Behavioral/Systems/Cognitive

Flexible Memories: Differential Roles for Medial Temporal Lobe and Prefrontal Cortex in Cross-Episode Binding

Dagmar Zeithamova and Alison R. Preston

Department of Psychology, Center for Learning and Memory, and Institute for Neuroscience, The University of Texas at Austin, Austin, Texas 78712

Episodic memory is characterized by rapid formation of new associations that bind information within individual episodes. A powerful aspect of episodic memory is the ability to flexibly apply and recombine information from past experience to guide new behavior. A critical question for memory research is how medial temporal lobe (MTL) and prefrontal cortex (PFC), regions implicated in rapid within-episode binding, further support cross-episode binding in service of mnemonic flexibility. We set to answer this question using an associative inference task in humans that required rapid binding of information across overlapping experiences (AB, BC) to enable successful transfer to novel test probes (AC). Within regions predicting subsequent associative memory for directly learned associations, encoding activation in MTL, including hippocampus and parahippocampal cortex, uniquely predicted success on novel transfer trials both within and across participants, consistent with an integrative encoding mechanism where overlapping experiences are linked into a combined representation during learning. In contrast, during retrieval, PFC activation predicted trial-by-trial transfer success while MTL predicted transfer performance across participants. Moreover, increased MTL-PFC coupling was observed during novel transfer trials compared with retrieval of directly learned associations. These findings suggest that inferential processes support transfer of rapidly acquired experiences to novel events during retrieval where multiple memories are recalled and flexibly recombined in service of successful behavior. Together, these results demonstrate distinct encoding and retrieval mechanisms that support mnemonic flexibility, revealing a unique role for MTL regions in cross-episode binding during encoding and engagement of interactive MTL-PFC processes during flexible transfer at test.

Introduction

Episodic memory—memory for individual events—requires rapid formation of new associations that bind information within episodes. The current environment, however, often differs from prior experience requiring flexible combination of information across events to guide new behavior. Neuroimaging studies have identified a network of regions including prefrontal cortex (PFC) and medial temporal lobe (MTL) that support associative memory formation (Giovanello et al., 2004; Kirwan and Stark, 2004; Köhler et al., 2005; Murray and Ranganath, 2007; Staresina and Davachi, 2009). A fundamental question concerns how episodic memory networks implicated in rapid within-episode binding further support cross-episode binding in service of mnemonic flexibility.

One form of mnemonic flexibility—associative inference—requires recombination of information across overlapping experiences (AB, BC) to enable novel transfer judgments (AC). Animal lesion studies have revealed a critical role for the MTL and PFC in mnemonic flexibility (Bunsey and Eichenbaum, 1996; Dusek and Eichenbaum, 1997; DeVito et al., 2010a,b), but

it is unknown whether MTL and PFC contribute to cross-episode binding by the same or different mechanisms.

One hypothesized mechanism for cross-episode binding proposes that individual experiences are separately, but flexibly encoded, and then recalled and recombined at retrieval to support transfer (Eichenbaum and Cohen, 2001). According to this view, within-episode binding processes engaged during individual events (e.g., AB, BC) would have a similar impact on later flexibility, but transfer judgments would evoke additional inferential processing not required for directly experienced events. An alternate hypothesis proposes that overlapping events (e.g., BC) trigger specialized encoding mechanisms that lead to retrieval of related experiences (AB) and the formation of an integrated representation that binds information across episodes (ABC) (O'Reilly and Rudy, 2000; Shohamy and Wagner, 2008). Such integrative encoding could then support subsequent transfer without the need for additional inference during retrieval.

Neuroimaging studies have observed increased MTL (Heckers et al., 2004; Preston et al., 2004) and PFC activation (Acuna et al., 2002) during inferential retrieval relative to retrieval of directly experienced events, suggesting an important role for these regions in the flexible expression of memory at test. However, scanning in these studies was limited to blocked-design retrieval making it difficult to adjudicate the hypothesized mechanisms for cross-episode binding as well as preventing a trial-by-trial analysis linking brain activation to successful performance.

Two recent neuroimaging studies observed increases in MTL encoding activation across repetitions of overlapping events that

predicted individual differences in transfer (Greene et al., 2006; Shohamy and Wagner, 2008). However, these studies used procedures where a limited number of associations were presented several times during learning. These designs prevent direct comparison of encoding activation related to individual associations with activation during overlapping events that might represent the integration of distinct episodes. Thus, the observed MTL encoding activation could be consistent with either integrative encoding or flexible encoding of individual associations.

By using single-trial learning combined with success-based, event-related analyses, the current study aims to isolate encoding and retrieval mechanisms that support cross-episode binding and mnemonic flexibility, and their relative reliance on MTL and PFC processing.

Materials and Methods

Participants. Twenty-eight healthy, right-handed volunteers (9 females), ages 19–27 years, participated in the experiment after giving consent in accordance with a protocol approved by the University of Texas at Austin Institutional Review Board. Participants received $25/h for their involvement. Data from five participants were excluded from the analysis due to failure to achieve above chance accuracy for directly learned associations.

Materials. Stimuli consisted of 360 grayscale images of common objects. The stimuli were used to construct 144 overlapping object pairs (72 AB pairs, 72 BC pairs) and 72 unique, nonoverlapping pairs (XY). Overlapping AB and BC pairs were constructed such that two objects (A, C) shared an association with a third overlapping object (B) (Fig. 1*A*). XY pairs were constructed of two unique objects that did not share an overlapping association with other pairings. The presentation of object stimuli across pair types and the order of presentation were randomized across participants by assigning each participant to one of six randomization groups.

Procedures. The encoding and transfer task was a modified version of the associative inference paradigm (Preston et al., 2004) and consisted of two phases: (1) an encoding phase during which participants intentionally learned associations between object pairs (AB, BC, XY) and (2) a test phase during which participants were tested on directly learned associations (AB, BC, XY) and transfer trials consisting of novel combinations of object pairings (AC; Fig. 1*B*). Functional magnetic resonance imaging (fMRI) data were collected during both encoding and test phases. Before scanning, participants were instructed that they would be tested on directly learned associations (AB, BC, XY) as well as indirectly related transfer (AC) associations. Therefore, participants were aware before scanning that they would be required to make inferential AC judgments and were given an opportunity to practice the encoding and test phases of the experiment.

