Spatial Reconstruction by Patients with Hippocampal Damage is Dominated by Relational Memory Errors

Patrick D. Watson,[1]* Joel L. Voss,[2] David E. Warren,[3] Daniel Tranel,[3,4] and Neal J. Cohen[1]

ABSTRACT: Hippocampal damage causes profound yet circumscribed memory impairment across diverse stimulus types and testing formats. Here, within a single test format involving a single class of stimuli, we identified different performance errors to better characterize the specifics of the underlying deficit. The task involved study and reconstruction of object arrays across brief retention intervals. The most striking feature of patients' with hippocampal damage performance was that they tended to reverse the relative positions of item pairs within arrays of any size, effectively "swapping" pairs of objects. These "swap errors" were the primary error type in amnesia, almost never occurred in healthy comparison participants, and actually contributed to poor performance on more traditional metrics (such as distance between studied and reconstructed location). Patients made swap errors even in trials involving only a single pair of objects. The selectivity and severity of this particular deficit creates serious challenges for theories of memory and hippocampus. © 2013 Wiley Periodicals, Inc.

KEY WORDS: binding; online processing; working memory; declarative memory; swap errors

INTRODUCTION

The precise role of the hippocampus in memory is a topic of much investigation. Observations of patients with amnesia following hippo-

campal damage reveal a complicated pattern of impaired and spared memory functions. The deficits include profound and pervasive impairment in learning and remembering new facts and events, preventing patients, for example, from normal learning of new routes, places, or people, and from keeping track of the appointments and events of daily life. However, other aspects of memory such as skill learning remain fully intact. Taken together with converging evidence using other neuroscience methods, the functional dissociations resulting from hippocampal damage illuminate the scope and limits of hippocampal involvement in memory (Cohen and Squire, 1980; Cohen and Eichenbaum, 1993; Schacter and Tulving, 1994; McClelland et al., 1995; Aggleton and Brown, 1999; Nadel et al., 2000; Eichenbaum and Cohen, 2001; Eichenbaum et al., 2007).

Some research has focused on the role of the hippocampus in processing spatial information and maintaining a dynamic, flexible "mental map" of space; highlighting deficits in numerous spatial tasks after hippocampal damage, along with evidence of place-sensitive cells in the hippocampus, and correlations between hippocampal volume and spatial ability (O'Keefe and Nadel, 1978; Hayes et al., 2004; Ryan et al., 2010). Another line of research points to the role of the hippocampus in managing declarative memory load—finding deficits following hippocampal damage when capacity limits are reached or when delays become sufficiently long (Stark and Squire 2003; Squire et al., 2004; Gold et al., 2006). Other research findings emphasize the nature of the representations generated by the hippocampus (Cohen and Eichenbaum, 1993; Henke, 2010). For example, one extensive body of work reports impairment following hippocampal damage for relational memory, showing deficits in representing the relationships among disparate elements of scenes or events (Eichenbaum and Cohen, 2001) or in representing cross-modal bindings (Marr, 1971; Damasio, 1989; Vargha-Khadem et al., 1997; Aggleton and Brown, 1999). Such deficits are manifested for all manner of accidental or arbitrary relations (Konkel et al., 2008; Konkel and Cohen, 2009), regardless of the timescale over which the relational information must be maintained (Hannula et al., 2006; Warren et al., 2012).

Across all of these different lines of research, hippocampal damage is seen to produce memory impair-

ment—and hence hippocampus is clearly engaged—in many different categories of stimuli and test formats. This highlights the broad scope and pervasiveness of hippocampal function in memory, but also makes identifying the critical factors(s) that tie these findings together challenging. What is the fundamental nature of the deficit, and hence the role of the hippocampus in memory?

Even within a single stimulus domain and test format, memory impairment following hippocampal damage can be difficult to interpret unambiguously. Deficits in learning to navigate among multiple locations in large spatial environments could be attributed to spatial, load/capacity, relational, or other demands. In the current experiment, we employed a simple memory test complemented by a set of performance analyses rich enough to identify various categories of errors arising from the different predicted deficits, permitting a more direct evaluation of various predictions within a single experimental paradigm.

Our goal was to determine whether hippocampal damage causes errors even in short-delay spatial reconstruction, and whether specific types of errors occur disproportionately, in a fashion that would be helpful in assessing the role of hippocampus in various types or aspects of memory.

MATERIALS AND METHODS

Participants

Behavioral data were collected from three individuals with amnesia subsequent to hippocampal damage and from four comparison participants with no known neurological impairments. Each comparison participant was matched to an amnesic participant in age (within 2 yr), educational attainment (within 1 yr), sex, and handedness. Table 1 summarizes each amnesic patient's etiology along with demographic, neuropsychological, and hippocampal volumetric measures where available.

Amnesic Etiology and Neuroanatomy

All amnesic participants suffered acute episodes in adulthood that rendered them memory impaired, and previous reports have established that each amnesic participant has substantial damage to the hippocampus bilaterally. Patient 2363 became amnesic after cardiac arrest and an accompanying anoxic episode that resulted in selective regional atrophy without lesion. The bilateral volume of his hippocampus has been measured and quantified, and found to be significantly less than normal for his age and sex based on a regression model fit to hippocampal volumes of healthy comparison participants (Allen et al., 2006), with a Studentized residual value of -2.64. 2636's cerebral gray matter has been characterized as less than normative (Allen et al., 2006) with a Studentized residual value of -2.47, which was driven in large part by a normatively small amount of parietal gray matter (Studentized residual value of -2.78). Gray matter volume in the frontal and temporal lobes was less than normative but unremarkable. Patient 1846 became amnesic after a combined, hour-long episode of status epilepticus and anoxia that resulted in selective regional atrophy without lesion. Her hippocampus is atrophied bilaterally, with the atrophy being greater on the left (Warren et al., 2012). Her bilateral hippocampal volume has also been measured and quantified, and found to be significantly less than normal, with a Studentized residual value of -4.23 (Allen et al., 2006). Outside of the MTL, 1846's brain has been described as normal except for "some evidence of cortical thinning in the paracentral lobule and precuneus" (Warren et al., 2012) that may be related to her anoxic etiology. Otherwise her brain volume (gray and white matter, both total and per-lobe) has been characterized as within the normal range (Allen et al, 2006). Patient 2308 became amnesic after an episode of herpes simplex encephalitis (HSE) that damaged significant portions of his left and right temporal lobes. Specifically, 2308 has bilateral damage to the medial temporal lobe (including the amygdalae and the anterior hippocampi in their entirety) and medial temporal poles along with unilateral damage to left ventral and lateral temporal lobe extending to the left temporal pole (Cavaco et al., 2012). The hippocampal lesions in 2308 are so extensive that it is not possible to measure meaningfully the remaining tissue and make a quantitative comparison to normative data. Beyond the temporal lobes, 2308 has left-lateralized damage to the insular cortex, basal forebrain, and the posterior portion of orbitofrontal cortex, and right-lateralized damage to the insular cortex.

