NEUROSCIENCE

Reinstatement and transformation of memory traces for recognition

Elias M. B. Rau[1]*, Marie-Christin Fellner[1], Rebekka Heinen[1], Hui Zhang[1], Qin Yin[2], Parisa Vahidi[3,4], Malte Kobelt[1], Eishi Asano[5], Olivia Kim-McManus[6,7], Shifteh Sattar[7], Jack J. Lin[8], Kurtis I. Auguste[9,10], Edward F. Chang[10], David King-Stephens[11,12], Peter B. Weber[11], Kenneth D. Laxer[11], Robert T. Knight[13], Elizabeth L. Johnson[14,15]†, Noa Ofen[2,3,16]†, Nikolai Axmacher[1]†

Episodic memory relies on the formation and retrieval of content-specific memory traces. In addition to their veridical reactivation, previous studies have indicated that traces may undergo substantial transformations. However, the exact time course and regional distribution of reinstatement and transformation during recognition memory have remained unclear. We applied representational similarity analysis to human intracranial electroencephalography to track the spatiotemporal dynamics underlying the reinstatement and transformation of memory traces. Specifically, we examined how reinstatement and transformation of item-specific representations across occipital, ventral visual, and lateral parietal cortices contribute to successful memory formation and recognition. Our findings suggest that reinstatement in temporal cortex and transformation in parietal cortex coexist and provide complementary strategies for recognition. Further, we find that generalization and differentiation of neural representations contribute to memory and probe memory-specific correspondence with deep neural network (DNN) model features. Our results suggest that memory formation is particularly supported by generalized and mnemonic representational formats beyond the visual features of a DNN.

INTRODUCTION

Episodic memories are the sediments of previous experiences—they allow us to recall specific details of our lives and to recognize familiar events or items. This crucial cognitive function depends on the formation of memory traces or engrams (*1*, *2*) that can be tracked via the coordinated activity of single cells and neural populations (*3*, *4*). Although the formation of engrams relies critically on the hippocampus (*5*, *6*), retrieving their full information content requires interactions with distributed neocortical networks that represent the various facets of an episode (*7*). In contrast to rodents, in whom engrams can be directly measured using optogenetics (*8*) and cellular imaging methods (*9*), cognitive neuroscience studies in humans have applied multivariate analysis methods such as representational similarity analysis (RSA) (*10*) to track item-specific activity patterns (*11*, *12*). These methods allow for a comparison of the neural features, for example, time-frequency features in electrophysiological recordings, which represent item-specific contents during encoding with those during retrieval (encoding-retrieval similarity, ERS) (*13*, *14*).

The cortical reinstatement hypothesis (*15*, *16*) states that memory retrieval requires the reinstatement of neural activity patterns that were present during encoding—i.e., that levels of ERS are higher for remembered than for forgotten items. Previous studies demonstrated that retrieval depends on reinstatement of encoding-related activity patterns in medial temporal lobe (MTL) (*11*, *17*) and neocortex (*13*, *18*–*20*) that are specific to individual items. Whereas neuroimaging studies using functional magnetic resonance imaging (fMRI) combine relatively high spatial resolution with full-brain coverage, human intracranial electroencephalography (iEEG) recordings have provided deeper insights into the temporal dynamics of the neural patterns that constitute reinstatement such as the contribution of neural oscillations at specific frequencies (*21*–*23*). These studies showed that increased power of gamma-band activity (>30 Hz) in the first second of encoding correlates with higher levels of reinstatement in hippocampus (*11*) and in regions of the lateral temporal cortex (*19*). Reinstatement has been observed during both recall and recognition memory (*24*) and is related to both recollection and familiarity (*16*).

However, in addition to results showing reinstatement, memory traces have also been found to undergo considerable representational transformations between encoding and retrieval (*25*–*27*). More specifically, it has been shown that regions of lateral parietal cortex (LPC) represent event-specific information during retrieval but not during encoding (*28*). This suggests that LPC supports retrieval not through reinstatement of encoding-related activity patterns but through activation of transformed mnemonic representations (*29*),

[1]Department of Neuropsychology, Institute of Cognitive Neuroscience, Ruhr University Bochum, Bochum, Germany. [2]Center for Vital Longevity, School of Behavioral and Brain Sciences, University of Texas at Dallas, Dallas, TX, USA. [3]Life-Span Cognitive Neuroscience Program, Institute of Gerontology, Wayne State University, Detroit, MI, USA. [4]Department of Psychology, College of Liberal Arts and Sciences, Wayne State University, Detroit, MI, USA. [5]Departments of Pediatrics and Neurology, Children's Hospital of Michigan, Detroit Medical Center, Wayne State University, Detroit, MI, USA. [6]Department of Neurosciences, University of California, San Diego, San Diego, CA, USA. [7]Division of Child Neurology, Rady Children's Hospital, San Diego, CA, USA. [8]Department of Neurology, University of California, Davis, Davis, CA, USA. [9]Department of Pediatric Neurosurgery, Benioff Children's Hospital, Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, USA. [10]Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA, USA. [11]Department of Neurology and Neurosurgery, California Pacific Medical Center, San Francisco, CA, USA. [12]Department of Neurology, Yale School of Medicine, New Haven, CT, USA. [13]Helen Wills Neuroscience Institute and Department of Psychology, University of California, Berkeley, Berkeley, CA, USA. [14]Departments of Medical Social Sciences and Pediatrics, Northwestern University, Chicago, IL, USA. [15]Department of Psychology, Northwestern University, Evanston, IL, USA. [16]Department of Psychology, School of Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, TX, USA.
*Corresponding author. Email: elias.rau@ruhr-uni-bochum.de
†These authors contributed equally to this work

i.e., reflecting a dissociation of neuronal networks involved in sensory and mnemonic functions. The LPC has both direct and indirect connections to temporal lobe structures (*30*) and may thus be recruited during attempts to retrieve item-specific representations from memory (*31*). Thus, complement to reinstatement of sensory representations, brain regions such as the LPC are associated with the retrieval of mnemonic representations that result from transformations of encoded information.

ERS quantifies the similarity of item-specific neurophysiological features between encoding and retrieval and, therefore, does not inform about how a given item is represented relative to other items. Hence, observing transformation at the level of ERS is insufficient to indicate changes in the underlying representational geometry (*32*). We nevertheless hypothesized that successful memory may also depend on measures of between-item similarity during encoding (encoding-encoding similarity, EES) or retrieval (retrieval-retrieval similarity, RRS), reflecting representational distances. While ERS and EES/RRS do not directly relate to each other because of different underlying cross-correlation matrices, they may still both be associated with successful memory (*33*, *34*). Specifically, the magnitude of similarity across items during encoding (EES) may relate to the magnitude of ERS in regions showing reinstatement, i.e., where items during encoding are represented in a way that benefits subsequent recognition. Accordingly, the magnitude of similarity between items during recognition (RRS) may relate to the magnitude of ERS in regions showing transformation, i.e., where item-specific representations need to acquire a format which differs from the one during encoding, to be recognized as old.

More specifically, the translation from sensory information to the formation of durable memory traces presumably affects the information that is available during retrieval, which may relate to semantization (*35*), i.e., the preferential retention of conceptual information at the expense of perceptual details. This process appears to start already during brief offline periods following encoding (*36*). Here, we conceptualize the different features that are available in a memory trace as different representational formats (*37*). For visual stimuli, these representational formats reflect the sensory processing hierarchy along the ventral visual stream (VVS) (*38*, *39*). Along the posterior-to-anterior extent of the VVS, neural assemblies represent increasingly complex information: Whereas posterior VVS regions represent low-level perceptual features such as edges or colors, anterior regions represent more complex and categorical features including faces (*40*) and places (*41*). Previous studies showed that the VVS can be broadly divided into early (occipital, Occ), mid-level (posterior ventral temporal cortex, pTC), and high-level (anterior ventral temporal cortex, aTC) regions of interest (ROIs), which carry distinct representational formats of increasing complexity (*42*). Analytically, these distinctions can be captured via deep neural networks (DNNs): DNNs trained to classify visual stimuli have become powerful tools to study the neural representations underlying visual perception (*38*, *43*, *44*) and have more recently been applied to assess the representational formats of mnemonic representations (*36*, *37*, *45*). Thus, the potential to investigate content-based representational changes following mnemonic transformation using DNNs and the high spatiotemporal resolution of iEEG are a promising combination to study how correspondences between neural and DNN representational formats along the VVS unfold during the first few hundred milliseconds after stimulus presentation (*39*, *46*).

Although reinstatement and transformation were primarily investigated in free or cued recall paradigms (*22*), they may also be relevant during recognition (*47*). While both types of memory tests involve the retrieval of mnemonic information, they differ in the types of cues that initiate the reactivation of previous experience. Whereas retrieval during free or cued recall is initiated via internal, partial, or associative cues, recognition involves the comparison of sensory inputs (identical to those during encoding) to mnemonic representations built during encoding, which allow recognizing that these sensory inputs had already been presented before. Hence, the concepts of representational reinstatement and transformation differ in that during recognition, reinstatement does not concern the active reproduction of individual sensory stimuli from memory but of information about the prior presentation of these stimuli. This information does not necessarily rely on item-specific sensory features but may rather reflect mnemonic information shared across multiple items. Vice versa, transformation during free or cued recall likely reflects deviations of mnemonic representations (during recall) from sensory-driven representations (during encoding); transformation during recognition presumably involves the change in processing of a previously seen item (during recognition) from the processing during the first encounter with that item (during encoding). During recognition, sensory input is identical between trials of successful and unsuccessful recognition, allowing the attribution of memory-specific differences to mnemonic effects. Both reinstatement and transformation may contribute to successful recognition, with differential functional relevance of representational formats of memory traces.

The functional compartmentalization of visual representational formats across the VVS exists already in children but continues to be refined through adolescence (*48*). Developmental changes in the representational format of a memory trace may affect the strategic usage of encoded information: Whereas encoding of event-specific perceptual formats may enhance the differentiation of memory traces and allow for a separation from memories of similar episodes (*49*), predominant retention of conceptual formats facilitates inferences to novel experiences with shared commonalities (*50*, *51*). These two memory functions, i.e., to represent either specific or generalized information, have been ascribed to complementary processes and brain regions (*52*, *53*) that undergo differential maturational trajectories (*54*–*56*), leading to an infantile bias toward generalization (*57*). Capitalizing on iEEG recordings, previous studies showed that while memory formation in children and adolescents depends on the same brain regions as in adults, there are developmental differences in the magnitude (*56*, *58*–*61*) and intricate temporal dynamics of memory processing (*62*–*64*) that mediate age-related gains during adolescence (*65*). Hence, we would not expect that the levels of performance, although generally lower as compared to adults, are due to the recruitment of qualitatively different neural processes. More likely, the age-related variability in performance may be partially explained by the interplay and functional relevance of the involved mechanisms. Thus, the study of recognition memory during development is well suited to identify possible age- or performance-related differences in brain-behavior relationships.

How the reinstatement and transformation of representations supports recognition memory, whether they relate to differentiated versus generalized representations and on which representational formats they rely is not well understood. In addition, it is unclear whether these effects vary across development. Therefore, we sought to identify brain regions exhibiting reinstatement or transformation that are associated with successful recognition memory. Further, we

aimed to relate these processes to generalized versus differentiated representations during encoding and retrieval. Last, using a convolutional DNN (cDNN) trained on a similar task as the participants performed during encoding (i.e., classify scenes into semantic categories), we aimed to characterize the relevance of specific visual representational formats in memory representations for reinstatement or transformation. To address these questions, we analyzed direct brain recordings in a sample of children, adolescents, and young adults (age range 6 to 32 years) with pharmaco-resistant epilepsy ($N = 46$) engaging in a simple scene encoding and recognition paradigm. Electrode coverage across the brain spanned large proportions of neocortex including early, middle, and late regions of the VVS and areas in the LPC putatively related to retrieval. Using RSA, we computed item-specific similarity estimates across electrophysiological features between pairs of items within (EES and RRS) and across (ERS) experimental phases and related these measures to memory success. ERS values were always calculated as the difference in similarity of same items during encoding and retrieval (ERS_{Same}) and different items during encoding and retrieval (ERS_{Diff}): $ERS_{Item} = ERS_{Same} - ERS_{Diff}$. These ERS_{Item} values were calculated separately for remembered and forgotten scenes and then subtracted: ERS_{Item}(remembered) − ERS_{Item}(forgotten). Higher values of ERS_{Item}(remembered) than ERS_{Item}(forgotten) values reflected a memory benefit of reinstatement, while higher values of ERS_{Item}(forgotten) than ERS_{Item}(remembered) values reflected a memory benefit of transformation.