During six encoding scans, participants encountered AB, BC, and XY pairs, in addition to X objects in isolation. Stimuli were presented in a mixed fMRI design. An encoding scan consisted of four cycles during which four condition blocks (AB, BC, X, XY) were each presented once. Within a cycle, condition blocks were presented in a pseudo-random order, with AB blocks always preceding BC blocks and X blocks always preceding XY blocks. Each block lasted 72 s, consisting of 12 object pairs each presented for 3 s followed by a 1 s response period. Critically, each object pair was presented only once during encoding, requiring rapid acquisition of object associations. This single-trial learning procedure allowed for identification of regions associated with cross-episode binding on a trial-by-trial basis, to test the critical prediction of integrative encoding hypothesis that reactivation of prior experience during encoding of overlapping events supports later flexibility. During the response phase, participants were asked to provide a judgment of learning (Kao et al., 2005), rating on a 3-point scale how well they believed they learned the object pair (1 = will remember, 2 = may remember, 3 = will forget). These judgments were collected to ensure participants' attention during the encoding phase and were not considered in the analysis of fMRI data. Trial onsets within each condition block were jittered with a variable number (range 0–3) of odd/even digit baseline trials (Stark and
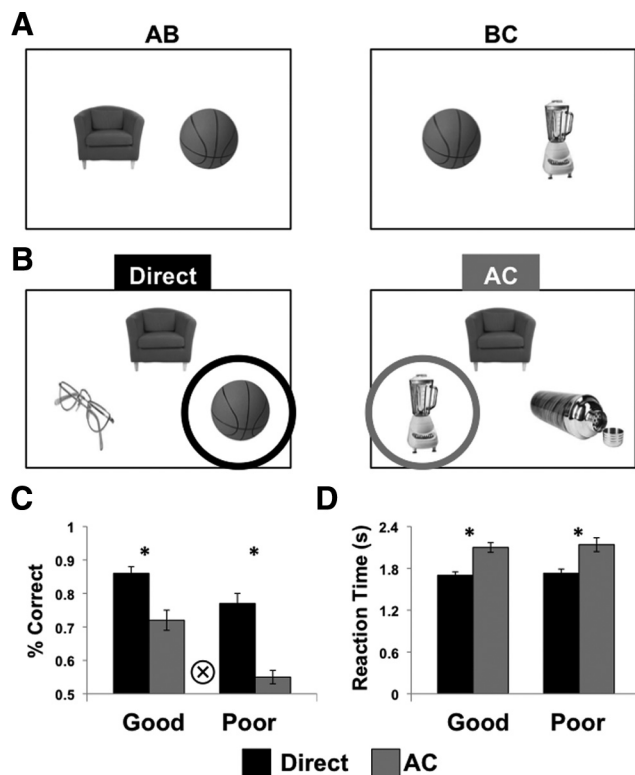

Figure 1. Task design and behavioral results. *A*, During encoding, participants intentionally learned associations (AB) and overlapping associations (BC) between two object stimuli. On each trial, object associations were presented for 3 s followed by a 1 s response period during which participants made a judgment of learning. *B*, After each encoding scan, participants were tested on directly learned associations and transfer trials (AC). On each test trial, a cue object was presented at the top of the screen, and participants made two-alternative forced-choice judgments selecting the correct association from two choice stimuli on the bottom of the screen. *C*, *D*, Proportion correct (*C*) and reaction times (*D*) for directly learned associations (black bars) and transfer trials (gray bars) for participants with good transfer performance and those with poor transfer performance. Error bars denote across-subject SEM. Asterisk denotes pairwise significance at $p < 0.05$, circled cross denotes trial type $\times$ group interaction at $p < 0.05$.

Squire, 2001) using a sequence optimization program (Dale, 1999), to allow for event-related analyses of encoding trials within each block. During each 2 s baseline trial, a single digit between 1 and 8 was presented on the screen and participants indicated whether the digit was odd or even. Additional 12 s blocks of the baseline task occurred at the beginning and end of each encoding scan and between each condition block.

After each encoding scan, participants were tested on directly learned associations (AB, BC, XY) and transfer trials (AC) (Fig. 1*B*) using event-related methods. During each 4 s test trial, a single cue object (e.g., an A object) was presented on the top of the screen and two choice objects were presented on the bottom of the screen (e.g., two B objects). Participants were required to make a forced-choice decision indicating which of the two choice objects was associated with the cue on the top. Participants were instructed that on transfer trials, the association between the cue (A) and the correct choice (C) was indirect, mediated through a third object that shared an association with both the cue and the correct choice during encoding (i.e., object B). The incorrect choice was another familiar C object studied at encoding, but that did not share a common association with the cue.

Individual transfer trials (AC) were presented before corresponding AB and BC trials to ensure that subjects could not form an association between A and C objects during the test. Importantly, for both directly learned and transfer associations, the incorrect choice objects were all familiar items that had been studied in the context of another object different from the cue. Thus, correct responses required the retrieval of learned associations and could not be made based on familiarity of the choice objects. Odd/even baseline trials were intermixed with test trials,

Table 1. Regions of interest demonstrating successful subsequent associative memory

Region	Left/right	Cluster size	Z value	X	Y	Z
Whole brain ($p < 0.05$ corrected)						
Inferior frontal gyrus	L	214	4.6	−46	28	14
Orbitofrontal gyrus	L	104	4.14	−38	42	−12
Posterior parietal cortex	L	55	3.83	−36	−80	34
Inferior temporal cortex	L	215	3.98	−38	−72	−6
Lateral occipital cortex	R	333	4.28	28	−94	16
Fusiform gyrus	R	86	3.92	30	−44	−18
Medial temporal lobe (voxel $p < 0.005$, cluster size > 10)						
Hippocampus	R	18	3.39	22	−30	−4
Parahippocampal cortex	L	91	3.37	−28	−40	−16
	R	73	3.44	34	−38	−14

with the order of conditions determined by a sequencing program to optimize the efficiency of event-related fMRI design (Dale, 1999).

fMRI data acquisition. Whole-brain imaging data were acquired on a 3.0T GE Signa MRI system (GE Medical Systems). Structural images were acquired using a T2-weighted flow-compensated spin-echo pulse sequence [repetition time (TR) = 3 s; echo time (TE) = 68 ms; 256 × 256 matrix, 1 × 1 mm in-plane resolution] with 31 3-mm-thick oblique axial slices (0.6 mm gap), ~20 degrees off the AC-PC line. Functional images were acquired using a multiecho GRAPPA parallel imaging echoplanar imaging (EPI) sequence using the same slice prescription as the structural images (TR = 2 s; TE = 30 ms; 2 shot; flip angle = 90°; 64 × 64 matrix; 3.75 × 3.75 mm in-plane resolution, interleaved slice acquisition). For each functional scan, the first six EPI volumes were discarded to allow for T1 stabilization. Head movement was minimized using foam padding.

fMRI data analysis. Data were preprocessed and analyzed using SPM5 (Wellcome Department of Imaging Neuroscience, University College London, London, UK) and custom Matlab (MathWorks) routines. Functional images were corrected to account for differences in slice acquisition times by interpolating the voxel time series using sinc interpolation and resampling the time series using the first slice as a reference point. Functional images were then realigned to the first volume in the time series to correct for motion. The T2-weighted structural image was coregistered to the mean T2*-weighted functional volume computed during realignment. The structural image was then spatially normalized into common stereotactic space using the MNI (Montreal Neurological Institute) template brain. The spatial transformation calculated during normalization of the structural image was then applied to the functional time series and resampled to 2 mm isotropic voxels. After normalization, the functional images were spatially smoothed using an 8 mm full-width half-maximum Gaussian kernel, high-pass filtered with a 128s filter, and converted to percent signal.

Voxelwise analysis was performed under the assumptions of the general linear model. Each phase (encoding and test) was analyzed separately. Regressor functions were constructed by modeling stimulus-related activation as a stick function convolved with a canonical hemodynamic response function together with the temporal derivative. Functional regions of interest were identified using a subsequent associative memory contrast and then interrogated using region of interest (ROI) analyses to test critical hypotheses regarding the contributions of these regions to flexible transfer.

