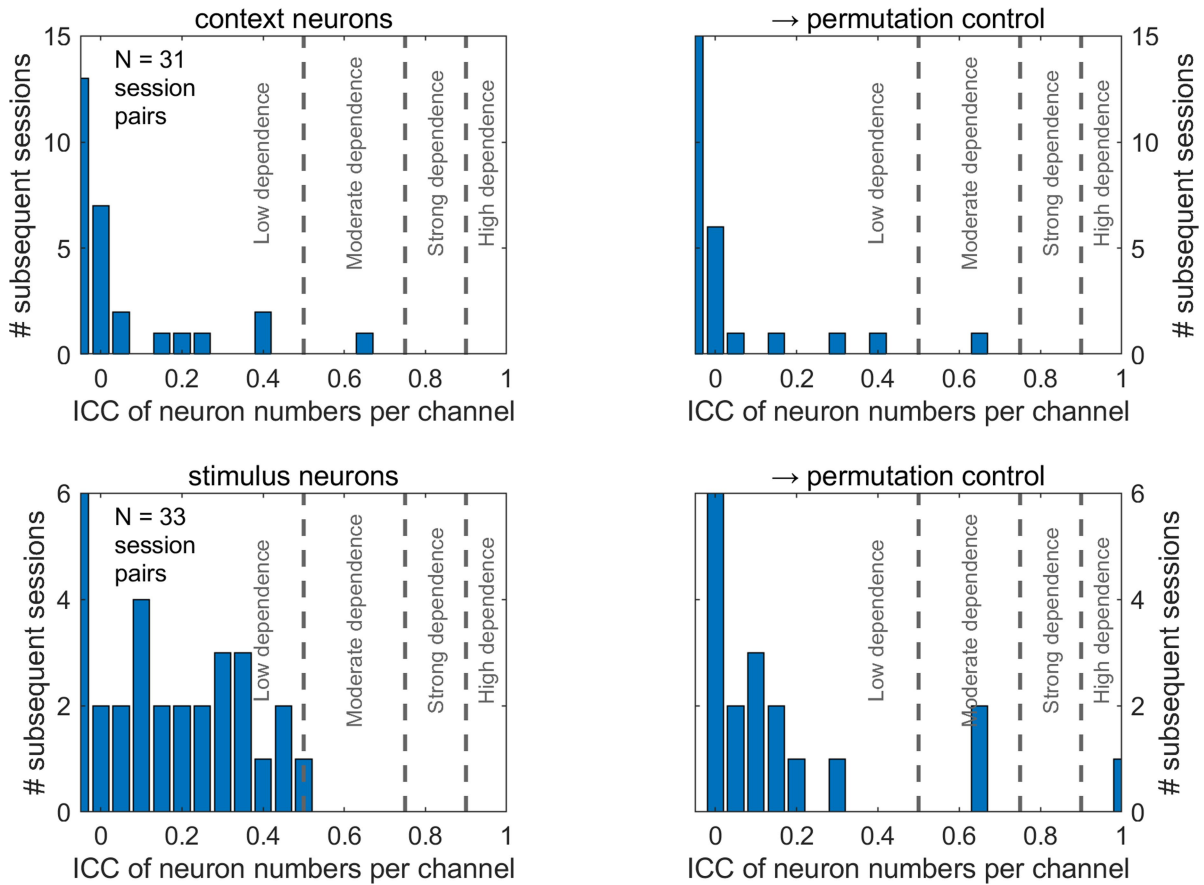
than between baseline and picture activity ($r = 0.03$, right in gray; $p = 6.13 \times 10^{-3}$; two-sided Wilcoxon signed-rank test). **d.** Shift-corrected cross correlations of entorhinal mere stimulus and hippocampal mere context neurons during picture presentations depending on whether a preferred stimulus (red, maximum response) vs. a non-preferred stimulus (gray, randomly drawn different stimuli) of mere stimulus neurons was presented. **e.** As in **d**, but depending on whether a preferred context (green, maximum response) vs. a non-preferred context (gray, randomly drawn different contexts) of mere context neurons was shown and averaged per session (only 6 sessions). Firing of stimulus neurons predicts that of context neurons more strongly (though not significantly) in preferred versus non-preferred contexts. **f.** As in **d**, but averaged per session rather than per neuron pair. Red horizontal lines indicate significant differences ($p < 0.05$, green vs gray, two-sided cluster permutation test).