Because of this operationalization, reinstatement and transformation may either rely on memory differences of ERS_{Same} and/or of ERS_{Diff} values, which were therefore compared separately in post hoc analyses: ERS_{Same}(remembered) values were compared to ERS_{Same}(forgotten) values, and ERS_{Diff}(remembered) values were compared to ERS_{Diff}(forgotten) values. Since ERS_{Diff} correlations reflect item-unspecific correlations between encoding and retrieval, we hypothesized that item-specific reinstatement and transformation effects are primarily reflected in differences of ERS_{Same} values of remembered versus forgotten items. Specifically, we hypothesized that reinstatement of item-specific representations would involve positive ERS_{Same} correlations of remembered trials, which should be absent for forgotten trials, and one may expect these memory effects not to occur for ERS_{Diff} correlations. In turn, a transformation of item-specific representations would lead to near-zero ERS_{Same} correlations of remembered trials but positive correlations of forgotten trials. Equivalently, we operationalized the generalization and differentiation of memory representations within experimental phases (encoding, EES; recognition, RRS) as the relative difference in pairwise similarity between (different) remembered and forgotten trials in each phase.

We hypothesized that regions of the VVS implicated in sensory processing of visual stimuli would engage in reinstatement, whereas parietal regions implicated in mnemonic functions would exhibit transformation. Therefore, we tested whether item-specific similarity between encoding and retrieval (ERS_{Item}) differed between remembered and forgotten items. We used time-frequency patterns across electrodes within predefined ROIs as features for the computation of representational similarities and investigated conditional differences in these features. Specifically, we expected effects of reinstatement and transformation to be located in ventral visual and parietal regions, respectively (*13*, *29*). Similarly, we conducted planned analyses on memory-specific generalization and differentiation of neural representations, as well as statistical tests for the matching of neural and cDNN model predictions, without concrete expectations concerning the direction of memory-related effects. Further, given the subtle differences across childhood development described in earlier studies (*56*, *61*–*64*, *66*), we did not have strong expectations concerning age- or performance-dependent differences in the recruitment of reinstatement and transformation processes. We considered both the presence and absence of differential contributions across development informative about the neural mechanisms underlying memory in children and adolescents.

We found that both reinstatement in aTC and transformation in LPC contribute to successful recognition memory. Further, we present evidence suggesting that reinstatement concerns mnemonic rather than purely visual representational formats and is related to generalized representational patterns during encoding. Our results provide evidence for the contribution of reinstatement and transformation of distinct representational formats to recognition memory in a sample of children, adolescents, and young adults.

RESULTS

Behavioral performance

During encoding, participants were instructed to memorize visual scenes while completing an indoor/outdoor classification task (Fig. 1A). Scene classification performance was overall high [accuracy, 93.67 ± 8.07% (mean ± SD)], with response times of 1590 ± 533 ms. All subsequent analyses were restricted to trials with correct classification responses. During retrieval, participants classified old and new (foil) scenes as old/new. We used corrected recognition scores *Pr* (hits-false alarms) as a measure of recognition performance. Two participants with negative *Pr* scores were excluded from further analyses. In the remaining sample ($N = 46$), *Pr* was reliably above zero (0.41 ± 0.19; Fig. 1B). Corrected recognition scores were larger for indoor ($Pr = 0.462 \pm 0.207$) as compared to outdoor scenes ($Pr = 0.358 \pm 0.224$; $T_{45} = 3.742$, $P < 0.001$). Response times during recognition were lower for hits (2116 ± 759 ms) as compared to misses (2552 ± 1.003 ms; $T_{45} = -5.108$, $P < 0.001$). Recognition memory performance was numerically higher in older participants (Fig. 1C). This effect did not reach statistical significance ($\rho = 0.234$, $P = 0.117$); however, we additionally compared levels of performance in this scene memory task with data from a larger nonclinical population of healthy individuals reported earlier (*56*, *61*, *66*) who engaged in an fMRI version of the paradigm. We observed comparable trajectories across age despite substantial variability during adolescence (fig. S1) [see also (*58*–*61*, *64*)]. This suggests that our data reflect typical memory development, allowing us to identify differences in the neuronal signatures of memory functions during childhood and adolescence.

Reinstatement of item-specific representations in anterior temporal cortex supports memory

Following the reinstatement hypothesis of episodic memory, we computed ERS across regions in Occ, ventral temporal and parietal ROIs (Fig. 1, D and E) and tested for a larger magnitude of item-level ERS (ERS_{Item}) between remembered versus forgotten scenes (Fig. 2A). We extracted the distribution of power values across a broad frequency range (1 to 150 Hz) in individual time windows (300-ms width and 50-ms step) and correlated these item-specific features across electrodes of each ROI both within and between encoding

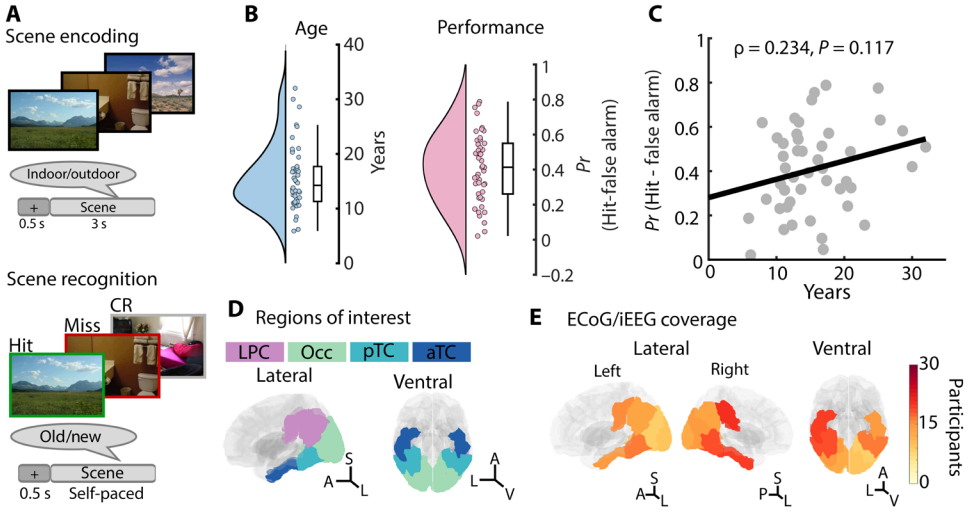

Fig. 1. Task, sample, behavioral performance, and electrode coverage. (**A**) Behavioral paradigm. Participants completed a visual scene memory recognition task with distinct encoding and retrieval blocks. (**B**) Sample distribution of age (15.61 ± 5.92 years; left) and recognition memory performance *Pr* (Hit – false alarm; 0.41 ± 0.19; right). (**C**) Corrected recognition memory performance (Hits – false alarms) across development. (**D**) Regions of interest (ROIs) based on Brodmann areas (BAs): Occ (BA17-19); pTC (BA37); aTC (BA20); LPC (BA39 + 40). (**E**) Heatmap of contributing electrodes in these ROIs. Coordinate axes labels: A, anterior; S, superior; L, lateral; P, posterior; V, ventral; ECoG, electrocorticography.

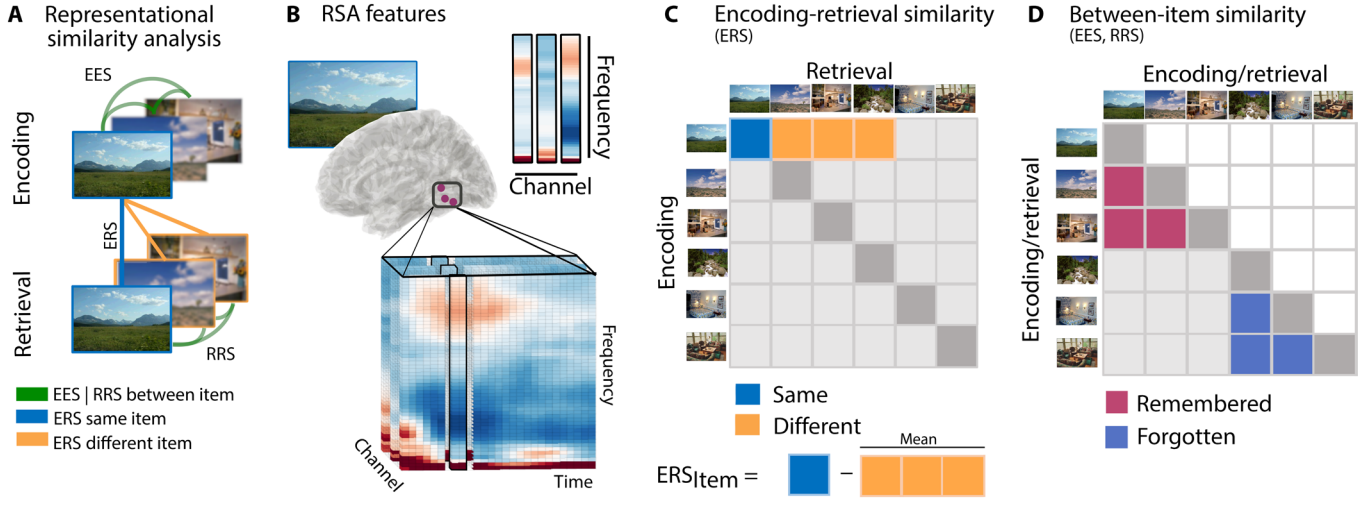

Fig. 2. Representational similarity analyses. (**A**) RSA. We computed pairwise similarities between different scenes within the same experimental phase (EES; RRS), and between the same and different scenes across experimental phases (ERS). (**B**) Time-frequency features used for RSA. We extracted item-specific distributions of power values across frequencies and electrodes within a given ROI in various time windows. (**C**) Schematic ERS matrix. Similarities during encoding and retrieval of the same scenes (ERS_{Same}) are depicted on the diagonal (blue) of this matrix, while similarities between different scenes (ERS_{Diff}) are depicted on the off-diagonal (orange). Item-specific ERS scores were computed by subtracting the average off-diagonal ERS score per row from its on-diagonal value: $ERS_{Item} = ERS_{Same} - avg.\ (ERS_{Diff})$. We compared ERS_{Item} values between remembered and forgotten trials, testing for a memory benefit of either reinstatement (remembered > forgotten) or transformation (forgotten > remembered). (**D**) Schematic pairwise similarity matrix during encoding (EES) or retrieval (RRS) reflecting pairwise representational similarities between different items. During both encoding and retrieval, we compared between-item similarities of remembered versus forgotten scenes. We compared the pairwise similarities between remembered and forgotten scenes, testing for memory benefits of generalization (remembered > forgotten) or differentiation (forgotten > remembered).

and retrieval (Fig. 2, B to D). Using the same time-frequency features in each ROI and experimental phase, we also tested for conditional differences associated with successful memory and their interaction with age and performance, which may influence trial-level ERS (for details, see the Supplementary Materials; fig. S2). In the aTC (Fig. 3A), we found higher ERS_{Item} for remembered versus forgotten scenes (Fig. 3B; $T_{sum} = 263.51$, $N_{bin} = 110$, $P_{cluster} = 0.018$). This reinstatement cluster extended from 500 to 1000 ms during encoding and from 250 to 1150 ms during recognition. Follow-up analyses on the distinct contributions of ERS_{Same} and ERS_{Diff} to ERS_{Item} values showed that memory effects were selectively found for ERS_{Same} ($T_{33} = 2.508$, $P = 0.017$) but not ERS_{Diff} ($T_{33} = -0.412$, $P = 0.682$; Fig. 3C). Further, ERS_{Same} correlations were significantly larger than zero for remembered items ($T_{33} = 3.687$, $P < 0.001$) but

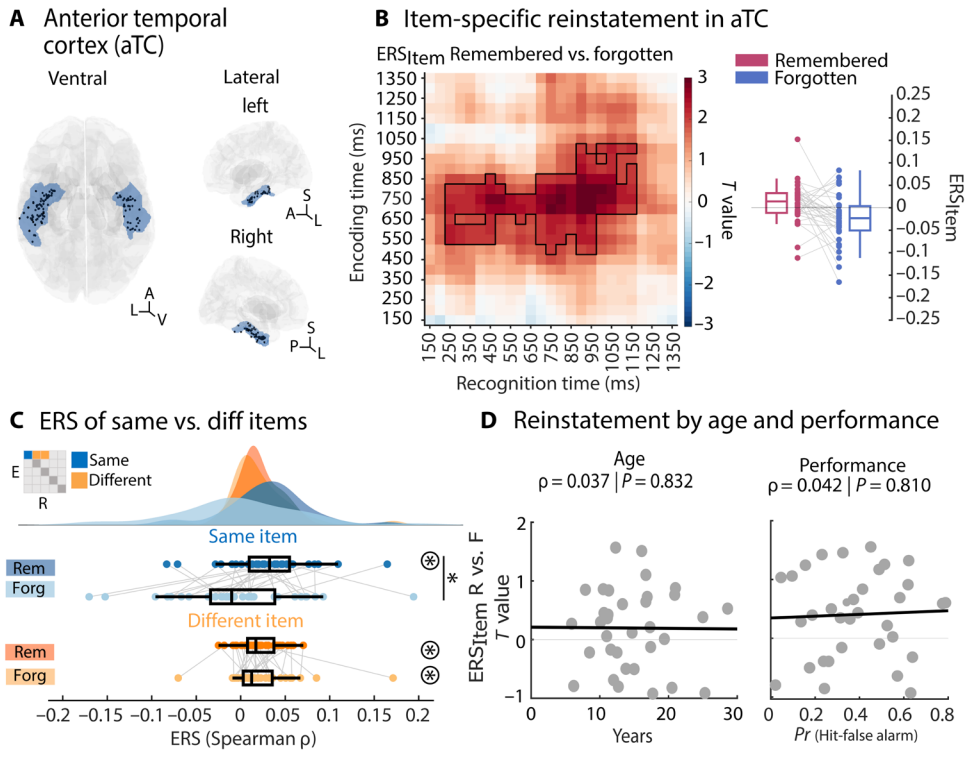

Fig. 3. Reinstatement of item-specific representations in anterior temporal cortex. (**A**) Contributing electrodes in the aTC region of interest (ROI) (34 participants, 107 channels). (**B**) Left, higher ERS_{Item} scores for remembered versus forgotten scenes. Color reflects *T* values across subjects for each time point. Black contour indicates cluster of time points during which ERS_{Item} values were significantly larger for remembered versus forgotten scenes after cluster-based correction for multiple comparisons across time points. Right, illustrative plot of ERS_{Item} similarities (Spearman's ρ) per subject for remembered (amaranth) versus forgotten (blue) scenes, averaged across time points of significant cluster. (**C**) ERS_{Same} (blue, diagonal) and ERS_{Diff} (orange, off-diagonal) values of remembered and forgotten items showing a memory effect of ERS_{Same} but not ERS_{Diff} scores. (**D**) No relationship between reinstatement effects on memory with either age (left) or performance (right). Circled * indicates $P < 0.05$ for one-sample *T* test. Line * indicates $P < 0.05$ for paired-sample *T* test.

not for forgotten items ($T_{33} = -0.454$, $P = 0.652$), indicating that reinstatement was indeed specific to same-item similarities during encoding and retrieval and not due to similarities between different items. ERS_{Diff} correlations were significantly larger than zero for both remembered ($T_{33} = 5.101$, $P < 0.001$) and forgotten ($T_{33} = 3.463$, $P = 0.002$) items.

