

Current Biology

Insight Reconfigures Hippocampal-Prefrontal Memories

Highlights

- Insight triggers reorganization of unrelated events into narratives
- Increased neural similarity in mPFC and hippocampus accompanies narrative integration
- Neural dissimilarity and hippocampal mismatch response signal event segregation
- De novo integrated representations emerge as a consequence of narrative insight

Authors

Branka Milivojevic, Alejandro Vicente-Grabovetsky, Christian F. Doeller

Correspondence

b.milivojevic@donders.ru.nl (B.M.), christian.doeller@donders.ru.nl (C.F.D.)

In Brief

Using fMRI and across-voxel correlations as a proxy of neural similarity, Milivojevic et al. demonstrate that gaining insight into how seemingly unrelated events fit together triggers a systematic reorganization of mnemonic representations of those events and leads to the formation of novel memories in the hippocampus and medial prefrontal cortex.



Insight Reconfigures Hippocampal-Prefrontal Memories

Branka Milivojevic,^{1,*} Alejandro Vicente-Grabovetsky,¹ and Christian F. Doeller^{1,*}

¹Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, 6525 EN Nijmegen, the Netherlands

*Correspondence: b.milivojevic@donders.ru.nl (B.M.), christian.doeller@donders.ru.nl (C.F.D.)

<http://dx.doi.org/10.1016/j.cub.2015.01.033>

SUMMARY

Our memories are remarkably dynamic and allow us to reinterpret the past once new information comes to light. Gaining novel insights can lead to mental reorganization of previously unrelated events, thus linking them into narratives. The hippocampus and medial prefrontal cortex (mPFC) support integration of partially overlapping events, but the neural mechanisms underlying the reorganization of memories for the formation of coherent narratives remain elusive. Here, we combine fMRI with *The Sims 3* videos of life-like animated events, which could either be integrated into narratives or not. We show that insight triggers the emergence of de novo mnemonic representations of the narratives and is associated with increased neural similarity between linked event representations in the posterior hippocampus, mPFC, and autobiographical-memory network. Simultaneously, events irrelevant to the newly established memory of the narrative were pruned out. This process was accompanied by increased neural dissimilarity between non-linked event representations in the posterior hippocampus and mPFC and was additionally signaled by a mismatch response in the anterior hippocampus. Our results demonstrate that insight leads to neural reconfiguration of representational networks within a memory space and have implications for knowledge acquisition in educational settings.

INTRODUCTION

Our rich autobiographical memories are woven into personal narratives that consist of multiple individual events. These narratives can be clear, as the case may be when we recall steps taken to plan a birthday party. At other times, however, a narrative may remain unclear until we gain novel insights that enable us to piece together a number of seemingly unrelated events. For instance, when someone throws us a surprise birthday party, we may only become aware of the significance of certain events once we realize that they were related to keeping the party a secret. In this case, gaining insight into the existence of the surprise party would lead to a reorganization of mental representations [1] of previously unrelated events in a memory space [2] and

to the construction of a continuous and linear narrative that contains only relevant events, while irrelevant events are pruned out.

But how does the brain accomplish this task? To reconcile the dual need of individuating and flexible recombining of memories into narratives, a neural system capable of both types of computations is required. Clear candidates for this are the hippocampus [1, 3], a region supporting both pattern separation and pattern completion [4–6], and the medial prefrontal cortex (mPFC), a region crucially involved in establishing unified representations of multiple events referred to as schemata [7, 8]. Both of these structures are critical components of the autobiographical-memory network (AMN) [9, 10], with the hippocampus playing a key role in individuating memories by establishing and maintaining separate event representations [2, 11–13], while also contributing to the formation of combined memory representations [14, 15] via its cross-talk with the mPFC [7, 16–18].

In addition to the polar question of *whether* the hippocampus [1, 3] and mPFC are involved in insight-triggered reconfiguration of mental event representations, an important open question is *how* these two regions might fulfil this role. Does the hippocampus create a novel, integrated memory of the newly constructed narrative and encode novel between-event associations by employing nodal [2] or conjunctive [6] representations, serving as a convergence zone [13]? Or are conjunctive representations, or schemata, established in the mPFC, with the hippocampus merely mediating the reactivation of associated event representations without creating a de novo conjunctive representation [16]? In addition to the mechanisms of insight-triggered event integration, an equally pressing issue includes the question of how the brain processes unrelated events. Does the hippocampus simply ignore these events or are they actively dissolved from the memory representation, akin to situations in which new information violates our predictions [7, 19, 20]?

Here we test the hypothesis that insights into narratives trigger reconfiguration of neural networks of interconnected event representations [2, 12, 13, 21], enabling a flexible recombination of information [9]. To determine whether this is indeed the case, we examined the emergence of novel, multi-event representations by combining fMRI with a novel behavioral “narrative-insight” task. We used the life-simulation game *The Sims 3* to create short videos depicting realistic, life-like events belonging to multiple different story lines, which could either be integrated (events A and B) into narratives or not (event X; see Figure 1 and Movie S1 for details of the experimental design). This enabled us to simulate some of the processes required for integration of events into narratives while retaining a great degree of experimental control. Multi-voxel pattern and adaptation analyses were used to test the overarching hypothesis that representations of

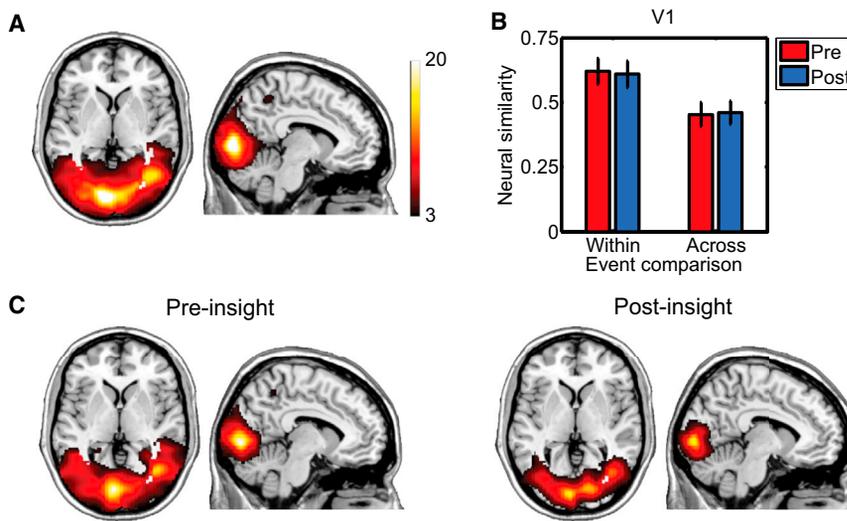


Figure 2. Stable Representation of Event-Specific Information across the Pre- and Post-insight Phases

(A) Brain regions showing higher within-event compared to across-event neural similarity. Whole-brain group effects, significant at $p < 0.05$, cluster corrected, are overlaid on a canonical structural template ($z = 2$, $x = -8$), thresholded at $p < 0.005$, uncorrected, for display purposes (only clusters with more than 50 contiguous voxels are displayed). The color bar indicates t statistic.