Extended Data Fig. 11 | Stimulus representations were maintained from pre-screenings to subsequent main experiments, particularly in micro-wires with stimulus neurons. Boxplots (Q1, median, Q3; whiskers: points within ± 1.5 IQR) of session-wise SVM picture decoding accuracies of the four pictures shown in the main experiment. SVM decoders were trained on main experiment activity and tested either on pre-screening (across in blue) or main experiment activity (within in green, leave-out cross-validation with matched trial counts). Dotted line in black denotes chance decoding of 0.25. For 45 recording sessions, all four pictures of the main experiment were part of the pre-screening before the experiment ($N = 45$). Picture decoding accuracies highly exceeded chance, particularly for micro-wire channels that contained stimulus neurons (left side; stimulus-across: 0.4854 ± 0.2243 SD, $p = 8.229 \times 10^{-7}$; stimulus-within: 0.5187 ± 0.1759 SD, $p = 5.676 \times 10^{-9}$) vs. those that did not (right side; non-stimulus-across: 0.3241 ± 0.1155 SD, $p = 3.017 \times 10^{-4}$; non-stimulus-within: 0.2991 ± 0.0734 SD, $p = 1.460 \times 10^{-4}$). Importantly, there were no significant differences in decoding accuracy when testing on pre-screening data (across) or on main experiment activity (within), both for stimulus neuron channels ($p_{\text{diff}} = 0.473$) and for non-stimulus neuron channels ($p_{\text{diff}} = 0.185$; two-sided Wilcoxon signed-rank test). Overall, stimulus representations were maintained across the relatively short time intervals ($4.71 \text{ h} \pm 2.02$ SD, min: 1.40 h, max: 9.15 h) between pre-screenings and subsequent main experiments, specifically in channels that contained stimulus neurons.

**** $p < 0.0001$; *** $p < 0.001$; ** $p < 0.01$ (uncorrected).

Neural independence of subsequent main recording sessions



Extended Data Fig. 12 | Numbers of stimulus and context neurons across micro-wires are largely independent between subsequent main recording sessions within patients. Bar plots of the distribution of intraclass correlation coefficients (ICCs) for the numbers of either context (top left) or stimulus neurons (bottom left) across channels (i.e. micro-wires) between subsequent main recording sessions. Corresponding channel permutation controls are

shown on the right. Vertical dashed lines in gray separate areas of differing effect sizes. Across the long time intervals between subsequent recording sessions ($46.94 \text{ h} \pm 24.19 \text{ SD}$, range: 16.57 h to 118.06 h), numbers of both context and stimulus neurons exhibit a low degree of dependence comparable to permutation controls.

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 (<i>n</i>) 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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ 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 <i>d</i> , Pearson's <i>r</i>), 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 recorded using Behnke-Fried depth electrodes (AdTech, Racine, WI) equipped with microwire bundles protruding from the tip of the electrodes (3-5 mm). Data were amplified and recorded using a 256-channel ATLAS amplifier (Neuralynx, Bozeman, MT) and the Pegasus software (Neuralynx, Bozeman, MT).
Data analysis	The experimental paradigm was implemented in Psychtoolbox3 (www.psychtoolbox.org) Spike extraction and sorting was performed using Combinato (https://github.com/jniediek/combinato) Statistical analyses were conducted in MATLAB 2021a including the Statistics and Machine Learning Toolbox (The MathWorks, Natick, MA) Code related to the main analyses of the manuscript is available at https://github.com/mabausch/ContentContextNeurons.git .

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](#)

All data supporting the findings of this study will be publicly available at <https://github.com/mabausch/ContentContextNeurons.git>.

Research involving human participants, their data, or biological material

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

Reporting on sex and gender	Our study included 17 participants, of whom 8 were female and 9 were male, according to the clinical reports. Our study did not include sex- or gender-specific analyses.
Reporting on race, ethnicity, or other socially relevant groupings	No socially constructed variables were used or analyzed in this study.
Population characteristics	In this study, no characteristics of the population were included as covariates.
Recruitment	All patients that underwent treatment for pharmacologically intractable epilepsy were asked to partake in the experiment and agreed. Therefore, there was no selection bias.
Ethics oversight	The study was approved by the Medical Institutional Review Board at the University of Bonn.

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

Field-specific reporting

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

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	Our analysis is based on a total of 3,109 units recorded in 49 experimental sessions from 16 epilepsy patients suffering from drug-resistant epilepsy and undergoing invasive seizure monitoring with the goal of subsequent neurosurgical resection. No a-priori sample-size calculation was performed. Our comprehensive dataset easily complies with or exceeds current standards in the field of human single unit recordings.
Data exclusions	One session of one subject (out of 17) was excluded based on a low behavioral performance (see manuscript).
Replication	Analyses were performed on the population of recorded units from 16 participants. Results are based on rigorous statistical analysis, including correction for multiple comparisons where appropriate. Replication in other subjects is not applicable owing to intervention constraints.
Randomization	There was no experimental group assignment in our study.
Blinding	The study did not include allocation of subjects to experimental groups so no blinding applies.