Next, we tested whether reinstatement varied by age and whether it depended on the participants' performance (Fig. 3D). We found no relationship between memory-specific ERS effects and age ($\rho = 0.037$, $P = 0.832$) or performance ($\rho = 0.042$, $P = 0.810$), indicating similar magnitudes of reinstatement effects on memory across development. To address influences of shared variance between our developmental measures, we also conducted partial correlation analyses between memory-specific ERS effects by controlling for mutual effects between age and performance. We found that neither age ($\rho = 0.024$, $P = 0.891$) nor performance ($\rho = 0.032$, $P = 0.858$) correlated with reinstatement. No significant memory effects of reinstatement were found in earlier VVS regions (i.e., Occ or pTC; all *P* values >0.187; fig. S3), suggesting that reinstatement of rather conceptual but not purely sensory representations is associated with recognition memory success.

Transformation of neural representations in LPC

We also investigated reinstatement or transformation effects in lateral parietal cortex (LPC) (Fig. 4A). In notable contrast to our results in aTC, we found that ERS_{Item} was reduced for remembered as compared to forgotten items in LPC ($T_{sum} = -685$, $N_{bin} = 257$, $P_{cluster} = 0.003$; Fig. 4B), suggesting transformation. This effect occurred from 0 to 1500 ms during encoding and from 0 to 900 ms during recognition. Notably, early onsets are most likely due to data smoothing, where the first time bin contains averaged data from a sliding window from 0 to 300 ms after stimulus onset. Again, we investigated the influence of same-item and different-item correlations and found memory differences in ERS_{Same} ($T_{37} = -2.488$, $P = 0.017$) but not ERS_{Diff} scores ($T_{37} = 1.127$, $P = 0.266$; Fig. 4C). Furthermore, ERS_{Same} scores were significantly larger than zero for forgotten ($T_{37} = 2.566$, $P = 0.015$) but not for remembered items ($T_{37} = 1.252$, $P = 0.218$). ERS_{Diff} correlations were significantly larger than zero for remembered items ($T_{37} = 3.499$, $P = 0.001$) but not forgotten ($T_{37} = 1.900$, $P = 0.065$) items.

We next tested whether subject-specific levels of transformation depended on age (Fig. 4D) or performance (Fig. 4E). While the relationship with age did not reach significance ($\rho = 0.237$, $P = 0.151$), we found a significant association between transformation and performance ($\rho = 0.372$, $P = 0.022$), indicating more pronounced transformation in low-performing participants. Partial correlation analyses confirmed our initial finding that when controlling for performance, the correlation of age and transformation remained nonsignificant ($\rho = 0.144$, $P = 0.396$), while the correlation of performance and transformation was significant when controlling for age ($\rho = 0.325$,

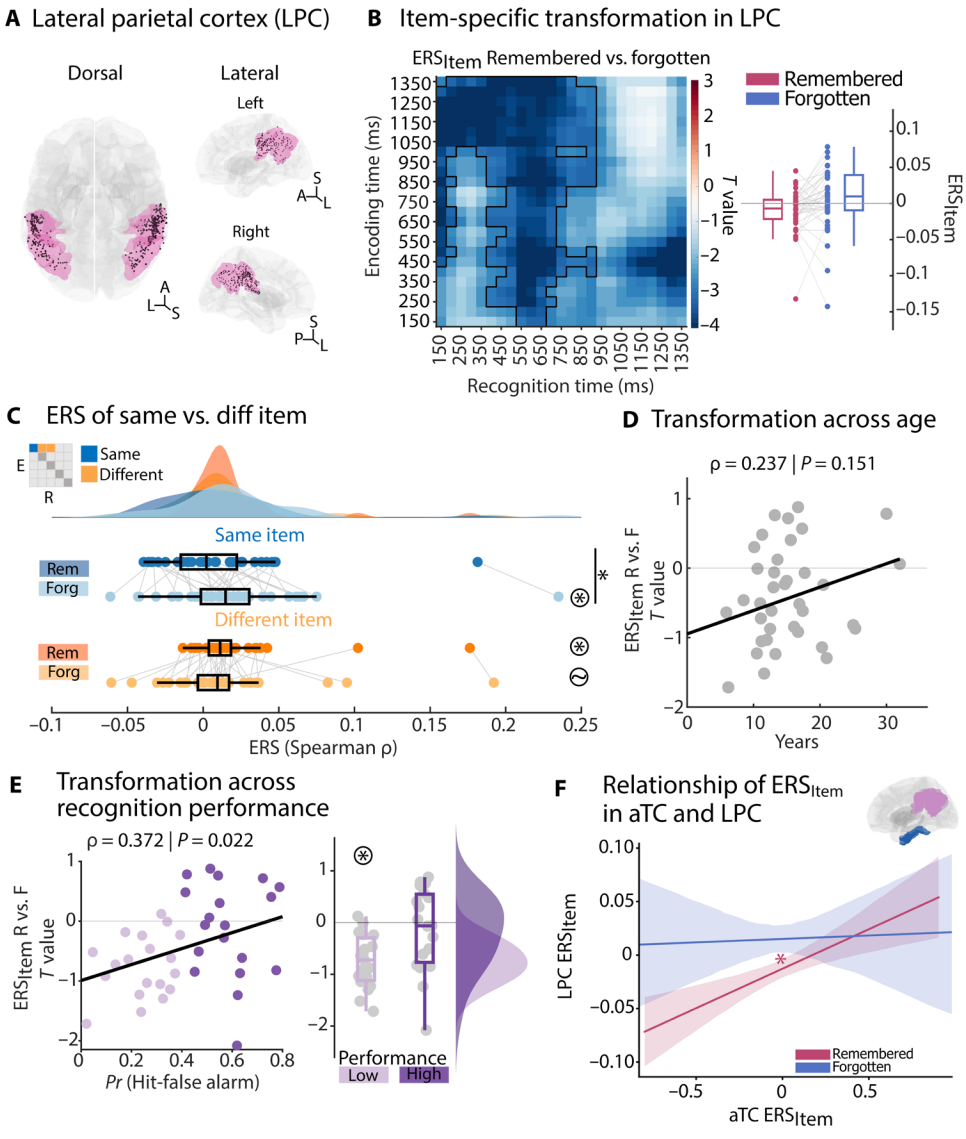

Fig. 4. Transformation of item-specific representations in lateral parietal cortex (LPC). (**A**) Contributing electrodes in LPC region of interest (ROI) (38 participants, 466 channels). D, dorsal. (**B**) Left, lower ERS_{Item} scores for remembered versus forgotten scenes. Color reflects *T* values across subjects for each time point. Black contour indicates cluster of time points during which ERS_{Item} values were significantly lower for remembered versus forgotten scenes after cluster-based correction for multiple comparisons across time points. Right, illustrative plot of ERS_{Item} similarities (Spearman's ρ) per subject for remembered (red) versus forgotten (blue) scenes, averaged across time points of significant cluster. (**C**) ERS_{Same} (blue, diagonal) and ERS_{Diff} (orange, off-diagonal) values of remembered and forgotten items showing a memory effect of ERS_{Same} but not ERS_{Diff} scores. (**D**) No significant relationship between transformation benefit for memory and participants' age. (**E**) Left, significant relationship between transformation benefit for memory and participants' performance. Colors indicate grouping following median split. Right, transformation effects in low- and high-performing participants (median split). (**F**) Relationship between ERS_{Item} magnitude in aTC and LPC: The magnitude of ERS_{Item} values in aTC is associated with ERS_{Item} magnitude in LPC. Circled * indicates $P < 0.05$ for one-sample *T* test. Line * indicates $P < 0.05$ for paired-sample *T* test. * indicates uncorrected $P < 0.05$ for main effect of LPC ERS_{Item} predicting aTC ERS_{Item}.

$P = 0.049$). A similar result was obtained when we applied a median split of corrected recognition performance rates *Pr* ($MD = 0.403$; transformation effect in low-performing participants: $T_{18} = -6.308$, $P < 0.001$; in high-performing participants: $T_{18} = -0.688$, $P = 0.499$; Fig. 4E and fig. S4).

Relationships between aTC reinstatement and LPC transformation

Given the observed dichotomy of aTC reinstatement and LPC transformation effects, we asked whether the two effects were related on a trial-by-trial level rather than both being associated with recognition memory success (Fig. 4F). We focused our analysis on the subgroup of $N = 28$ participants with electrodes in both ROIs ($N = 90$ aTC electrodes; $N = 362$ LPC electrodes). We extracted ERS_{Item} values averaged across time windows in which the ERS effects in aTC and LPC overlapped (see Materials and Methods) and compared these values across trials using mixed linear models with trial-wise aTC ERS_{Item} as criterion; LPC ERS_{Item}, memory success, stimulus category, and the interaction of LPC ERS_{Item}, and memory as predictors; and participants as random slopes. We found a significant

main effect of memory ($F_{1354} = 5.569$, $P = 0.018$), no main effect of LPC ERS_{Item} ($F_{1354} = 0.042$, $P = 0.837$), but a significant interaction of LPC ERS_{Item} with memory ($F_{1354} = 5.658$, $P = 0.017$). Follow-up analyses revealed a significant effect of LPC ERS_{Item} on aTC ERS_{Item} for remembered ($F_{872} = 14.795$, $P < 0.001$) but not forgotten items ($F_{482} = 0.044$, $P = 0.833$). This relationship indicates that the magnitude of ERS_{Item} values in LPC is associated with lower values of ERS_{Item} in aTC, suggesting that the two processes are related across individual trials.

Generalization and differentiation of memory representations during encoding and retrieval

Next, we sought to test the relevance of generalized versus differentiated representations for memory. First, we estimated the similarity of scene representations to all other scenes within the same experimental phase, i.e., the between-scene similarity during encoding and retrieval (EES and RRS; Fig. 2D) and tested whether similarities between remembered scenes differed from similarities between forgotten scenes, indicating differences in representational similarity. We found memory effects of EES in all VVS ROIs (Fig. 5), indicating more similar representations of subsequently remembered compared to forgotten scenes (Occ: $T_{sum} = 63.05$, $N_{bin} = 22$, $P_{cluster} < 0.001$, onset 300 ms, $EES_{rem} = 0.235$, $EES_{forg} = 0.205$; pTC: $T_{sum} = 53.32$, $N_{bin} = 18$, $P_{cluster} < 0.001$, onset at 500 ms, $EES_{rem} = 0.104$, $EES_{forg} = 0.086$; aTC: $T_{sum} = 7.14$, $N_{bin} = 3$, $P_{cluster} < 0.001$, onset at 600 ms, $EES_{rem} = 0.028$, $EES_{forg} = 0.013$). EES in the LPC was not related to memory.