TABLE 1.

Descriptive Data for Patients Including Age (Yr), Education (Yr), Sex, Hippocampal Volume (see below), WAIS-III Full-Scale IQ (FSIQ), WMS-III General Memory Index (GMI), and WMS-III Working Memory Index (WMI)

Patient	Age	Edu.	Sex	Hipp.	Etiology	WAIS-III FSIQ	WMS-III GMI	WMS-III WMI
2,308	55	16	M	[a]	HSE	98	45	85
2,363	55	16	M	-2.6	Anoxia	98	73	91
1,846	48	14	F	-4.2	Anoxia	84	57	88

Hipp.: Residual hippocampal volume (Studentized residual value of volume, see Allen et al., 2006 for details) and etiology (HSE = Herpes Simplex Encephalitis).
[a]The hippocampus is damaged bilaterally in patient 2308, and the lesion extends to other medial and lateral temporal structures, most notably in the left hemisphere.

Hippocampus

Amnesic Neuropsychology

Neuropsychological examination confirmed severe declarative memory impairment in each amnesic participant, with performance on the Wechsler Memory Scale—Third edition (WMS-III, Wechsler, 1997a) at least 25 points lower than their performance on the Wechsler Adult Intelligence Scale—Third edition (WAIS-III, Wechsler, 1997b), and the average delay score on the WMS-III more than two standard deviations below the population mean. Memory impairments were selective, in that, for example, none of the amnesic participants showed any systematic impairment on a battery of neuropsychological tests of executive function, including trail making, Wisconsin card-sorting, controlled oral word associations, and the tower of London (Konkel et al., 2008).

Experimental Paradigm

We used a spatial reconstruction task (Huttenlocher and Presson 1979; Smith and Milner, 1981; Jeneson et al., 2010). During each trial the participant studied an object array (containing between two and five objects), arranged on a 100 cm-by-100 cm white tabletop, and then reconstructed the spatial layout of the objects after a brief eyes-closed delay. During the "study" portion of the trial, the participant picked up each object with his or her dominant hand, named it, and immediately placed it back in the same location. When the participant had finished, he or she covered his or her eyes for approximately 4 s while the location of the objects was recorded in a digital photograph and the objects were cleared from the table. After the 4 s "blind" period, the participant attempted to place the objects back into the original configuration (reconstruction). The final location of the objects was then recorded in a second digital photograph, and the next trial began after a short break. Some trials involved familiar, nameable objects (e.g., a pen, a button, a toy car, etc.) and others involved novel objects carved out of white foam blocks into various complex shapes and covered in patterns of simple lines and other shapes ("Greebles" c.f. James et al., 2005). All of the materials composing the novel shapes were of the same composition and color for each stimulus, such that stimuli could not be distinguished based on simple features. Blocks involving novel objects were interleaved with familiar-object blocks, with an equal number of each block type in each experimental session. During the study portion of each trial, participants picked up each object (as for the trials involving familiar objects), but counted integers aloud instead of providing names, given the obvious difficulties that would be associated with attempting to name these novel objects.

Another condition was also included that varied from the main paradigm in the instructions given to the participants. In this condition, participants were instructed to create the initial configuration of the objects (whether novel or familiar) on each trial themselves, rather than studying locations selected by the experimenter (i.e., objects were self-placed rather than experimenter-placed). This condition was administered in blocks randomly interposed with the main experimental blocks described here. Data from this condition are not reported here because amnesic participants self-positioned objects in grossly different patterns than did comparison participants, thus confounding comparisons of subsequent relational memory performance.

The digital photographs taken after the study and reconstruction portions of each trial were analyzed offline using MATLAB software (MATLAB version 7.9 Natick, Massachusetts: The MathWorks Inc., 2009). The edges of the table were identified via a semiautomated algorithm and were used to warp the coordinate space of the table into a common, Cartesian coordinate system via linear deformation. There was no more than 1 cm of displacement in the position of the table edges for the reconstruction image relative to the study image for any trial, indicating that table and camera movements did not contribute significantly to measures of reconstruction errors. The location of the center of each object was marked prior to deformation and object coordinates in the common reference frame were used for analysis.

Memory Measures

This task permitted multiple error types, and assessed reconstruction performance using five metrics capable of capturing this error heterogeneity. Because this task involved spatial reconstruction performance could be evaluated with respect to spatial theories. Because the task involved variable set sizes, including as few as two objects, performance could be evaluated for the effects of memory load. Finally, our use of unique error analyses permitted a rich evaluation of performance with respect to the different types of representation required.

Different object-configuration schemes are sensitive to different types of reconstruction errors. Spatial reconstruction experiments have historically used the "item misplacement" measure, which is simply the distance (in cm) between each item's studied location and the location where each item was placed during reconstruction (Huttenlocher and Presson, 1979, Smith and Milner, 1981; Jeneson et al., 2010). Although it is a simple, intuitively appealing analytic approach, it assumes that the underlying representation is of each item's location in a grid-like Cartesian coordinate system. Although some theories of hippocampal spatial processing might endorse such a map-like representational scheme, it is not clear that all do (e.g., some spatial theories might argue that the representation is more like a path than a map). Moreover, such an item-based approach does not take into account the possibility that performance might be driven by memory representations of the configuration of the objects or the relations among the objects. Accordingly, we also measured spatial reconstruction using "edge resizing" and "edge deflection" metrics, which measure reconstructed changes in the length and direction (in cm and radians respectively) of vectors between each pair of items. These metrics assume that each items' location in the underlying representation serves as a landmark for each other item's location, with polar coordinate-like vectors between them. Moreover, we also measured memory for the overall arrangement of items

with a rearrangement metric. This metric assumes that the underlying representation has no fine-grained representation of distance or angle, but rather reduces the configuration of studied items to simple shape via perceptual closure and measures the frequency of categorical changes in shape (e.g., a square changing to a rhombus, or a line changing its orientation). This array of measures applied to amnesic performance allowed us to better characterize the hippocampal contribution to spatial representation.

One other new metric assessed is swaps, which is the rate at which any pair of objects "swap" places between study and reconstruction (i.e., when the correct locations were filled but with mis-assignment of particular objects to particular locations). This metric was applied for each possible pair of objects in a given reconstruction of two to five objects. Because the number of possible pairwise swaps increases combinatorially while set size increases linearly, our metric here was swaps-per-pairwise-relation, thereby avoiding confounding the increase in relational complexity with the increase in the number of items. We measured such swaps by counting the frequency that the vector connecting each pair of objects reversed direction (i.e., the sign of the vector's x and y components changed simultaneously between study and reconstruction). This metric assumes an underlying representation that involves binding each trial's set of object-identities onto each trial's set of locations. The experimenters' assignment of particular objects to particular locations was random. Thus, successful performance required memory for arbitrary relations, and the incidence of swaps in patients with hippocampal damage could be used to assess the role of hippocampus in relational memory.

