Title: A cortico-hippocampal scaffold for representing and recalling lifelike events

Zachariah M. Reagh[1] & Charan Ranganath[2,3]

[1] Department of Psychological & Brain Sciences, Washington University in St. Louis
[2] UC Davis Center for Neuroscience, University of California, Davis
[3] Department of Psychology, University of California, Davis

Correspondence: zreagh@wustl.edu

Summary

Real-world events are complex, featuring elements that may be unique to, or shared across, multiple situations. In the present study, we used fMRI to identify how different event components are represented in real-time and during memory retrieval. Twenty participants viewed and recalled eight videos depicting real-world events, combining people, contexts, and context types. Multi-voxel pattern similarity analyses revealed specific 'person' representations, persistent across contexts, in regions of an Anterior-Temporal Network. Conversely, we found specific 'context' representations, persistent across people, in regions of a Posterior-Medial Network. We also found schema-like generalization across contexts in medial prefrontal cortex, and episodic specificity in the hippocampus. Event patterns were reinstated during recall, and hippocampal reinstatement predicted the number of details retrieved. Finally, we observed distinct representational timescales across the hippocampus and cortical regions. These findings reveal mechanisms for scaffolding different aspects of lifelike event representations in cortico-hippocampal networks as experiences are observed and recalled.

Introduction

Understanding real-world events requires us to be able to extract key features from a dynamic, ongoing stream of complex information. We have the capability to remember these events as unique to a particular place and time, but we also can flexibly use information from events to make generalizations about the people, places, and situations that comprise them. Research on human memory has generally focused on the ability to recognize or recall specific items in isolation, or simple associations between items. This work has largely focused on the role of the hippocampus (HPC) and connected medial temporal lobe structures in these processes (Eichenbaum et al., 2007; Gold et al., 2006; Konkel & Cohen, 2009; Rugg et al., 2012). However, little is known about how the brain pulls apart and flexibly recombines elements across events.

Some insights into this question have come from recent studies investigating the representation of naturalistic events in human brain networks. Most work in this vein has focused on the structure of events, and has revealed multiple scales of representation in the human brain. In particular, recent evidence has emerged that the brain's 'default mode network' (DMN) carries patterns of activity that distinguish one event from another in an ongoing narrative, such as a story or television episode. Based on these studies, activity patterns in the DMN seem to reflect a broad understanding of an event, rather than the stimuli that are presented. For instance, DMN regions carry activity patterns over entire scenes during viewing of a movie, which are similar to those evoked when hearing a story about the movie or when recalling that movie (Bird et al., 2015; Chen et al., 2017; Oedekoven et al., 2017). Moreover, activity patterns in DMN regions remain stable as an event unfolds, but these patterns shift abruptly when one infers that a new event is taking place (Baldassano et al., 2017; Ben-Yakov & Henson, 2018). Other studies have suggested that certain components of the DMN – in particular, the medial prefrontal cortex (mPFC) – represent abstract event structures, or schemas, which generalize across particular instances of similar situations (Baldassano et al., 2017). Related work suggests that DMN regions may carry information related to prior schematic knowledge about television sitcoms (Raykov et al., 2018), and that schematic knowledge may drive representations of particular people in those sitcoms (Raykov et al., 2020). However, to date, there have been few investigations of how specific bits of information are extracted from complex events, whether these bits can be flexibly recombined in partially overlapping situations, and how this is reflected in neural representations.

Prior findings support the idea that the global content and structure of events may be represented in an abstract manner, and in a similar way across individuals. However, it remains an open question as to whether the brain differentially represents different kinds of content, or if events are represented in a more integrated manner throughout the DMN. In part, this stems from the fact that many such studies use feature films or television shows as stimuli, and as such, it is difficult to control which poses major challenges for systematic examination of how specific content may be represented across distinct events. Much of the DMN is comprised of parietal regions which have been proposed to support retrieved episodic content, irrespective of information type or modality (Bonnici et al., 2016; Kuhl & Chun, 2014; Lee & Kuhl, 2016; Thakral et al., 2017). Parietal components of the DMN have also been proposed to represent abstractions of semantic category across many exemplars (Binder & Desai, 2011; Devereux et al., 2013; Rugg & King, 2018). Finally, recent work has suggested a continuum between episodic and semantic memory in the DMN (Irish & Vatansever, 2020). This class of views, in general, would suggest that regions of the DMN may uniformly represent all aspects of an event, with the strength of the representation perhaps teasing apart the contributions of particular regions (Richter et al., 2016). At present, many findings from studies using naturalistic stimuli could be taken to support this perspective.

On the other hand, a wealth of findings suggest that the DMN can be divided into at least two subnetworks that interact closely with the HPC (Ranganath & Ritchey, 2012; Reagh & Ranganath,

2018; Ritchey, Libby, et al., 2015). One is a Posterior-Medial (PM) Network of brain regions, which interfaces with the HPC to represent contextual and situational information. Moreover, there is some evidence to suggest that multiple levels of abstraction exist within the PM Network (Ritchey & Cooper, 2020), with mPFC being increasingly thought to represent contextual information at the most general levels (i.e., 'schemas') (Ghosh & Gilboa, 2014; Gilboa & Marlatte, 2017; Preston & Eichenbaum, 2013; Spalding et al., 2015; van Kesteren et al., 2013). This contrasts with an Anterior-Temporal (AT) Network of brain regions, which interfaces with the HPC to represent local entities – such as people – that occupy those contexts. For example, consider a situation in which you are having dinner with friends. Regions of the PM Network may carry information about a particular restaurant, or the general situation of dining at a restaurant, while regions of the AT Network may carry information about objects or people in a restaurant. This contrasts markedly with more domain-general views of DMN contributions to event processing and memory. Notably, prior investigations of naturalistic events have largely focused on boundaries within ongoing experiences or global representations of whole events, but have not closely evaluated representations of specific event content in the face of nonoverlap with other similar events. That is, we do not know whether representations of local entities (such as people) can be flexibly constructed and reconstructed despite changes in contextual information, or vice versa, in complex, continuous real-world situations. It is thus an important open question whether dissociations based on information content hold in the face of complex experiences that unfold in real time.

The present study was designed to address the question of whether cortico-hippocampal networks provide a representational scaffold for different components of dynamic experiences. That is: do representations in these regions dissociate on the basis of information content when encoding and recalling lifelike events? Recent studies suggest that the DMN is prominently involved in representing naturalistic events, which may include all aspects of that event, shared across brain regions. If this is the case, we might expect to see uniform representation of event content across the entire DMN, and that the extent of nonoverlap from one event to another – irrespective of information type – would reduce the similarity of those representations across regions. Conversely, it is possible that the PM Network represents only some components of complex events (contextual information), whereas the AT Network represents others (local entities), and that neural representations across regions will critically depend on the kind of information which does not overlap when comparing two events. To address this, we designed brief video clips depicting everyday events. Critically, we systematically combined information about local entities (i.e., two people) and contexts (i.e., four locations), and further manipulated contextual specificity (i.e., two classes of contexts with two exemplars of each class). This contextual specificity manipulation allowed us to examine whether any regions of the PM Network – for example, mPFC – would merge across contexts which share abstract, categorical similarity but low perceptual similarity (i.e., consistent with representation of schemas). Participants underwent fMRI scanning during encoding of these video clips, as well as during spoken recall, and multi-voxel activity patterns were analyzed to assess representational similarity across events.

Results

Participants were presented with 8 35-second video clips systematically combining a focal person (Tommy or Lisa) each appearing in 1 of 4 contexts (Fig. 1A). To address the issue of context specificity or generalization, the design furthermore incorporated 2 different *types* of contexts (2 distinct cafes and 2 distinct grocery stores). During encoding, participants viewed each of the 8 video clips once per run, in randomized order, over 3 total encoding runs (i.e., each video was viewed a total of 3 times). Each clip consisted of a 5-second title screen, followed by the 35-second event, and a 10-second interval between events (Fig. 1B). During the single recall run, participants viewed the title screen for each event for 40 seconds (i.e., no video clip was played), during which they verbally

recalled the event in as much detail as possible, followed by a 10-second interval between recall of events (Fig. 1C). See Methods for additional details about the experimental design.

Accurate, detailed, and unbiased memory for events across participants

Memory for videos shown in the scanner was tested in two ways. Participants were scanned during free recall of each film, and following the scan session, they completed a recognition memory test consisting of statements about the events. A standardized protocol was used to score recall data (Cohn-Sheehy et al., 2020) for the number of verifiable details with respect to content in the video clips (see Methods for further details). Participants generally recalled many details about each movie, and each participant recalled information from each event. Across participants, recall scores were significantly above zero (average = 14.81, $t(19) = 11.737$, $p = 1.095e^{-10}$). Recall performance did not differ as a function of context (one-way ANOVA; $F(3,172) = 0.803$, $p = 0.412$) or person (paired t-test; $t(19) = 0.108$, $p = 0.893$) (Table 1). For recognition, we observed significantly greater than chance discrimination of true from false statements about the events (average d' = 1.49, $t(19) = 14.963$, $p = 3.675e^{-12}$), and there were no significant differences as a function of context (one-way ANOVA; $F(3,172) = 0.445$, $p = 0\ .739$) or person (paired t-test; $t(19) = 0.385$, $p = 0.522$). Thus, participants encoded and remembered events in considerable detail, but were not significantly biased to remember or forget details about particular events or particular individuals within those events.

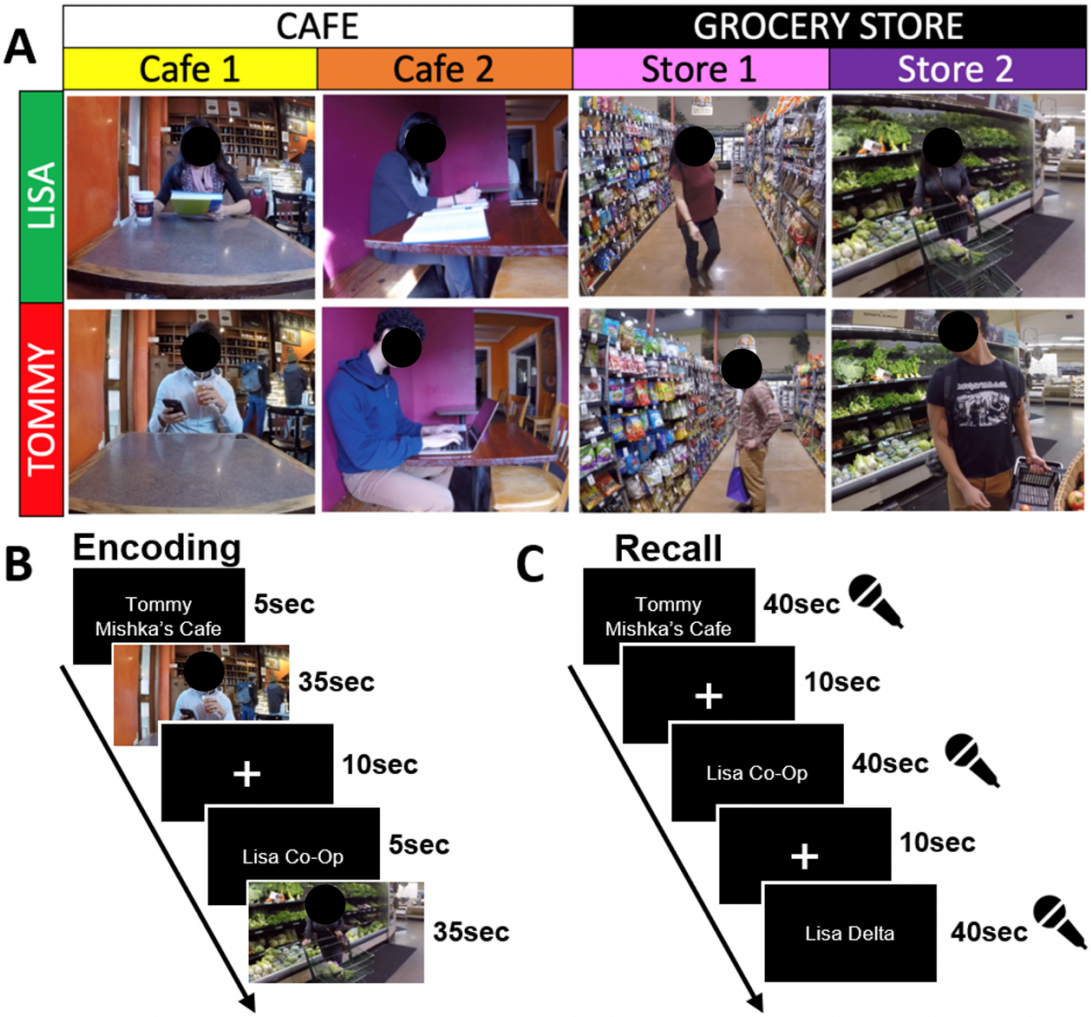

Figure 1: Experimental design and trial structure. (A) 8 videos were designed to systematically combine information about local entities (i.e., central person) and contexts (i.e., specific location), with an additional layer of context *type* (i.e., café vs. grocery store). (B) Encoding trial structure (3 runs, randomized event order). (C) Recall trial structure (1 run, randomized event order).