During recognition, between-item similarities (RRS) showed memory-specific generalization in pTC ($T_{sum} = 13.08$, $N_{bin} = 5$, $P_{cluster} < 0.001$) but not in Occ or aTC (Fig. 5). Effects in pTC during retrieval occurred from 300 to 500 ms after scene onset and thus during an earlier period as compared to encoding. LPC representations showed an opposite effect between 500 and 600 ms ($T_{sum} = -6.999$, $N_{bin} = 3$, $P_{cluster} < 0.001$, $RRS_{rem} = 0.026$, $RRS_{forg} = 0.037$): Pairwise

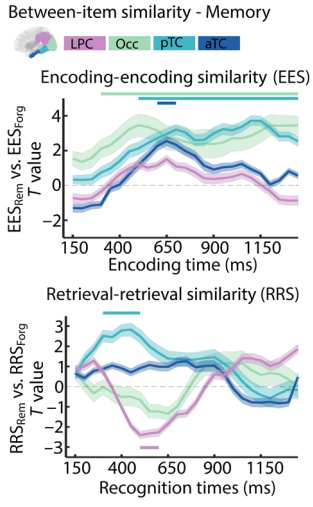

Fig. 5. Generalization and differentiation of memory representations. Memory effects of generalization (i.e., EES/RRS remembered > forgotten) or differentiation (i.e., EES/RRS remembered < forgotten) during encoding (EES, top) and retrieval (RRS, bottom). Colored horizontal lines indicate time points of significant memory effects after cluster-based correction for multiple comparisons across time points, separately for each region of interest (ROI).

similarities of remembered items were lower compared to forgotten items, indicating that differentiation but not generalization of LPC representations is associated with successful recognition. Further, we found no association of generalization or differentiation effects with development but with trial-level magnitudes of ERS_{Item} (for details, see the Supplementary Materials; fig. S5).

Representational formats during reinstatement and transformation

Last, we sought to quantify the occurrence of visual representational formats and their association with memory in aTC and LPC during timepoints where we previously reported reinstatement and transformation effects. For this, we used the neural representations between items encountered during encoding and retrieval. We capitalized on the representations in individual layers of the PlacesNet DNN (Fig. 6A and fig. S6) and compared them to the representations observed in the neural data (Fig. 6B). We specifically tested whether the correspondence between DNN and neural representations during encoding and retrieval was related to memory success.

To do so, we compared neural-DNN correspondences of remembered and forgotten items for all time points separately and performed cluster-based correction for multiple comparisons across time points. During encoding in aTC, we found higher correspondence for remembered versus forgotten items between 850 and 1000 ms with representations of layer fc7 ($T_{sum} = 9.472$, $N_{bin} = 4$, $P_{cluster} = 0.032$) and between 750 and 1000 ms with representations of layer fc8 ($T_{sum} = 14.702$, $N_{bin} = 6$, $P_{cluster} = 0.008$) (Fig. 6C, left). Both clusters fell into the encoding time periods showing reinstatement effects on memory. During retrieval, we did not observe any memory effects in neural-DNN correspondence (fig. S7). To better understand the memory-specific differences in neural-DNN correspondence in aTC, we repeated the analysis separately for remembered and forgotten items. For remembered items, we did not observe a significant correspondence between neural and cDNN representations across the entire encoding time period (Fig. 6C, middle, no time points with $P < 0.05$). In contrast, for forgotten items, most cDNN layers showed a significant negative correspondence (Fig. 6C, right): conv3 ($T_{sum} = -27.893$, $N_{bin} = 10$, $P_{cluster} < 0.001$), conv4 ($T_{sum} = -26.657$, $N_{bin} = 10$, $P_{cluster} = 0.005$), conv5 ($T_{sum} = -29.464$, $N_{bin} = 11$, $P_{cluster} = 0.010$), fc6 ($T_{sum} = -16.321$, $N_{bin} = 6$, $P_{cluster} = 0.048$), fc7 ($T_{sum} = -18.209$, $N_{bin} = 7$, $P_{cluster} = 0.008$), and fc8 ($T_{sum} = -18.925$, $N_{bin} = 7$, $P_{cluster} = 0.012$). These results indicate that correspondence of aTC representations to sensory representational formats as captured by the cDNN are associated with recognition memory failure, suggesting that other formats in this region are required for successful memory formation (for details, see the Supplementary Materials; fig. S8). Further, analyses in the other VVS regions showed that correspondence of DNN representations to pTC representations related to subsequent memory as well, while correspondence of DNN representations to Occ cortex representations were not related to memory (fig. S7). In LPC, we found no significant difference in the neural-DNN correspondence of remembered and forgotten trials during encoding (fig. S7). During recognition, we found that neural-DNN correspondence was significantly larger for forgotten as compared to remembered trials in layer conv5 ($T_{sum} = -11.689$, $N_{bin} = 5$, $P_{cluster} = 0.022$) between 400 and 600 ms (for details, see the Supplementary Materials).

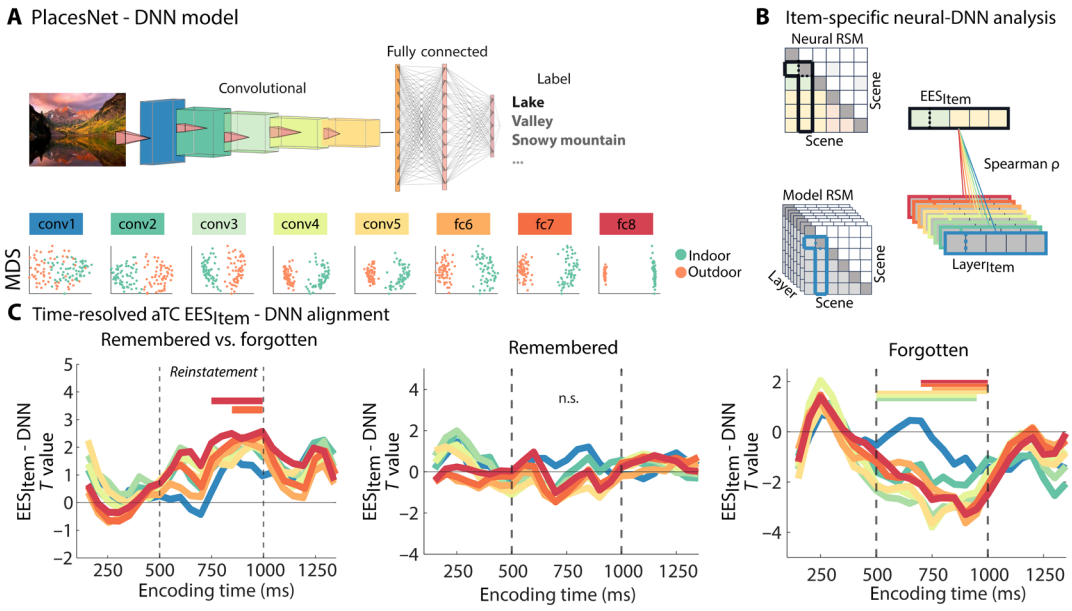

Fig. 6. Deep neural networks (DNNs) reveal memory-relevant representational formats. (**A**) Top, schematic of cDNN architecture trained to classify scenes (PlacesNet). Bottom, multidimensional scaling (MDS) plots for scene stimuli used in the experiment showing more pronounced categorical (indoor/outdoor) clustering of representations in higher versus lower DNN layers. (**B**) RSA comparing neural representational similarity matrices (RSMs) during encoding (EES) and retrieval (RRS) with RSMs from the individual DNN layers. During both encoding and retrieval, we correlated trial-wise RSMs observed in neural data with RSMs in each DNN layer and then conducted a follow-up test for differences of neural-DNN similarities depending on memory (i.e., remembered versus forgotten trials). (**C**) Time-resolved analysis of encoding-DNN correspondence for remembered versus forgotten items (left) and separately for only remembered (middle) or forgotten (right) items. Dashed vertical lines indicate on- and offset of encoding and retrieval time periods relevant for reinstatement in aTC. Horizontal colored bars indicate time points of cluster-corrected effects across time points, separately for each layer of the DNN.

Similar findings of neural-DNN correspondence in aTC during encoding and LPC during recognition were obtained when performing post hoc tests specifically on those time points showing reinstatement and transformation effects (fig. S9). In addition, we tested whether the negative neural-DNN correspondence of subsequently forgotten trials in aTC was associated with the magnitude of ERS_{Same} correlations in the ERS analysis. We found that the negative neural-DNN correspondence was more pronounced in the subgroup of forgotten trials that yielded low ERS_{Same} correlations, pointing toward shared influences across analyses (for details, see the Supplementary Materials; fig. S10). Notably, given the exploratory nature of these analyses, interpretation of effects based on individual layers should not be overemphasized. Nevertheless, the overall pattern of results of memory-specific alignment of separable visual formats during encoding and retrieval may provide additional insights into the nature of the underlying memories and how they unfold during encoding and recognition.

DISCUSSION

The formation and retrieval of memory traces is at the heart of adaptive behavior and involves the recruitment of multiple distinct neural mechanisms acting upon specific item-related feature formats (*67*, *68*). Here, we used direct electrophysiological recordings to investigate the functional roles of reinstatement and transformation for recognition memory, the relevance of distinct representational formats, and the unfolding of these processes in a sample of children, adolescents, and young adults.

We tested whether item-specific representations between encoding and retrieval (ERS) indicated whether memory is supported by reinstatement or transformation, reflected in higher or lower levels of ERS between remembered and forgotten trials. In line with previous findings, we found beneficial effects of reinstatement for memory but also showed that these effects were specific to the aTC. Reinstatement related to temporally restricted encoding activity patterns (500 to 1000 ms), pointing toward the relevance of specific representational formats (*36*). Consistent with findings in previous studies (*69*), we found that reinstatement occurred during earlier time windows during retrieval as compared to encoding, in line with a reversed processing hierarchy in the VVS between perception (posterior-to-anterior) and memory (anterior-to-posterior) (*70*, *71*). We did not observe any memory effects of reinstatement in early- and mid-level VVS regions, suggesting that only the reinstatement of mnemonic representations is associated with successful recognition memory. Further analyses on representational formats in aTC using a cDNN indicated a diminished relevance of sensory formats compared to pTC and that recognition memory decisions may be associated with the reinstatement of rather mnemonic but not purely sensory representations, whereas regions presumably carrying low-level sensory content were not associated with memory-specific reinstatement.

Further, reinstatement of item-specific representations in aTC as reflected in higher ERS_{Item} correlations of remembered as compared to forgotten scenes could potentially relate to recollection-based processes of recognition that have been shown in temporal cortex regions before (*11*, *16*) and that presumably reflect the reactivation of contextual or item-specific information. Notably, the reinstatement

effect appeared to be driven predominantly by a drop of ERS_{Same} correlations of forgotten items rather than particularly high ERS_{Same} correlations of remembered items—although the former did not differ from zero while the latter did, ERS_{Diff} correlations of both remembered and forgotten items were larger than zero as well, suggesting that the ERS of remembered items was at least partially due to reoccurrence of unspecific processes. In addition, follow-up analyses of the forgotten items suggest that the low ERS_{Same} values were related to their negative match with DNN representational formats, suggesting that they were ambiguously processed in the aTC. However, ERS_{Diff} correlations did not differ between remembered and forgotten items, and, indeed, the initial finding of reinstatement was defined by a significant interaction between ERS_{Same} versus ERS_{Diff} correlations of remembered versus forgotten items, pointing toward the specificity of memory effects to activity patterns elicited by specific items. Thus, while reinstatement appeared to be related to a peculiar drop in ERS_{Same} values of some forgotten items that showed negative correlations with DNN representational formats, these effects did not occur for remembered items, preserving the similarity of the activity patterns of these remembered items during encoding with their patterns during retrieval. In other words, while the successful engagement of memory processes for the remembered items is not associated with prominent item-specific representations in the aTC that reoccur during retrieval, the putative "mnemonic" representational formats of these items correspond to similar activity patterns during encoding and retrieval of individual items that are absent for forgotten items. Future studies should aim to further characterize the putative mnemonic representational formats of remembered items and to clarify the representational mechanisms underlying the unexpected negative alignment of forgotten items to visual representational formats captured by DNNs.

In contrast to the memory effects of reinstatement in aTC, we found evidence for a transformation of item-specific representations in LPC (*28*, *29*). Stimulus-related activity patterns during retrieval resembled those during encoding if participants did not remember having seen the scene before, while successful remembering elicited activity patterns that were unrelated to those during encoding. LPC has emerged as a prominent region implicated in retrieval-related accounts of episodic memory, reflecting retrieval success (*30*, *72*, *73*). Accordingly, some studies reported negative subsequent memory effects in this area (*74*). Being centrally located at the temporo-parietal junction, the LPC receives inputs from various sources including the dorsal and ventral visual processing streams as well as prefrontal and MTL structures (*30*). Two theoretical frameworks on LPC function, the AtoM (attention-to-memory) (*75*) and the CoBRA (cortical binding of relational representations) (*73*) models, both emphasize its close link with MTL structures but diverge in their interpretation concerning its underlying function. The AtoM model focuses on the relevance of bottom-up versus top-down attentional accounts (*76*) and ascribes bottom-up–driven orientation toward mnemonic representations to the ventral posterior parietal cortex. However, this model does not assume that the parietal cortex represents specific contents itself, which has been challenged by recent findings of content-based information decoding (*28*, *77*) and the representation of engrams (*78*) in this region.

The CoBRA model argues for an integrative role of LPC serving the binding of object-based, semantic, and contextual information into multimodal formats. This may fit to our findings since the higher difference between encoding and retrieval activity patterns of remembered items may reflect the integration of multimodal stimulus features that ultimately results in transformed representations during retrieval, which may be selectively recruited to aid recognition memory decisions. Following earlier regard that reinstatement in aTC may relate to recollection-based processes during recognition, transformation in LPC may possibly reflect familiarity-based decisions of recognition memory, which reflect a modified response to an item upon its repeated encounter in a signal-detection process (*79*). Presumably, familiarity would reflect low- or near-zero similarities between encoding and retrieval, a pattern that we observe for ERS_{Same} correlations. Further, earlier studies have already described LPC to be involved in familiarity-based decisions of recognition memory (*80*, *81*) as well as post-retrieval mechanisms of cognitive control through interactions with temporal cortices (*82*).