Two regressors were constructed representing all associative encoding trials that were later remembered and those that were forgotten. Single-object (X) trials and trials in which a participant failed to make the response within the allocated time were modeled separately from correct and incorrect associative encoding trials and were not considered in the fMRI analyses. For each participant, the linear contrast of subsequently correct and subsequently incorrect associative trials (AB, BC, XY) was computed, generating statistical maps representing differences in brain activation between the two conditions using a fixed effect analysis. Contrast images generated in the individual participant analysis were analyzed across participants using a mixed effects general linear model, treating participants as a random effect to allow for population inference. The whole-brain statistical maps were thresholded at voxel threshold of FDR corrected $p < 0.05$ and cluster extent threshold of 46 voxels (cluster
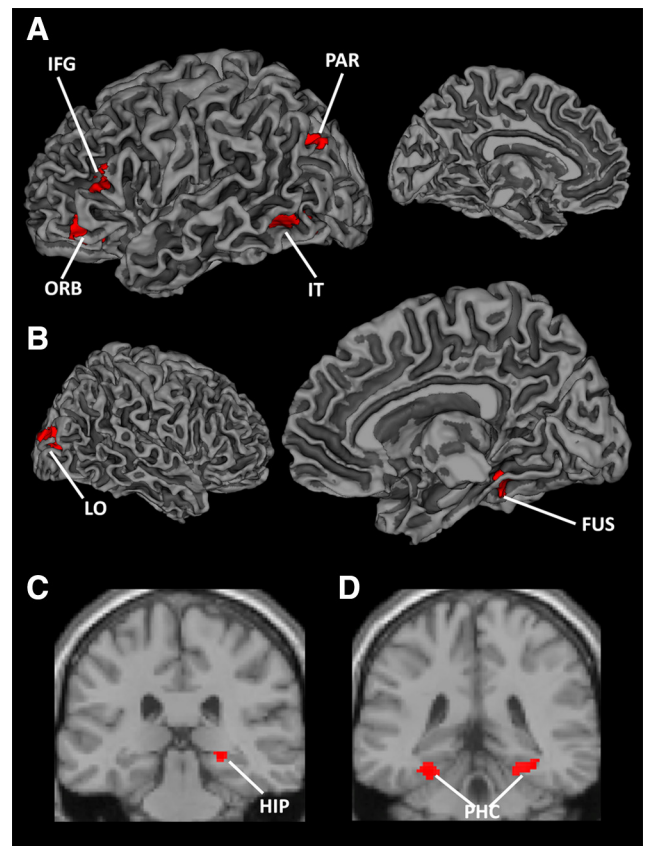

Figure 2. Brain regions showing a subsequent associative memory effect. Brain regions whose activity during encoding of later remembered associations was greater than during later forgotten associations. These areas were used as regions of interest for all transfer success analyses. Activation maps are shown on a single-subject MNI template brain. *A*, Surface rendering of left hemisphere. *B*, Surface rendering of right hemisphere. *C*, Coronal slice ($y = -30$) displaying hippocampal activation. *D*, Coronal slice ($y = -38$) displaying activation in parahippocampal cortex. PAR, Posterior parietal cortex; IT, inferior temporal cortex; LO, lateral occipital cortex; FUS, fusiform; HIP, hippocampus.

extent $p < 0.05$). This contrast provided six clusters demonstrating subsequent associative memory effects (Table 1, Fig. 2 *A*,*B*), serving as regions of interest for the critical analysis of transfer performance. In addition to the whole brain analysis, subsequent associative memory effects were assessed within MTL, our a priori anatomical region of interest. As in previous studies (Eldridge et al., 2000; Davachi and Wagner, 2002; Strange et al., 2002; Preston and Gabrieli, 2008), a threshold of 10 or more contiguous voxels exceeding an uncorrected threshold of $p < 0.005$ was adopted to assess activation in MTL. A MTL mask was derived from the Harvard-Oxford MNI Atlas (Center for Morphometric Analysis, Neuroscience Center, MGH-East, Charlestown, MA) and

encompassed hippocampus and parahippocampal gyrus. This contrast provided right hippocampal cluster (Fig. 2C) and right and left parahippocampal clusters (Fig. 2D) to serve as additional functional regions of interest (Table 1). The left and right parahippocampal clusters showed the same pattern of activation across all analyses and were treated as a single ROI.

The functional regions of interest were interrogated for their contribution to cross-episode binding during encoding and test. Two models were constructed to test the hypotheses regarding how these regions contribute to flexible acquisition and use of memory during the encoding and test phase respectively. For the encoding phase, AB and BC trials were separately sorted based on whether the corresponding transfer (AC) trial was later correct or incorrect. For the test phase, the model consisted of four regressors representing correct and incorrect directly learned associations and correct and incorrect transfer associations. These models specifically examined the open questions regarding how integrative encoding and inferential retrieval processes support mnemonic flexibility, by examining the specific contributions of MTL and PFC to these processes during encoding and retrieval. It is important to note that the functional ROI definition based on subsequent memory for all directly learned associations is unbiased with respect to the comparisons interrogating cross-episode binding. For each ROI, the deconvolved signal was extracted using a finite impulse response function implemented in Mars-Bar (http://marsbar.sourceforge.net/). Integrated percent signal change for both encoding and test phases was calculated as the sum of percent signal change across the time points 4–8 s poststimulus onset, allowing for comparison of ROI responses across conditions.

To investigate whether interactions between MTL and PFC underlie mnemonic flexibility during inferential retrieval, we assessed changes in inter-regional coupling during transfer. To estimate these changes, we compared inter-regional correlations across participants during retrieval of directly learned associations and during transfer trials using bootstrapping methods, obtaining a reliability index (p-value) for each connection. During bootstrap simulations, we used 10,000 simulated samples from our subject population, computed the pairwise inter-regional correlation during transfer and direct retrieval trials, and stored the difference between them. For each pair of regions, the proportion of simulations with difference greater than zero determines the one-tailed reliability (p-value) of the increase in inter-regional coupling and the proportion of simulations with difference smaller than zero determines the one-tailed p-value of the decrease in inter-regional coupling. To compute a two-tailed p-value, the smaller of the two p-values is doubled, meaning that a two-tailed test is significant at a $p = 0.05$ level when the one-tailed probability of the null hypothesis is <0.025.

Results

Behavioral results

All participants successfully acquired directly learned object associations (mean = 82% correct, SD = 9%; $t_{(22)} = 16.36$, $p < 0.001$). Mean performance on novel transfer trials was significantly above chance ($t_{(22)} = 5.81$, $p < 0.001$) averaging 64% correct (SD = 12%). Interestingly, large individual differences in transfer performance (range 47–90%) were observed, with a binomial test revealing that 12 of the 23 participants reached above chance accuracy on transfer trials (44 of 72 correct, $p < 0.05$). Dividing participants into those with good ($n = 12$) versus poor ($n = 11$) transfer performance revealed significant group differences in recognition accuracy for transfer ($t_{(21)} = 5.52$, $p < 0.001$) and directly learned associations ($t_{(21)} = 2.54$, $p = 0.02$). Importantly, performance on directly learned associations in the poor transfer group was well above chance (mean = 77%, SD = 8%, $t_{(10)} = 10.66$, $p < 0.001$), and accuracy differences between groups were significantly greater for transfer trials relative to directly learned associations (group × condition interaction, $F_{(1,21)} = 5.89$, $p = 0.02$; Fig. 1C). These observed behavioral differences enabled us to explore how individual differences in

performance are reflected in trial-by-trial encoding and retrieval processing within different ROIs.