Patient to Comparison ratios

Patient to comparison ratios were simply calculated as patient performance over healthy comparison performance on each metric. Standard error bars were obtained via propagation of uncertainty: For

$$f = \frac{a}{b}$$

$$(\sigma_f)^2 = \left(\frac{\sigma_a}{a}\right)^2 + \left(\frac{\sigma_b}{b}\right)^2$$

where a corresponds to patient performance, b corresponds to healthy comparison performance, and a and b are independent and uncorrelated.

Random Performance

Random performance was calculated by assuming a pair of objects was placed at two random locations (indexed by random, unique pairs of x- and y-coordinates drawn from our 100×100 spatial grid) during the study phase, and that these two objects were placed at two random locations during the test phase. The mean of the resulting misplacement, edge resizing, edge deflection, rearrangment, and swaps were calculated by applying our measures above. For any pair of objects placed

at random, mean item misplacement ought to be 52 cm, mean edge resizing 28 cm, mean edge deflection $\pi/2$ radians, and the expected probability for rearrangements and swaps 50 and 25% respectively.

Note that this definition of random performance does not take into account biases present in our experimenter or participants. For instance, during the study phase no object was positioned less than 10 cm from the outer edge of the table, meaning that the utilized area of the table was closer to 90×90. However, since making use of this information would imply some level of memory on the part of our participants, we chose to use the less constrained definition of randomness above.

RESULTS

What Kinds of Errors Do Amnesic Participants Make?

We first consider reconstruction performance for displays that contained familiar, everyday objects (Fig. 1).

Memory Load

Relative to healthy comparison participants, at all set sizes, and on all metrics, amnesic participants were impaired in reconstructing object configurations after an approximately 4 s delay. ANOVAs with factors Group (amnesic/comparison) and Set Size (2, 3, 4, and 5) were conducted on each measure. Although every measure yielded a main effect of group (all $P < 0.01$, see below), only rearrangements yielded a main effect of set size ($F(3,6) = 4.11$ $P < 0.01$). Thus, after accounting for relational complexity, we found no additional effect of memory load on patient performance on any measure except rearrangements.

Spatial Measures

We report mean performance on the item misplacement metric, collapsed across the number of objects that were studied (2, 3, 4, and 5), as well as the mean performance for each object set size (Fig. 1a). A mixed 2-by-4 ANOVA with factors of group and set size with yielded a main effect of Group ($F(1,6) = 137.74$, $P < 0.0001$), indicating reliably poorer placement among amnesics for all set sizes. The visual trend for more misplacement with increasing set sizes did not reach significance ($F(3,6) = 1.47$, $P > 0.22$), nor was there an interaction between the two factors ($F(3,6) = 0.47$, $P > 0.69$). On the edge resizing (Fig. 1b) and edge deflection (Fig. 1c) metrics that assess the reconstruction of the magnitude and direction of vectors between object pairs, we found similar effects. For edge resizing, a mixed 2-by-4 ANOVA with group and set size showed a significant main effect of group ($F(1,6) = 48.45$, $P < 0.0001$) but there was no significant effect of set size ($F(3,6) = 1.49$, $P > 0.21$), nor was there an interaction between the two factors ($F(3,6) = 2.05$, $P > 0.10$). Likewise
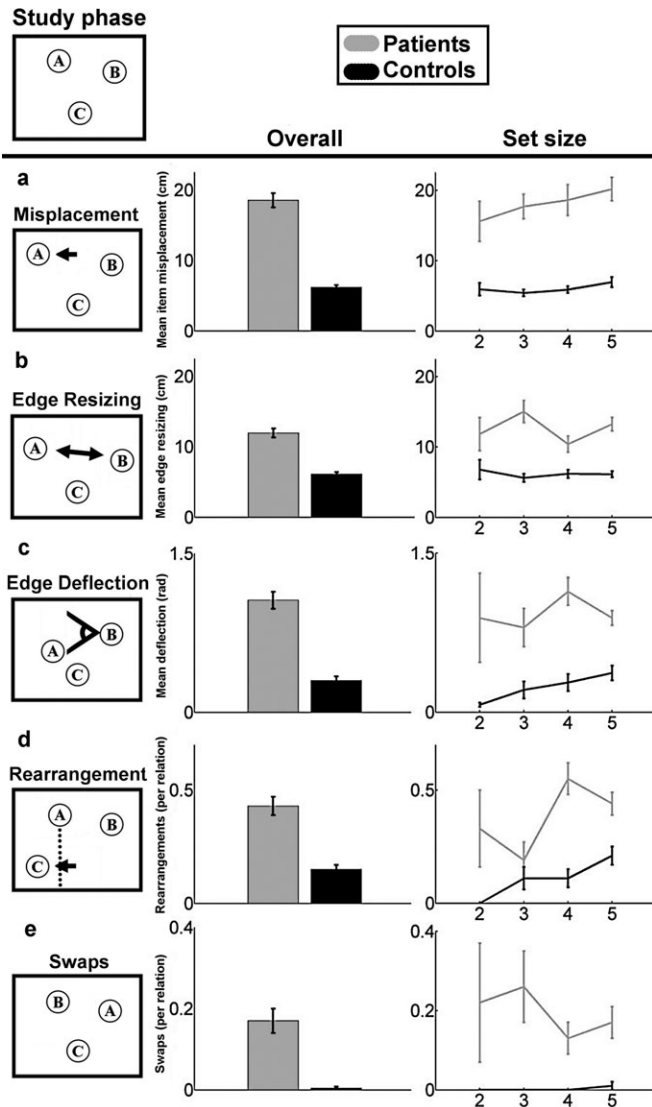
FIGURE 1. Patient and comparison participant performance quantified using five metrics of error quality. An example study configuration is provided at the top of the figure. (A–E) Each of the five reconstruction error types is demonstrated, with overall error (left, collapsed across the number of item in the study configuration) and error as a function of studies object set size (right) is provided for each error type. Error bars indicate SE.

for edge deflection, there was a significant main effect of group ($F(1,6) = 36.66$, $P < 0.0001$), a nonsignificant main effect of set size ($F(3,6) = 1.64$, $P > 0.17$), and nonsignificant interaction ($F(3,6) = 1.3$, $P > 0.27$). On the rearrangements metric (Fig. 1d), measuring performance at reconstructing an undistorted overall shape, showed a significant main effect of group ($F(1,6) = 22.53$, $P < 0.0001$), a significant main effect of set size ($F(3,6) = 4.11$, $P < 0.01$), and a significant interaction between the two factors ($F(3,6) = 3.43$, $P < 0.02$). Thus, amnesics were impaired overall, while both amnesics and comparisons made more rearrangement errors as set size increased, with this trend being disproportionately greater for amnesics.