Transformation of item-specific representations in LPC also contradicts previous ERS findings: Using fMRI to test emotional and neutral scene recognition memory, a previous study found that parietal reactivation rather than transformation predicted successful recognition memory (*18*). Nevertheless, meta-analytic results on parietal involvement in episodic memory suggest that inferior lateral parietal cortices are more strongly involved in retrieval- versus encoding-related processes (*83*) in line with distinct involvement during encoding and retrieval, suggesting transformation. On a more general level, an alternative account on the pattern of results in LPC (i.e., a lack of ERS_{Same} correlations for remembered items, positive ERS_{Same} correlations for forgotten items, and no link to DNN representational formats for either items) may be that activity patterns in this region do not reflect stimulus-specific memory traces in the first place but are rather due to item-unspecific cognitive processes that only show overall differences between remembered and forgotten items. For example, all forgotten items may induce similar activity patterns during encoding and retrieval that reflect, e.g., lack of attentional engagement or mind-wandering to unspecific associative contents, resulting in positive ERS_{Same} correlations. By contrast, activity patterns of remembered items may reflect successful engagement of encoding and retrieval modes during the respective experimental phases; these modes are associated with different activities in LPC (*28*), which may explain the low and nonsignificant ERS_{Same} correlations. However, while we cannot entirely rule out this interpretation, we believe that it is less likely to fully account for our data for several reasons. First, we did not observe memory-related differences in overall spectral power in the LPC, and, thus, the engagement of unspecific processes would need to manifest selectively in the distribution of power across electrodes, which may appear unlikely (though not impossible) given that different patients were implanted in different subregions of LPC. Second, although activity patterns of forgotten items did not match DNN representational formats, we only considered trials with correct category classification during encoding, ensuring allocation of at least some attentional resources toward the stimulus. Third, if encoding and retrieval processes were indeed entirely unspecific to individual items, one would expect similar differences between remembered and forgotten items for ERS_{Diff} correlations as well, which we did not find—our main finding was based on a significant interaction between ERS_{Same} and ERS_{Diff} correlations of remembered versus forgotten items.

Our findings show that both reinstatement and transformation promote successful recognition but are supported by distinct brain regions, suggesting the involvement of dissociable neural mechanisms. The aTC is located at the apex of the VVS and represents

sensory information at a conceptual level, whose reinstatement has previously been shown to benefit memory (*36*, *45*). By contrast, the integrated multimodal representations in LPC may be primarily accessible during retrieval rather than encoding, in line with our finding that this region contains information about scenes in a transformed representational format. Evidence for item-specific representations (ERS_{Item}) in LPC occurred particularly when the magnitude of reinstatement in aTC was low, indicating interrelated processes between the two regions. Further, comparisons of trial-level ERS_{Item} values in aTC and LPC revealed an association between the magnitudes of item-specific information specifically for remembered items: The magnitude of item-specific representations (ERS_{Item}) in LPC depended on aTC, suggesting interrelated processes between the two regions. Thus, temporal and parietal cortices may represent distinct types of information with unique contributions to successful recognition, which may flexibly and distinctly influence memory decisions. We did not find memory-related differences in spectral power in these two regions. This likely reflects content-specific representations of individual items within the distribution of power across frequencies quantifiable using RSA rather than conditional differences in predefined frequency bands. In addition, the lack of a memory effect in spectral power analyses suggest minor effects of differences in response times, which would equally affect memory-specific findings at all levels of analysis.

Are reinstatement and transformation processes necessary for memory? Of course, our study does not provide direct information on the causal role of any of the observed processes. Our findings suggest that reinstatement is a general and beneficial mechanism for memory retrieval independent of developmental gains in memory, while recognition memory may depend on transformation. Although we find that transformation of item-specific representations is associated with successful recognition on the trial-level, this does not necessarily imply that this association exists in every participant. We found this effect specifically in a subgroup of participants with low recognition memory performance. This suggests that recruitment of memory transformation in the LPC is a less effective and possibly compensatory strategy compared to reinstatement in aTC. Recognition memory can rely on different processes of recollection and familiarity (*31*, *79*, *82*, *84*), which may be differently associated with VVS reinstatement and LPC transformation effects (*11*, *16*) and differently recruited by high versus low-performing participants. Nevertheless, given the relationship between the two effects, we speculate that neither of them may be sufficient for memory retrieval alone but that they may flexibly aid retrieval by task-relevant recruitment of distinct representational formats. Here, the characteristics of the memory trace determine how and by which features a familiar scene may be recognized. Alternatively, both reinstatement and transformation enable the representation of information in complementary formats that contribute differently to memory decisions depending on performance. According to this perspective, transformation processes may, in principle, occur in high-performing participants as well but are less relevant for them to conduct the relatively simple recognition memory task in our study. Notably, performance in our sample is likely influenced by a diverse set of factors, such as age-related differences in metacognitive capacities, the utilization of mnemonic strategies, or demands on compensatory mechanisms to overcome limited memory capacity. Although structural or functional deficits due to epilepsy cannot be ruled out, previous findings point toward comparable memory processes in epilepsy and healthy control groups (*64*, *85*, *86*).

We hypothesized that because of reinstatement and transformation processes, aTC and LPC contain distinct representational formats that differ between the two regions. Although we found a subsequent memory effect in Occ high-frequency broadband activity, and previous research links Occ gamma to early visual DNN formats (*39*), correspondence of representations in the Occ lobe and in the DNN was functionally irrelevant to memory in our study. However, our DNN results show a contribution of intermediate DNN formats to memory traces in pTC and more abstract formats in aTC. These results suggest a dissociation between the role of overall processing in early visual cortex (that may, e.g., reflect selective attention and thereby support memory indirectly) and the formation of the memory trace itself (which appears to rely on representational formats of higher-level visual areas). Further, similar to our findings in aTC, a previous study applied DNNs to investigate perirhinal cortex, an MTL region situated at the apex of the VVS, and found that it enables perceptual behaviors beyond VVS contributions, arguing for more abstract and potentially mnemonic representations (*87*) in this region. During retrieval, we found no memory-specific DNN correspondence with representations in aTC or upstream VVS regions. One may speculate that this is due to a diminished comparability between behavioral and network representations due to different task demands in encoding (scene classification) and retrieval (recognition). However, the correspondence between neural and DNN representations in the LPC during recognition tentatively suggested a memory effect that was influenced by a negative matching of neural and cDNN representations. We are currently unable to provide a full explanation to this effect. While negative correlations between related items have previously been described as "repulsion" of mnemonic representations (*14*, *88*–*90*), these effects were associated with a functional benefit for memory, different from our data. Alternative explanations should thus be considered, e.g., that later forgetting is associated with processing of noninformative features during encoding. Notably, from our aTC ERS and EES-DNN analyses, we cannot directly infer that the described representational formats were also relevant for ERS because both types of analyses consider different cross-correlation matrices: Whereas ERS compared neural features of same versus different items, EES reflects representational distances between different items. Nevertheless, we observe a temporal (latency) and spatial (ROI) overlap in memory-specific differences in both types of analyses, which suggest that both effects are supported by overlapping representational and neurophysiological features. In the future, it may be promising to compare neural representations during retrieval to representations in generative DNNs such as variational autoencoders (*91*) or in recurrent networks trained on memory tasks (*92*).

How are reinstatement and transformation related to the generalization or differentiation of memory representations? We find that regions in the VVS showed memory-specific generalization during encoding. Further, we additionally found a reinstatement of memory representations in aTC. In LPC, we observed memory-specific differentiation during recognition and a transformation of representations between encoding and retrieval. According to theoretical accounts, both generalization and differentiation could be beneficial: While generalized representations allow inferences to novel experiences (*33*), distinct representations improve their separation from the representations of similar events (*93*). We found that generalization but not differentiation of VVS representations during encoding related to memory success. This effect may be explained by pronounced

extraction of conceptual information that overlaps across multiple items, in line with pronounced correspondence to visual representational formats of subsequently forgotten but not remembered trials. Alternatively, generalization of representations may happen either because of the integration of all subsequently remembered items into a unified episodic context, i.e., reflecting episodic binding (*94*–*97*), or because of increased similarity of memory representations between all studies items, reflecting global matching (*33*). Accordingly, it has been shown that increased between-item similarity in high-, but not low-level, sensory areas during encoding is associated with successful memory and vividness during recall and levels of cortical reinstatement (*34*). By contrast, during retrieval, we found that representations of remembered items in the LPC were more differentiated than those of forgotten items. Notably, we observe generalization and differentiation effects in regions and during task phases not showing conditional differences in spectral power (i.e., in aTC and pTC during encoding and in LPC during recognition), suggesting memory-specific differences at the level of item representations beyond domain-general effects.

There are a few limitations to our study that require mentioning and that could be addressed in future research. First, although we have described the functional relevance of distinct representational formats during successful memory formation and recognition, our paradigm did not involve the experimental manipulation of low- and high-level sensory formats. Explicit manipulation of such formats could deliver more direct evidence for which formats are necessary for reinstatement or transformation and where they are represented in the brain. Second, although recordings from grid electrodes offer extensive sampling from cortical surface structures, our dataset did not allow fine-grained analyses of deeper brain structures in MTLs such as the hippocampus. More extensive MTL recordings would provide valuable insights into the involvement of hippocampal activity and its mechanistic function in the coordination of reinstatement and transformation-related processes in neocortical areas. This would especially be interesting to further characterize the differential contribution of sensory and mnemonic effects during recognition memory. Third, sensory and mnemonic effects could be addressed in future studies using only partial cues, item-based, and source memory tests during recognition, more detailed behavioral responses in the form of continuous memory ratings or recollection/familiarity ratings.

To conclude, we demonstrate that successful recognition memory is associated with both the reinstatement and transformation of item-specific representations, which are located in the ventral temporal and lateral parietal cortices, respectively. Our findings further suggest that the generalization and differentiation of item representations across these regions contribute to successful memory formation and retrieval, likely due to the relevance of distinct representational formats. Our study sheds light on the multifaceted nature of recognition memory and shows that this seemingly simple cognitive function involves multiple operations acting on different aspects of the memory trace.

MATERIALS AND METHODS

Patient sample

In total, $N = 48$ patients participated in the experiment. All patients suffered from pharmacoresistant epilepsy and were implanted with invasive electrocorticography (ECoG) or stereo-EEG (sEEG) electrodes for presurgical localization of epileptic foci. Implantations were conducted at the Children's Hospital of Michigan, hospitals of the University of California (UC) San Diego, UC Irvine, UC San Francisco, and the California Pacific Medical Center. Sites of implantation were chosen by medical staff and based solely on clinical needs of the patient. Two participants were excluded from all analyses because of poor memory performance. The final sample consisted of $N = 46$ patients (20 female, 26 male) aged 5.9 to 32 years (15.61 $\pm$ 5.92 years; for details, see Fig. 1). The institutional review boards of the Wayne State University (no. 048404MP2E), UC Irvine and UC San Diego (no. HS# 2014-1522), UC San Francisco (no. 10-03842), and the California Pacific Medical Center (no. 666687-17) approved the study in accordance with the Declaration of Helsinki. Written informed consent was obtained from patients aged ≥18 years and from the guardians of patients <18 years; written assent was obtained from patients aged 13 to 17 years; oral assent was obtained from younger children.

Behavioral task

Participants completed a scene recognition memory paradigm that has been used to study memory with pediatric samples (*58*–*65*). During encoding, the participants were presented with equal numbers of indoor/outdoor scenes (each scene presented for 3 s) and instructed to verbally indicate indoor/outdoor as well as to try and memorize the scenes for a later memory test.

Each trial started with a fixation cross of 500 ms, followed by 3 s of scene presentation. For all subsequent analyses, we only considered trials with correct indoor/outdoor classification responses during encoding. During the subsequent recognition memory test, the participants were presented with old and new (foil) scenes and instructed to verbally indicate their memory judgement (i.e., "old" or "new"). Each trial started with a fixation cross of 500 ms, followed by self-paced duration of scene presentation.

Each encoding block consisted of $N = 40$ trials and was followed by a recognition memory test in which the participants saw all 40 old scenes intermixed with 20 new (foil) scenes. Forty-three percent ($N = 20$) of the participants completed one block, and 57% ($N = 26$) completed two encoding-recognition blocks. As a measure of memory performance, we computed corrected recognition scores *Pr* (*98*), i.e., rate of hits – rate of false alarms. $N = 2$ participants with negative recognition scores were excluded from further analyses. For all analyses, we only considered trials with artifact-free neural data during both encoding and recognition (49 $\pm$ 18 trials per participant).