At test, mean reaction times for correct trials were significantly faster for directly learned associations (1714 ms, SD = 192 ms) relative to transfer trials (2119 ms, SD = 274 ms; $t_{(22)} = 11.87$, $p < 0.001$), suggesting that a more effortful process was associated with retrieval on transfer trials. Reaction times for directly learned and transfer probes did not differ between participants with good versus poor transfer performance (all p-values > 0.7; Fig. 1D).

Regions of interest involved in successful associative binding within an episode

Because flexible transfer across episodes depends on having acquired the individual associations (binding within an episode), we first identified brain regions that tracked subsequent memory (Brewer et al., 1998; Paller and Wagner, 2002) for directly learned associations, to use as functional ROIs. To identify these ROIs, we sorted all associative encoding trials (AB, BC, XY) based on whether these directly learned associations were successfully recognized or forgotten at test. This contrast identified two PFC regions [orbitofrontal (ORB) and inferior frontal gyrus (IFG)], one posterior parietal region, three regions in object-sensitive extrastriate cortex (lateral occipital cortex, fusiform, and inferior temporal cortex), and two MTL regions [right hippocampus and bilateral parahippocampal cortex (PHC)] demonstrating subsequent memory effects for learned associations (Table 1, Fig. 2). These regions were then interrogated for their contribution to associative inference (associative binding across episodes) during encoding and retrieval. While these regions have been previously implicated in the formation of individual associative memories (Sperling et al., 2003; Jackson and Schacter, 2004; Kirwan and Stark, 2004; Köhler et al., 2005; Murray and Ranganath, 2007; Staresina and Davachi, 2009), it is currently unknown whether, or by what mechanism, these regions contribute to cross-episode binding and the flexible expression of memory.

Encoding: MTL supports flexibility through integration across experiences

The regions that support associative binding within an episode may support subsequent transfer via two hypothesized mechanisms. First, successful associative encoding of individual experiences may be essential to subsequent flexible recombination of learned information. Accordingly, for the flexible recombination at test to be successful, both associations (AB, BC) must be successfully encoded and accessible from partial cues at retrieval. Therefore, successful within-episode binding of individual associations at encoding would be expected to predict subsequent AC performance, with encoding of both AB and BC associations having a similar impact on later transfer. In complement to this separate but flexible associative encoding mechanism, a second, integrative encoding process engaged during the presentation of the overlapping event has been hypothesized to further support transfer performance. The integrative encoding hypothesis proposes that during encoding of an event with feature overlap to an existing representation (e.g., BC), the overlapping element (B) serves as a cue for reactivation of previous experience (AB), leading to the formation of a new association (AC) that was not directly experienced. This newly encoded association could then be used to support novel AC decisions without the need for retrieval of each individual association at test. Importantly, this view makes a different prediction about encoding activation than the associative mechanism described above in that BC encoding
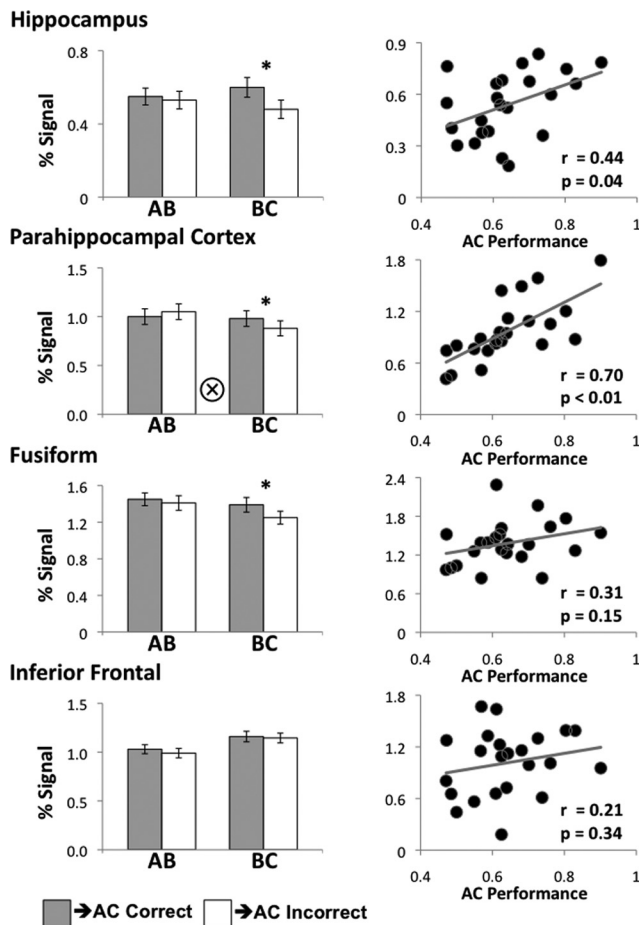
Figure 3. Relationship between encoding activation and subsequent transfer. Left column, Activation during encoding of AB trials and BC trials, sorted based on subsequent transfer success of the corresponding AC trial. BC activation was found to predict subsequent transfer in hippocampus, parahippocampal cortex, and fusiform. Error bars denote across-subject SEM. Asterisk denotes pairwise significance at $p < 0.05$, circled cross denotes trial type $\times$ transfer success interaction at $p < 0.05$. Right column, Across-subject correlation between transfer accuracy and encoding activation. Only encoding activation in hippocampus and parahippocampal cortex was found to track individual differences in performance.

activation should have a greater (or unique) impact on subsequent transfer judgments than AB encoding because of the additional cross-episode binding evoked on these trials.

To test for these hypothesized patterns, we separately sorted AB and BC encoding trials based on whether the corresponding AC trial was subsequently correct or incorrect at test, and interrogated encoding activation during these four conditions within the ROIs. For each ROI, we ran a 2 $\times$ 2 within-subject ANOVA with trial type (AB, BC) $\times$ transfer success (correct, incorrect) analysis, followed by a subsequent simple effect analysis (paired t test), separately for AB and BC trials. Three regions showed subsequent transfer effects as revealed by the ANOVA—hippocampus, PHC, and fusiform (Fig. 3, left column). Hippocampus showed a main effect of transfer success ($F_{(1,22)} = 7.42$, $p = 0.012$), no effect of trial type ($p > 0.9$), and no interaction ($p > 0.18$), suggesting a role for this region in flexible encoding of individual associations. Fusiform showed a main effect of transfer success ($F_{(1,22)} = 6.15$, $p = 0.021$, a main effect of trial type (greater AB than BC activation, $F_{(1,22)} = 10.7$, $p = 0.003$) and no interaction ($p > 0.15$). The main effect of trial type may reflect novelty response in this region (Kirchhoff et al., 2000). Finally, PHC showed a main effect of trial type (greater AB than BC

activation, $F_{(1,22)} = 10.1$, $p = 0.004$), no main effect of transfer success ($p > 0.4$), and a significant transfer success $\times$ trial type interaction ($F_{(1,22)} = 4.38$, $p = 0.048$), with BC trials, but not AB trials, predicting transfer success, as predicted by the integrative encoding hypothesis. The simple effects analysis revealed no region predicting subsequent transfer for AB trials, but the same regions indicated by the ANOVA—hippocampus, PHC and fusiform—demonstrated a subsequent transfer effect for BC trials (Fig. 3, left column). The pattern of simple effects for BC trials in hippocampus, PHC, and fusiform further suggest that additional processes are engaged during BC trials that promote successful transfer performance, consistent with the integrative encoding hypothesis. However, the lack of a significant transfer success $\times$ trial type interaction in hippocampus suggests that processing in this region could reflect flexible encoding of individual associations, integrative encoding evoked during BC trials, or a combination of both mechanisms. Future studies using high-resolution fMRI methods may provide a means to differentiate distinct encoding signatures within hippocampus.