These findings replicate previous reports of worse performance for amnesics versus comparison participants using item misplacement measures of reconstruction (Smith and Milner, 1981, Jeneson et al., 2010). In addition, they demonstrate that amnesic participants are also impaired relative to comparisons at reconstructing the positions of objects relative to each other and are also less likely than comparisons to reconstruct an accurate version of the general shape they observed during the study phase.

Relational Measure

Errors made in reconstructing the relative positions of objects, resulting in objects "swapping" positions, are summarized (Fig. 1e). The swaps measure showed a significant main effect of group ($F(1,6) = 29.78$, $P < 0.0001$), but no significant main effect of set size ($F(3,6) = 0.96$, $P > 0.4$) and no significant interaction between the two factors ($F(3,6) = 1.01$, $P > 0.3$). Thus patients were no less likely to swap a pair of items that appeared alone (set size 2) than they were to swap any pair of items that appeared in set sizes of three, four, or five. This finding shows that hippocampal amnesics are also more likely than comparisons to fail at binding item identities to locations.

Do Any Metrics Show Disproportionate Impairment?

All five measures were highly intercorrelated, with R^2 between 0.95 and 0.69 (all $P < 0.02$, Table 2). This was expected since some reconstruction errors might lead to high values on more than one metric (e.g., a swapped item will also be misplaced relative to its original location). To identify if any of these error types disproportionately impaired to overall performance we examined the relative proportion of reconstruction errors committed by amnesic versus comparison participants. Relative to the other metrics, amnesics produced a strikingly disproportionate rate of errors for the swap metric (Fig. 2). For the other four metrics, amnesics performed between two and four times worse than comparisons. However, on the swap metric, amnesics were more than 40 times worse than comparisons. Amnesics made a swap error on 17% of pairwise relations (31 swaps in 182 pairs) whereas comparisons made the error on only 0.4% (one swap in 242 pairs) of pairwise relations. There was only a single swap error made by any of the comparisons in any of the familiar-object conditions. By comparison, all three amnesic participants made numerous swap errors in the familiar object condition, with patient 2363 making an average of 0.17 swap errors per relation (10 swaps in 60 pairs), patient 1846 making 0.14 (11 swaps in 80 pairs), and patient 2308 making 0.25 (10 swaps in 40 pairs). For trials in which amnesic participants committed swap errors, 66% involved one swap error, 24% involved two errors, 5% involved three errors, and 5% involved four errors. Thus, more than one swap error in a single trial was a frequent occurrence for amnesics (34%), while never occurring in the performance of any of the comparison participants (0%). Additionally, when adjusted for number of

TABLE 2.

Cross Correlations Between Each of the Five Measures

Metrics R (P)	Misplacement	Edge resizing	Edge deflection	Rearrangement	Swaps
Misplacement	1 (1)				
Edge resizing	0.90 (<0.01)	1 (1)			
Edge Deflection	0.94 (<0.01)	0.81 (<0.01)	1 (1)		
Rearrangement	0.90 (<0.01)	0.95 (<0.01)	0.85 (<0.01)	1 (1)	
Swaps	0.89 (<0.01)	0.71 (<0.01)	0.79 (<0.01)	0.70 (<0.01)	1 (1)
Mean R	0.91	0.84	0.85	0.85	0.77

All five measures are highly intercorrelated (R between 0.6984 and 0.9490), but the swaps metric was least well correlated with the other measures.

pairwise relations amnesic participants made swap errors at an approximately equal rate for all set sizes while comparisons made the error only once and only in a five-item set. Swap errors were thus an essentially unique identifier of amnesic participants.

Are Swap Errors the Primary Deficit in Amnesia?

Given the strikingly disproportionate prevalence of swap errors relative to the other error types, we next asked whether swap errors constitute the primary deficit in amnesia. In other words, what is the causal relation between swap errors and errors on the other metrics?

One possible explanation for the high incidence of swaps errors made by amnesic participants is simply that they made large misplacement errors. That is, what appeared to be swap errors actually could have resulted from misplacement errors wherein item locations were reconstructed so inaccurately that items were actually placed in another object's studied location. Thus, we examined the likelihood of swaps given the study-time distance between pairs of items, and the magnitude of patients' misplacement errors. We tested this with a Monte Carlo simulation that utilized the item misplacement values and the study-time interitem distances collected from patients in our observed data. The simulation randomly drew a pair of misplacements and an interitem distance each iteration, controlling for set size. This had the effect of creating a new set of data in which pairs of objects moved randomly according to the distribution of the actual misplacement data. Each run of the simulation produced a number of data points equal to those observed in the real experiment, and we calculated the mean number of swaps present in the simulated data over 1,000 runs. For each run of the simulation, we performed a one-way ANOVA on the observed versus the simulated data, and here report the mean incidence of swaps, and mean p values produced by this simulation (Fig. 3). The Monte Carlo simulation showed that based on item misplacements alone, patients should have made only 0.045 (SD = 0.015) swaps per pairwise relation, far less than the 0.17 swaps per relation actually observed. This meant that on average, the Monte-Carlo simulation produced approximately 8 (SD = 2.73) swaps, while the empirical data contained 31. This difference was significant: Given the observed level of misplacement error, the
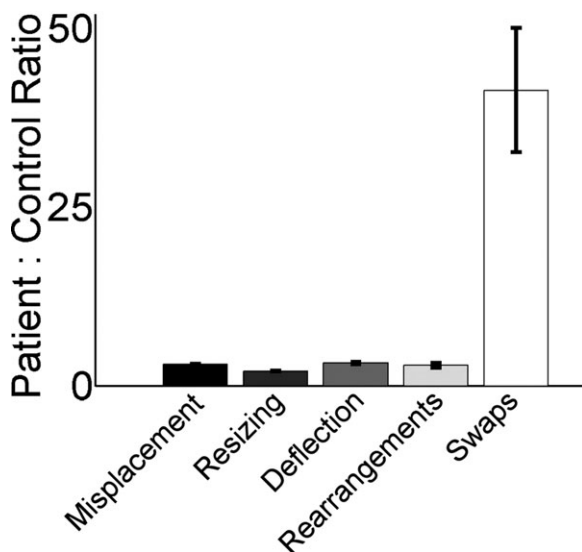