Behavioural & social sciences study design

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

Study description	Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).
Research sample	State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic

Research sample	<i>information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.</i>
Sampling strategy	<i>Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.</i>
Data collection	<i>Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.</i>
Timing	<i>Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Non-participation	<i>State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.</i>
Randomization	<i>If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.</i>

Ecological, evolutionary & environmental sciences study design

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

Study description	<i>Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.</i>
Research sample	<i>Describe the research sample (e.g. a group of tagged <i>Passer domesticus</i>, all <i>Stenocereus thurberi</i> within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.</i>
Sampling strategy	<i>Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.</i>
Data collection	<i>Describe the data collection procedure, including who recorded the data and how.</i>
Timing and spatial scale	<i>Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Reproducibility	<i>Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.</i>
Randomization	<i>Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.</i>
Blinding	<i>Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.</i>

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	<i>Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).</i>
Location	<i>State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).</i>
Access & import/export	<i>Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).</i>

Disturbance

Describe any disturbance caused by the study and how it was minimized.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

- n/a Involved in the study
- ☒ ☐ Antibodies
- ☒ ☐ Eukaryotic cell lines
- ☒ ☐ Palaeontology and archaeology
- ☒ ☐ Animals and other organisms
- ☒ ☐ Clinical data
- ☒ ☐ Dual use research of concern
- ☒ ☐ Plants

Methods

- n/a Involved in the study
- ☒ ☐ ChIP-seq
- ☒ ☐ Flow cytometry
- ☒ ☐ MRI-based neuroimaging

Antibodies

Antibodies used

Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.

Validation

Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.

Authentication

Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.

Mycoplasma contamination

Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines
(See [ICLAC](#) register)

Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology and Archaeology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

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

Animals and other research organisms

Policy information about [studies involving animals](#); ARRIVE [guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	<i>For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.</i>
Wild animals	<i>Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.</i>
Reporting on sex	<i>Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.</i>
Field-collected samples	<i>For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.</i>
Ethics oversight	<i>Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.</i>

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

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	<i>Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.</i>
Study protocol	<i>Note where the full trial protocol can be accessed OR if not available, explain why.</i>
Data collection	<i>Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.</i>
Outcomes	<i>Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.</i>

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes
☐	☐
☐	☐
☐	☐
☐	☐
☐	☐
☐	☐
☐	☐
☐	☐
☐	☐

☐ Demonstrate how to render a vaccine ineffective
☐ Confer resistance to therapeutically useful antibiotics or antiviral agents
☐ Enhance the virulence of a pathogen or render a nonpathogen virulent
☐ Increase transmissibility of a pathogen
☐ Alter the host range of a pathogen
☐ Enable evasion of diagnostic/detection modalities
☐ Enable the weaponization of a biological agent or toxin
☐ Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.
Novel plant genotypes	Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.
Authentication	Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links <i>May remain private before publication.</i>	For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.
Files in database submission	Provide a list of all files available in the database submission.
Genome browser session (e.g. UCSC)	Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates	Describe the experimental replicates, specifying number, type and replicate agreement.
Sequencing depth	Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.
Antibodies	Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.
Peak calling parameters	Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.
Data quality	Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.
Software	Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	<i>Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.</i>
Instrument	<i>Identify the instrument used for data collection, specifying make and model number.</i>
Software	<i>Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.</i>
Cell population abundance	<i>Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.</i>
Gating strategy	<i>Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.</i>

- ☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	<i>Indicate task or resting state; event-related or block design.</i>
Design specifications	<i>Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.</i>
Behavioral performance measures	<i>State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).</i>

Acquisition

Imaging type(s)	<i>Specify: functional, structural, diffusion, perfusion.</i>
Field strength	<i>Specify in Tesla</i>
Sequence & imaging parameters	<i>Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.</i>
Area of acquisition	<i>State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.</i>
Diffusion MRI	☐ Used ☐ Not used

Preprocessing

Preprocessing software	<i>Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).</i>
Normalization	<i>If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.</i>
Normalization template	<i>Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.</i>
Noise and artifact removal	<i>Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).</i>

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

(See [Eklund et al. 2016](#))

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a | Involved in the study

☐ ☐ Functional and/or effective connectivity

☐ ☐ Graph analysis

☐ ☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.