The same trial type $\times$ transfer success ANOVA revealed two additional regions whose activation was modulated by trial type but not subsequent transfer. The IFG (Fig. 3) and the posterior parietal cortex demonstrated greater activation during BC trials than AB trials (IFG: $F_{(1,22)} = 11.5$, $p = 0.003$; posterior parietal cortex: $F_{(1,22)} = 5.80$, $p = 0.025$), regardless of subsequent transfer success (all other $p > 0.3$). The success-based analysis suggest that the pattern observed in these regions does not reflect integrative encoding during overlapping events, as the BC activation did not predict subsequent transfer success. Rather, this pattern may reflect the previously established role of IFG in overcoming interference (Badre and Wagner, 2007; Oztekin et al., 2009). Importantly, these results demonstrate that MTL, but not PFC, contribute to cross-episode binding during learning, with this dissociation being supported by significant region $\times$ condition interactions (IFG-PHC: $F_{(3,66)} = 12.3$, $p < 0.001$; IFG-hippocampus: $F_{(3,66)} = 4.95$, $p = 0.004$).

After establishing the relationship between trial-by-trial encoding activation and subsequent transfer in the MTL and fusiform, we investigated whether these or other ROIs tracked individual differences in transfer success across subjects. We found a significant correlation between individual differences in hippocampal and PHC activation during encoding and transfer accuracy across participants (hippocampus: $r(23) = 0.44$, $p = 0.04$; PHC: $r(23) = 0.70$, $p < 0.001$), with participants who showed greater MTL activation during encoding demonstrating greater transfer accuracy (Fig. 3, right column). No other region predicted individual differences in performance (all $p > 0.15$). These findings provide additional evidence for the role of hippocampus and PHC in cross-episode binding during learning.

To provide an additional test for the contribution of each region to individual differences in transfer success across participants (while controlling for accuracy on directly learned associations), we used a stepwise regression using transfer accuracy as the dependent variable and accuracy on the directly learned associations and task related encoding activation in each ROI as the independent variables. We found that activation in PHC was most strongly predictive of the transfer success ($\beta = 0.518$, $p = 0.004$), followed by the accuracy on directly learned associations ($\beta = 0.383$, $p = 0.027$). The addition of other regions to the regression model did not further account for variance in transfer performance (all $p > 0.38$).

Retrieval: PFC and MTL support inferential retrieval

While our encoding findings provide evidence for an MTL-based integrative learning mechanism triggered by overlapping events, it is possible that the two hypothesized mechanisms for cross-episode binding are not mutually exclusive, but rely on distinct networks of brain regions. If flexible transfer additionally relies on inferential processing during retrieval, AC trials should evoke unique processes at test not observed during retrieval of directly learned associations. Such specialized processes during transfer judgments may be manifested by a pattern of activation where retrieval activation differentiates correct and incorrect AC trials, but not correct and incorrect retrieval of directly learned associations. Additionally, it may be represented by greater activation during AC trials compared with direct retrieval trials, where such a difference is linked to a better transfer performance. To investigate how regions supporting successful within-episode binding further support flexible transfer at the time of retrieval, we again used trial-by-trial and across-participant analyses of retrieval activation within each ROI identified from the encoding time period.

First, for the trial-by-trial analysis, we sorted retrieval trials based on success (correct, incorrect), separately for directly learned associations and transfer trials (AC), and interrogated retrieval activation during these four conditions. For each ROI, we ran a 2 × 2 within-subject ANOVA with factors for trial type (direct, AC) and retrieval success (correct, incorrect), followed by a simple retrieval success analysis (paired t test), separately for direct and AC trials. Only the two PFC regions—IFG (Fig. 4) and ORB—showed trial-by-trial accuracy effects. Both regions showed the same pattern, with no main effect of trial type, no effect of retrieval success (all $p > 0.2$) and significant trial type × retrieval success interaction (IFG: $F_{(1,22)} = 9.15$, $p = 0.006$; ORB: $F_{(1,22)} = 4.96$, $p = 0.037$), with greater activation during correct relative to incorrect transfer trials but greater activation for incorrect relative to correct direct retrieval. This pattern of results cannot be explained by differences in memory strength between the directly learned and transfer trials, as activation would be expected to be greater for correct relative to incorrect trials for both trial types. Instead, these results provide a strong evidence for additional processing during transfer trials that uniquely determines trial-by-trial transfer success. No other regions showed significant retrieval success effects in either ANOVA or simple effects analysis. These findings highlight the unique role of the PFC in predicting trial-by-trial transfer success at test, but not at encoding, and are consistent with the notion of inferential processing during retrieval (DeVito et al., 2010b; Wendelken and Bunge, 2010).

It is possible that other regions, such as the MTL, differentially contribute to inferential retrieval through the successful recall of previously learned associations. Specifically, greater activation during transfer trials relative to retrieval of directly learned associations would indicate the role of a region in flexible memory use. To be behaviorally relevant, this difference should only be observed in the good transfer group, but not the poor transfer group. To test for this pattern, we interrogated activation within each ROI during direct retrieval and transfer (AC) trials. We then compared activation during these two types of trials for each group of participants, using a mixed design ANOVA with trial type (direct, AC) as a within-subject factor and group (poor transfer, good transfer) as a between-subject factor. Two regions showed a significant trial type × group interaction—the hippocampus and PHC (Fig. 4)—that tracked individual differences in performance. The hippocampus showed no main effects (both
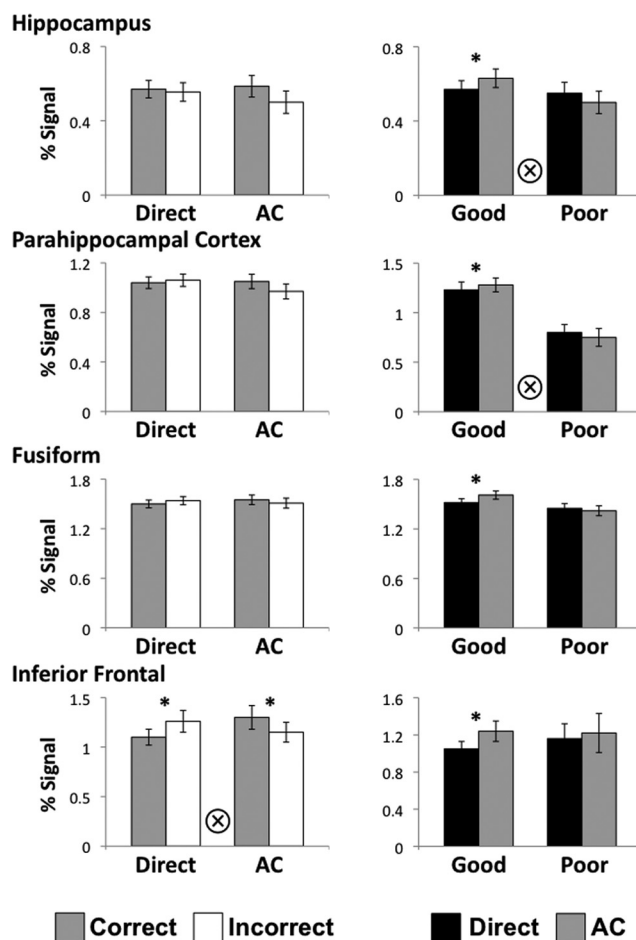