FIGURE 2.   Disproportionately high swap errors in patients. The ratios of mean performance to mean comparison performance are provided for each of the five performance metrics. Error bars indicate SE, calculated by error propagation.
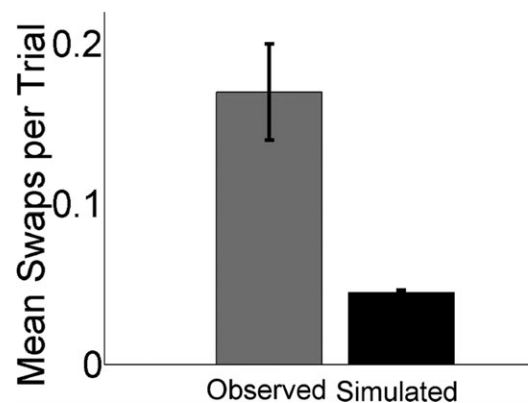

FIGURE 3.   Swap errors in patients were more numerous than would be expected by chance. Mean actual swap errors per trial made by patients are plotted along with the number of errors expected based on chance given actual misplacement error, as determined by Monte-Carlo simulation.

mean probability of observing the number of swaps in the actual data due to item misplacement alone was 0.007. Therefore, overall poor spatial positioning of items individually was not the cause of the high level of swap errors actually observed (i.e., item misplacement was not primary to swap errors).

Do Swap Errors Contribute to Amnesics' Poor Performance on Item Misplacement?

We next tested the opposite direction of causality (i.e., determining whether the high incidence of swap errors might be contributing to the overall poor spatial positioning performance of amnesic participants). We recalculated item misplacement after removing the error values introduced by swaps (i.e., calculating item misplacement only for unswapped items). If swap errors were primary to item-placement errors, we would expect a significant reduction in simple spatial errors after removing the effects of swaps. This was confirmed (Fig. 4). Although amnesic participants still performed worse than comparisons after removal of item misplacement due to swaps (main effect of group, $F(1,3) = 5.15$, $P < 0.03$), removing swaps led to a significant reduction in item misplacement [one way ANOVA with swaps-present vs. swaps-removed, $F(1,3) = 31.49$, $P < 0.0001$]. This suggests that the amnesic participants' deficits on the standard item-misplacement measure can be at least partially attributed to their poor performance on the swap metric. We also used an "unswapping" algorithm (see Supporting Information), the results of which converge with the above analysis showing that improvements in item misplacement are not simply due to discarding poor trials.

Do Patients Simply Perform Randomly?

It is illustrative to compare participants' performance across the measures to an objective benchmark: random performance.
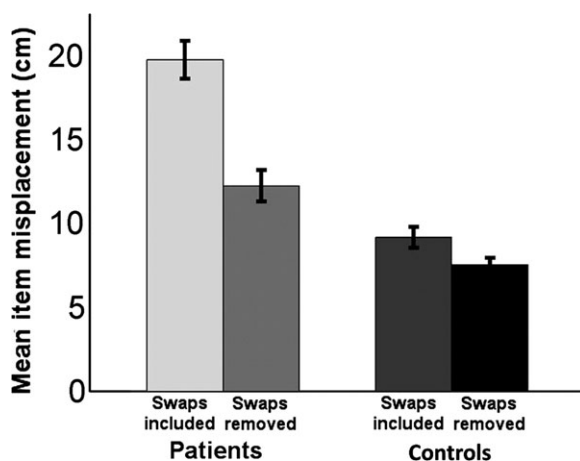

FIGURE 4.   Swap errors were partially responsible for patient misplacement errors. Misplacement error values for patients are shown averaged for all objects in each trail as well as only for the objects for which swap errors did not occur. Removing swap errors in this manner led to a significant reduction in misplacement error for patients. Error bars indicate SE.

Supposing neither patients nor comparisons were allowed to see the study phase, but were still able to place objects in random positions. Their performance could only be randomly related to the experimenter-positioned objects (based on the premise that both the experimenter and participants are equally likely to place any object at any coordinate on a 100×100 cm^2 grid, see Methods).

On all measures except swaps, both patients and comparisons perform far above chance across all set sizes (the 95% confidence interval does not include chance), demonstrating that, in an information processing sense, they possessed some useful information about the original configuration. However on swaps, comparisons' performance for familiar objects was nearly perfect: the 95% confidence interval included 0%. For amnesics, swap performance was nearly at chance: the 95% confidence interval included 25%. This quantitative difference cannot be much larger. For reconstructing the object-identity-to-vertex-position bindings, comparisons behaved as if they had nearly perfect information, while patients behaved as if they had nearly no information about such bindings.

Do Novel-Object Displays Have a Disproportionate Impact on Patients?

The same reconstruction was also performed using a set of 14 novel objects, composed of white foam blocks carved into various complex shapes and covered in patterns of simple lines and other shapes (i.e., "Greebles" James et al., 2005).

We performed the same series of analyses (i.e., misplacement, edge resizing, edge deflection, rearrangements, and swaps) of reconstruction performance for arrays composed of novel objects (Fig. 5), as well as an ANOVA for each measure with a factor of item type (familiar vs. novel). Here, we describe the main findings averaged across all set sizes to facilitate comparisons with the effects identified using familiar objects.

Performance was significantly worse for novel objects relative to familiar objects for three of the five metrics, as indicated by significant main effects of the object-type factor (for the item-misplacement, edge-deflection, and rearrangement metrics, ($F(1,6) = 7.48$, 6.6, and 4.932, respectively, Ps < 0.01, 0.02, and 0.04), although not for the edge-resizing [$F(1,6) = 0.3$, $P = 0.60$) and swap ($F(1,6) = 3.25$, $P = 0.72$) metrics]. As was the case for familiar objects, for novel objects amnesic participants showed worse performance than comparison participants for all five metrics, as indicated by significant main effects of group (for the item-misplacement, edge-resizing, edge-deflection, rearrangement, and swap metrics, [$F(1,6) = 220.2$, 136.7, 129.9, 60.3, and 84.2, respectively, all Ps < 0.001]. However, there was no evidence that novel objects impaired performance to a greater extent in amnesic than in comparison participants for any of the metrics; the interactions of object-type by group were all nonsignificant (Ps between 0.15 and 0.65).