(B) Bars show neural similarity for the peak voxel in visual cortex (Montreal Neurological Institute [MNI] coordinate for the peak voxel: $x = -10$, $y = -90$, $z = 0$) averaged across participants \pm SEM, separately for within-event and across-event comparisons and for pre-insight (red) and post-insight (blue) phases, respectively.

(C) Brain regions showing higher within-event compared to across-event neural similarity, separately for pre-insight and post-insight phases. Event-specific effects were stable over experimental phases (no significant phase-by-event interaction, $F(1,18) = 4.18$, $p = 0.056$).

See also [Figure S2](#) and [Table S1](#).

interaction between these two factors (link and phase) by generating a contrast image comparing AB correlations and AX correlations in pre-insight phases to the correlations between the same event pairs in post-insight phases. This contrast is henceforth denoted as the “insight analysis.” Here, we observed that neural similarity selectively increased for AB events relative to AX events in both the posterior hippocampus (pHPC) and the mPFC after narrative insight was gained ([Figure 3](#)).

Pairwise comparisons in posterior hippocampus and mPFC indicated that the interaction can be attributed to a difference in similarity between linked and non-linked events in the post-insight phase, whereas no such difference was apparent in the pre-insight phase. This divergence in similarity between linked and non-linked events was due to an increase in similarity from the pre-insight phase to the post-insight phase for linked AB events in the mPFC and a similar trend in the hippocampus, and a simultaneous decrease in similarity for AX events, which was apparent in both regions ([Figure 3](#)). Crucially, these effects were not due to visual-feature similarity ([Figure S2](#), [Table S2](#), and the [Supplemental Experimental Procedures](#)) or to univariate signal differences ([Figure S3](#)). This suggests that insight into narratives is accompanied by an increase in similarity between linked representations and a decrease in similarity between non-linked representations. Note that we controlled for spurious effects of visual similarity between linked and non-linked events by between-subject counterbalancing of storyline versions and by using the 2×2 factorial design with the “insight effect” expressed in the interaction term.

Event Segregation Is Signaled by Narrative Mismatch in Anterior Hippocampus

Changes of the event representations ensued after participants gained insight into the narratives, but how are these changes induced? What are the underlying neural computations that give rise to this plasticity? One likely candidate mechanism is novelty signaling [[19](#), [22](#)], which is associated with increased hippocampal activity [[19](#), [23](#)]. We expected the novelty responses

[[19](#)] to signal a mismatch, or incongruity [[7](#)], between the inputs for the unrelated events and the mnemonic representation of the narrative, reflecting a violation of a narrative prediction rather than novelty per se [[20](#)]. Thus, we hypothesized that event segregation should elicit lag-dependent fMRI novelty responses for non-linked events (AX), compared to linked events (AB), selectively for the post-insight versus pre-insight phase (see [Figure 4](#) for details of analysis logic). This interaction was exactly the pattern of results we observed in a cluster in the anterior medial temporal lobe extending between amygdala and the anterior hippocampus (aHPC; see [Figure 4](#) and [Table S4](#) for details). Post hoc pairwise comparisons confirmed this observation: the mismatch response was present for non-linked events in the post-insight phase only, reflecting an increase in mismatch from pre- to post-insight phases. Interestingly, we observed a greater mismatch response for *linked* than non-linked events in the pre-insight phase that decreased, albeit non-significantly, in the post-insight phase. No narrative mismatch was observed in regions where we observed representational change in the insight analysis, and vice versa: the insight-reconfiguration effect was absent in regions showing narrative mismatch, suggesting that the mismatch response is unrelated to representational changes (see [Figure S4](#)). In sum, these data indicate that the anterior hippocampus treats non-linked events as more incongruent than linked events during the post-insight phase.

Interplay between Insight Reconfiguration and Narrative Mismatch

Are mechanisms of insight-triggered narrative formation and mismatch-signaled event segregation independent of each other or do they interact? To answer this question, we performed across-subject correlation analyses between the effects observed in the narrative-mismatch and the insight analyses. We found a significant correlation between the mismatch effect in aHPC and the insight effect in mPFC, but not in the pHPC ([Figure 5](#)). In sum, participants with the strongest insight-induced

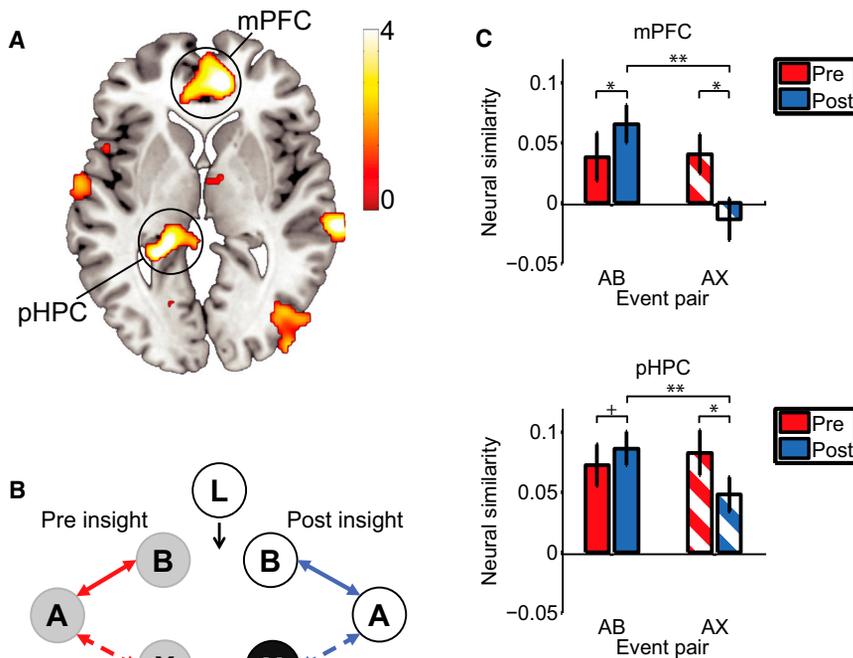


Figure 3. Dynamic Narrative Integration and Event Segregation in the Posterior Hippocampus and mPFC in the Post-insight Phase

(A) Brain regions that show increases in event similarity for AB versus AX event pairs after compared to before the presentation of the link event L (phase-by-link interaction). mPFC, medial prefrontal cortex; pHPIC, posterior hippocampus. Effects are significant at $p < 0.05$, cluster corrected; for display purposes, the group-level t statistics image is thresholded at $p < 0.005$, uncorrected (cluster extent: 50 contiguous voxels).