Figure 4. Relationship between retrieval activation and behavioral performance. Left column, Activation during successful (gray bars) and unsuccessful (white bars) retrieval of directly learned associations and transfer trials. Activation in the inferior frontal gyrus was predictive of trial-by-trial transfer success. Error bars denote across-subject SEM. Asterisk denotes pairwise significance at $p < 0.05$; circled cross denotes trial type × retrieval success interaction at $p < 0.05$. Right column, Activation during direct retrieval trials (black bars) and transfer trials (gray bars) for the good and poor transfer groups. Retrieval activation in hippocampus and parahippocampal cortex tracked individual differences in transfer performance. Error bars denote SEM. Asterisk denotes pairwise significance at $p < 0.05$, circled cross denotes trial type × performance group interaction at $p < 0.05$.

$p > 0.4$), but a significant interaction ($F_{(1,21)} = 4.43$, $p = 0.048$), with good transfer group, but not poor transfer group, showing greater activation during transfer trials relative to direct retrieval. The PHC showed no main effect of trial type ($p > 0.9$), a significant main effect of group ($F_{(1,21)} = 17.4$, $p < 0.001$), with the good transfer group showing overall greater parahippocampal activation than the poor transfer group, and a significant interaction ($F_{(1,21)} = 4.56$, $p = 0.045$). These findings emphasize the relationship between MTL activation and individual differences in transfer performance.

Similar to encoding, we performed an additional test of the contribution of each region's retrieval activation to the individual differences in transfer success across participants, taking into account accuracy on directly learned associations. We used a stepwise regression with transfer accuracy as the dependent variable and accuracy on directly learned associations and activation during transfer trials in each ROI as the independent variables. As observed during encoding, we found that activation in PHC was strongly predictive of transfer success ($\beta = 0.544$, $p = 0.002$), as
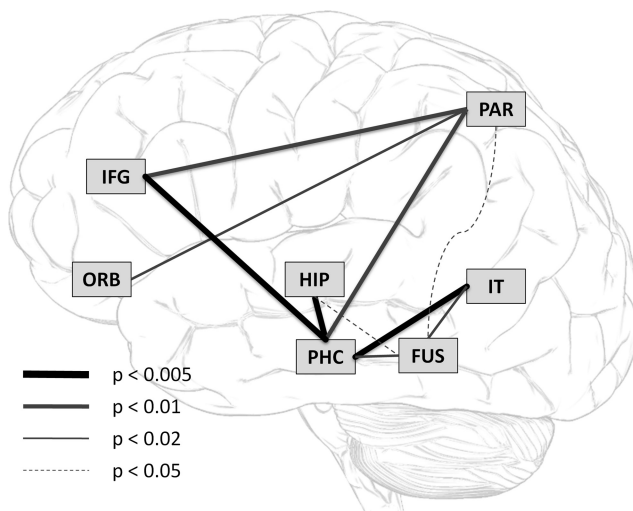
Figure 5. Inter-regional connectivity during inferential retrieval. All inter-regional connections that showed a significant increase during transfer trials relative to retrieval of directly learned associations. The most robust increase in connectivity was observed between parahippocampal cortex and three other regions— hippocampus, inferior frontal gyrus, and inferior temporal cortex—suggesting interactions in the episodic memory network mediated by parahippocampal cortex in service of successful inferential retrieval.

was accuracy on directly learned associations ($\beta = 0.389$, $p = 0.018$). No other region accounted for a significant amount of variance in transfer performance (all $p > 0.19$).

Inter-regional coupling: MTL-mediated interactions between regions support flexibility during test

The complementary roles of PFC and MTL during inferential retrieval suggest an interaction between these regions in service of flexible transfer. To explicitly test whether transfer trials evoke changes in network connectivity relative to recognition of learned associations, we measured inter-regional coupling separately during retrieval of directly learned associations and transfer trials, using across-subject correlations. To identify significant changes in connectivity, we computed the reliability of the connection change (p-value) from retrieval of directly learned associations to transfer trials using bootstrapping methods (Efron and Tibshirani, 1993). The results are schematically presented in Figure 5, depicting all inter-regional connections that showed significant increase during transfer trials (two-tailed $p = 0.05$, uncorrected for multiple comparisons). The results demonstrate a significant increase in connectivity across the network, with 10 connections significantly increasing during transfer trials relative to recognition of learned associations ($p < 10^{-6}$; binomial test for 10 of 28 possible connections at $p = 0.05$). No connection significantly decreased (all $p > 0.6$); the coupling between lateral occipital cortex and other regions did not change. The connectivity analysis depicted in Figure 5 suggests that PHC plays an important role in mediating the interactions between hippocampus, PFC, and the posterior visual areas in service of associative inference.

Discussion

The ability to rapidly extract knowledge from a single experience and use this knowledge flexibly is a fundamental characteristic of episodic memory. In the current study, we investigated flexible application of episodic experience using an associative inference task that required integration of information across overlapping experiences (AB, BC) to allow for successful transfer (AC). Using

single-trial learning and success-based, event-related analyses, we provide the first evidence linking trial-by-trial encoding and retrieval activation within an episodic memory network with the flexible acquisition and use of memory. We found distinct mechanisms supporting mnemonic flexibility during encoding and retrieval, with a unique role for MTL regions in cross-episode binding during learning and an interactive role of MTL and PFC during flexible transfer at test.

At encoding, PFC activation was associated with successful binding within individual experiences, predicting later retrieval success for directly learned associations, but did not predict flexible transfer. Rather, IFG demonstrated greater activation during BC trials relative to AB trials, consistent with the role of IFG in overcoming interference from overlapping memories (Badre and Wagner, 2007; Oztekin et al., 2009).

In contrast to PFC, encoding activation in MTL regions— hippocampus and PHC—predicted subsequent transfer within and across participants. The current results demonstrate the recruitment of specialized processes during encoding of overlapping BC events, where hippocampal and PHC activation during encoding of BC, but not AB, associations predicted subsequent flexible transfer. This finding suggests that hippocampus and PHC both make important contributions to cross-episode binding at encoding. We propose that in the presence of overlapping contextual information (**B**C), hippocampal pattern completion mechanisms (O'Reilly and Rudy, 2001) elicit the reactivation of prior event details (i.e., **A**B) represented by PHC (Bar et al., 2008) thus enabling the formation of a new association between current and prior experience (AC). Importantly, MTL associative binding, in contrast to PFC encoding mechanisms, may support this integrative process through the elemental, or conjunctive, encoding of episodic information that allow memories to be addressable from partial input (Eichenbaum et al., 1996; O'Reilly and Rudy, 2001). The pattern of results observed in hippocampus may reflect such flexible conjunctive encoding as well as integration during overlapping events. Because hippocampal-PHC interactions are essential to successful cross-episode binding, encoding activation in both regions would be expected to predict individual differences in transfer performance, as we observe in the current study.