Data acquisition and preprocessing

ECoG and sEEG data were acquired using a Nihon Kohden, Natus, or Tucker-Davis Technologies system at a minimum sampling rate of 1 kHz, and data acquired at higher sampling rates were resampled to 1 kHz after filtering. Data were filtered offline using a finite impulse response filter (0.1-Hz high-pass and 300-Hz low-pass). Line noise (60 Hz) and its harmonics (120, 180, and 240 Hz) were removed by applying a narrow-band notch filter to a discrete Fourier transform as implemented in fieldtrip (*99*). We excluded electrode contacts that showed pronounced epileptiform activity, overlapped with clinically defined seizure onset zones, or contained distinct artefactual signals such as poor contact. Trial-level artefact rejection was done manually by trained experts. Continuous data were then epoched into individual trials from −1 to +3 s around scene onset, or longer if response times exceeded the duration of scene presentation.

Electrode implantation, localization, and bipolar referencing

ECoG patients ($N = 34$) were implanted with subdural platinum electrodes (10-mm inter-contact distance; 4-mm diameter) placed directly onto the surface of the cortex. Here, individual electrode contacts were contained in larger grids (e.g., 8 by 8 or 2 by 5) or in one-dimensional (1D) strips of various length (4 to 12 contacts). SEEG patients ($N = 12$) were implanted with stereotactic depth electrodes (5-mm inter-contact distance) of various lengths (4 to 16 contacts). Implantation sites varied between patients based on clinical considerations. Implantations were performed in left ($N = 17$), right ($N = 18$), or both ($N = 11$) hemispheres. In all analyses reported here, we collapsed across ROIs of both hemispheres and treated ECoG/subdural and sEEG/depth electrodes in all preprocessing and analysis steps equally.

For localization of electrode positions, 3D reconstructions were created by coregistering postimplantation planar CT images of the cortical surface with preoperative T1 MR images (*100*). After reconstruction of implantation schemes, we created patient-specific templates for bipolar referencing. To do so, we interpolated locations of virtual electrodes from pairs of adjacent electrodes lying on the same depth electrode or on the same strip of grid electrodes. For 2D-grid electrodes, we sampled across both rows and columns of the grid. Only virtual electrodes based on two artefact-free neighbors were included into further analyses. We excluded bipolar electrodes where both neighbors were in white matter or ventricles. Montreal Neurological Institute (MNI) coordinates of all virtual electrodes were used to derive labels of Brodmann's areas using BioImage Suite MNI2TAL toolbox (*101*). The final coverage contained $N = 3691$ bipolar referenced, artefact-free virtual electrodes, distributed over 32 unique Brodmann areas. Whole sample coverage for all ROIs analyzed throughout this work is displayed in Fig. 1E.

We grouped Brodmann areas into distinct ROIs for further analysis: Occ cortex (Occ; Brodmann areas 17, 18, and 19) including calcarine sulcus with area striata (BA17; V1), as well as secondary (BA18; V2) and tertiary (BA19; V3-5) association cortices; pTC (Brodmann area 37), which corresponds to posterior parts of the fusiform gyrus; aTC (Brodmann area 20) posited rostral to pTC covering large parts of the ventral temporal cortex and bound medially but excluding hippocampus, peri-, entorhinal, and parahippocampal cortices (BA34, 35, 36); LPC (Brodmann areas 39, 40) corresponding to gyrus angularis (BA39) and gyrus supramarginalis (BA40). Note that our selection of ROIs was relatively broad to allow sufficient sampling of electrodes across participants. Similar divisions of VVS into functionally distinct subregions have been used in earlier studies (*39*, *46*).

Spectral decomposition and time-frequency features for RSA

First, we applied a band-pass filter (0.1-Hz high-pass, 150-Hz low-pass) to raw and artefact-free trial segments to restrict the data to the frequency range of interest. To avoid edge artefacts in spectral decomposition, we mirrored individual time-series data segments by appending flipped trial-data to the beginning (from −5 to −1 s) and end (from 3 to 7 s) of each segment. For time-frequency decomposition, we applied a Morlet wavelet transformation to data segments from −500 to 1500 ms. We used wavelets from 1:1:150 Hz with a linear increase from 3 to 6 cycles for frequencies 1 to 30 Hz, and 6 to 12 cycles for frequencies 31 to 150 Hz. We normalized poststimulus periods by their relative change to a prestimulus baseline from −500 to −200 ms before stimulus onset. We then extracted the power of frequencies from 1 to 30 Hz in 1 Hz steps and averaged subsequent frequencies in 5 Hz bins from 31 to 150 Hz (31 to 35 Hz, 36 to 40 Hz, etc.), resulting in 54 distinct frequencies. We then averaged data in temporal windows of 300-ms length, sliding in 50-ms steps from 0 to 1500 ms. For all reports, each window is assigned to its middle time point (e.g., the first window expands from 0 to 300 ms and is indicated as 150 ms). For the range of 0 to 1500 ms, this resulted in 25 distinct windows. Electrode-specific time-frequency representations of individual trials resulted in n (channels) × 54 (frequencies) × 25 (time steps) values that were entered into subsequent RSA.

Representational similarity analysis

Here, we report results from two different sets of RSA (*10*), which are computed with the same neural features derived from preprocessed electrophysiological data. We estimated the pairwise similarity of trials either within the same experimental phase (EES; RRS) or between experimental phases (ERS) using Spearman correlation (Fig. 2, A to D). As features for similarity estimation, we used the distribution of power values across frequencies and electrodes in each ROI. If, for example, a participant contributed three (bipolar) electrodes to a certain Brodmann area, the trial-specific spectral power across frequencies of all three channels (3 electrodes × 54 frequencies) from encoding were concatenated into a 1D vector of $3 \times 54 = 162$ data points and correlated with the neural data of the same and of all different scenes in either the same or a different experimental phase. This analysis was done for various time windows (see above).

Cluster-based permutation statistic

To correct for multiple comparisons in all RSA, we applied cluster-based permutation statistics (*102*). We repeated each first-level computation during RSA (e.g., for a specific time bin of RSA feature data) after shuffling the trial-specific labels (remembered/forgotten trials) while keeping the amount of data per category intact. For example, when performing a first-level independent-samples T test of ERS_{Item} between remembered and forgotten trials, in each permutation, we randomly assigned the same ERS_{Item} values to each memory type while keeping the original number of trials per category intact. This procedure was conducted for 1000 iterations. We then corrected for multiple comparisons across time by extracting clusters of adjacent time points with uncorrected group-level P values <0.05 within the time (1D) or time × time (2D) maps. Depending on the type of analysis, we looked for temporally contiguous clusters in 1D temporal progression (EES, RRS, and DNN-EES) or in 2D temporally generalized data (ERS). For each identified empirical or surrogate cluster, we summed all T values contributing to it. Last, we computed the probability of observing the empirical cluster by extracting its rank from the distribution of all clusters in the same direction (positive or negative) in the surrogate data (P = rank/1000). Permutation statistics for memory-dependent differences in time-frequency spectra was conducted using fieldtrip toolbox. RSA analyses used custom MATLAB scripts.

Between-item similarity (EES/RRS)

For the between-item similarity analysis (see Fig. 3D), we computed the pairwise similarity of scenes within the same experimental phase (EES/RRS). Separately for each time point of the sliding window, we extracted trial-specific time-frequency features across electrodes and

estimated their similarity using Spearman correlation. This yielded a symmetric identity stimulus × stimulus matrix where the diagonal is the similarity of a scene to itself. Obtained matrices were Fisher-*Z* transformed before further analysis. All analyses were performed on the lower-diagonal values of the symmetric similarity matrices. For each matrix and time point, we tested whether pairwise similarities of remembered versus forgotten scenes differed. To do so, we computed an independent-samples *T* test of pairwise similarities of scenes from the same memory category (all remembered-remembered and forgotten-forgotten scene pairs) for each participant. On the second level, we tested whether subject-specific first-level *T* values consistently deviated from zero. In the time series of group-level effects, we then looked for adjacent time points forming temporal clusters and estimated their statistical significance using cluster-based permutations statistics. Surrogate first-level data were obtained by shuffling the labels of trials (remembered versus forgotten) and repeating the analysis with shuffled data 1000 times.

Encoding-retrieval similarity

In the second line of RSA, we investigated memory-specific representational reinstatement or transformation. Specifically, we estimated the neural similarity between the representation of a specific scene during encoding and recognition (ERS). This analysis yielded a cross-correlation matrix of scenes (dimension: stimulus × stimulus). Matrices were Fisher-*Z* transformed before further analyses. In this matrix, the diagonal elements denote similarities between the same scene during encoding and recognition, while the off-diagonal elements denote similarities between different scenes. Since we applied a temporal sliding-window approach, the stimulus × stimulus matrix was calculated for all possible combinations of time points, yielding a 4D matrix of ERS values (dimensions: stimulus × stimulus × encoding time × recognition time). This time-resolved approach allowed us to investigate whether neural representations occurring at a specific time point during encoding correlated with representations at a specific—possibly different—time point during recognition.

Next, we quantified the item-specific neural similarity between encoding and recognition by taking the similarity score of the same scene during encoding and recognition (same item, diagonal element) and subtracting the average similarity of this scene during encoding with all other scenes during recognition (different item, row-wise off-diagonal elements), i.e. $ERS_{Item} = ERS_{Same} - avg.(ERS_{Diff})$. Crucially, ERS_{Diff} considered values from the same scene type (indoor/outdoor) and memory category (remembered/forgotten). We replicated ERS results for aTC and LPC when considering only column-wise off-diagonal elements, as well as row and column-wise off-diagonal elements (fig. S11).

Does the degree of scene-specific similarity between encoding and recognition support memory performance? To answer this question, we performed an independent-samples *T* test of ERS_{Item} scores, comparing their magnitude across all remembered and forgotten trials within each participant. This yielded one map of *T* values for each participant across the 25 encoding time points and the 25 recognition time points, which was used for subsequent group level analysis. On the group level, we performed a one-sample *T* test across the *T* values of all participants against zero. This analysis was conducted for all 25 × 25 encoding/recognition time points. We then extracted all *T* values corresponding to *P* values <0.05 and summed them across all adjacent time points for estimation of statistical significance (empirical cluster value).

Item specificity in aTC and LPC (ERS – ERS)

We tested the relationship of the magnitude of ERS_{Item} between aTC and LPC using hierarchical linear mixed models. This analysis was conducted on the subset of participants ($N = 28$) contributing electrodes to both regions. We averaged ERS_{Item} values across overlapping time points, i.e., encoding-retrieval time bins with significant ERS effects in aTC (Fig. 3B) and LPC (Fig. 4B), $N_{bins} = 33$, encoding time range 500 to 1000 ms, retrieval time range 400 to 900 ms) of reinstatement and transformation effects and then predicted the magnitude of aTC ERS_{Item} by LPC ERS_{Item}, included a main effect of memory and modeled subject-specific slopes [aTC ~ LPC * memory + (1|subject)].

DNN model

To classify scene images shown throughout the experiment, we used a pretrained convolutional neural network "PlacesNet (*103*)." PlacesNet was trained on the Places database (*104*), incorporating 2.5 million scene images of 205 category labels. cDNNs with the same network architecture [e.g., AlexNet (*105*)] trained on a variety of object classes have been found to exhibit representational geometries corresponding to neural representations observed throughout the VVS (*38*, *39*, *43*, *106*). PlacesNet consists of five convolutional and three fully connected layers (Fig. 6A). Information is initially processed on pixel-by-pixel level in convolutional layer 1 and is then unidirectionally transformed to downstream layers, ultimately leading to the assignment of semantic labels in the output layer (softmax). We averaged the values in each convolutional layer for each image over the spatial dimension (*107*) (retinotopic units) resulting in one value per feature (images × features). Extracted pairwise Spearman's ρ estimates based on internal feature activations yielded symmetric layer-specific representational similarity matrices (RSMs) of all 120 by 120 scene images, resulting in 8 RSMs of 120 by 120 images (see fig. S6C). RSMs were Fisher *Z* transformed for further analyses.

Item-specific alignment of neural and DNN model features

RSMs with corresponding structure to RSMs from DNN layers were obtained from EES and RRS analyses (scene*scene), thus allowing for a direct comparison of representational structure between iEEG ROIs and PlacesNet. EES/RRS matrices were obtained for each individual time point during encoding and retrieval, resulting in 25 scene*scene matrices per participant, ROI, and experimental phase. Since internal features of feed-forward models are insensitive to temporal dimensions, we obtained one RSM for each layer of the network, resulting in eight model matrices equally used for encoding- and retrieval related neural data.

To match the item-specific approach in the previous ERS analysis, we computed scene-specific alignment of neural features with each of the network layers [see (*36*, *45*) for a similar approach]. For each scene, we extracted the vector of similarity values with all other scenes (vector of length N_{scenes}-1, e.g., one row in matrix) from neural and DNN matrices. These vectors describe the representational geometry of the current scene to all other scenes. We then computed the Spearman correlation of neural and model vectors separately for each time point and network layer, yielding a score describing the item-specific similarity of neural and network representations. The subsequent statistical procedure was similar to ERS analysis: First, we computed independent-samples *T* tests between the alignment scores of remembered and forgotten items to test whether shared similarities differed depending on memory at any time point. On

the second level, we tested across participants whether these T values differed significantly from zero. Then, we identified adjacent time points with group-level T values corresponding to P values <0.05 and performed correction for multiple comparisons across adjacent time points by comparing the clusters in the empirical data with clusters from surrogate data, where the order of scene-specific similarity vectors from the neural data was randomly shuffled before comparison with scene-specific vectors from individual DNN layer matrices.