(B) Schematic illustration of similarity measures of interest for AB pairs (full lines) and AX pairs (dashed lines) in pre-insight (left) and post-insight (right) phases.

(C) Phase-by-link interaction: bars show neural similarity averaged across participants \pm SEM for AB pairs (filled bars) and AX pairs (striped bars) in the pre-insight and post-insight phases at peak voxels in pHPIC ($-24, -42, 2$; peak z value = 3.64; $p < 0.05$, cluster corrected) and mPFC ($10, 46, 6$; $z = 3.88$; $p < 0.05$, cluster corrected; see the Supplemental Experimental Procedures for details and Table S3 for full list of regions). This interaction can be attributed to a difference in similarity between linked and non-linked events in the post-insight phase (pHPIC: mean difference [Md] = 0.038, SEM difference [SEMd] = 0.008, $t(18) = 4.76$,

$p < 0.001$; mPFC: Md = 0.079, SEMd = 0.012, $t(18) = 6.87$, $p < 0.001$), whereas no such difference was apparent in the pre-insight phase (pHPIC: Md = -0.011 , SEMd = 0.012, $t(18) = -0.89$, $p = 0.384$; mPFC: Md = -0.002 , SEMd = 0.014, $t(18) = -0.16$, $p = 0.874$). This post-insight divergence in similarity was due to an increase in similarity from the pre-insight phase to the post-insight phase for linked AB events in the mPFC (Md = 0.027, SEMd = 0.011, $t(18) = 2.40$, $p = 0.028$) and the hippocampus (Md = 0.014, SEMd = 0.008, approaching significance at $t(18) = 1.819$, $p = 0.086$) and a simultaneous decrease in similarity for AX events, which was apparent in both regions (pHPIC: Md = -0.035 , SEMd = 0.012, $t(18) = -3.00$, $p = 0.008$; mPFC: Md = -0.054 , SEMd = 0.015, $t(18) = -3.53$, $p = 0.002$). ** $p < 0.001$, * $p < 0.05$, * $p = 0.086$.

See also Figure S3 and Table S3.

representational change in the mPFC show the strongest narrative-mismatch response in aHPC, suggesting an interplay between insight-induced reorganization of event representations and strength of mismatch signal to newly segregated events.

Insight Reconfiguration and Narrative Mismatch within the Autobiographical-Memory Network

Finally, we were interested whether other regions comprising the AMN [9, 10] were involved in representing de novo narrative representations and whether they were sensitive to the newly segregated events. In addition to hippocampus and mPFC reported above, we observed a striking overlap between the AMN [9, 10] and regions showing a significant effect in the insight analysis, including the anterior and posterior midline regions (mPFC, middle cingulate, and precuneus), thalamus, ventrolateral and dorsolateral prefrontal cortices, and left temporal pole, as well as areas exhibiting the narrative-mismatch effect, including the precuneus, left temporo-parietal junction, and ventrolateral prefrontal cortex (Figure 6).

DISCUSSION

In this fMRI study, we used a unique combination of a novel realistic life-like episodic-memory task, functional neuroimaging, and representation-based analysis methods to map how the “neural neighborhood” structure of memory representations changes as a consequence of insight. We show that gaining

new insights about how two seemingly unrelated events relate to each other triggers systematic reconfiguration of memory networks in the hippocampus and mPFC. This reconfiguration leads to formation of a novel unified narrative-level representation, expressed by the increased neural similarity of linked event representation. At the same time, the representations of non-linked events also undergo reconfiguration with respect to the novel narrative network, as indexed by an increase in neural *dissimilarity* for those events. Non-linked events additionally elicit a neural mismatch response signaling incongruence with the newly formed narrative context. In sum, we show for the first time that we can establish a novel mnemonic representation in the hippocampus and mPFC by providing new information about how two seemingly unrelated events relate to each other (see Figure 7 for a schematic illustration of insight-triggered representational reconfiguration).

Insight-Triggered Event Reconfiguration in the Brain

Insight can lead to flexible recombination of mental representations [1], but it remains unclear what neural mechanisms underlie such changes. Here, after gaining insight into the linking event, participants were able to determine confidently and accurately which events belonged together. Our results are consistent with the idea that event representations are organized into dynamic networks of related mnemonic representations within a memory space [2], which can be reconfigured through insight