Table 1: Memory performance. Recall (verifiable details, top row) and true/false recognition memory (d', bottom row) and per event are significantly above zero for every context and person depicted, but do not differ as a function of context or person.

	Context				Person	
	Cafe1	**Cafe2**	**Store1**	**Store2**	**Tommy**	**Lisa**
Recall (details)	14.893	14.225	14.317	14.601	14.773	14.594
Recognition (d')	1.493	1.445	1.457	1.518	1.733	1.704

Regional differences in context, person, schema, and episode-specific patterns at encoding

Having established that participants successfully formed detailed memories for each event, we next addressed whether different cortico-hippocampal networks represent distinct components of these events during encoding. Regions of Interest (ROIs; Supplemental Fig. 1) of the PM Network – angular gyrus (ANG), posterior-medial cortex (PMC), and parahippocamapal cortex (PHC) – were selected based on theoretical involvement in context representation, as well as their roles in event representation in recent studies. Given accumulating evidence for representation of abstract schemas in mPFC, we considered this region separately from the PM Network. ROIs in the AT Network – perirhinal cortex (PRC) and temporal poles (TP) – were selected due to demonstrated sensitivity to item-level information. We additionally included a HPC ROI, based on previously published anatomical tracings (Ritchey, Montchal, et al., 2015). Briefly, we separately modeled the unique activity pattern for each video clip. Each video was shown once in each encoding run, for a total of three exposures. To characterize event representations during encoding, we systematically correlated patterns of activity across runs in accord with the factorial design of our study.

To test direct a priori hypotheses about representational content across ROIs, event-by-event correlation matrices for each region were compared to model matrices depicting Person, Context, Schema, and Episode-Specific (or Episodic) representations (Fig. 2). In our experimental design, schemas were operationalized as similarity across a *type* of context (e.g., Cafe1 + Cafe2, or Store1 + Store2). Given an expectation of similar response profiles across individual PM Network and AT Network regions (corroborated by analyses over individual ROIs; see Supplemental Information), we collapsed within-network across ROIs for the main analyses. Further analytical details can be found in the Methods section, and full regional results can be found in Supplemental Information.
Event-by-event correlation matrices were then compared to model matrices to test specific hypothesized representational profiles. We first analyzed PM Network regions. Pattern similarity values were separately estimated for each PM Network ROI (PMC, ANG, and PHC), and then averaged across ROIs to enable hypothesis tests at the network level. In regions of the PM Network (Fig. 3A), the average pattern similarity correlation matrix across participants was most related to the Context model matrix ($r = 0.842$, $p = 2.88e^{-18}$) with significant correlations also found for the Schema ($r = 0.629$, $p = 2.675e^{-08}$) and Episodic ($r = 0.603$, $p = 1.326\ e^{-07}$) matrices (Fig. 3B). The Person matrix fit was not significant ($r = 0.095$, $p = 0.453$). There was a significant difference among model fits ($F(3,57) = 16.187$, $p < 0.001$), and the Context matrix fit was significantly stronger than each of the other three ($p_{Tukey} < 0.05$ corrected). Follow-up analyses, shown in the Supplemental Results (Individual ROI encoding-encoding model fits), confirmed that similar patterns of results were seen across the different PM Network ROIs.

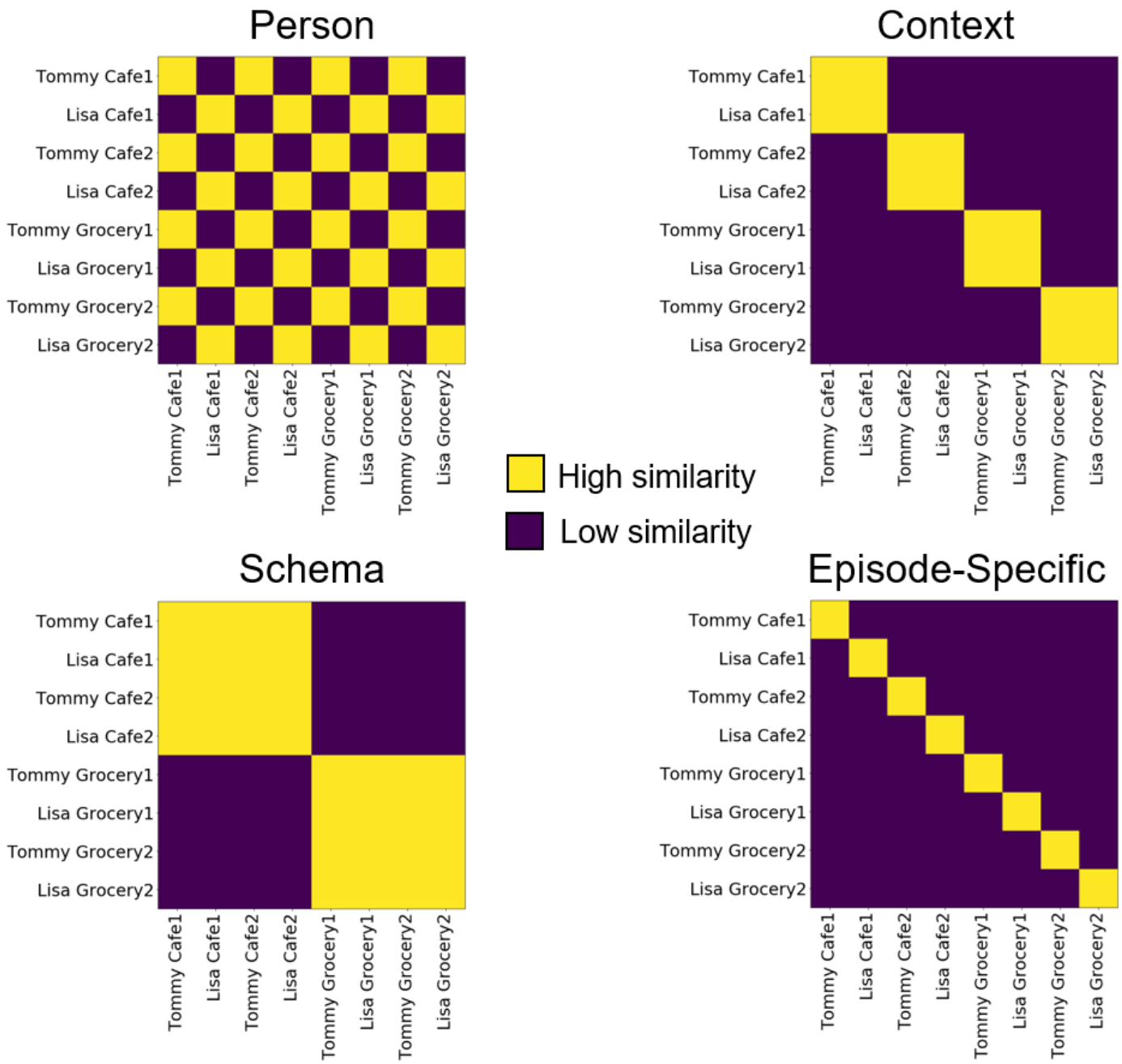

Figure 2: Model matrices based on hypothesized representational profiles. Event-by-event correlation matrices resulting from pattern similarity analyses were compared to hypothesized model matrices depicting Character, Context, Schema, and Episode-Specific representations via point-biserial correlations.

We next compared representational profiles in these AT Network regions to hypothesized model matrices. In AT Network regions (Fig. 3C), we found the largest correlation between the event-by-event pattern similarity matrix and the Person model matrix ($r = 0.845$, $p = 1.722e^{-18}$), though fits for all other matrices were also significant: Context ($r = 0.274$, $p = 0.028$), Schema ($r = 0.366$, $p = 0.003$), Episodic ($r = 0.575$, $p = 6.664e^{-07}$) (Fig. 3D). Model fits differed significantly ($F(3,57) = 18.9$, $p < 0.001$), and the Person matrix was a significantly stronger fit than all other model matrices ($p_{Tukey} < 0.05$ corrected). Similar to PM Network analyses, AT Network regions (PRC, TP) showed highly similar response profiles (Supplemental Results: Individual ROI encoding-encoding model fits).

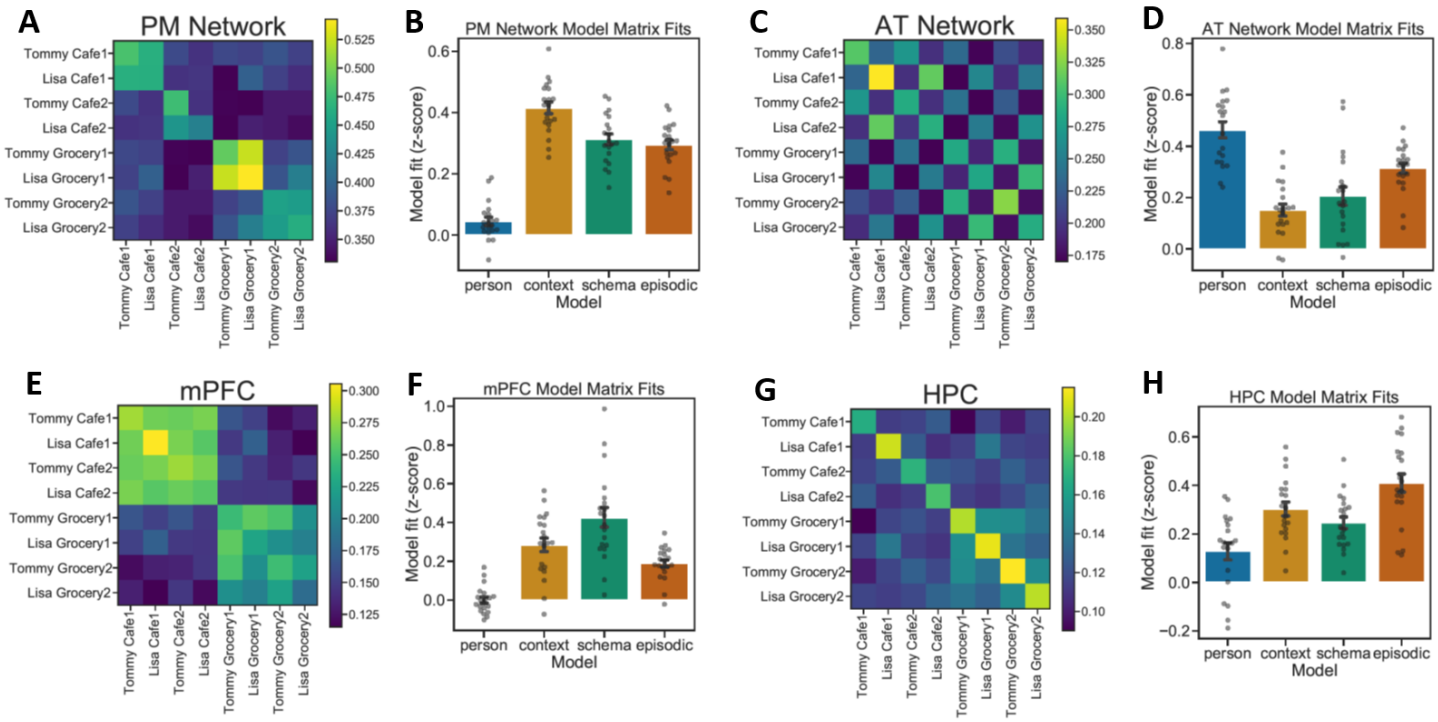

Figure 3: Across-event pattern similarity at encoding. (A,B) PM Network pattern similarity results most strongly fit the Context model matrix. (C,D) AT Network pattern similarity results most strongly fit the Person model matrix. (E,F) mPFC pattern similarity results most strongly fit the Schema model matrix. (G,H) HPC pattern similarity results most strongly fit the Episodic model matrix. (Error bars = SEM.)