It is possible, however, that BC encoding activation does not reflect integrative encoding, but instead reflects strengthening of individual AB and BC associations through the same within-episode binding mechanisms evident during nonoverlapping events. In such a case, BC encoding activation should then predict performance on AB associations (as seen by Kuhl et al., 2010). To test this alternative, we interrogated BC encoding activation with respect to subsequent AB retrieval and did not find evidence for this associative strengthening hypothesis (hippocampus, $p = 0.64$; PHC, $p = 0.37$). This result suggests that an AC association is formed during BC encoding, serving as a basis for subsequent transfer performance. It is important to note that this finding may result from the intentional nature of our inference task, as participants were aware during encoding that they would be tested on indirect associations. An important avenue for future research will be to determine how these mechanisms operate under incidental learning conditions. Regardless of the specific interpretation of MTL encoding signatures during overlapping events, we demonstrate a unique role for the hippocampus and PHC in encoding processes that support the later flexible use of experience that were not evident in PFC.

The current findings also support the hypothesis that additional retrieval processing promotes flexible transfer. Partici-

pants took significantly longer to respond to transfer trials than directly learned associations. This finding contrasts a recent study where overlapping associations were acquired across repeated exposures that found no reaction time difference between learned and transfer probes in participants who successfully transferred (Shohamy and Wagner, 2008). We hypothesize that integrative encoding predominately supports transfer when learning occurs across multiple repetitions, by providing multiple opportunities for cross-episode binding. In contrast, rapid single-trial learning requires additional processing during retrieval, namely the flexible recombination of stored representations (Eichenbaum, 2000; Preston et al., 2004).

Our imaging findings provide additional support for such flexible recombination during retrieval. During test, PFC regions—IFG and ORB— uniquely predicted trial-by-trial transfer success, with greater activation for successful relative to unsuccessful transfer, but the opposite pattern for directly learned associations. These findings demonstrate that transfer trials do not simply differ from directly learned trials in the memory strength, but rather, they recruit qualitatively different processes. IFG has been implicated in non-mnemonic relational reasoning where multiple relationships must be considered to infer an unknown relationship (Christoff et al., 2001; Wendelken and Bunge, 2010), suggesting that PFC contributions to flexible AC transfer may be through inferential judgments from premise associations (AB and BC, therefore AC). Animal lesion studies have demonstrated that PFC is essential for transitive inference in rats (DeVito et al., 2010b). However, because lesions to PFC were performed before learning, the exact mechanism by which PFC contributes to flexibility could not be determined. The current findings suggest a key contribution of PFC to inferential retrieval rather than cross-episode binding during encoding.

We also found evidence for MTL contributions to transfer during retrieval, demonstrating hippocampal and PHC activation that tracked individual differences in transfer performance. In contrast, PFC activation did not predict individual differences. These findings suggest that differences in transfer performance were primarily the result of individual differences in memory processing, such as the strength of encoded memories or ease of their retrieval, rather than differences in inferential reasoning. We observed an increase in inter-regional coupling during transfer compared with directly learned associations, primarily between MTL regions, IFG, and extrastriate cortex, indicating an important role for PFC-MTL interactions during inferential retrieval. We propose that during transfer, MTL and PFC regions play complementary roles, where MTL regions support retrieval of individual events from partial cues that can be flexibly recombined by PFC (Ranganath, 2010). This retrieval event may also provide another opportunity for cross-episode binding in the case that such binding was not achieved during encoding, as was reflected in increased MTL activation in participants with good transfer.

In conclusion, the current findings answer several open questions regarding flexible episodic memory. The current findings provide the first demonstration of integrative encoding on a trial-by-trial basis and highlight the unique role of MTL in associative binding not only within an episode but also across episodes, to promote the formation of new knowledge extending beyond individual events. Leading theories of MTL function suggest that the PHC supports encoding and recollection of contextual (both spatial and nonspatial) information associated with specific items, whereas the hippocampus binds those items to context information within individual events (Davachi, 2006; Diana et

al., 2007; Eichenbaum et al., 2007). In the current study, object A can be considered a nonspatial context for object B that can be recalled during the encoding of the overlapping BC event to support cross-episode binding. Our findings build upon these theories to suggest that the inherent flexibility of episodic memory can result from binding mechanisms that enable the integration of information from multiple contexts.

The flexibility to combine experiences in novel ways to infer new knowledge is fundamental to many forms of goal-directed behavior. Recent evidence has implicated MTL regions in acquisition and use of categorical knowledge (Zeithamova et al., 2008; Kumaran et al., 2009). The current findings suggest that cross-episode binding may be the key mechanism by which MTL regions contribute to such behaviors to allow for the abstraction of common features across multiple related experiences and flexible generalization to similar future events.

Our findings provide further evidence for the complementary roles of MTL and PFC regions during flexible transfer where MTL regions support the retrieval of individual events and PFC enables the flexible recombination of those experiences through inferential processing. While memory is most often considered in terms of its ability to recall the past, imagining and predicting future events relies on much of the same neural network (Okuda et al., 2003; Addis et al., 2007, 2009; Hassabis et al., 2007; Szpunar et al., 2007). Some studies have shown greater recruitment of MTL and PFC in future event simulation than in episodic recollection (Okuda et al., 2003; Addis et al., 2007), that may reflect constructive recombination and binding of disparate elements from past experiences into a novel coherent representation (Hassabis and Maguire, 2007, 2009; Schacter et al., 2008; Addis et al., 2009). The same reconstructive processes are likely recruited in the associative inference task during transfer, and the present study suggests the mechanism of such future simulation, with complementary roles of MTL and PFC in flexible recombination of existing representations to support the predictive capacity of memory. Together, the current findings shed light on the individual contributions and functional interactions of MTL and PFC in service of mnemonic flexibility, providing important evidence for the mechanisms that support the flexible use of experience beyond the original learning circumstances.

References

Acuna BD, Eliassen JC, Donoghue JP, Sanes JN (2002) Frontal and parietal lobe activation during transitive inference in humans. Cereb Cortex 12:1312–1321.

Addis DR, Wong AT, Schacter DL (2007) Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. Neuropsychologia 45:1363–1377.

Addis DR, Pan L, Vu MA, Laiser N, Schacter DL (2009) Constructive episodic simulation of the future and the past: distinct subsystems of a core brain network mediate imagining and remembering. Neuropsychologia 47:2222–2238.

Badre D, Wagner AD (2007) Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia 45:2883–2901.

Bar M, Aminoff E, Ishai A (2008) Famous faces activate contextual associations in the parahippocampal cortex. Cereb Cortex 18:1233–1238.

Brewer JB, Zhao Z, Desmond JE, Glover GH, Gabrieli JD (1998) Making memories: brain activity that predicts how well visual experience will be remembered. Science 281:1185–1187.