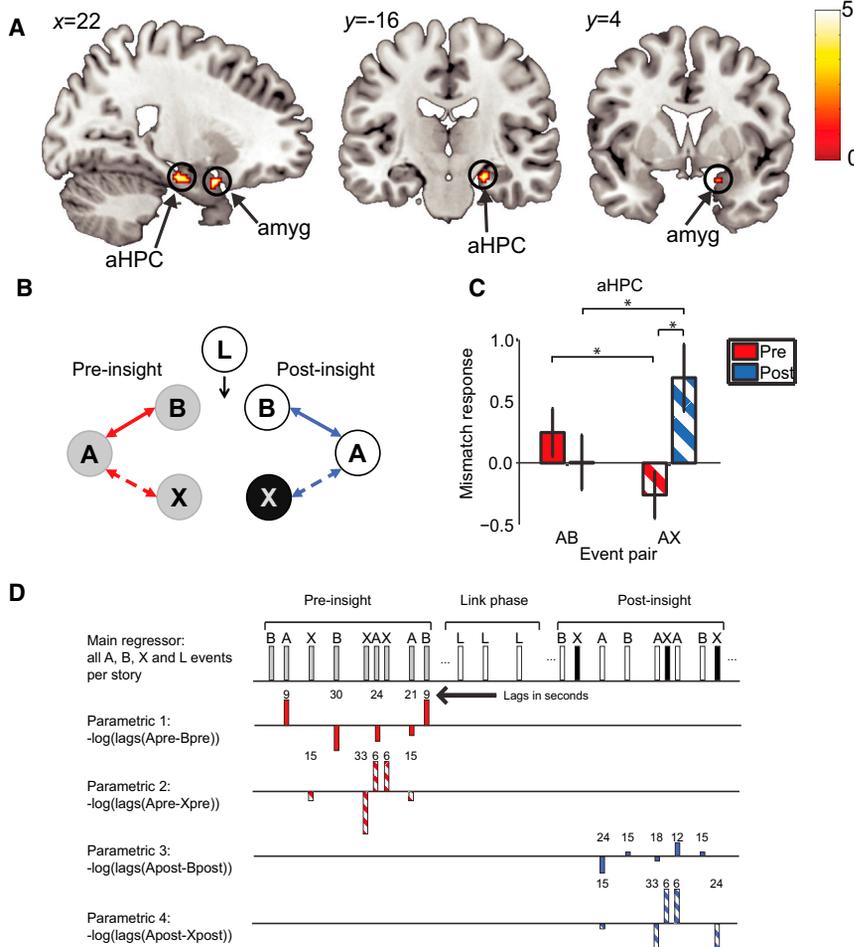


Figure 4. Narrative-Mismatch Signaling in the Anterior Medial Temporal Lobe

(A) Brain regions that show the narrative-mismatch response selectively for AX pairs in the post-insight phase (phase-by-link interaction), extending between amygdala (amyg, cluster peak: 18, -6, -20; $z = 3.84$; voxel-level $p < 0.001$; cluster significant at $p < 0.05$, cluster corrected) and anterior hippocampus (aHPC: 22, -16, -18; $z = 3.42$; voxel-level $p < 0.001$). Effects were significant at $p < 0.05$, cluster corrected; for display purposes, the group-level t statistics image is thresholded at $p < 0.005$, uncorrected (cluster extent: 50 contiguous voxels). (B) Schematic illustration of lag-dependent fMRI analysis for linked (AB; full lines) and non-linked (AX; dashed lines) events, separately for the pre-insight (left) and post-insight (right) phase. (C) Bars show strength of mismatch response (parameter estimates, averaged across participants, \pm SEM) for AB pairs (filled bars) and AX pairs (striped bars) in the pre-insight and post-insight phase at peak voxel in the anterior hippocampus ($x = 22$, $y = -16$, $z = -18$). Post hoc pairwise comparisons were performed: mismatch response was greater for non-linked than linked events in the post-insight phase ($AX_{\text{post}} > AB_{\text{post}}$; $Md = 0.686$, $SEMd = 0.209$, $t(18) = 3.30$, $p = 0.004$) and reflected an increase in mismatch response from pre- to post-insight phases ($AX_{\text{post}} > AX_{\text{pre}}$; $Md = 0.950$, $SEMd = 0.353$, $t(18) = 2.69$, $p = 0.015$). A significantly greater mismatch effect was observed for linked than non-linked events in the pre-insight phase ($AB_{\text{pre}} > AX_{\text{pre}}$; $Md = 0.507$, $SEMd = 0.209$, $t(18) = 2.42$, $p = 0.026$) and decreased, albeit non-significantly, in the post-insight phase ($AB_{\text{post}} > AB_{\text{pre}}$; $Md = -0.242$, $SEMd = 0.313$, $t(18) = -0.773$, $p = 0.449$). No narrative mismatch was observed in regions where we observed representational change in the insight analysis (pHPC: $F(1,18) = 2.27$, $p = 0.149$; mPFC: $F(1,18) = 1.23$, $p = 0.282$;

$F(1,18) = 0.22$, $p = 0.643$), and vice versa: the insight-reconfiguration effect was absent in regions showing narrative mismatch (aHPC: $F(1,18) = 1.23$, $p = 0.282$; amygdala: $F(1,18) = 0.698$, $p = 0.414$); see Figure S4 and Table S4.

(D) Logic for mismatch analysis. The generalized linear model consisted of a single main regressor per story that modeled all four event types (A, B, X, and L) across the three phases (pre-insight, link, and post-insight). For modeling of the amplitude change associated with mismatch responses, four parametric regressors were added. These four parametric regressors corresponded to two event-pair categories (AB or AX) in pre- and post-insight phases. The parametric regressors were modulated by the lags between two stimuli from the same event-pair category (either AB or AX pairs) using the equation $-\log(\text{lags}(\text{Event 1} - \text{Event 2}))$ and were demeaned.

into their relationship. Modification of a memory representation through the link event L also dovetails with previous reports, which show that a reminder cue can change subsequent memories [24]. Here, we were able to visualize this reorganization of event representations leading to the formation of novel integrated narrative representations. Our data further suggest that those novel narrative representations are constructed from multiple events akin to how novel event representations are constructed from multiple interrelated elements [25], indicating that mnemonic representations may have hierarchical organization [21, 26]. The dynamic nature of event representations shown here also accords with proposals that flexible recombination of stored representations is pertinent for shaping and reshaping of map-like representations of past events [13] embedded within their spatial context in the hippocampal formation [13, 27]. Importantly, our findings have implications for educational settings [28] as they suggest that teachers can aid learning not

only by explicitly providing the context for new material [8], but also by pointing out relationships between previously learned items, thus eliciting reorganization of neural representations and the accompanying formation of an integrated “big picture.”