Event correlations in mPFC (Fig. 3E), were best described by the Schema model matrix ($r = 0.914$, $p = 5.089e^{-26}$), with significant fits also to the Context ($r = 0.0.591$, $p = 2.759e^{-07}$) and Episodic ($r = 0.385$, $p = 0.002$) matrices (Fig. 3F). The fit with the Person matrix was not significant ($r = 0.001$, $p = 0.995$). The model fits differed significantly ($F(3,57) = 42.916$, $p < 0.001$), with the Schema matrix fitting significantly better than the other three ($p_{Tukey} < 0.05$ corrected). Thus, mPFC showed a degree of contextual sensitivity, but a significantly different profile from PM Network regions, such that schema-level representation was more prominent than representation of specific contexts.

Event correlations in HPC (Fig. 3G), were best described by the Episodic model matrix ($r = 0.588$, $p = 1.305e^{-18}$) (Fig. 3H). Significant fits were also observed with the Context ($r = 0.67$, $p = 1.38e^{-09}$), Schema ($r = 0.535$, $p = 5.377e^{-06}$), and Person ($r = 0.247$, $p = 0.049$) matrices. There was a significant difference among model fits ($F(3,57) = 29.6975$, $p < 0.001$). The Episodic model matrix was a stronger fit to the data than the Person or Schema matrices ($p_{Tukey} < 0.05$ corrected), though there was not a significant pairwise difference between the Episodic and Context matrices. The data suggest that HPC is best characterized as representing event-specific information.

A different approach of analyzing the data is to collapse across events of a particular kind, and compare to other events of a different kind (e.g., [same character + same context] versus [same character + different context]). Per this analysis, we can factorially cross character (same vs. different) and context (same, similar, or different). In the PM Network, we found a significant effect of Context ($F(2,114) = 6.325$, $p = 0.002$), with significant effects in each individual region (see Supplemental Information). This effect was driven by greater pattern similarity when participants viewed events that occurred within the same context compared to similar (pairwise contrasts: $p < 0.05$ corrected) or different contexts (pairwise contrasts: $p < 0.05$ corrected) (Fig. 4A). Pattern similarity did

not differ between similar and different contexts. Conversely, we did not find an effect of Person ($F(1,114) = 0.24$, $p = 0.625$) nor an interaction ($F(2,114) = 0.018$, $p = 0.982$). In the AT Network, we found a significant effect of Person ($F(1,114) = 20.139$, $p < 0.001$), but no effect of Context ($F(2,114) = 1.617$, $p = 0.203$) nor an interaction ($F(2,114) = 0.194$, $p = 0.824$). This was driven by significantly greater pattern similarity when participants viewed events depicting the same person compared to a different person (pairwise contrasts: $p < 0.05$ corrected) (Fig. 4B). In mPFC, we found a significant effect of Context ($F(2,114) = 13.569$, $p < 0.001$), but neither an effect of Person ($F(1,114) = 0.108$, $p = 0.742$) nor an interaction ($F(2,114) = 0.09$, $p = 0.914$). Like PM Network regions, we found higher pattern similarity for same versus different contexts and between similar versus different contexts (pairwise contrasts: $p < 0.05$ corrected), but not between same and similar contexts (Fig. 4C). Finally, in HPC, we found a significant effect of Context ($F(2,114) = 5.018$, $p = 0.008$) and a trending interaction ($F(2,114) = 2.997$, $p = 0.054$), but no effect of Person ($F(1,114) = 1.96$, $p = 0.164$). In line with model matrix analyses, post-hoc contrasts revealed that pattern similarity was highest in HPC when viewing the same event (i.e., same person + same context; pairwise contrasts: $p < 0.05$ corrected) (Fig. 4D).

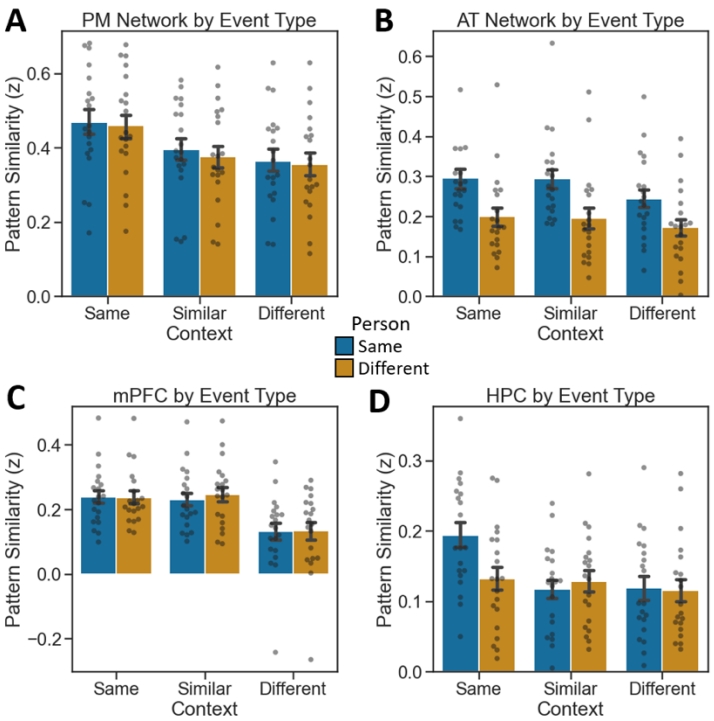

Figure 4: Encoding pattern similarity results by event type. (A) PM Network patterns show context-specificity. (B) AT Network patterns show person-specificity. (C) mPFC patterns generalize across similar contexts. (D) HPC patterns are episode-specific. (Error bars = SEM.)

Content-selective neural reinstatement of event components during spoken recall

We next asked whether neural patterns associated with these events that were present across encoding runs would be reinstated during spoken recall of each event. Importantly, we wanted to examine whether these reinstated patterns would be modality-selective (i.e., would PM Network regions flexibly re-establish context-specific patterns, and would AT Network regions flexibly re-establish character-specific patterns?). Furthermore, we asked whether this reinstatement would be related to participants' recall performance. For encoding-recall pattern similarity analyses, multi-voxel patterns associated with each encoding run were compared to patterns associated with retrieval of each event. Importantly, the comparisons of interest are identical to encoding-encoding analyses, except the question now turns to event pattern reinstatement in the absence of the perceptual inputs associated with those events. Similar to our approach for encoding-encoding pattern similarity analyses, we collapsed across PM Network regions (PMC, ANG, and PHC) and AT Network regions (PRC and TP). Full regional results can be found in Supplemental Information.

In the PM Network (Fig. 5A), we observed the largest correlation between encoding-recall pattern similarity data and the Context matrix ($r = 0.592$, $p = 2.604e^{-07}$), though a significant correlation was also observed with the Schema matrix ($r = 0.426$, $p = 0.0004$). The fit to the Episodic model was trending, but nonsignificant ($r = 0.243$, $p = 0.053$), and there was a significant negative fit to the Person model ($r = -0.458$, $p = 0.0001$). Model fits differed significantly ($F(3,57) = 20.475$, $p < 0.001$), driven by a poorer fit to the Person matrix than other matrices ($p_{Tukey} < 0.05$ corrected) (Fig. 5B).

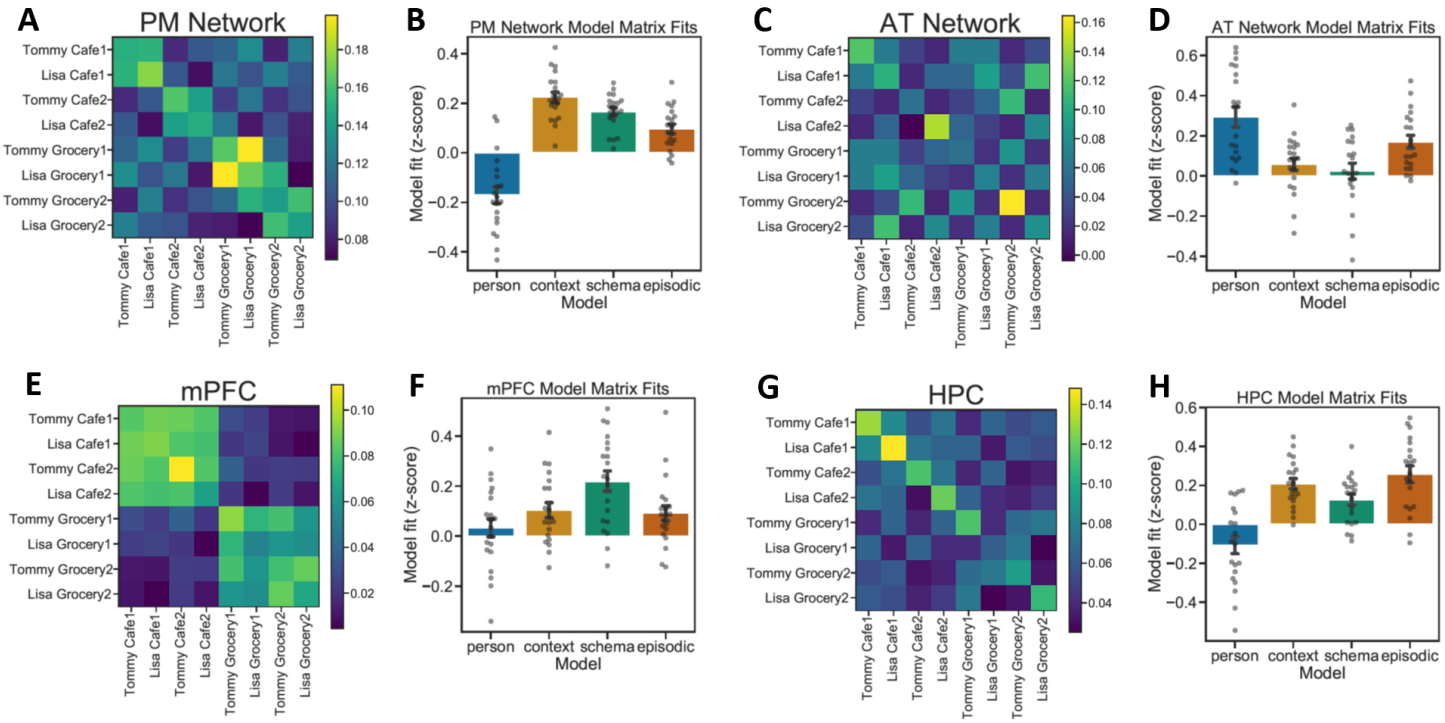

Figure 5: Content-selective pattern reinstatement across ROIs during spoken recall. (A,B) PM Network pattern reinstatement most strongly fits the Context model matrix. (C,D) AT Network pattern reinstatement most strongly fits the Person model matrix. (E,F) mPFC pattern reinstatement most strongly fits the Schema model matrix. (G,H) HPC pattern reinstatement most strongly fits the Episodic model matrix. (Error bars = SEM.)