As for familiar objects, a striking characteristic of amnesic participants' performance with novel objects was the highly disproportionate incidence of swap errors compared to all other
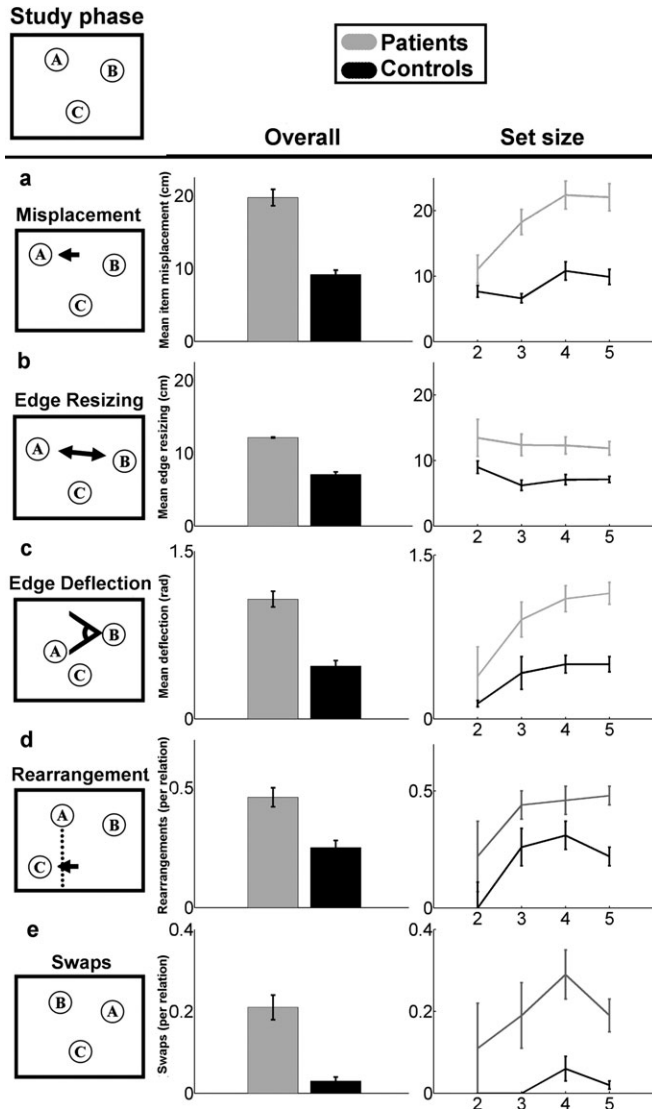
FIGURE 5. Patients were also impaired for all metrics in novel-object arrays. Once again, an example study configuration is provided at the top of the figure (A–E). Each of the five reconstruction error types is demonstrated, and overall error (collapsed across the number of items in the study configuration) and error as a function of studied object set size is provided for each type. Error bars indicate SE.

error types. Amnesics made a swap error on 21.4% of the opportunities they had to do so (39 swaps in 182 pairs), whereas comparisons made the error on only 3% of their opportunities (seven swaps in 249 pairs). All three amnesic participants made numerous swap errors, with patient 2363 making an average of 0.18 swap errors per pairwise relation (11 swaps in 60 pairs), patient 1846 making 0.28 (22 swaps in 80 pairs), and patient 2308 making 0.15 (six swaps in 40 pairs). For trials on which any swap error occurred, amnesics frequently made multiple swaps: 54% involved one swap error, but 19% involved two errors, 11% involved three errors, 5% involved four errors, and 11% involved five errors. By contrast, multiple swaps were much less frequent in comparison partici-
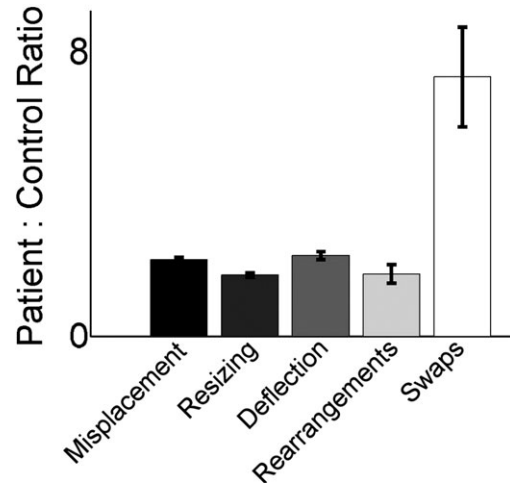
FIGURE 6. Disproportionately high swaps in patients in the novel object condition despite nonceiling comparison performance. The ratios of mean patient performance to mean comparison performance are provided for each of the five performance metrics. Error bars indicate SE calculated by error propagation.

pants: 85% involved only one swap error and the other 15% involved only two. As was the case for familiar objects, amnesics made swap errors for all set sizes whereas comparisons made these errors only for four- and five-item sets.

In addition, since comparison participants were no longer at "ceiling" on the swap metric, we once again examined patient to comparison relative performance across the five metrics (Fig. 6). Once again, the swap measure showed the most disproportionate deficit, with patients performing 7.2 times worse than comparisons, with relative performance on the other four measures between 1.7 and 2.3.

Finally, Monte Carlo simulations showed the same direction of causality for the various error types for novel objects as was observed with familiar objects. The prevalence of swap errors was significantly higher than would have been expected by chance if they were due entirely to pure item-misplacement error (mean simulated swaps per relation = 0.05, mean $P < 0.01$). Furthermore, removal of all item-misplacement error due to swap error led to a significant reduction in item-misplacement error for both groups [$F(1,6) = 216.31$, $P < 0.001$], and a disproportionately greater reduction in error for amnesics. Thus, for novel objects as for familiar ones, swap errors cause an over-estimate of the memory deficits suggested by item-misplacement errors.

DISCUSSION

Individuals with hippocampal amnesia displayed impaired performance in reconstruction of spatial locations of small arrays of objects over a short delay interval. Impairments were present both in the standard measure in such paradigms, involving the degree of item misplacements (mean distance between objects' position at study versus reconstructed position

as placed by the participant at test), and for all other metrics we used to examine aspects of the memory representations needed to support reconstruction of object locations. Strikingly, the observed deficit was markedly disproportionate for errors involving object-for-object swapping, which evaluated object-identity-to-relative-location bindings. For arrays of familiar objects, amnesic participants committed swap errors at a rate more than 40 times that of comparison participants, who almost never committed this kind of error; making swaps nearly a unique identifier of amnesia in our sample.

The high incidence of swap errors in amnesia was shown via simulation analysis not to arise from larger-than-normal item misplacement. Instead, the causal relationship between these error types was in the other direction. Findings showed that removal of swap errors from the analysis led to a significant reduction in the estimates of item misplacement error, suggesting that a significant proportion of the overall poor performance resulted from the inability to track object-identity-to-relative-location bindings. Notably, the prevalence of swap errors in the performance of participants with hippocampal amnesia was seen for two independent stimulus categories (familiar objects and novel objects), and held across all set sizes.