Nature of Integrated Representations

Here we report that reconfiguration of mental representations following insight leads to formation of integrated representations evident in the post-insight phase. But what is the nature of those integrated representations? One possibility, depicted in Figure 7, is that the insight established a strong association between all of the linked events (A, B, and L) such that the increase in similarity between A and B in the post-insight phase reflects reactivation of an associated event, as reported by Schapiro and colleagues [29]. This “direct-associative hypothesis” would predict that the representation of A would become more similar to how B was encoded in the pre-insight phase, and vice versa. The

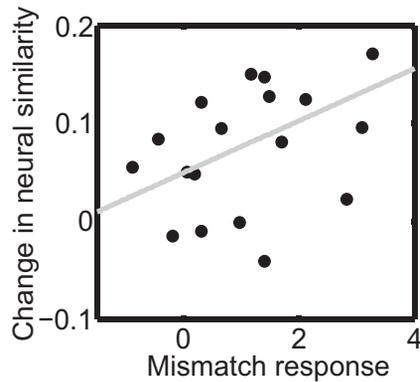


Figure 5. The Strength of Insight-Triggered Representational Change Relates to the Strength of Narrative Mismatch Response

Scatter plot showing significant relationship between change in neural similarity in medial prefrontal cortex (mPFC; interaction effect for peak voxel of phase-by-link interaction effect from Figure 3) and mismatch response in anterior hippocampus (aHPC; parameter estimate for peak voxel of phase-by-link interaction effect from Figure 4; $r(19) = 0.466$, $p = 0.044$). No such relationship was evident between change in neural similarity in the posterior hippocampus and the mismatch response in the anterior hippocampus ($r(19) = 0.034$, $p = 0.891$). Each dot represents one participant.

increase in similarity may arise as a consequence of retrieval of both events A and B during the encoding of the L event, resulting in AB associations through this shared temporal context [30]. Another possibility, which we term the “indirect-associative hypothesis,” is that the linked events are sequentially linked through the narrative in such a way that A and B are not directly associated with each other, but both are associated with the linking event L. According to this hypothesis, the increase in similarity between A and B in the post-insight phase reflects reactivation of the associated event L. The third possibility is that the representational change of the AB pairs could be attributed to an emergence of a novel mnemonic representation of the narrative, which would be unlike the initial representations of individual events A and B (or, indeed, L), termed the “de novo representation hypothesis.”

To compare these three possibilities, we performed two separate analyses, presented in Figure S5. The results do not support either the direct- or the indirect-associative hypotheses, but rather suggest that the insights into narrative meanings triggered the emergence of de novo mnemonic representations of the narratives during the post-insight phase that were unlike the initial representations of individual A and B events. Note, however, that the increased temporal distance for across-phase, relative to within-phase, comparisons might have reduced the sensitivity of these analyses.

Nevertheless, emergence of de novo representations would be consistent with the absence of event-specific representations in the hippocampus and the mPFC (see Figure 2). Namely, posterior brain regions, including the ventral and dorsal visual streams, seem to represent event-specific information that is stable across time. Concurrent representations of individual events and integrated narratives in distinct brain regions are suggestive of hierarchical organization of mnemonic representations [21, 26], in which the roles of the hippocampus and mPFC would be to combine event-specific representations

into stable narrative-specific representations. The emergence of integrated representations in this context is consistent with this view and suggests that the de novo representations contain information pertinent to conjunctive [6] or nodal [2] coding of events in the hippocampus and wider event context [18] or narrative schemas [7, 8] in the mPFC.

The newly emerging integrated representations in hippocampus and mPFC most likely contain an amalgamation of previously unrelated events, rather than simply reflecting associations between the events. This interpretation stands in contrast to several previous reports suggesting that, during memory integration, hippocampal involvement is short lived or limited to reactivation of associated event representations [16, 26], whereas the mPFC represents novel multi-event networks [7, 16]. One possibility is that our design may have brought about the formation of a new network, whereas in other paradigms, such as prior knowledge [7], paired-associate inference task [16], or transfer learning [26], the new representations were sufficiently similar to already existing representations and that they facilitated incorporation of new information into pre-existing networks. Although the current paradigm is similar to the transitive inference paradigms reported in the literature [16, 31–33], there are several crucial differences. First, the stimuli used in the current paradigm are more complex and thus allow for multiple associations to be created within and across events which is likely to strengthen the integrated memory representation. Second, although the A, B, and L events are similar to “AB,” “CD,” and “BC” events in transitive inference studies [31, 33], the simplified paired-associate inference task [16, 32] only uses partially overlapping pairs (AB and BC), but not the crucial non-overlapping pairs (AB and CD). Since the simplified version of the paradigm has been used to explore the nature of the integrated representations, we feel that this crucial difference between the current paradigm and previously published findings may account for the differences reported here. Since we imaged the initial formation of a novel representational network, it is possible that persisting integrated representations in the hippocampus reflect an early stage in the life cycle of a mnemonic network. With time, the integrated representation in the hippocampus may gradually change, whereas the mPFC representation may remain [34].

Event Integration and Segregation in the Hippocampus

Previous reports have demonstrated that the hippocampus is involved in insight [1, 3]. Here, we provide a mechanistic explanation for this involvement by demonstrating for the first time that hippocampal representations are reconfigured as a consequence of gaining insight, consistent with its involvement in memory integration [31–33]. Our results dovetail with previous reports on the integrative function of the hippocampus: mediating inferential reasoning about relations which were not directly encoded [2, 5, 16, 17, 26], encoding associations between pre-existing and novel information [14–16], learning regularities across episodes [35], and bridging spatiotemporal gaps between events in both humans [15, 36] and rats [37].

In addition, similarly to [26], we show that the hippocampus plays a dual role in event reconfiguration. We demonstrate that, during the recombination process, events are also dissolved from mnemonic networks when they are not relevant to the narrative, similarly to how memories of opposing significance

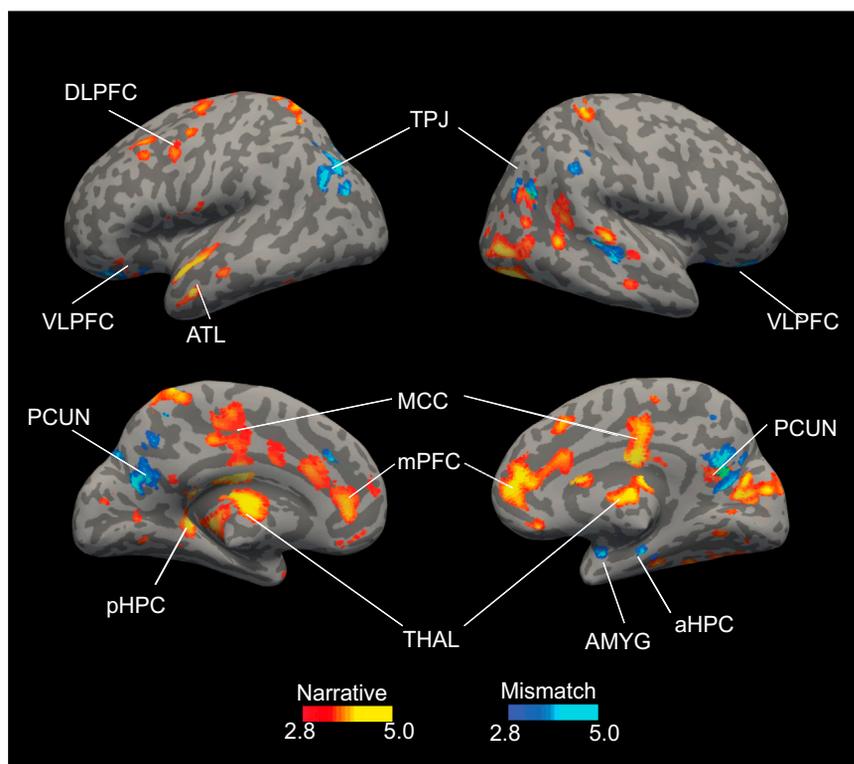


Figure 6. The Autobiographical-Memory Network Is Involved in Insight Reconfiguration and Narrative Mismatch