Though the Context matrix was numerically the strongest fit to PMN reinstatement patterns at recall, this fit did not differ statistically from Schema and Episodic matrix fits at corrected thresholds.

Encoding-retrieval pattern similarity in the AT Network (Fig. 5C), was best described by the Person matrix ($r = 0.539$, $p = 2.904e^{-06}$), with a significant fit also to the Episodic matrix ($r = 0.307$, $p = 0.014$). The AT Network did not show significant correlations with the Context ($r = 0.053$, $p = 0.677$) or Schema ($r = 0.144$, $p = 0.255$) matrices. Model fits differed significantly ($F(3,57) = 12.625$, $p < 0.001$), driven by a significantly better fit between the Person matrix and the Context and Schema matrices ($p_{Tukey} < 0.05$ corrected) (Fig. 5D). This suggests that person-specific patterns present at encoding are reinstated during recall of events.

In mPFC (Fig. 5E), the encoding-recall pattern similarity data were best described by the Schema matrix ($r = 0.774$, $p = 5.258e^{-14}$), with significant fits also for the Context ($r = 0.424$, $p = 0.0001$) and Episodic matrices ($r = 0.369$, $p = 0.003$). The Person matrix fit was not significant ($r = 0.008$, $p = 0.944$). Model fits differed significantly ($F(3,57) = 7.276$, $p < 0.001$), driven by significantly stronger fit between the data and the Schema matrix than other model matrices ($p_{Tukey} < 0.05$ corrected) (Fig. 5F). Thus, in line with comparisons across encoding patterns for the depicted events, mPFC appears to reinstate generalized contextual representations consistent with schemas at recall.

Finally, HPC (Fig. 5G) encoding-retrieval pattern was most strongly characterized by the Episodic matrix ($r = 0.615$, $p = 6.548e^{-08}$), though fits to all 3 other model matrices were also significant: Context ($r = 0.521$, $p = 1.035e^{-05}$); Person ($r = 0.299$, $p = 0.016$); Schema ($r = 0.321$, $p = 0.009$). There was a significant different among model fits ($F(3,57) = 27.727$, $p < 0.001$). The Episodic matrix was a significantly stronger fit to the data than the Schema or Person matrices ($p_{Tukey} < 0.05$ corrected), though the Episodic and Context matrices did not differ significantly from one another (Fig.

5H). HPC activity patterns thus carry information about specific events not only across distinct encoding epochs, but also during spoken recall.

Similar to our approach at encoding (Fig. 4), we analyzed recall-related pattern reinstatement as a function of event type. In the PM Network, we found a significant effect of Context ($F(2,114) = 3.887$, $p = 0.023$), with significant effects in PMC and ANG, but not PHC (see Supplemental Information) (Fig. 6A). There was neither a significant effect of Person ($F(1,114) = 0.382$, $p = 0.536$), nor an interaction ($F(2,114) = 0.005$, $p = 0.995$). The AT Network showed a significant effect of Person ($F(1,114) = 4.844$, $p = 0.029$), which was found individually in PRC, but not TP (see Supplemental Information). We observed no effect of Context ($F(2,114) = 0.016$, $p = 0.984$) nor an interaction ($F(2,114) = 0.029$, $p = 0.971$) in the AT Network (Fig. 6B). In mPFC, we found a significant effect of Context ($F(2,114) = 5.455$, $p = 0.005$), but neither an effect of Person ($F(1,114) = 0.441$, $p = 0.508$) nor an interaction ($F(2,114) = 0.007$, $p = 0.993$) (Fig. 6C). Finally, HPC did not show a significant effect of either Person ($F(1,114) = 2.273$, $p = 0.134$) or Context ($F(2,114) = 1.164$, $p = 0.316$), but did show a significant interaction ($F(2,114) = 3.597$, $p = 0.031$). Reinstated patterns were weaker and individual mean differences less pronounced than comparisons across encoding epochs, and no post-hoc pairwise contrasts were significant at corrected thresholds.

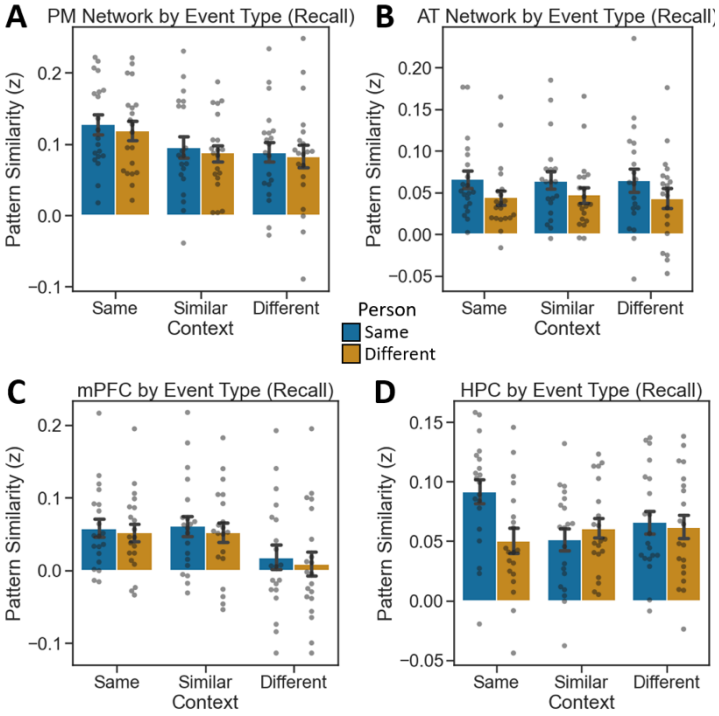

Figure 6: Recall pattern similarity by event type. (A) PM Network patterns show context-specificity. (B) AT Network patterns show person-specificity. (C) mPFC patterns generalize across similar contexts. (D) HPC patterns are episode-specific. (Error bars = SEM.)

Finally, we asked whether neural pattern reinstatement related to recall success. Prior studies have shown that encoding-retrieval similarity, particularly in or mediated by the HPC, correlate with memory performance (Gordon et al., 2014; Ritchey et al., 2013; Tompary et al., 2016; Wing et al., 2015). Among ROIs we examined, we only observed a relationship between hippocampal reinstatement for specific events (i.e., same-person + same-context) and the overall number of verifiable details recalled. Specifically, we found that the extent of pattern similarity between encoding and recall for the same events was significantly correlated with the total number of verifiable details retrieved for all events across participants ($r = 0.515$, $p = 0.02$) (Fig. 7). Taken together, these results indicate that neural patterns associated with complex events across multiple encoding episodes are, to some extent, brought back online during recall, and that this reinstatement of neural patterns follows information content dissociations present during encoding. Furthermore, the relationship between encoding and retrieval patterns for specific events in the HPC is correlated with recall success.

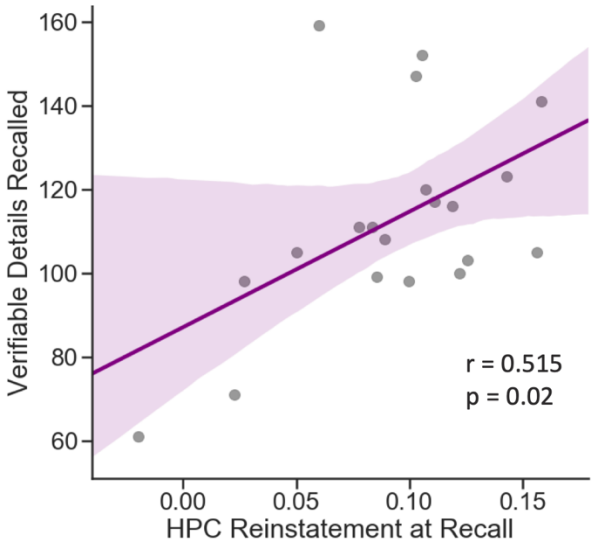

Figure 7: Correlation between HPC reinstatement at recall and the overall number of verifiable details retrieved across all events.

Distinct timescales of pattern similarity across regions for specific events

In the analyses described above, activity patterns were averaged across the entire duration of each movie, similar to other recent studies of naturalistic events (Bird et al., 2015; Chen et al., 2017; Oedekoven et al., 2017). Our video stimuli were fairly brief and our experiment was not explicitly designed to tease apart representations in terms of their temporal evolution as an event unfolds. However, given that event content unfolds dynamically over the course of an event, we next investigated the temporal dynamics of voxel pattern information during each video differed across regions. Prior studies suggest a hierarchy of representational timescales across the brain, with 'fast' timescale regions shifting their activity patterns frequently and 'slow' timescale regions shifting their activity patterns less frequently during an ongoing experience (Chen et al., 2015; Hasson et al., 2008; Hasson et al., 2015; Honey et al., 2012; Lerner et al., 2011). These findings suggest that regions of the DMN have among the slowest representational timescales, which are sensitive to narrative understanding (Simony et al., 2016) and have been shown to match human judgments of event transitions (Baldassano et al., 2017).

In order to assess each ROI in a comparable situation, we chose events in which the same individual was depicted in the same context, as all regions showed increases in pattern similarity for these comparisons. For these analyses, we modeled each TR in the run individually, in line with beta series approaches (Rissman et al., 2004). Each TR from a given event in one encoding run was correlated with its corresponding TR in that same event in a different run (corresponding to the Episode-Specific model matrix analysis). In order to simplify analyses, we chose to bin event epochs. Given prior evidence for the importance of HPC representations at event boundaries (Ben-Yakov et al., 2013; Lu et al., 2020), 40-second events were binned into event Onset (the first 8 TRs), event Offset (the last 8 TRs), and Mid event (the intermediate 17 TRs). For each ROI, each epoch was contrasted against its mean pattern similarity value across all events, across all runs (i.e., the grand mean). Each ROI was then subjected to a one-way ANOVA and post hoc Tukey's HSD tests, evaluating differences across epochs within each ROI (See Methods for further details).

In the PMN (Fig. 8A), pattern similarity for the same event across runs was significantly above the grand mean at Mid event ($t(19) = 2.723$, $p = 0.014$) and event Offset ($t(19) = 2.903$, $p = 0.009$), but not at event Onset ($t(19) = 1.518$, $p = 0.145$). However, event epochs did not differ significantly from one another ($F(2,38) = 0.877$, $p = 0.042$). In the ATN (Fig. 8B), pattern similarity was significantly above the grand mean at event Onset ($t(19) = 2.115$, $p = 0.048$) and Mid event ($t(19) = 4.916$, $p = 9.594e^{-05}$), but not at event Offset ($t(19) = 1.269$, $p = 0.219$). Event epochs differed significantly ($F(2,38) = 4.903$, $p = 0.013$), and Mid event pattern similarity was significantly higher than event Offset ($p_{Tukey} < 0.05$ corrected). In