One possible explanation of the pattern of performance across the various metrics is that both patients and comparisons were able to represent the object arrays as simple "shapes" formed by perceptual closure (with each object corresponding to a vertex of the shape, c.f., Uttal and Chiong, 2004), and/or as a motoric sequence indexing each location within the array (as in the Corsi block tapping task, which is partially spared in patients with hippocampal damage, c.f. Corsi 1972, Kessels et al., 2000). However, whatever spared representation underlies patient performance; it seems to lack the cross-domain binding information about which items occupy which spatial indices. Thus, comparison participants were highly successful both at reconstructing the array outline and at placing each object at the specific vertex position at which it was studied, as demonstrated by their relatively successful reconstruction performance measured using all metrics. Amnesic participants were somewhat less successful at reconstructing the geometry of object arrays (as indicated by their deficits in our first four edge metrics), but showed strikingly disproportionate impairment in representing the arbitrary object-to-vertex mappings required to replace the correct objects in their specific vertex positions, instead swapping object-to-vertex relations. These swap errors were nearly diagnostic of hippocampal amnesia; in our sample if a swap error was observed on a familiar object trial, it was 97% likely that a patient was positioning the objects.

This is best illustrated by one especially striking feature of patient performance: the presence of swap errors by patients on two-object trials. Even in our task's simplest condition, with only single binding between a single pair of familiar objects, requiring maintenance for a few seconds, patients still reversed the positions of the two objects approximately once every five opportunities—a rate of swapping similar to that which would be produced without any knowledge of the proper arrangement of objects. Intuitively, and as observed in the performance of healthy comparisons, errors of this kind should be vanishingly rare in neurologically intact participants.

The deficits observed here were for brief retention intervals and short lags traditionally associated with working memory. This is consistent with other findings of relational memory deficits in amnesia at short retention intervals (Ryan and Cohen, 2004; Hannula et al., 2006, 2007). and also with recent evidence that the human hippocampus is essential for the expression of memory even with no interposed study-test delay (e.g., when memory is used online to guide exploration behavior, as in Voss et al., 2011a,b, or to assemble and maintain complex representations as in Warren et al., 2012). It also converges with fMRI findings of hippocampal activation for relational memory over the same very short timescale (Mitchell et al., 2000; Piekema et al., 2006; Hannula and Ranganath, 2009) as well as with imaging data implicating the hippocampus more generally on the short-term/working memory timescale (Ranganath and D'Esposito, 2001; Stern et al., 2001; Ranganath and Blumenfeld, 2005).

One of our goals in using a simple but open-ended test complemented by a suite of performance metrics was to test different theoretical accounts of hippocampal deficits. Our analysis provides the strongest support for theories that emphasize arbitrary relational bindings as the primary hippocampal representation (indexed by our swaps metric). Because removal of swap errors did not entirely ameliorate patients' reconstruction deficits, our analysis also provides partial support for theories that emphasize geometric, spatial, hippocampal representations (at least where these spatial representations correspond to simple Cartesian coordinate maps, the vectors and landmarks of a polar coordinate representation, or unitized, shapes formed by perceptual/motor processes), as indexed by our misplacement, edge resizing, edge deflection, and rearrangement metrics. However, we were able to explain the preponderance of these spatial deficits as the secondary consequences of swap errors. Although the remaining deficit does partly support spatial theories of hippocampal function, it may also be a consequence of our pair-wise swap measure which counts only two-dimensional rotations of inter-item vectors in a continuous, Cartesian space. Swap measures that take into account multi-item swaps (e.g., rotation of trios of objects), categorical swaps (e.g., swapping the left and right halves of a figure), or nonspatial swaps (e.g., perseveration of a previously reconstructed shape) could perhaps account for some additional portion of the deficit and are an intriguing avenue of future study. Finally, we found little evidence that hippocampal amnesia was best explained by a deficit in transferring information from a limited capacity working-memory system to long-term memory (indeed patients made swap errors over short delays even in the two object condition, and the error rate did not increase for larger set sizes after accounting for the increase in relational complexity). We would explain findings that amnesics make disproportionally greater misplacement errors on arrays with large numbers of items (Jeneson et al., 2010) as a natural consequence of linear growth in item counts producing combinatorially more opportunities to commit swap errors.

However, our goal was not to adjudicate competing theories, but to identify the primary memory deficit resulting from hippocampal damage. As the experiment examines only spatial reconstruction, it is impossible to infer if swap errors are a consequence of a deficit specific to item-identity-to-location bindings (c.f., Lee et al., 2005; Hartley et al., 2007), or if they arise from a deficit in a more domain-general binding system. Our findings are compatible with any theory that proposes that the hippocampus is critical to performance that relies upon flexible, reconfigurable bindings that index locations, and that it is the disruption of such bindings that causes generally poor spatial performance. Our assertion that these findings more strongly support a representation scheme based upon arbitrary relational binding than a scheme that emphasizes spatial relations arises from the fact that spatial representations often carry connotations of geometric properties such as coordinates, distances, angles, and shapes which, according to the measures reported here, were less disrupted by hippocampal damage than item-identity-to-location-bindings. Four of the metrics (edge resizing, edge deflection, rearrangement, and swaps) are spatial relational measures, but only the relations supporting performance on the swap metric are disproportionately impaired by hippocampal damage. The spatial reconstruction paradigm we used provides rich behavioral records, and avoids many confounds between relational and item memory (e.g., test format, test difficulty, retention interval, etc.). We showed that while the hippocampus is certainly involved in spatial representations of all types, there is one kind of representation (which binds item identities to their relative locations) for which an intact hippocampus is the difference between chance level performance and perfect performance. Furthermore, this binding deficit is primary to more traditionally measured spatial reconstruction impairments. The precision of this result supports a similarly precise theoretical account that emphasizes that general memory impairments result from specific binding deficits. Rather than simply highlighting tasks which are impaired by hippocampal damage, theoretical accounts should be able to explain why a measure uniquely sensitive to binding errors (e.g., our swap metric) is most indicative of hippocampal impairment, and why representations of arbitrarily assigned, reconfigurable, item-identity-to-relative-location-bindings so critically depend on hippocampal function.

REFERENCES

Aggleton JP, Brown MW. 1999. Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. Behav Brain Sci 22:425–44; discussion 444–89.

Allen JS, Tranel D, Bruss J, Damasio H. 2006. Correlations between regional brain volumes and memory performance in anoxia. J Clin Exp Neuropsychol 28:457–476.

Cavaco S, Feinstein JS, Van Twillert H, Tranel D. 2012. Musical memory in a patient with severe anterograde amnesia. J Clin Exp Neuropsychol 34:1089–100; doi:10.1080/13803395.2012.728568.

Cohen N, Eichenbaum H. 1993. Memory, Amnesia, and the Hippocampal System. Cambridge: MIT Press.

Cohen N, Squire L. 1980. Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. Science 210;207–210.

Corsi P. 1972. Human memory and the medial temporal region of the brain. Dissertation Abstracts International, 34 (02), 891B. (University Microfilms No. AAI05–77717).

Damasio AR. 1989. The brain binds entities and events by multiregional activation from convergence zones. Neural Comput 1:123–132.

Eichenbaum H, Cohen N. 2001. From Conditioning to Conscious Recollection: Memory Systems of the Brain. New York: Oxford University Press.