Significant t maps for the insight effect (hot colors) and narrative-mismatch effect (cool colors), with the overlap between the two t-maps maps in green, overlaid on a rendered inflated cortical surface. Effects were significant at $p < 0.05$, cluster corrected; for display purposes, the group-level t statistics images are thresholded at $p < 0.005$, uncorrected (cluster extent: 50 contiguous voxels). Critical regions associated with autobiographical memory [10] are highlighted and labeled. mPFC, medial prefrontal cortex; MCC, middle cingulate; PCUN, precuneus; TPJ, temporo-parietal junction; ATL, anterior temporal lobe; DLPFC, dorsolateral prefrontal cortex; VLPFC, ventrolateral prefrontal cortex; THAL, thalamus; AMYG, amygdala; pHPC, posterior hippocampus; aHPC, anterior hippocampus.

result in separation of neural representations [26]. This dual function of integration and segregation of events parallels the well-established role of the hippocampus in both pattern completion and pattern separation [4]. We also show that anatomically distinct regions along the hippocampal long axis (cf. [38, 39, 40]) contribute differently to event integration and segregation. More specifically, we show that the posterior hippocampus plays a role in both integration and segregation of events, similarly to [26], while the anterior hippocampus seems to be involved in event segregation, although based on the data it is unclear whether one process precedes or follows the other or whether they occur simultaneously. Unlike some of the other associative memory tasks, our paradigm requires both functions, and this dissociation between posterior and anterior hippocampus might provide a clue as to how the hippocampus can perform such complementary operations within a similar time frame.

Narrative Formation Depends on the Interaction of mPFC and Hippocampus and a Wider Autobiographical Memory Network

We propose that the involvement of the mPFC in insight-triggered reconfiguration of mental representations reflects the need for the formation of a wider event context or narrative [18]. After insight into the linking event, a novel narrative context is established in the mPFC by flexibly recombining stored representations of individual events. At the same time, the stored representations of the non-linked events are dissolved from the novel contextual representations. As mentioned above, this reconfiguration of mnemonic representations probably reflects complex dynamics of multiple processes that may be active dur-

ing all, or only some, of the multiple repetitions of each trial type required to improve the power of the experiment. Nevertheless, the reconfiguration is most likely to occur during the post-insight phase, rather than during the insight phase (see Figure S5 and the Supplemental Experimental Procedures for details).

Our data accord with reports that the mPFC plays a key role in integrating information over multiple episodes [7, 16–18] and memory-guided behavior and decision making [17, 41]. Our results also parallel previous findings that learning-related adaptation in anterior hippocampus is associated with stronger integrated representations [16].

What might be the mechanisms underlying the across-subject relationship between the mismatch response in the anterior hippocampus and the integrated representation in the mPFC? One possibility is that anterior medial temporal lobe triggers segregation of irrelevant events by signaling the mismatch between the integrated narrative representation and temporally proximal unrelated events [7, 19, 42]. This mismatch signal might support the reconfiguration of event representations in memory by inducing neural plasticity [20, 43] and might relate to findings that the connectivity between anterior hippocampus and the mPFC increases for schema-incongruent representations [7]. Another possibility is that the mismatch signaling may be a response to a newly established narrative context or schema in the mPFC. Namely, before the linking event is introduced, shared temporal context of the pre-insight phase is likely to lead to certain degree of baseline relatedness between the events ([30], cf. [36, 44]). However, once insight is gained through the linking event, a new context emerges for events that are incorporated into a narrative schema, while non-linked events are assigned a distinct context, probably during the post-insight phase, as no increase in similarity or dissimilarity was observed with the link event (Figure S5). Nevertheless, the emergence of this new context may be a consequence of the co-occurrence (i.e., shared temporal context [30]) of individual event features from events A and B during