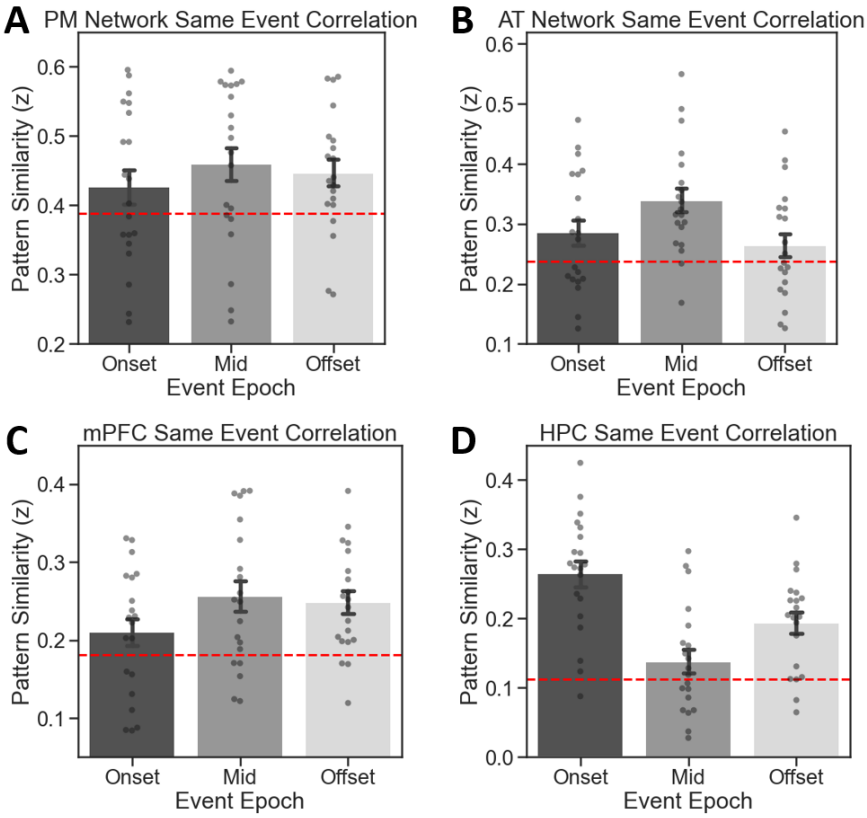

Figure 8: Pattern similarity for same-event comparisons by event epoch. (A) PM Network, (B) AT Network, (C) mPFC, (D) HPC. Red lines indicate the mean pattern similarity value across all events and across all runs for that ROI. (Error bars = SEM.)

mPFC (Fig. 8C), pattern similarity was significantly above the grand mean at Mid event ($t(19) = 3.672$, $p = 0.002$) and event Offset ($t(19) = 4.427$, $p = 0.0002$), but not event Onset ($t(19) = 1.611$, $p = 0.124$). Event epochs did not differ significantly ($F(2,18) = 1.753$, $p = 0.187$). In HPC (Fig. 8D), pattern similarity was significantly different from the grand mean at event Onset ($t(19) = 7.983$, $p = 1.729e^{-07}$) and event Offset ($t(19) = 5.079$, $p = 6.656e^{-05}$), but not at the Mid event epoch ($t(19) = 1.491$, $p = 0.153$). Event epochs differed significantly from one another ($F(2,38) = 18.836$, $p = 2.07e^{-06}$), driven by significantly higher pattern similarity at event Onset than Mid event or event Offset ($p_{Tukey} < 0.05$ corrected). Thus, while all ROIs showed significant increases in pattern similarity for same-event comparisons across runs, these increases occurred at different parts of the event across regions. Notably, HPC showed a dramatic increase in PS at event onset, and was the only ROI showing no significant difference from the grand mean at Mid event timepoints.

Discussion

Recent research has strongly implicated the DMN as being important for the ongoing perception of and memory for naturalistic events. However, there is also strong evidence that the DMN does not operate as one unitary entity (Ritchey & Cooper, 2020). A wealth of studies has established that the DMN may in actuality be comprised of at least two cortico-hippocampal networks in support of memory: a PM Network supporting contextual or situational details, and an AT Network supporting information about local entities, such as objects or people (Ranganath & Ritchey, 2012; Reagh & Ranganath, 2018; Ritchey, Libby, et al., 2015). However, the question has remained unresolved whether these networks differentially process information in real time, in complex real-world situations. We addressed this question by creating video stimuli depicting lifelike events, which were viewed and recalled by participants during fMRI scanning. Across multiple types of neural pattern similarity analyses, we demonstrate that regions of the PM Network flexibly use and re-use representations of specific contexts irrespective of which specific focal individual is being observed. Like the PM Network, mPFC was sensitive to contextual features. However, in line with a growing body of evidence, we found that its context representations were generalized across similar situations rather than being context-specific. This is consistent with a role in representing generalized schemas (Ghosh & Gilboa, 2014; Gilboa & Marlatte, 2017; Preston & Eichenbaum, 2013; Spalding et al., 2015; van Kesteren et al., 2013). Regions of the AT Network also feature flexible representations, but of individual people irrespective of their contextual associations. HPC patterns were highly event-specific, such that pattern similarity across encoding of events was highest when the same event, featuring the same person in the same context, was viewed. At recall, HPC pattern reinstatement correlated with retrieved verifiable details about events. Finally, we show evidence for distinct representational timescales for a given event, with HPC being unique in showing the highest pattern similarity across instances of the same event at event onset.

Previous characterizations of the DMN have consistently pointed to an important role in memory (Andrews-Hanna et al., 2010; Buckner et al., 2008; Gilmore et al., 2015; Kim, 2012; Sestieri et al., 2011; Spreng & Grady, 2010). However, the precise role of the DMN and the extent to which regional contributions may be shared or unique have been matters of ongoing debate. There is strong evidence that the DMN can be broken into smaller sub-networks on the basis of functional attributes and differences in content-sensitivity (Andrews-Hanna et al., 2010, 2014; Barnett et al., 2020; Braga et al., 2019; Ranganath & Ritchey, 2012), and there is growing evidence that sub-networks may exist even within the PM Network (Ritchey & Cooper, 2020). A recent study by Morton and colleagues (Morton et al., 2021) provides especially compelling evidence for this, and in line with our findings, the authors show pattern similarity among famous people in the AT Network and famous spatial locations in the PM Network on the basis of semantic knowledge. The view of the DMN as consisting of at least two sub-networks is seemingly at odds with views ascribing broad roles in domain-agnostic episodic

(Bonnici et al., 2016; Kuhl & Chun, 2014; Lee & Kuhl, 2016; Thakral et al., 2017) or semantic (Binder & Desai, 2011; Devereux et al., 2013; Rugg & King, 2018) memory retrieval across DMN regions, raising the question: what underlies this discrepancy? Many typical laboratory experiments use information of a single modality (e.g., objects or words, or simple associations between these), which may cause otherwise heterogeneous brain networks to act in a more unitary fashion. Relatedly, even if participants recall precise item-level information in these experiments, this may be accompanied by strong reinstatement of abstract representations (e.g., an image of a specific cup may trigger an abstract 'cup' representation) or contextual associations (e.g., an image of a cup may trigger representations of a familiar coffee shop) which are very difficult to control or even assess across participants. A clear difference between the present study and many related prior experiments is that the stimuli in question were rich, dynamic individuals in numerous contexts which were themselves rich and dynamic. This may have permitted the construction and use of event models – mental representations of characters and situations that are strongly tied to the events depicted on the screen, but which also tap into more abstract processes required to understand events like those that occur in the real world. Importantly, the use of event models is a latent, uncontrolled variable in typical lab memory studies. While the present experiment certainly did not entirely control this factor, the present design may have constrained their use in such a way that specific representational components could be identified. Finally, testing participants on stimuli with a rich event structure may have itself been a driving force in eliciting strong representations of event content.

Although our data provide strong support for content-based dissociations of the PM and AT Networks, we do not take this to suggest that all regions of these networks are doing exactly the same thing. Our results demonstrate that some information content is indeed shared across these networks, but more fine-grained variation of event content than is present in our design could further dissociate regional differences. Our findings allude to this possibility in contrasting fits to the Context and Schema model matrices, where mPFC (which is strongly affiliated with PM Network regions) showed more schematic representations as opposed to the more context-specific PM Network regions. However, we note that we observed fairly strong fits between PM Network regions and the Schema model matrix, and between mPFC and the Context model matrix. Thus, context-specificity and generalization likely exist in varying degrees, distributed across the PM Network. This is consistent with other findings from Baldassano and colleagues (Baldassano et al., 2018), who used stereotypical progressions of events through airports and restaurants in movie and audio clips to test shared neural responses across different stories featuring schema-congruent event sequences. Baldassano and colleagues reported schema-level representations not only in mPFC, but across other PM Network regions such as PMC and PHC. Their study differs markedly from our approach in that they sought to examine shared neural patterns across highly distinct events sharing a common schematic theme, whereas our approach was to systematically recombine characters and contexts to interrogate representations of specific event content. However, these findings are highly complementary.

Though the hippocampus is also universally agreed to be critical for mnemonic processes, the exact nature of its role is also a matter of ongoing debate. One widely held view of the hippocampus is that it is important for creating and maintaining unique representations of specific events. This has often been linked to its theoretical role in pattern separation (Marr, 1971; O'Reilly & McClelland, 1994; Yassa & Stark, 2011), which has been bolstered by empirical research in animal models (Hunsaker & Kesner, 2013; Leutgeb et al., 2007; Neunuebel & Knierim, 2014) and humans (Bakker et al., 2008; Berron et al., 2016; LaRocque et al., 2013). Our results, indicating that neural patterns are specific to particular person-context associations, are broadly consistent with this view. Our findings also align with studies showing a relationship between reinstatement of specific hippocampal patterns and retrieval success (Gordon et al., 2014; Ritchey et al., 2013; Tompary et al., 2016; Wing et al., 2015). In contrast to representational specificity, other studies strongly suggest that the hippocampus has a role in our ability to generalize across similar situations (Berens & Bird, 2017; Bowman & Zeithamova, 2018; Kumaran & McClelland, 2012; Schlichting et al., 2015). Indeed, the hippocampus has the

capacity to support specific pattern-separated representations as well as to support generalization via pattern completion (Rolls, 2013). Though our data do not provide clear support for hippocampal generalization, it is possible that the relative lack of specific instructions during event encoding resulted in a stronger bias to create and use specific hippocampal representations rather than to generalize across similar instances. That is, if individuals were incentivized to generalize across similar events (e.g., strategically collapse across events featuring the character Lisa), hippocampal representations may be driven away from specificity and more toward generalization. This possibility can be tested by follow-up experiments.

Closely related to the topic of event memory is that of event structure. Recent studies using naturalistic stimuli such as films have primarily focused on event boundaries rather than the content of events per se. A finding with increasing support is that the hippocampus and certain PM Network regions are transiently more active at event boundaries (Baldassano et al., 2017; Ben-Yakov & Henson, 2018; Reagh et al., 2020), and may moreover strengthen their functional coupling at event boundaries (Cooper et al., 2021). These boundary-evoked responses are not observed in AT Network regions (Reagh et al., 2020). The present study, in contrast, was not designed to examine event segmentation, but rather to examine the way that content being shared or not shared across events may result in more or less representational overlap, and how this may differ as a function of information modality. However, some insight may be gained from our results. Coarse event boundaries (which are fairly linked to story-related changes rather than individual actions) are reliably elicited by changes in spatiotemporal context (Zacks, 2010; Zacks, Speer, et al., 2009; Zwaan et al., 1995). Based on the present findings, it is reasonable to expect that a shift in spatiotemporal context in an ongoing narrative would drive changes in PM Network regions, and would not do so in AT Network regions. This raises the interesting question of whether character or item-driven event boundaries would reverse this network-level pattern of engagement at event boundaries.

Relating to event structure, our data reveal not only evidence of distinct content-selective contributions of different brain regions to representing complex, naturalistic events, but also that these representations may come online at distinct epochs as events unfold. The PMN and mPFC showed a more gradual increase in pattern similarity, not differing from the grand mean at event onset, but sustaining above the mean at mid-event timepoints and throughout event offset. In contrast, the ATN showed above-mean pattern similarity at event onset and mid-event timepoints, but fell at event offset. Finally, HPC showed a dramatic increase in pattern similarity at event onset, falling markedly at mid-event timepoints, and increasing again at event offset. This suggests that representational timescales may be another important factor which differs across regions, broadly consistent with previous work suggesting relatively faster and slower pattern changes during temporally extended stimuli such as movies (Chen et al., 2015; Hasson et al., 2008; Hasson et al., 2015; Honey et al., 2012; Lerner et al., 2011). With regards to HPC, however, our result of relatively higher pattern similarity at event onset and offset than mid-event aligns with the idea of HPC being particularly important for encoding information at event boundaries (Ben-Yakov et al., 2013; Lu et al., 2020). Lu et al. used a computational modeling approach to find that selectively storing information at event boundaries (but not mid-event) was beneficial to storing and later retrieving that information (Lu et al., 2020). Relatedly, a recent study by Cohn-Sheehy and colleagues suggests that event boundaries may act not only as moments in which an event transition occurs, but moments in which related information can be integrated across distinct experiences (Cohn-Sheehy et al., 2020). Future studies should continue to investigate representational timescales with respect to specific event components.