Bunsey M, Eichenbaum H (1996) Conservation of hippocampal memory function in rats and humans. Nature 379:255–257.

Christoff K, Prabhakaran V, Dorfman J, Zhao Z, Kroger JK, Holyoak KJ, Gabrieli JD (2001) Rostrolateral prefrontal cortex involvement in relational integration during reasoning. Neuroimage 14:1136–1149.

Dale AM (1999) Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109–114.

Davachi L (2006) Item, context and relational episodic encoding in humans. Curr Opin Neurobiol 16:693–700.

Davachi L, Wagner AD (2002) Hippocampal contributions to episodic encoding: insights from relational and item-based learning. J Neurophysiol 88:982–990.

DeVito LM, Kanter BR, Eichenbaum H (2010a) The hippocampus contributes to memory expression during transitive inference in mice. Hippocampus 20:208–217.

DeVito LM, Lykken C, Kanter BR, Eichenbaum H (2010b) Prefrontal cortex: Role in acquisition of overlapping associations and transitive inference. Learn Mem 17:161–167.

Diana RA, Yonelinas AP, Ranganath C (2007) Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends Cogn Sci 11:379–386.

Dusek JA, Eichenbaum H (1997) The hippocampus and memory for orderly stimulus relations. Proc Natl Acad Sci U S A 94:7109–7114.

Efron B, Tibshirani RJ (1993) An introduction to bootstrap. New York.: Chapman and Hall.

Eichenbaum H (2000) A cortical-hippocampal system for declarative memory. Nat Rev Neurosci 1:41–50.

Eichenbaum H, Cohen NJ (2001) From conditioning to conscious recollection: memory systems of the brain. New York: Oxford UP.

Eichenbaum H, Schoenbaum G, Young B, Bunsey M (1996) Functional organization of the hippocampal memory system. Proc Natl Acad Sci U S A 93:13500–13507.

Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. Annu Rev Neurosci 30:123–152.

Eldridge LL, Knowlton BJ, Furmanski CS, Bookheimer SY, Engel SA (2000) Remembering episodes: a selective role for the hippocampus during retrieval. Nat Neurosci 3:1149–1152.

Giovanello KS, Schnyer DM, Verfaellie M (2004) A critical role for the anterior hippocampus in relational memory: evidence from an fMRI study comparing associative and item recognition. Hippocampus 14:5–8.

Greene AJ, Gross WL, Elsinger CL, Rao SM (2006) An fMRI analysis of the human hippocampus: inference, context, and task awareness. J Cogn Neurosci 18:1156–1173.

Hassabis D, Maguire EA (2007) Deconstructing episodic memory with construction. Trends Cogn Sci 11:299–306.

Hassabis D, Maguire EA (2009) The construction system of the brain. Philos Trans R Soc Lond B Biol Sci 364:1263–1271.

Hassabis D, Kumaran D, Vann SD, Maguire EA (2007) Patients with hippocampal amnesia cannot imagine new experiences. Proc Natl Acad Sci U S A 104:1726–1731.

Heckers S, Zalesak M, Weiss AP, Ditman T, Titone D (2004) Hippocampal activation during transitive inference in humans. Hippocampus 14:153–162.

Jackson O 3rd, Schacter DL (2004) Encoding activity in anterior medial temporal lobe supports subsequent associative recognition. Neuroimage 21:456–462.

Kao YC, Davis ES, Gabrieli JD (2005) Neural correlates of actual and predicted memory formation. Nat Neurosci 8:1776–1783.

Kirchhoff BA, Wagner AD, Maril A, Stern CE (2000) Prefrontal-temporal circuitry for episodic encoding and subsequent memory. J Neurosci 20:6173–6180.

Kirwan CB, Stark CE (2004) Medial temporal lobe activation during encoding and retrieval of novel face-name pairs. Hippocampus 14:919–930.

Köhler S, Danckert S, Gati JS, Menon RS (2005) Novelty responses to relational and non-relational information in the hippocampus and the para-hippocampal region: a comparison based on event-related fMRI. Hippocampus 15:763–774.

Kuhl BA, Shah AT, DuBrow S, Wagner AD (2010) Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. Nat Neurosci 13:501–506.

Kumaran D, Summerfield JJ, Hassabis D, Maguire EA (2009) Tracking the emergence of conceptual knowledge during human decision making. Neuron 63:889–901.

Murray LJ, Ranganath C (2007) The dorsolateral prefrontal cortex contributes to successful relational memory encoding. J Neurosci 27:5515–5522.

Okuda J, Fujii T, Ohtake H, Tsukiura T, Tanji K, Suzuki K, Kawashima R, Fukuda H, Itoh M, Yamadori A (2003) Thinking of the future and past: the roles of the frontal pole and the medial temporal lobes. Neuroimage 19:1369–1380.

O'Reilly RC, Rudy JW (2000) Computational principles of learning in the neocortex and hippocampus. Hippocampus 10:389–397.

O'Reilly RC, Rudy JW (2001) Conjunctive representations in learning and memory: principles of cortical and hippocampal function. Psychol Rev 108:311–345.

Oztekin I, Curtis CE, McElree B (2009) The medial temporal lobe and the left inferior prefrontal cortex jointly support interference resolution in verbal working memory. J Cogn Neurosci 21:1967–1979.

Paller KA, Wagner AD (2002) Observing the transformation of experience into memory. Trends Cogn Sci 6:93–102.

Preston AR, Gabrieli JD (2008) Dissociation between explicit memory and configural memory in the human medial temporal lobe. Cereb Cortex 18:2192–2207.

Preston AR, Shrager Y, Dudukovic NM, Gabrieli JD (2004) Hippocampal contribution to the novel use of relational information in declarative memory. Hippocampus 14:148–152.

Ranganath C (2010) Binding items and contexts: the cognitive neuroscience of episodic memory. Curr Dir Psychol Sci 19:131–137.

Schacter DL, Addis DR, Buckner RL (2008) Episodic simulation of future events: concepts, data, and applications. Ann N Y Acad Sci 1124:39–60.

Shohamy D, Wagner AD (2008) Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. Neuron 60:378–389.

Sperling R, Chua E, Cocchiarella A, Rand-Giovannetti E, Poldrack R, Schacter DL, Albert M (2003) Putting names to faces: successful encoding of associative memories activates the anterior hippocampal formation. Neuroimage 20:1400–1410.

Staresina BP, Davachi L (2009) Mind the gap: binding experiences across space and time in the human hippocampus. Neuron 63:267–276.

Stark CE, Squire LR (2001) When zero is not zero: the problem of ambiguous baseline conditions in fMRI. Proc Natl Acad Sci U S A 98:12760–12766.

Strange BA, Otten LJ, Josephs O, Rugg MD, Dolan RJ (2002) Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. J Neurosci 22:523–528.

Szpunar KK, Watson JM, McDermott KB (2007) Neural substrates of envisioning the future. Proc Natl Acad Sci U S A 104:642–647.

Wendelken C, Bunge SA (2010) Transitive inference: distinct contributions of rostrolateral prefrontal cortex and the hippocampus. J Cogn Neurosci 22:837–847.

Zeithamova D, Maddox WT, Schnyer DM (2008) Dissociable prototype learning systems: evidence from brain imaging and behavior. J Neurosci 28:13194–13201.