Eichenbaum H, Yonelinas AP, Ranganath C. 2007. The medial temporal lobe and recognition memory. Annu Rev Neurosci 30:123–52.

Gold JJ, Smith CN, Bayley PJ, Shrager Y, Brewer JB, Stark CE, Hopkins RO, Squire LR. 2006. Item memory, source memory, and the medial temporal lobe: Concordant findings from fMRI and memory-impaired patients. Proc Natl Acad Sci USA 103:9351–9356.

Hannula DE, Ryan JD, Tranel D, Cohen NJ. 2007. Rapid onset relational memory effects are evident in eye movement behavior, but not in hippocampal amnesia. J Cogn Neurosci 19:1690–1705.

Hannula DE, Tranel D, Cohen NJ. 2006. The long and the short of it: Relational memory impairments in amnesia, even at short lags. J Neurosci 26:8352–8359.

Hannula DE, Ranganath C. 2009. The eyes have it: Hippocampal activity predicts expression of memory in eye movements. Neuron 63:592–599.

Hartley T, Bird CM, Chan D, Cipolotti L, Husain M, Vargha-Khadem F, Burgess N. 2007. The hippocampus is required for short-term topographical memory in humans. Hippocampus 17:34–48; doi:10.1002/hipo.20240.

Hayes SM, Ryan L, Schnyer DM, Nadel L. 2004. An fMRI study of episodic memory: Retrieval of object, spatial, and temporal information. Behav Neurosci 118:885–96.

Henke K. 2010. A model for memory systems based on processing modes rather than consciousness. Nat Rev Neurosci 11:523–532.

Huttenlocher J, Presson C. 1979. The coding and transformation of spatial information. Cogn Psychol 11:375–394.

James TW, Shima DW, Tarr MJ, Gauthier I. 2005. Generating complex three-dimensional stimuli. Greebles) for haptic expertise training. Behav Res Methods Instrum Comput 37:353–358

Jeneson A, Mauldin KN, Squire LR. 2010. Intact working memory for relational information after medial temporal lobe damage. J Neurosci 30:13624–13629.

Kessels RP, Van Zandvoort MJ, Postma A, Kappelle LJ, De Haan EH. 2000. The Corsi Block-Tapping Task: Standardization and normative data. Appl Neuropsychol 7:252–8.

Konkel A, Cohen NJ. 2009. Relational memory and the hippocampus: Representations and methods. Front Neurosci 3:166–174.

Konkel A, Warren DE, Duff MC, Tranel DN, Cohen NJ. 2008. Hippocampal amnesia impairs all manner of relational memory. Front Hum Neurosci 2:15.

Konkel A, Warren DE, Duff MC, Tranel DN, Cohen NJ. 2008. Hippocampal amnesia impairs all manner of relational memory. Front Hum Neurosci 2:15.

Lee ACH, Buckley MJ, Pegman SJ, Spiers H, Scahill VL, Gaffan D, Bussey TJ, Davies RR, Kapur N, Hodges JR, Graham KS. 2005. Specialization in the medial temporal lobe for processing of objects and scenes. Hippocampus 15:782–797; doi:10.1002/hipo.20101.

Marr D. 1971. Simple memory: A theory for archicortex. Philos Trans R Soc B Biol Sci 262:23–81.

McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychol Rev 102:419–457.

Mitchell KJ, Johnson MK, Raye CL, D'Esposito M. 2000. fMRI evidence of age-related hippocampal dysfunction in feature binding in working memory. Brain Res Cogn Brain Res 10:197–206.

Nadel L, Samsonovich A, Ryan L, Moscovitch M. 2000. Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. Hippocampus 10:352–368.

O'Keefe J, Nadel L. 1978. The Hippocampus as a Cognitive Map. Oxford University Press.

Piekema C, Kessels RP, Mars RB, Petersson KM, Fernandez G. 2006. The right hippocampus participates in short-term memory maintenance of object-location associations. Neuroimage 33:374–382.

Ranganath C, Blumenfeld RS. 2005. Doubts about double dissociations between short- and long-term memory. Trends Cogn Sci 9:374–380.

Ranganath C, D'Esposito M. 2001. Medial temporal lobe activity associated with active maintenance of novel information. Neuron 31:865–873.

Ryan JD, Cohen NJ. 2004. Processing and short-term retention of relational information in amnesia. Neuropsychologia 42:497–511.

Ryan L, Lin C-Y, Ketcham K, Nadel L. 2010. The role of medial temporal lobe in retrieving spatial and nonspatial relations from episodic and semantic memory. Hippocampus 20:11–18.

Schacter DL, Tulving E. 1994. Memory Systems, First Edition. MIT Press, Cambridge, MA.

Smith ML, Milner B. 1981. The role of the right hippocampus in the recall of spatial location. Neuropsychologia 19:781–793.

Squire LR, Stark CE, Clark RE. 2004. The medial temporal lobe. Annu Rev Neurosci 27:279–306.

Stark CE, Squire LR. 2003. Hippocampal damage equally impairs memory for single items and memory for conjunctions. Hippocampus 13:281–292.

Stern CE, Sherman SJ, Kirchhoff BA, Hasselmo ME. 2001. Medial temporal and prefrontal contributions to working memory tasks with novel and familiar stimuli. Hippocampus 11:337–346.

Uttal DH, Chiong C. 2004. Seeing space in more than one way: Children's use of higher order patterns in spatial memory and cognition. In: Human Spatial Memory: Remembering Where. Allen G.L., Earlbaum, L.; Editors. pp125–142.

Vargha-Khadem F, Gadian D, Watkins K, Connelly A, Van Paesschen W, Mishkin M. 1997. Differential effects of early hippocampal pathology on episodic and semantic memory. Science 277:376–380.

Voss JL, Gonsalves BD, Federmeier KD, Tranel D, Cohen NJ. 2011a. Hippocampal brain-network coordination during volitional exploratory behavior enhances learning. Nat Neurosci 14:115–120.

Voss JL, Warren DE, Gonsalves BD, Federmeier KD, Tranel D, Cohen NJ. 2011b. Spontaneous revisitation during visual exploration as a link among strategic behavior, learning, and the hippocampus. Proc Natl Acad Sci USA 108;E402–E409.

Warren DE, Duff MC, Jensen U, Tranel D, Cohen NJ. 2012. Hiding in plain view: Lesions of the medial temporal lobe impair online representation. Hippocampus 22;1577–1588; doi:10.1002/hipo.21000.

Wechsler D. 1997a. Weschler Adult Intelligence Scale,Third edition. New York, New York: The Psychological Corporation.

Wechsler D. 1997b. Weschler Memory Scale,Third edition, New York, New York: The Psychological Corporation.