We note that, although the present study tested depictions of lifelike situations, the events themselves were fairly limited in scope and in time. Each event was 40 seconds in duration, limiting analyses of the way different representations of event components may evolve over long timescales. We also did not systematically vary features of characters such as actions, intentions, and other similar factors which could influence event understanding (Richmond & Zacks, 2017; Zacks et al., 2007; Zacks,

Kumar, et al., 2009). However, having established what may be core components of event representations, future studies can more closely examine these factors with respect to the content of event representations. Another topic of interest is that of schema congruency, which has been found to affect recognition and recall of information (Brod et al., 2015; A. Greve et al., 2019; Richter et al., 2019; van Kesteren et al., 2012). However, few studies manipulating and testing effects of schema congruency per se have done so testing memory for lifelike situations (Bonasia et al., 2018). An approach such as the one used in the present design could be modified to manipulate congruency between people or their actions and their contextual associations.

In sum, our results address a gap in existing studies of event cognition and memory by establishing that content dissociations classically observed in studies of memory extend to naturalistic events. Spanning multiple lifelike events with content overlap being systematically varied, we observed content-selective representations in PM and AT cortico-hippocampal networks, schematic representation in mPFC, and episode-specific representation in HPC consistent with a role in binding different information streams into unique episodic codes. Together, these components provide the neural substrates to scaffold representations on the basis of information modality and specificity, allowing us to flexibly represent and remember complex, real-world events.

Methods

Participants

Participants were recruited from the University of California, Davis and the Davis, California community via flyer and email advertisement. Participants were paid $25 per hour, and gave written informed consent in accordance with the University of California, Davis Institutional Review Board under an approved protocol. Exclusion criteria included magnetic bodily implants, claustrophobia, a history of major psychiatric or neurological disorders, a concussion in the past 6 months, untreated diabetes or hypertension, current drug or alcohol abuse, left-handedness, age outside a range of 18-35 years, and 4 or fewer hours of sleep on the day of data collection.

We initially collected 24 participants, and excluded 1 due to falling asleep during MRI scanning, 2 due to excessive motion in the MRI scanner, and 1 due to equipment malfunction during data collection. The final sample of 20 participants had a mean age of 23.4 years (SD = 3.1 years), consisting of 13 self-identified females and 7 self-identified males.

Stimuli

8 video clips were developed for use in the present experiment. These video clips were recorded at locations in Davis, California using a GoPro HERO+LCD camera, at 1080p 60fps resolution. Two members of the UC Davis Dynamic Memory Lab served as central characters. Several ‘takes’ were recorded for each video. Videos were then edited in the GoPro editing software to be 35 seconds in duration, and to include a 5 second title screen with white text against a black background for each event depicted (making each video 40 seconds in total).

Prior to video viewing, participants viewed a series of still images depicting the characters and locations subsequently viewed in the event videos. These data were not analyzed here. Videos were presented in the MRI scanner environment using PsychoPy version 3.1.5, and stimulus onset was synchronized with MRI pulse sequences via fiber optic trigger pulses sent to the stimulus computer. Videos were downsampled to from 1080p to 720p in the experiment to reduce the time taken to load

and buffer the files in PsychoPy. After video viewing, participants were cued by on-screen text via PsychoPy to recall the events using an MRI-compatible microphone (see Procedure).

After in-scanner tasks, participants were given a True/False recognition test outside the scanner consisting of a series of single-sentence statements about video content. False statements were designed to probe for accuracy of memory, but were not designed to test fine-grained mnemonic discrimination (i.e., they were not designed to be highly similar to factual events). Single sentences were presented in white text against a black background on a 2015 MacBook Pro 13” laptop.

Procedure

Participants arrived at the UC Davis Center for Neuroscience, and were escorted into the Dynamic Memory Lab testing space where they were briefed on the experiment, and gave written informed consent. Participants then completed an MRI screening form and demographics questionnaires. Upon successful screening for the experiment, participants were escorted to the Imaging Research Center, and were prepped to undergo scanning.

Scanning took place in a Siemens 3T Skyra with a 32-channel head coil. Participants were fitted with MRI-compatible earbuds with replaceable foam inserts (MRIaudio), and were provided with additional foam padding inside the head coil for hearing protection and to mitigate head motion. Participants were additionally given bodily padding, blankets, and corrective lenses as needed. An MRI-compatible microphone (Optoacoustics FOMRI-III) with bidirectional noise-cancelling was affixed to the head coil, and the receiver (covered by a disposable sanitary pop screen) was positioned over the participant’s mouth. Participants were given a description of strategies to remain still while speaking during functional image acquisition. During MRI data acquisition, an eye tracker was operational to monitor participants’ wakefulness and head motion during spoken recall, but no eye tracking data were recorded.

After testing the earbuds and microphone, high-resolution T1-weighted structural images were acquired using a magnetization prepared rapid acquisition gradient echo (MPRAGE) pulse sequence (FOV = 256mm, image matrix = 256 x 256, sagittal slices = 208, thickness = 1mm). Participants then completed a functional imaging run in which they were shown still images of the central people and contexts subsequently shown in the videos (these data are not analyzed or discussed here). Participants then completed 3 runs in which all 8 event videos were shown, in random order across runs and across participants. Video stimuli (title screen + event) were 40 seconds, with a 10 second inter-stimulus interval displaying a fixation cross. Instructions were to remain still and closely attend to the videos, as memory for the videos would be tested later in the experiment. After the 3 encoding runs were completed, participants completed spoken recall of each event. Event titles were displayed in random order, in white text against a black background for 40 seconds, with a 10 second inter-stimulus interval displaying a fixation cross. Participants were instructed to begin recalling the named event, in as much detail as possible and in order to the extent possible, and to stop recall either when finished or when the event title transitioned to the fixation cross. Functional images were acquired using a multi-band gradient echo planar imaging (EPI) sequence (TR = 1220ms, TE = 24ms, FOV = 192mm, image matrix = 64 x 64, flip angle = 67, multi-band factor = 2, axial slices = 38, voxel size = 3mm isotropic, P >> A phase encoding, AC-PC alignment). A single 4 TR functional scan of reverse phase encoding polarity (A >> P) was acquired for unwarping (see fMRI data preprocessing below).

Behavioral analyses

In-scanner recall was scored using an adapted version of the Autobiographical Memory Interview ((Diamond et al., 2020; Levine et al., 2002; see Cohn-Sheehy et al., 2020 for a highly similar adapted approach). Three scorers transcribed the audio recording for each event into text, and segmented the document into meaningful detail units (Z.M.R. and two research assistants). Detail units refers to the smallest possible meaningful unit, and labels were assigned to those details on the basis of their content. These details were then classified as “verifiable” if they referred to a factually accurate piece of information pertaining to the events depicted in the videos, and were not preceded by statements of uncertainty (e.g., “maybe”). Redundant information was not counted. Once scoring was completed, interrater reliability was assessed, and was fairly high across the three raters overall (Pearson r = 0.86), and in terms of details scored as being verifiable (Pearson r = 0.84). Verifiable details were compared across characters and across contexts using one-way ANOVAs, and were correlated with pattern similarity and recall-driven reinstatement using Pearson correlations.

Out-of-scanner recognition test results for each participant considered distinction of true statements from false statements about the viewed events. To quantify this, we calculated a d’ score for each participant: z(Hit rate) – z(False Alarm rate). Similar to in-scanner recall performance above, recognition memory scores were compared across characters and across contexts using one-way ANOVAs, and standard Pearson correlations were used to assess relationships with neural data across events. The ultimate goal of these analyses was to ensure that recognition and recall of event information were significantly nonzero, and unbiased across the events.

fMRI data preprocessing

Results included in this manuscript come from preprocessing performed using FMRIPREP version stable (Esteban et al., 2019, 2020, RRID:SCR_016216), a Nipype (Gorgolewski et al., 2011, 2017, RRID:SCR_002502) based tool. Each T1w (T1-weighted) volume was corrected for INU (intensity non-uniformity) using N4BiasFieldCorrection v2.1.0 (Tustison et al., 2010) and skull-stripped using antsBrainExtraction.sh v2.1.0 (using the OASIS template). Brain surfaces were reconstructed using recon-all from FreeSurfer v6.0.1 (Dale et al., 1999, RRID:SCR_001847), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (Klein et al., 2017, RRID:SCR_002438). Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2009, RRID:SCR_008796) was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 (Avants et al., 2008, RRID:SCR_004757), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (Zhang et al., 2001) (FSL v5.0.9, RRID:SCR_002823).

Functional data was slice time corrected using 3dTshift from AFNI v16.2.07 (Cox, 1996, RRID:SCR_005927) and motion corrected using mcflirt (FSL v5.0.9, Jenkinson et al., 2002). Distortion correction was performed using an implementation of the TOPUP technique (Andersson et al., 2003) using 3dQwarp (AFNI v16.2.07, Cox, 1996). This was followed by co-registration to the corresponding T1w using boundary-based registration (D. N. Greve & Fischl, 2009) with six degrees of freedom, using bbregister (FreeSurfer v6.0.1). Motion correcting transformations, field distortion correcting warp, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos interpolation.

Physiological noise regressors were extracted applying CompCor (Behzadi et al., 2007). Principal components were estimated for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). A mask to exclude signal with cortical origin was obtained by eroding the brain mask, ensuring it only contained subcortical structures. Six tCompCor components were then calculated including only the top 5% variable voxels within that subcortical mask. For aCompCor, six components were calculated within the intersection of the subcortical mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run. Frame-wise displacement (Power et al., 2014) was calculated for each functional run using the implementation of Nipype.

Many internal operations of FMRIPREP use Nilearn (Abraham et al., 2014, RRID:SCR_001362), principally within the BOLD-processing workflow. For more details of the pipeline see https://fmriprep.readthedocs.io/en/stable/workflows.html.

fMRI data analysis – representational similarity analysis

All representational similarity analyses (RSA) were run on unsmoothed native-space functional images after the preprocessing steps described above. Each of the 3 encoding runs and the recall run of the raw data were entered into a general linear model (GLM) in AFNI using the 3dDeconvolve function. Nuisance regressors for linear scanner drift (first order polynomial), head motion (6 directions plus their 6 derivatives), and the first two principal components of a combined white matter and CSF mask (via aCompCor above). Given the relatively long 40 second duration of the 'trials' being modeled and the 10 second gap between the videos, a least-squares all (LSA) approach was used, such that each run was modeled to produce beta estimate images for either encoding or recall of each of the 8 events. Though a number of prior studies examining naturalistic encoding and recall have not used GLMs to model data, and have instead used a time-shifted averaging approach, we chose to model the present data as this allowed us to more directly mitigate influences of signal drift, motion, and global nuisance signal in evaluating representational patterns. As events, including title screens, were 40 seconds in duration and the TR for our EPI sequence was 1.22 seconds, there were approximately 33 TRs covering each event.

Resultant beta images were masked using a region-of-interest (ROI) approach, based on theoretical interest in areas of the Posterior-Medial and Anterior-Temporal Networks. ROIs in the Posterior-Medial Network included PMC, ANG, and PHC. ROIs in the Anterior-Temporal Network included PRC and TP (lateral orbitofrontal cortex is theoretically a part of the AT Network, but was excluded from these analyses due to poor MRI signal quality in the region). We additionally included mPFC and hippocampal ROIs. Hippocampal, parahippocampal, and perirhinal ROIs were derived from in-house hand-traced template images in MNI space, which were warped to each subject's native space using ANTs. All other ROIs are derived from the FreeSurfer Desikan atlas. Multi-voxel patterns were extracted from each ROI mask and written to CSV files for subsequent analyses using the 3dmaskdump function in AFNI. Voxels with null values in any scanning run were excluded.