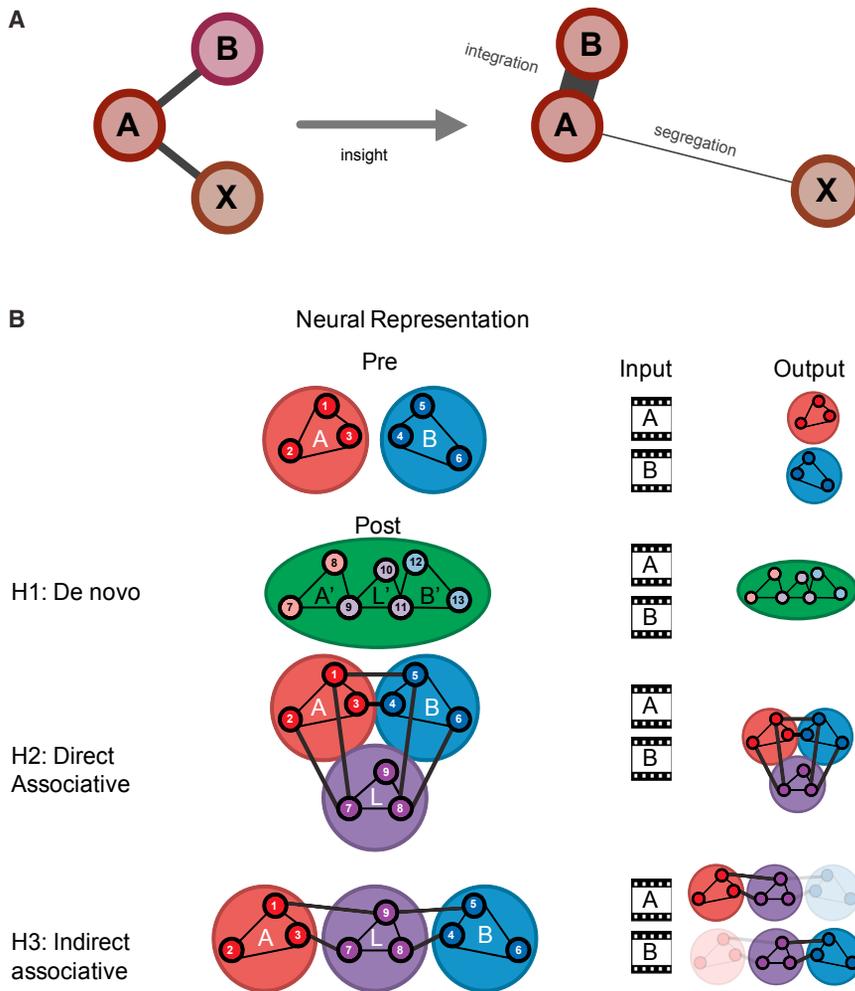


Figure 7. Schematic of Alternative Possibilities of the Structure of Integrated Representations after Insight into the Narrative Structure

(A) Gaining insight into how some events are linked through a narrative triggers neural event reconfiguration. Relative to the pre-insight phase (left), in which events are equidistant in mnemonic space, the topography of this representational space is systematically reconfigured in the post-insight phase (right). Integration of related events is accompanied by increased neural similarity, whereas segregation of an unrelated event is associated with increased neural dissimilarity and signaled by a hippocampal mismatch response.

(B) Pre: The initial encoding of events A and B results in formation of event-specific memories, represented by distinct neural populations (A, neurons 1–3; B, neurons 4–6). These neural populations would be activated in response to encountering event A or B. Post: After encoding of the event L, which links together events A and B, a change in neural representations of events A and B is hypothesized to occur. Three alternative hypotheses can be described: H1, or the de novo hypothesis, suggests that a novel, integrated representation emerges as a consequence of linking together the three events (A, B, and L). This novel representation is unlike the initial representations A and B and is likely to recruit new neuronal populations to support its formation. H2, or the direct associative hypothesis, suggests that a consequence of linking together the three events is the formation of across-event connections, whereas the neural populations underlying individual event representations themselves remain unchanged. These new connections can serve to reactivate related events, and thus encountering event A may reactivate associated events B and L. H3, or the indirect associative hypothesis, is similar to the direct associative hypothesis and

suggests that a consequence of linking together the three events is the formation of across-event connections, whereas the neural populations underlying individual event representations themselves remain unchanged. However, in contrast to the direct associative hypothesis, the indirect associative hypothesis would suggest that only events adjacent within a narrative structure would become strongly associated and that encountering event A may reactivate the adjacent event L, but not the subsequent event B. See also [Figure S5](#).

the event L or due to retrieval of both events A and B during event L. After the insight phase, the non-linked events would become incongruent with this newly established narrative context and thus elicit a mismatch signal ([7, 19, 20, 39], cf. [43]), which would work against similarity that is simply based on temporal proximity of events.

In addition to the hippocampus and mPFC, we observed insight-related representational changes in a wider network of brain regions involved in autobiographical memory, including anterior and posterior midline areas, temporo-parietal junction, lateral and medial temporal cortices, dorsolateral prefrontal cortex, ventrolateral prefrontal cortex, thalamus, and amygdala [9, 10]. Therefore, memory reconfiguration associated with insights into narratives most likely depends on the same neural network that underlies construction of personal narratives within autobiographical memory [9, 10], although participants' knowledge that the stories will be completed later in time (i.e., during the insight phase), coupled with repeated presentation of events,

make the direct comparison to episodic memory more difficult. Consistent with previous reports that different regions within the AMN are involved in different processes [9, 40], we show that the insight effect and narrative-mismatch effect are somewhat differently distributed. The precuneus and temporo-parietal junction, regions typically found to be sensitive to narrative length and coherence [45], are sensitive to mismatch of individual events from newly constructed narratives but play a smaller role in de novo narrative representation. On the other hand, both the medial temporal lobes and anterior regions of the AMN are involved in constructing de novo narrative representations. This difference between anterior and posterior regions may enable the autobiographical memory system to integrate distant events into coherent personal narratives while at the same time smoothing out inconsistencies. Therefore, our results further elucidate the neural mechanisms that enable everyday events to be recombined, both for the formation of our rich personal memories and for the simulation of novel

scenarios based on recombination of prior experiences for future thinking [9].

Conclusions

When new information comes to light, people seldom have difficulties reinterpreting the relationships between previously experienced events to form coherent narratives. Here, we elucidate the neural mechanisms underlying such insight-triggered reorganization of memory representations. Our data provide compelling evidence for the notion that events appear to be represented in dynamic memory networks [2] and that those networks can be systematically reconfigured into de novo representations once new information comes to light [1], by bringing together relevant information and pruning irrelevant connections. The results suggest that neural plasticity triggered through insight leads to new mnemonic representations of broader context of the narratives in the hippocampus, mPFC and a wider autobiographical-memory network. These results answer important questions regarding how the hippocampus and mPFC contribute to insight-related narrative construction and may provide a more general neural mechanism for how insight affects cognition, with possible applications in educational settings [28].

EXPERIMENTAL PROCEDURES

The experimental procedures are summarized briefly in the [Introduction](#) and [Figure 1](#) and are presented in detail in the [Supplemental Experimental Procedures](#), [Figure S1](#), and [Movie S1](#).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, four tables, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2015.01.033>.

AUTHOR CONTRIBUTIONS

C.F.D. and B.M. conceived and designed the experiment. B.M. developed the experimental task. B.M. and A.V.-G. performed the experiment. B.M. analyzed the data; A.V.-G. and C.F.D. gave advice on data analyses. B.M. and C.F.D. wrote the paper.

ACKNOWLEDGMENTS

This work was supported by the European Research Council (ERC-StG 261177) and the Netherlands Organisation for Scientific Research (NWO-Vidi 452-12-009). The authors would like to thank A. Backus, M.A.S. Boksem, S. Collin, L. Deuker, G. Fernandez, R. Kaplan, P. Meyer, and R. Willems for useful discussions and three anonymous reviewers for their helpful comments and suggestions.