Supplementary Materials

This PDF file includes:

Supplementary Text
Figs. S1 to S11
References

REFERENCES AND NOTES

1. R. Semon, *(1921). The Mneme.* (London: Allen, 1921).
2. S. A. Josselyn, S. Köhler, P. W. Frankland, Heroes of the Engram. *J. Neurosci.* **37**, 4647–4657 (2017).
3. M. J. Ison, R. Quian Quiroga, I. Fried, Rapid encoding of new memories by individual neurons in the human brain. *Neuron* **87**, 220–230 (2015).
4. U. Rutishauser, T. Aflalo, E. R. Rosario, N. Pouratian, R. A. Andersen, Single-neuron representation of memory strength and recognition confidence in left human posterior parietal cortex. *Neuron* **97**, 209–220.e3 (2018).
5. L. R. Squire, S. Zola-Morgan, The medial temporal lobe memory system. *Science* **253**, 1380–1386 (1991).
6. S. A. Josselyn, S. Tonegawa, Memory engrams: Recalling the past and imagining the future. *Science* **367**, eaaw4325 (2020).
7. Z. M. Reagh, C. Ranganath, Flexible reuse of cortico-hippocampal representations during encoding and recall of naturalistic events. *Nat. Commun.* **14**, 1279 (2023).
8. X. Liu, S. Ramirez, P. T. Pang, C. B. Puryear, A. Govindarajan, K. Deisseroth, S. Tonegawa, Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature* **484**, 381–385 (2012).
9. N. T. M. Robinson, L. A. L. Descamps, L. E. Russell, M. O. Buchholz, B. A. Bicknell, G. K. Antonov, J. Y. N. Lau, R. Nutbrown, C. Schmidt-Hieber, M. Häusser, Targeted activation of hippocampal place cells drives memory-guided spatial behavior. *Cell* **183**, 1586–1599.e10 (2020).
10. N. Kriegeskorte, Representational similarity analysis – connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
11. B. P. Staresina, S. Michelmann, M. Bonnefond, O. Jensen, N. Axmacher, J. Fell, Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *eLife* **5**, e17397 (2016).
12. L. Kunz, L. Deuker, H. Zhang, N. Axmacher, Chapter 26 - Tracking human engrams using multivariate analysis techniques in *Handbook of Neural Plasticity Techniques*, D. Manahan-Vaughan, Ed. (Elsevier, 2018), pp. 481–508.
13. D. Pacheco Estefan, M. Sánchez-Fibla, A. Duff, A. Principe, R. Rocamora, H. Zhang, N. Axmacher, P. F. M. J. Verschure, Coordinated representational reinstatement in the human hippocampus and lateral temporal cortex during episodic memory retrieval. *Nat. Commun.* **10**, 2255 (2019).
14. D. Pacheco Estefan, R. Zucca, X. Arsiwalla, A. Principe, H. Zhang, R. Rocamora, N. Axmacher, P. F. M. J. Verschure, Volitional learning promotes theta phase coding in the human hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021238118 (2021).
15. J. F. Danker, J. R. Anderson, The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychol. Bull.* **136**, 87–102 (2010).
16. J. D. Johnson, M. D. Rugg, Recollection and the reinstatement of encoding-related cortical activity. *Cereb. Cortex* **17**, 2507–2515 (2007).
17. B. P. Staresina, R. N. A. Henson, N. Kriegeskorte, A. Alink, Episodic reinstatement in the medial temporal lobe. *J. Neurosci.* **32**, 18150–18156 (2012).
18. M. Ritchey, E. A. Wing, K. S. LaBar, R. Cabeza, Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cereb. Cortex* **23**, 2818–2828 (2013).
19. R. B. Yaffe, M. S. D. Kerr, S. Damera, S. V. Sarma, S. K. Inati, K. A. Zaghloul, Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18727–18732 (2014).
20. H. Zhang, J. Fell, B. P. Staresina, B. Weber, C. E. Elger, N. Axmacher, Gamma power reductions accompany stimulus-specific sepresentations of dynamic events. *Curr. Biol.* **25**, 635–640 (2015).
21. E. L. Johnson, R. T. Knight, Intracranial recordings and human memory. *Curr. Opin. Neurobiol.* **31**, 18–25 (2015).
22. B. P. Staresina, M. Wimber, A neural chronometry of memory recall. *Trends Cogn. Sci.* **23**, 1071–1085 (2019).
23. E. L. Johnson, J. W. Y. Kam, A. Tzovara, R. T. Knight, Insights into human cognition from intracranial EEG: A review of audition, memory, internal cognition, and causality. *J. Neural Eng.* **17**, 051001 (2020).
24. S. M. Polyn, M. J. Kahana, Memory search and the neural representation of context. *Trends Cogn. Sci.* **12**, 24–30 (2008).
25. Y. Dudai, A. Karni, J. Born, The consolidation and transformation of memory. *Neuron* **88**, 20–32 (2015).
26. M. J. Sekeres, G. Winocur, M. Moscovitch, The hippocampus and related neocortical structures in memory transformation. *Neurosci. Lett.* **680**, 39–53 (2018).
27. S. E. Favila, H. Lee, B. A. Kuhl, Transforming the concept of memory reactivation. *Trends Neurosci.* **43**, 939–950 (2020).
28. B. A. Kuhl, M. M. Chun, Successful remembering elicits event-specific activity patterns in lateral parietal cortex. *J. Neurosci.* **34**, 8051–8060 (2014).
29. X. Xiao, Q. Dong, J. Gao, W. Men, R. A. Poldrack, G. Xue, Transformed neural pattern reinstatement during episodic memory retrieval. *J. Neurosci.* **37**, 2986–2998 (2017).
30. A. D. Wagner, B. J. Shannon, I. Kahn, R. L. Buckner, Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* **9**, 445–453 (2005).
31. I. G. Dobbins, H. J. Rice, A. D. Wagner, D. L. Schacter, Memory orientation and success: Separable neurocognitive components underlying episodic recognition. *Neuropsychologia* **41**, 318–333 (2003).
32. I. K. Brunec, J. Robin, R. K. Olsen, M. Moscovitch, M. D. Barense, Integration and differentiation of hippocampal memory traces. *Neurosci. Biobehav. Rev.* **118**, 196–208 (2020).
33. G. Xue, The neural representations underlying human episodic memory. *Trends Cogn. Sci.* **22**, 544–561 (2018).
34. M. Hebscher, W. A. Bainbridge, J. L. Voss, Neural similarity between overlapping events at learning differentially affects reinstatement across the cortex. *Neuroimage* **277**, 120220 (2023).
35. J. Lifanov, J. Linde-Domingo, M. Wimber, Feature-specific reaction times reveal a semanticisation of memories over time and with repeated remembering. *Nat. Commun.* **12**, 3177 (2021).
36. J. Liu, H. Zhang, T. Yu, D. Ni, L. Ren, Q. Yang, B. Lu, D. Wang, R. Heinen, N. Axmacher, G. Xue, Stable maintenance of multiple representational formats in human visual short-term memory. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 32329–32339 (2020).
37. R. Heinen, A. Bierbrauer, O. T. Wolf, N. Axmacher, Representational formats of human memory traces. *Brain Struct. Funct.* **229**, 513–529 (2024).
38. U. Guclu, M. A. J. van Gerven, Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *J. Neurosci.* **35**, 10005–10014 (2015).
39. I. Kuzovkin, R. Vicente, M. Petton, J.-P. Lachaux, M. Baciu, P. Kahane, S. Rheims, J. R. Vidal, J. Aru, Activations of deep convolutional neural networks are aligned with gamma band activity of human visual cortex. *Commun. Biol.* **1**, 107 (2018).
40. K. Grill-Spector, N. Knouf, N. Kanwisher, The fusiform face area subserves face perception, not generic within-category identification. *Nat. Neurosci.* **7**, 555–562 (2004).
41. R. Epstein, N. Kanwisher, A cortical representation of the local visual environment. *Nature* **392**, 598–601 (1998).
42. E. J. Allen, G. St-Yves, Y. Wu, J. L. Breedlove, J. S. Prince, L. T. Dowdle, M. Nau, B. Caron, F. Pestilli, I. Charest, J. B. Hutchinson, T. Naselaris, K. Kay, A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nat. Neurosci.* **25**, 116–126 (2022).
43. R. M. Cichy, A. Khosla, D. Pantazis, A. Torralba, A. Oliva, Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Sci. Rep.* **6**, 27755 (2016).
44. A. Doerig, R. P. Sommers, K. Seeliger, B. Richards, J. Ismael, G. W. Lindsay, K. P. Kording, T. Konkle, M. A. J. van Gerven, N. Kriegeskorte, T. C. Kietzmann, The neuroconnectionist research programme. *Nat. Rev. Neurosci.* **24**, 431–450 (2023).
45. S. W. Davis, B. R. Geib, E. A. Wing, W.-C. Wang, M. Hovhannisyan, Z. A. Monge, R. Cabeza, Visual and semantic representations predict subsequent memory in perceptual and conceptual memory tests. *Cereb. Cortex* **31**, 974–992 (2021).
46. T. C. Kietzmann, C. J. Spoerer, L. K. A. Sörensen, R. M. Cichy, O. Hauk, N. Kriegeskorte, Recurrence is required to capture the representational dynamics of the human visual system. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 21854–21863 (2019).
47. B. P. Staresina, L. Davachi, Differential encoding mechanisms for subsequent associative recognition and free recall. *J. Neurosci.* **26**, 9162–9172 (2006).
48. M. Nordt, K. Semmelmann, E. Genç, S. Weigelt, Age-related increase of image-invariance in the fusiform face area. *Dev. Cogn. Neurosci.* **31**, 46–57 (2018).
49. E. Tulving, Episodic memory: From mind to brain. *Annu. Rev. Psychol.* **53**, 1–25 (2002).
50. D. Zeithamova, A. R. Preston, Flexible memories: Differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *J. Neurosci.* **30**, 14676–14684 (2010).