For model matrix comparisons (see below), each event pattern was correlated with each other event pattern. For conditional comparisons across RSA results, we collapsed events across 4 key conditions: (1) same event (i.e., person-in-context), (2) same person, different context, (3) same person, similar context, (4) same context, different person), and (5) different person, different context. Pairwise comparisons across each event were conducted within subjects, and resulting correlation coefficients were z-scored. For each ROI, events satisfying each condition were averaged, and conditional averages were statistically compared against one another using two-way ANOVAs, with person and context information as factors. Significant effects were evaluated using Tukey's HSD for post-hoc comparisons.

fMRI data analysis – model matrix comparisons

To probe specific hypotheses about representation of person, context, schema, and episodic specificity in our ROIs, we compared event-by-event correlation matrices for pattern similarity data to model matrices. Beta images resulting from the preprocessing steps described above underwent pairwise comparisons across events for each participant and for each ROI, resulting in an event-by-event correlation matrix. This was compared to four model matrices: (1) an episode-specific matrix, in which events were hypothesized to correlate only with themselves, (2) a person matrix, in which events were hypothesized to correlate with those events which featured the same person, (3) a context matrix, in which events were hypothesized to correlate with those events which featured the same context, and (4) a schema matrix, in which events were hypothesized to correlate with those events which featured a *similar* context (e.g., two similar but not identical cafes). Model matrices were constructed simply to feature a value of 1 in hypothesized high-correlation cells, and a value of 0 in hypothesized low-correlation cells. In general, model fit was evaluated using a point-biserial correlation between the pattern similarity matrix and the model matrices for each ROI. Model fits were evaluated in two ways. First, for each ROI, we computed the biserial correlation between group average event-wise pattern similarity data and each of the 4 model matrices, yielding overall model matrix correlations per region. Second, to incorporate sensitivity to variability across participants and to directly compare different model fits per region, we assessed the model fits for each participant per ROI. For this approach, each participant's model fit correlations were z-scored, and correlations were aggregated across participants. Finally, one-way ANOVAs and post-hoc Tukey's HSD comparisons were used to relate the overall fits of the empirical data to the models, and to directly compare fits across models within each ROI. We note that, with this approach, some model matrices will share mutual information (e.g., the Episodic matrix is a component of all other matrices), but we can nonetheless assess the relative fit of each matrix for a given region.

fMRI data analysis – time series correlations

Additional analyses were conducted examining pattern similarity across events on the order of TRs. Preprocessed data were modeled distinctly from prior analyses: rather than a beta image being generated for each *event*, a beta image was generated for each *time step* comprising an event (similar to a beta-series correlation approach (Rissman et al., 2004). Confound regressors were unchanged from the other GLM which modeled whole events with single regressors. The analysis consisted of a TR-by-TR pattern similarity analysis within an event pair (rather than event-by-event pattern similarity analysis) across ROIs and collapsed across participants. Given that different ROIs responded to different conditions, we chose to compare only same-event neural patterns across runs for this analysis, as this is the only condition in which all ROIs reliably showed pattern similarity increases across participants. Furthermore, given evidence for distinct mechanisms of hippocampal encoding at event boundaries (Ben-Yakov et al., 2013; Lu et al., 2020), we binned the 33-TR time series into event Onset (the first 8 TRs), event Offset (the final 8 TRs), and Mid event timepoints (the intermediate 17 TRs). This additionally simplified analyses of the data while nonetheless allowing for evaluation of a temporally-evolving change in pattern similarity. We conducted two sets of analyses for these data. The first analysis was a series of one-sample t-tests against the mean pattern similarity value across all encoding runs, evaluating whether pattern similarity for same-event comparisons was significantly different from the baseline pattern similarity of that ROI across all events, at each of the 3 epochs we defined. The second analysis was a one-way ANOVA with post hoc Tukey's HSD to test whether the 3 event epochs differed significantly from one another.

Acknowledgments

We thank Jessica Macaluso and Ryan Bugsch for assisting with scoring of recall transcripts. We thank current and former members of the Dynamic Memory Lab for crucial feedback, with special thanks to Brendan Cohn-Sheehy, Jordan Crivelli-Decker, and Derek Huffman. We thank Alex Barnett and Kamin Kim for acting in the videos in addition to providing helpful input. We thank Sarah Morse for helpful input on the manuscript. We thank Delta of Venus, Mishka's Café, The Nugget, and Davis Food Co-Op in Davis, California for allowing us to record videos featuring their businesses. Finally, we thank our funding sources supporting this work: ONR Grant N00014-15-1-0033 to C.R. and NIA T32AG050061 to Z.M.R.

Author Contributions

Z.M.R. and C.R. designed the research and wrote the paper. Z.M.R. conducted the experiments and analyzed the data.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*. https://doi.org/10.3389/fninf.2014.00014

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*, *20*(2), 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-Anatomic Fractionation of the Brain's Default Network. *Neuron*, *65*(4), 550–562. https://doi.org/10.1016/j.neuron.2010.02.005

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*(1), 29–52. https://doi.org/10.1111/nyas.12360

Avants, B., Epstein, C., Grossman, M., & Gee, J. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern Separation in the Human Hippocampal CA3 and Dentate Gyrus. *Science*, *319*(5870), 1640–1642. https://doi.org/10.1126/science.1152882

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709-721.e5. https://doi.org/10.1016/j.neuron.2017.06.041

Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of Real-World Event Schemas during Narrative Perception. *The Journal of Neuroscience*, *38*(45), 9689–9699. https://doi.org/10.1523/JNEUROSCI.0251-18.2018

Barnett, A. J., Reilly, W., Dimsdale-Zucker, H., Mizrak, E., Reagh, Z., & Ranganath, C. (2020). Organization of cortico-hippocampal networks in the human brain. *BioRxiv*, 2020.06.09.142166. https://doi.org/10.1101/2020.06.09.142166

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Ben-Yakov, A., Eshel, N., & Dudai, Y. (2013). Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. *Journal of Experimental Psychology: General*,

142(4), 1255–1263. https://doi.org/10.1037/a0033558
Ben-Yakov, A., & Henson, R. N. (Eds.). (2018). The Hippocampal Film Editor: Sensitivity and Specificity to Event Boundaries in Continuous Experience. *The Journal of Neuroscience*, *38*(47), 10057–10068. https://doi.org/10.1523/JNEUROSCI.0524-18.2018
Berens, S. C., & Bird, C. M. (2017). The role of the hippocampus in generalizing configural relationships. *Hippocampus*, *27*(3), 223–228. https://doi.org/10.1002/hipo.22688
Berron, D., Schutze, H., Maass, A., Cardenas-Blanco, A., Kuijf, H. J., Kumaran, D., & Duzel, E. (2016). Strong Evidence for Pattern Separation in Human Dentate Gyrus. *Journal of Neuroscience*, *36*(29), 7569–7579. https://doi.org/10.1523/JNEUROSCI.0518-16.2016
Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–536. https://doi.org/10.1016/j.tics.2011.10.001
Bird, C. M., Keidel, J. L., Ing, L. P., Horner, A. J., & Burgess, N. (2015). Consolidation of Complex Events via Reinstatement in Posterior Cingulate Cortex. *Journal of Neuroscience*, *35*(43), 14426–14434. https://doi.org/10.1523/JNEUROSCI.1774-15.2015
Bonasia, K., Sekeres, M. J., Gilboa, A., Grady, C. L., Winocur, G., & Moscovitch, M. (2018). Prior knowledge modulates the neural substrates of encoding and retrieving naturalistic events at short and long delays. *Neurobiology of Learning and Memory*, *153*, 26–39. https://doi.org/10.1016/j.nlm.2018.02.017
Bonnici, H. M., Richter, F. R., Yazar, Y., & Simons, J. S. (2016). Multimodal Feature Integration in the Angular Gyrus during Episodic and Semantic Retrieval. *Journal of Neuroscience*, *36*(20), 5462–5471. https://doi.org/10.1523/JNEUROSCI.4310-15.2016
Bowman, C. R., & Zeithamova, D. (2018). Abstract Memory Representations in the Ventromedial Prefrontal Cortex and Hippocampus Support Concept Generalization. *The Journal of Neuroscience*, *38*(10), 2605–2614. https://doi.org/10.1523/JNEUROSCI.2811-17.2018
Braga, R. M., Van Dijk, K. R. A., Polimeni, J. R., Eldaief, M. C., & Buckner, R. L. (2019). Parallel distributed networks resolved at high resolution reveal close juxtaposition of distinct regions. *Journal of Neurophysiology*, *121*(4), 1513–1534. https://doi.org/10.1152/jn.00808.2018
Brod, G., Lindenberger, U., Werkle-Bergner, M., & Shing, Y. L. (2015). Differences in the neural signature of remembering schema-congruent and schema-incongruent events. *NeuroImage*, *117*, 358–366. https://doi.org/10.1016/j.neuroimage.2015.05.086
Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The Brain's Default Network. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. https://doi.org/10.1196/annals.1440.011
Chen, J., Hasson, U., & Honey, C. J. (2015). Processing Timescales as an Organizing Principle for Primate Cortex. *Neuron*, *88*(2), 244–246. https://doi.org/10.1016/j.neuron.2015.10.010
Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450
Cohn-Sheehy, B. I., Delarazan, A. I., Reagh, Z. M., Crivelli-Decker, J. E., Zacks, J. M., & Ranganath, C. (2020). Bridges, not walls: the hippocampus builds narrative memories across distant events. *BioRxiv*, 2020.12.11.422162. https://doi.org/10.1101/2020.12.11.422162
Cooper, R. A., Kurkela, K. A., Davis, S. W., & Ritchey, M. (2021). Mapping the organization and dynamics of the posterior medial network during movie watching. *BioRxiv*, 2020.10.21.348953. https://doi.org/10.1101/2020.10.21.348953
Cox, R. W. (1996). AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014
Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395
Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational Similarity Analysis Reveals Commonalities and Differences in the Semantic Processing of Words and Objects. *Journal of Neuroscience*, *33*(48), 18906–18916. https://doi.org/10.1523/JNEUROSCI.3809-13.2013

Diamond, N. B., Armson, M. J., & Levine, B. (2020). The Truth Is Out There: Accuracy in Recall of Verifiable Real-World Events. *Psychological Science*, *31*(12), 1544–1556. https://doi.org/10.1177/0956797620954812

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Annual Review of Neuroscience*, *30*(1), 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Esteban, O., Markiewicz, C. J., Goncalves, M., DuPre, E., Kent, J. D., Salo, T., Ciric, R., Pinsard, B., Blair, R. W., Poldrack, R. A., & Gorgolewski, K. J. (2020). *fMRIPrep: a robust preprocessing pipeline for functional MRI*. https://doi.org/10.5281/ZENODO.4252786

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*, 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Gilboa, A., & Marlatte, H. (2017). Neurobiology of Schemas and Schema-Mediated Memory. *Trends in Cognitive Sciences*, *21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013

Gilmore, A. W., Nelson, S. M., & McDermott, K. B. (2015). A parietal memory network revealed by multiple MRI methods. *Trends in Cognitive Sciences*, *19*(9), 534–543. https://doi.org/10.1016/j.tics.2015.07.004

Gold, J. J., Smith, C. N., Bayley, P. J., Shrager, Y., Brewer, J. B., Stark, C. E. L., Hopkins, R. O., & Squire, L. R. (2006). Item memory, source memory, and the medial temporal lobe: Concordant findings from fMRI and memory-impaired patients. *Proceedings of the National Academy of Sciences*, *103*(24), 9351–9356. https://doi.org/10.1073/pnas.0602716103

Gordon, A. M., Rissman, J., Kiani, R., & Wagner, A. D. (2014). Cortical Reinstatement Mediates the Relationship Between Content-Specific Encoding Activity and Subsequent Recollection Decisions. *Cerebral Cortex*, *24*(12), 3350–3364. https://doi.org/10.1093/cercor/bht194