Received: September 17, 2014

Revised: December 13, 2014

Accepted: January 14, 2015

Published: February 26, 2015

REFERENCES

- Wagner, U., Gais, S., Haider, H., Verleger, R., and Born, J. (2004). Sleep inspires insight. *Nature* 427, 352–355.
- Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. (1999). The hippocampus, memory, and place cells: is it spatial memory or a memory space? *Neuron* 23, 209–226.
- Luo, J., and Niki, K. (2003). Function of hippocampus in “insight” of problem solving. *Hippocampus* 13, 316–323.
- Bakker, A., Kirwan, C.B., Miller, M., and Stark, C.E.L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science* 319, 1640–1642.
- Kumaran, D., and McClelland, J.L. (2012). Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychol. Rev.* 119, 573–616.
- Norman, K.A., and O’Reilly, R.C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol. Rev.* 110, 611–646.
- van Kesteren, M.T.R., Ruiters, D.J., Fernández, G., and Henson, R.N. (2012). How schema and novelty augment memory formation. *Trends Neurosci.* 35, 211–219.
- Tse, D., Takeuchi, T., Takekuma, M., Kajii, Y., Okuno, H., Tohyama, C., Bito, H., and Morris, R.G.M. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science* 333, 891–895.
- Schacter, D.L., Addis, D.R., and Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* 8, 657–661.
- Spreng, R.N., Mar, R.A., and Kim, A.S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* 21, 489–510.
- Squire, L.R., and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science* 253, 1380–1386.
- Eichenbaum, H.D., and Cohen, N.J. (2001). *From Conditioning to Conscious Recollection: Memory Systems of the Brain*. (Oxford University Press).
- O’Keefe, J., and Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. (Oxford University Press).
- Shohamy, D., and Wagner, A.D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* 60, 378–389.
- Staresina, B.P., and Davachi, L. (2009). Mind the gap: binding experiences across space and time in the human hippocampus. *Neuron* 63, 267–276.
- Zeithamova, D., Dominick, A.L., and Preston, A.R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron* 75, 168–179.
- Kumaran, D., Summerfield, J.J., Hassabis, D., and Maguire, E.A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron* 63, 889–901.
- Preston, A.R., and Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* 23, R764–R773.
- Kumaran, D., and Maguire, E.A. (2006). An unexpected sequence of events: mismatch detection in the human hippocampus. *PLoS Biol.* 4, e424.
- Lisman, J.E., and Grace, A.A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46, 703–713.
- Milivojevic, B., and Doeller, C.F. (2013). Mnemonic networks in the hippocampal formation: from spatial maps to temporal and conceptual codes. *J. Exp. Psychol. Gen.* 142, 1231–1241.
- Ranganath, C., and Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nat. Rev. Neurosci.* 4, 193–202.
- Bunzeck, N., and Düzal, E. (2006). Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron* 51, 369–379.
- Hupbach, A., Gomez, R., Hardt, O., and Nadel, L. (2007). Reconsolidation of episodic memories: a subtle reminder triggers integration of new information. *Learn. Mem.* 14, 47–53.
- Horner, A.J., and Burgess, N. (2014). Pattern completion in multielement event engrams. *Curr. Biol.* 24, 988–992.
- McKenzie, S., Frank, A.J.J., Kinsky, N.R.R., Porter, B., Rivière, P.D.D., and Eichenbaum, H. (2014). Hippocampal representation of related and

- opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron* 83, 202–215.
27. Doeller, C.F., Barry, C., and Burgess, N. (2010). Evidence for grid cells in a human memory network. *Nature* 463, 657–661.
 28. Sigman, M., Peña, M., Goldin, A.P., and Ribeiro, S. (2014). Neuroscience and education: prime time to build the bridge. *Nat. Neurosci.* 17, 497–502.
 29. Schapiro, A.C., Kustner, L.V., and Turk-Browne, N.B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Curr. Biol.* 22, 1622–1627.
 30. Howard, M.W., and Kahana, M.J. (2002). A distributed representation of temporal context. *J. Math. Psychol.* 46, 269–299.
 31. Dusek, J.A., and Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proc. Natl. Acad. Sci. USA* 94, 7109–7114.
 32. Preston, A.R., Shrager, Y., Dudukovic, N.M., and Gabrieli, J.D.E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus* 14, 148–152.
 33. Heckers, S., Zalesak, M., Weiss, A.P., Ditman, T., and Titone, D. (2004). Hippocampal activation during transitive inference in humans. *Hippocampus* 14, 153–162.
 34. Frankland, P.W., and Bontempi, B. (2005). The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130.
 35. Doeller, C.F., Opitz, B., Krick, C.M., Mecklinger, A., and Reith, W. (2005). Prefrontal-hippocampal dynamics involved in learning regularities across episodes. *Cereb. Cortex* 15, 1123–1133.
 36. Ezzyat, Y., and Davachi, L. (2014). Similarity breeds proximity: pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron* 81, 1179–1189.
 37. Gupta, A.S., van der Meer, M.A.A., Touretzky, D.S., and Redish, A.D. (2010). Hippocampal replay is not a simple function of experience. *Neuron* 65, 695–705.
 38. Poppenk, J., Evensmoen, H.R., Moscovitch, M., and Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends Cogn. Sci.* 17, 230–240.
 39. Komorowski, R.W., Garcia, C.G., Wilson, A., Hattori, S., Howard, M.W., and Eichenbaum, H. (2013). Ventral hippocampal neurons are shaped by experience to represent behaviorally relevant contexts. *J. Neurosci.* 33, 8079–8087.
 40. Ranganath, C., and Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nat. Rev. Neurosci.* 13, 713–726.
 41. Barron, H.C., Dolan, R.J., and Behrens, T.E.J. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat. Neurosci.* 16, 1492–1498.
 42. Kaplan, R., Horner, A.J., Bandettini, P.A., Doeller, C.F., and Burgess, N. (2014). Human hippocampal processing of environmental novelty during spatial navigation. *Hippocampus* 24, 740–750.
 43. Kim, G., Lewis-Peacock, J.A., Norman, K.A., and Turk-Browne, N.B. (2014). Pruning of memories by context-based prediction error. *Proc. Natl. Acad. Sci. USA* 111, 8997–9002.
 44. Schapiro, A.C., Rogers, T.T., Cordova, N.I., Turk-Browne, N.B., and Botvinick, M.M. (2013). Neural representations of events arise from temporal community structure. *Nat. Neurosci.* 16, 486–492.
 45. Lerner, Y., Honey, C.J., Silbert, L.J., and Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* 31, 2906–2915.