51. D. Zeithamova, A. L. Dominick, A. R. Preston, Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron* **75**, 168–179 (2012).
52. E. S. Buchberger, A.-K. Joechner, C. T. Ngo, U. Lindenberger, M. Werkle-Bergner, Age differences in generalization, memory specificity and their overnight fate in childhood. PsyArXiv (2022). https://doi.org/10.31234/osf.io/y6ndr.
53. E. S. Buchberger, A. M. Brandmaier, U. Lindenberger, M. Werkle-Bergner, C. Ngo, The process structure of memory abilities in early and middle childhood. PsyArXiv (2023). https://doi.org/10.31234/osf.io/p84a6.
54. A. Keresztes, C. T. Ngo, U. Lindenberger, M. Werkle-Bergner, N. S. Newcombe, Hippocampal maturation drives memory from generalization to specificity. *Trends Cogn. Sci.* **22**, 676–686 (2018).
55. A. M. Daugherty, R. Flinn, N. Ofen, Hippocampal CA3-dentate gyrus volume uniquely linked to improvement in associative memory from childhood to adulthood. *Neuroimage* **153**, 75–85 (2017).
56. L. Tang, P. J. Pruitt, Q. Yu, R. Homayouni, A. M. Daugherty, J. S. Damoiseaux, N. Ofen, Differential functional connectivity in anterior and posterior hippocampus supporting the development of memory formation. *Front. Hum. Neurosci.* **14**, 204 (2020).
57. A. I. Ramsaran, M. L. Schlichting, P. W. Frankland, The ontogeny of memory persistence and specificity. *Dev. Cogn. Neurosci.* **36**, 100591 (2019).
58. N. Ofen, Y.-C. Kao, P. Sokol-Hessner, H. Kim, S. Whitfield-Gabrieli, J. D. E. Gabrieli, Development of the declarative memory system in the human brain. *Nat. Neurosci.* **10**, 1198–1205 (2007).
59. X. Chai, N. Ofen, L. F. Jacobs, J. D. E. Gabrieli, Scene complexity: Influence on perception, memory, and development in the medial temporal lobe. *Front. Hum. Neurosci.* **4**, 21 (2010).
60. N. Ofen, The development of neural correlates for memory formation. *Neurosci. Biobehav. Rev.* **36**, 1708–1717 (2012).
61. L. Tang, A. T. Shafer, N. Ofen, Prefrontal cortex contributions to the development of memory formation. *Cereb. Cortex* **28**, 3295–3308 (2018).
62. E. L. Johnson, L. Tang, Q. Yin, E. Asano, N. Ofen, Direct brain recordings reveal prefrontal cortex dynamics of memory development. *Sci. Adv.* **4**, eaat3702 (2018).
63. Q. Yin, E. L. Johnson, L. Tang, K. I. Auguste, R. T. Knight, E. Asano, N. Ofen, Direct brain recordings reveal occipital cortex involvement in memory development. *Neuropsychologia* **148**, 107625 (2020).
64. E. L. Johnson, Q. Yin, N. B. O'Hara, L. Tang, J.-W. Jeong, E. Asano, N. Ofen, Dissociable oscillatory theta signatures of memory formation in the developing brain. *Curr. Biol.* **32**, 1457–1469.e4 (2022).
65. N. Ofen, L. Tang, Q. Yu, E. L. Johnson, Memory and the developing brain: From description to explanation with innovation in methods. *Dev. Cogn. Neurosci.* **36**, 100613 (2019).
66. L. Tang, Q. Yu, R. Homayouni, K. L. Canada, Q. Yin, J. S. Damoiseaux, N. Ofen, Reliability of subsequent memory effects in children and adults: The good, the bad, and the hopeful. *Dev. Cogn. Neurosci.* **52**, 101037 (2021).
67. N. Axmacher, A useful code for sequences. *Nat. Neurosci.* **19**, 1276–1277 (2016).
68. R. A. Cowell, M. D. Barense, P. S. Sadil, A roadmap for understanding memory: Decomposing cognitive processes into operations and representations. *eNeuro* **6**, 10.1523/ENEURO.0122-19.2019 (2019).
69. R. B. Yaffe, A. Shaikhouni, J. Arai, S. K. Inati, K. A. Zaghloul, Cued memory retrieval exhibits reinstatement of high gamma power on a faster timescale in the left temporal lobe and prefrontal cortex. *J. Neurosci.* **37**, 4472–4480 (2017).
70. J. Linde-Domingo, M. S. Treder, C. Kerrén, M. Wimber, Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nat. Commun.* **10**, 179 (2019).
71. M. S. Treder, I. Charest, S. Michelmann, M. C. Martín-Buro, F. Roux, F. Carceller-Benito, A. Ugalde-Canitrot, D. T. Rollings, V. Sawlani, R. Chelvarajah, M. Wimber, S. Hanslmayr, B. P. Staresina, The hippocampus as the switchboard between perception and memory. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2114171118 (2021).
72. B. A. Kuhl, J. Rissman, A. D. Wagner, Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia* **50**, 458–469 (2012).
73. A. P. Shimamura, Episodic retrieval and the cortical binding of relational activity. *Cogn. Affect. Behav. Neurosci.* **11**, 277–291 (2011).
74. M. R. Uncapher, A. D. Wagner, Posterior parietal cortex and episodic encoding: Insights from fMRI subsequent memory effects and dual-attention theory. *Neurobiol. Learn. Mem.* **91**, 139–154 (2009).
75. R. Cabeza, E. Ciaramelli, I. R. Olson, M. Moscovitch, The parietal cortex and episodic memory: An attentional account. *Nat. Rev. Neurosci.* **9**, 613–625 (2008).
76. M. Corbetta, G. L. Shulman, Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
77. H. Lee, B. A. Kuhl, Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. *J. Neurosci.* **36**, 6069–6082 (2016).
78. S. Brodt, S. Gais, J. Beck, M. Erb, K. Scheffler, M. Schönauer, Fast track to the neocortex: A memory engram in the posterior parietal cortex. *Science* **362**, 1045–1048 (2018).
79. A. P. Yonelinas, Components of episodic memory: The contribution of recollection and familiarity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **356**, 1363–1374 (2001).
80. M. de Chastelaine, J. T. Mattson, T. H. Wang, B. E. Donley, M. D. Rugg, Independent contributions of fMRI familiarity and novelty effects to recognition memory and their stability across the adult lifespan. *Neuroimage* **156**, 340–351 (2017).
81. M. de Chastelaine, S. Srokova, M. Hou, A. Kidwai, S. S. Kafafi, M. L. Racenstein, M. D. Rugg, Cortical thickness, gray matter volume, and cognitive performance: A crosssectional study of the moderating effects of age on their interrelationships. *Cereb. Cortex* **33**, 6474–6485 (2023).
82. A. P. Yonelinas, L. J. Otten, K. N. Shaw, M. D. Rugg, Separating the brain regions involved in recollection and familiarity in recognition memory. *J. Neurosci.* **25**, 3002–3008 (2005).
83. J. Spaniol, P. S. R. Davidson, A. S. N. Kim, H. Han, M. Moscovitch, C. L. Grady, Event-related fMRI studies of episodic encoding and retrieval: Meta-analyses using activation likelihood estimation. *Neuropsychologia* **47**, 1765–1779 (2009).
84. A. P. Yonelinas, The nature of recollection and familiarity: A review of 30 years of research. *J. Mem. Lang.* **46**, 441–517 (2002).
85. P. F. Hill, D. R. King, B. C. Lega, M. D. Rugg, Comparison of fMRI correlates of successful episodic memory encoding in temporal lobe epilepsy patients and healthy controls. *Neuroimage* **207**, 116397 (2020).
86. E. L. Johnson, R. T. Knight, How can iEEG be used to study inter-individual and developmental differences? in *Intracranial EEG: A Guide for Cognitive Neuroscientists*, N. Axmacher, Ed. (Springer International Publishing, Cham, 2023; in Neuroscience, Psychology and Behavioral Economics, pp. 143–154; https://doi.org/10.1007/978-3-031-20910-9_10)Studies.
87. T. Bonnen, D. L. K. Yamins, A. D. Wagner, When the ventral visual stream is not enough: A deep learning account of medial temporal lobe involvement in perception. *Neuron* **109**, 2755–2766.e6 (2021).
88. M. L. Schlichting, J. A. Mumford, A. R. Preston, Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat. Commun.* **6**, 8151 (2015).
89. A. J. H. Chanales, A. Oza, S. E. Favila, B. A. Kuhl, Overlap among spatial memories triggers repulsion of hippocampal representations. *Curr. Biol.* **27**, 2307–2317.e5 (2017).
90. L. Kunz, L. Wang, D. Lachner-Piza, H. Zhang, A. Brandt, M. Dümpelmann, P. C. Reinacher, V. A. Coenen, D. Chen, W.-X. Wang, W. Zhou, S. Liang, P. Grewe, C. G. Bien, A. Bierbrauer, T. Navarro Schröder, A. Schulze-Bonhage, N. Axmacher, Hippocampal theta phases organize the reactivation of large-scale electrophysiological representations during goal-directed navigation. *Sci. Adv.* **5**, eaav8192 (2019).
91. Z. Fayyaz, A. Altamimi, C. Zoellner, N. Klein, O. T. Wolf, S. Cheng, L. Wiskott, A model of semantic completion in generative episodic memory. *Neural Comput.* **34**, 1841–1870 (2022).
92. J. L. Breedlove, G. St-Yves, C. A. Olman, T. Naselaris, Generative feedback explains distinct brain activity codes for seen and mental images. *Curr. Biol. CB* **30**, 2211–2224.e6 (2020).
93. M. A. Yassa, C. E. L. Stark, Pattern separation in the hippocampus. *Trends Neurosci.* **34**, 515–525 (2011).
94. N. Burgess, E. A. Maguire, J. O'Keefe, The human hippocampus and spatial and episodic memory. *Neuron* **35**, 625–641 (2002).
95. K.-H. Bäuml, B. Pastötter, S. Hanslmayr, Binding and inhibition in episodic memory—Cognitive, emotional, and neural processes. *Neurosci. Biobehav. Rev.* **34**, 1047–1054 (2010).
96. A. P. Yonelinas, C. Ranganath, A. D. Ekstrom, B. J. Wiltgen, A contextual binding theory of episodic memory: Systems consolidation reconsidered. *Nat. Rev. Neurosci.* **20**, 364–375 (2019).
97. A. Bierbrauer, M.-C. Fellner, R. Heinen, O. T. Wolf, N. Axmacher, The memory trace of a stressful episode. *Curr. Biol.* **31**, 5204–5213.e8 (2021).
98. J. G. Snodgrass, J. Corwin, Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *J. Exp. Psychol. Gen.* **117**, 34–50 (1988).
99. R. Oostenveld, P. Fries, E. Maris, J.-M. Schoffelen, FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869.
100. A. Stolk, S. Griffin, R. van der Meij, C. Dewar, I. Saez, J. J. Lin, G. Piantoni, J.-M. Schoffelen, R. T. Knight, R. Oostenveld, Integrated analysis of anatomical and electrophysiological human intracranial data. *Nat. Protoc.* **13**, 1699–1723 (2018).
101. C. M. Lacadie, R. K. Fulbright, R. T. Constable, X. Papademetris, More accurate talairach coordinates for neuroimaging using nonlinear registration. *Neuroimage* **42**, 717–725 (2008).
102. E. Maris, R. Oostenveld, Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
103. B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, A. Oliva, Learning deep features for scene recognition using places database. *Neural Inf. Process. Syst.* **27**, 9 (2014).

104. B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, A. Torralba, Places: A 10 million image database for scene recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**, 1452–1464 (2018).
105. A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2017).
106. M. Eickenberg, A. Gramfort, G. Varoquaux, B. Thirion, Seeing it all: Convolutional network layers map the function of the human visual system. *Neuroimage* **152**, 184–194 (2017).
107. D. Lindh, I. G. Sligte, S. Assecondi, K. L. Shapiro, I. Charest, Conscious perception of natural images is constrained by category-related visual features. *Nat. Commun.* **10**, 4106 (2019).
108. J. R. Manning, J. Jacobs, I. Fried, M. J. Kahana, Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *J. Neurosci. Off. J. Soc. Neurosci.* **29**, 13613–13620 (2009).
109. J. F. Burke, N. M. Long, K. A. Zaghloul, A. D. Sharan, M. R. Sperling, M. J. Kahana, Human intracranial high-frequency activity maps episodic memory formation in space and time. *Neuroimage* **85**, 834–843 (2014).
110. J. F. Burke, A. G. Ramayya, M. J. Kahana, Human intracranial high-frequency activity during memory processing: Neural oscillations or stochastic volatility? *Curr. Opin. Neurobiol.* **31**, 104–110 (2015).
111. B. S. Katerman, Y. Li, J. K. Pazdera, C. Keane, M. J. Kahana, EEG biomarkers of free recall. *Neuroimage* **246**, 118748 (2022).
112. B. J. Griffiths, O. Jensen, Gamma oscillations and episodic memory. *Trends Neurosci.* **46**, 832–846 (2023).
113. P. J. Uhlhaas, F. Roux, W. Singer, C. Haenschel, R. Sireteanu, E. Rodriguez, The development of neural synchrony reflects late maturation and restructuring of functional networks in humans. *Proc. Natl. Acad. Sci.* **106**, 9866–9871 (2009).
114. W. Gaez, T. P. L. Roberts, K. D. Singh, S. D. Muthukumaraswamy, Functional and structural correlates of the aging brain: Relating visual cortex (V1) gamma band responses to age-related structural change. *Hum. Brain Mapp.* **33**, 2035–2046 (2012).
115. B. T. Hutchinson, K. Pammer, B. Jack, Pre-stimulus alpha predicts inattentional blindness. *Conscious. Cogn.* **87**, 103034 (2021).
116. N. S. Dehaghani, B. Maess, R. Khosrowabadi, R. Lashgari, S. Braeutigam, M. Zarei, Pre-stimulus alpha activity modulates face and object processing in the intra-parietal sulcus, a MEG study. *Front. Hum. Neurosci.* **16**, 831781 (2022).
117. W. Klimesch, P. Sauseng, S. Hanslmayr, EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Res. Rev.* **53**, 63–88 (2007).

Acknowledgments

Funding: This work was supported by grants from the European Research Council (grant: CoG 864164 to N.A.), the German Israeli Foundation (grant: I-1478-418.13/2018 to N.A.), the German Research Foundation (grant: 419049386 to N.A.) and the US National Institutes of Health (NIMH grant R01MH107512 to N.O.; NINDS grant R00NS115918 to E.L.J., and NINDS grant R01NS064033 to E.A.). **Author contributions:** Conceptualization: N.A., N.O., E.L.J., E.M.B.R., R.T.K., O.K.-M., and M.-C.F. Methodology: N.A., N.O., E.L.J., E.M.B.R., R.H., H.Z., and M.-C.F. Software: E.M.B.R., H.Z., E.L.J., and M.K. Validation: N.A., N.O., E.L.J., and M.K. Formal analysis: E.M.B.R., M.K., H.Z., and R.H. Investigation: N.O., E.L.J., E.A., O.K.-M., R.T.K., E.M.B.R., S.S., J.J.L., K.I.A., E.F.C., D.K.-S., K.D.L., and P.B.W. Resources: N.O., E.L.J., E.M.B.R., O.K.-M., and R.H. Data curation: N.O., E.L.J., E.M.B.R., Q.Y., O.K.-M., and P.V. Writing—original draft: E.M.B.R. and N.O. Writing—review and editing: N.A., E.L.J., E.M.B.R., M.K., Q.Y., O.K.-M., and R.H. Visualization: E.M.B.R. and R.H. Supervision: N.A., N.O., E.L.J., and E.M.B.R. Project administration: N.A., N.O., E.L.J., E.M.B.R., and O.K.-M. Funding acquisition: N.A., N.O., E.L.J., R.T.K., and E.A. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Deidentified raw data have been deposited at an NIMH database (https://nda.nih.gov/study.html?id=2938). Processed data and analysis code to reproduce the main results presented in the manuscript are publicly available at (https://zenodo.org/records/13866456).

Submitted 24 April 2024
Accepted 16 January 2025
Published 19 February 2025
10.1126/sciadv.adp9336