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics*, *5*. https://doi.org/10.3389/fninf.2011.00013

Gorgolewski, K. J., Esteban, O., Ellis, D. G., Notter, M. P., Ziegler, E., Johnson, H., Hamalainen, C., Yvernault, B., Burns, C., Manhães-Savio, A., Jarecka, D., Markiewicz, C. J., Salo, T., Clark, D., Waskom, M., Wong, J., Modat, M., Dewey, B. E., Clark, M. G., … Ghosh, S. (2017). *Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. 0.13.1*. https://doi.org/10.5281/ZENODO.581704

Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology: General*, *148*(2), 325–341. https://doi.org/10.1037/xge0000498

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A Hierarchy of Temporal Receptive Windows in Human Cortex. *Journal of Neuroscience*, *28*(10), 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008

Hasson, Uri, Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences*, *19*(6), 304–313. https://doi.org/10.1016/j.tics.2015.04.006

Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., Doyle, W. K., Rubin, N., Heeger, D. J., & Hasson, U. (2012). Slow Cortical Dynamics and the Accumulation of

Information over Long Timescales. *Neuron*, *76*(2), 423–434. https://doi.org/10.1016/j.neuron.2012.08.011
Hunsaker, M. R., & Kesner, R. P. (2013). The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neuroscience & Biobehavioral Reviews*, *37*(1), 36–58. https://doi.org/10.1016/j.neubiorev.2012.09.014
Irish, M., & Vatansever, D. (2020). Rethinking the episodic-semantic distinction from a gradient perspective. *Current Opinion in Behavioral Sciences*, *32*, 43–49. https://doi.org/10.1016/j.cobeha.2020.01.016
Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132
Kim, H. (2012). A dual-subsystem model of the brain's default network: Self-referential processing, memory retrieval processes, and autobiographical memory retrieval. *NeuroImage*, *61*(4), 966–977. https://doi.org/10.1016/j.neuroimage.2012.03.025
Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology*, *13*(2), e1005350. https://doi.org/10.1371/journal.pcbi.1005350
Konkel, A., & Cohen, N. J. (2009). Relational memory and the hippocampus: Representations and methods. *Frontiers in Neuroscience*, *3*(2), 166–174. https://doi.org/10.3389/neuro.01.023.2009
Kuhl, B. A., & Chun, M. M. (2014). Successful Remembering Elicits Event-Specific Activity Patterns in Lateral Parietal Cortex. *Journal of Neuroscience*, *34*(23), 8051–8060. https://doi.org/10.1523/JNEUROSCI.4328-13.2014
Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*(3), 573–616. https://doi.org/10.1037/a0028681
LaRocque, K. F., Smith, M. E., Carr, V. A., Witthoft, N., Grill-Spector, K., & Wagner, A. D. (2013). Global Similarity and Pattern Separation in the Human Medial Temporal Lobe Predict Subsequent Memory. *Journal of Neuroscience*, *33*(13), 5466–5474. https://doi.org/10.1523/JNEUROSCI.4293-12.2013
Lee, H., & Kuhl, B. A. (2016). Reconstructing Perceived and Retrieved Faces from Activity Patterns in Lateral Parietal Cortex. *The Journal of Neuroscience*, *36*(22), 6069–6082. https://doi.org/10.1523/JNEUROSCI.4286-15.2016
Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. *Journal of Neuroscience*, *31*(8), 2906–2915. https://doi.org/10.1523/JNEUROSCI.3684-10.2011
Leutgeb, J. K., Leutgeb, S., Moser, M.-B., & Moser, E. I. (2007). Pattern Separation in the Dentate Gyrus and CA3 of the Hippocampus. *Science*, *315*(5814), 961–966. https://doi.org/10.1126/science.1135801
Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: dissociating episodic from semantic retrieval. *Psychology and Aging*, *17*(4), 677–689. http://www.ncbi.nlm.nih.gov/pubmed/12507363
Lu, Q., Hasson, U., & Norman, K. A. (2020). Learning to use episodic memory for event prediction. *BioRxiv*, 2020.12.15.422882. https://doi.org/10.1101/2020.12.15.422882
Marr, D. (1971). Simple memory: a theory for archicortex. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *262*(841), 23–81. https://doi.org/10.1098/rstb.1971.0078
Morton, N. W., Zippi, E. L., Noh, S. M., & Preston, A. R. (2021). Semantic Knowledge of Famous People and Places Is Represented in Hippocampus and Distinct Cortical Networks. *The Journal of Neuroscience*, *41*(12), 2762–2779. https://doi.org/10.1523/JNEUROSCI.2034-19.2021
Neunuebel, J. P., & Knierim, J. J. (2014). CA3 Retrieves Coherent Representations from Degraded Input: Direct Evidence for CA3 Pattern Completion and Dentate Gyrus Pattern Separation. *Neuron*, *81*(2), 416–427. https://doi.org/10.1016/j.neuron.2013.11.017

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, *4*(6), 661–682. https://doi.org/10.1002/hipo.450040605

Oedekoven, C. S. H., Keidel, J. L., Berens, S. C., & Bird, C. M. (2017). Reinstatement of memory representations for lifelike events over the course of a week. *Scientific Reports*, *7*(1), 14305. https://doi.org/10.1038/s41598-017-13938-4

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 320–341. https://doi.org/10.1016/j.neuroimage.2013.08.048

Preston, A. R., & Eichenbaum, H. (2013). Interplay of Hippocampus and Prefrontal Cortex in Memory. *Current Biology*, *23*(17), R764–R773. https://doi.org/10.1016/j.cub.2013.05.041

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, *13*(10), 713–726. https://doi.org/10.1038/nrn3338

Raykov, P. P., Keidel, J. L., Oakhill, J., & Bird, C. M. (2018). Shared contextual knowledge strengthens inter-subject synchrony and pattern similarity in the semantic network. *BioRxiv*, 276683. https://doi.org/10.1101/276683

Raykov, P. P., Keidel, J. L., Oakhill, J., & Bird, C. M. (2020). The brain regions supporting schema-related processing of people's identities. *Cognitive Neuropsychology*, *37*(1–2), 8–24. https://doi.org/10.1080/02643294.2019.1685958

Reagh, Z. M., Delarazan, A. I., Garber, A., & Ranganath, C. (2020). Aging alters neural activity at event boundaries in the hippocampus and Posterior Medial network. *Nature Communications*, *11*(1), 3980. https://doi.org/10.1038/s41467-020-17713-4

Reagh, Z. M., & Ranganath, C. (2018). What does the functional organization of cortico-hippocampal networks tell us about the functional organization of memory? *Neuroscience Letters*, *680*, 69–76. https://doi.org/10.1016/j.neulet.2018.04.050

Richmond, L. L., & Zacks, J. M. (2017). Constructing Experience: Event Models from Perception to Action. *Trends in Cognitive Sciences*, *21*(12), 962–980. https://doi.org/10.1016/j.tics.2017.08.005

Richter, F. R., Bays, P. M., Jeyarathnarajah, P., & Simons, J. S. (2019). Flexible updating of dynamic knowledge structures. *Scientific Reports*, *9*(1), 2272. https://doi.org/10.1038/s41598-019-39468-9

Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *ELife*, *5*. https://doi.org/10.7554/eLife.18260

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*, *23*(2), 752–763. https://doi.org/10.1016/j.neuroimage.2004.06.035

Ritchey, M., & Cooper, R. A. (2020). Deconstructing the Posterior Medial Episodic Network. *Trends in Cognitive Sciences*, *24*(6), 451–465. https://doi.org/10.1016/j.tics.2020.03.006

Ritchey, M., Libby, L. A., & Ranganath, C. (2015). Cortico-hippocampal systems involved in memory and cognition. *Progress in Brain Research*, *219*, 45–64. https://doi.org/10.1016/bs.pbr.2015.04.001

Ritchey, M., Montchal, M. E., Yonelinas, A. P., & Ranganath, C. (2015). Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. *ELife*, *4*. https://doi.org/10.7554/eLife.05025

Ritchey, M., Wing, E. A., LaBar, K. S., & Cabeza, R. (2013). Neural Similarity Between Encoding and Retrieval is Related to Memory Via Hippocampal Interactions. *Cerebral Cortex*, *23*(12), 2818–2828. https://doi.org/10.1093/cercor/bhs258

Rolls, E. T. (2013). The mechanisms for pattern completion and pattern separation in the hippocampus. *Frontiers in Systems Neuroscience*, *7*. https://doi.org/10.3389/fnsys.2013.00074

Rugg, M. D., & King, D. R. (2018). Ventral lateral parietal cortex and episodic memory retrieval. *Cortex*, *107*, 238–250. https://doi.org/10.1016/j.cortex.2017.07.012

Rugg, M. D., Vilberg, K. L., Mattson, J. T., Yu, S. S., Johnson, J. D., & Suzuki, M. (2012). Item memory, context memory and the hippocampus: fMRI evidence. *Neuropsychologia*, *50*(13), 3070–3079. https://doi.org/10.1016/j.neuropsychologia.2012.06.004

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*(1), 8151. https://doi.org/10.1038/ncomms9151

Sestieri, C., Corbetta, M., Romani, G. L., & Shulman, G. L. (2011). Episodic Memory Retrieval, Parietal Cortex, and the Default Mode Network: Functional and Topographic Analyses. *Journal of Neuroscience*, *31*(12), 4407–4420. https://doi.org/10.1523/JNEUROSCI.3335-10.2011

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., & Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature Communications*, *7*(1), 12141. https://doi.org/10.1038/ncomms12141

Spalding, K. N., Jones, S. H., Duff, M. C., Tranel, D., & Warren, D. E. (2015). Investigating the Neural Correlates of Schemas: Ventromedial Prefrontal Cortex Is Necessary for Normal Schematic Influence on Memory. *The Journal of Neuroscience*, *35*(47), 15746–15751. https://doi.org/10.1523/JNEUROSCI.2767-15.2015

Spreng, R. N., & Grady, C. L. (2010). Patterns of Brain Activity Supporting Autobiographical Memory, Prospection, and Theory of Mind, and Their Relationship to the Default Mode Network. *Journal of Cognitive Neuroscience*, *22*(6), 1112–1123. https://doi.org/10.1162/jocn.2009.21282

Thakral, P. P., Madore, K. P., & Schacter, D. L. (2017). A Role for the Left Angular Gyrus in Episodic Simulation and Memory. *The Journal of Neuroscience*, *37*(34), 8142–8149. https://doi.org/10.1523/JNEUROSCI.1319-17.2017

Tompary, A., Duncan, K., & Davachi, L. (2016). High-resolution investigation of memory-specific reinstatement in the hippocampus and perirhinal cortex. *Hippocampus*, *26*(8), 995–1007. https://doi.org/10.1002/hipo.22582

Tustison, N. J., Avants, B. B., Cook, P. A., Yuanjie Zheng, Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

van Kesteren, M. T. R., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, *51*(12), 2352–2359. https://doi.org/10.1016/j.neuropsychologia.2013.05.027

van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219. https://doi.org/10.1016/j.tins.2012.02.001

Wing, E. A., Ritchey, M., & Cabeza, R. (2015). Reinstatement of Individual Past Events Revealed by the Similarity of Distributed Activation Patterns during Encoding and Retrieval. *Journal of Cognitive Neuroscience*, *27*(4), 679–691. https://doi.org/10.1162/jocn_a_00740

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Zacks, J. M. (2010). The brain's cutting-room floor: segmentation of narrative cinema. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/fnhum.2010.00168

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition*, *112*(2), 201–216. https://doi.org/10.1016/j.cognition.2009.03.007

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*(2), 307–327. https://doi.org/10.1037/a0015305

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57. https://doi.org/10.1109/42.906424

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The Construction of Situation Models in

Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, *6*(5), 292–297. http://www.jstor.org/stable/40063035