HIPPOCAMPUS 22:1451–1466 (2012)

The Nature and Time-Course of Medial Temporal Lobe Contributions to Semantic Retrieval: An fMRI Study on Verbal Fluency

Signy Sheldon[1,2]* and Morris Moscovitch[1,2]

ABSTRACT: Recent investigations have shown that the medial temporal lobe (MTL), a region thought to be exclusive to episodic memory, can also influence performance on tests of semantic memory. The present study examined further the nature of MTL contributions to semantic memory tasks by tracking MTL activation as participants performed category fluency, a traditional test of semantic retrieval. For categories that were inherently autobiographical (e.g. names of friends), the MTLs were activated throughout the time period in which items were generated, consistent with the MTLs role in retrieving autobiographical memories. For categories that could not benefit from autobiographical or spatial/context information (e.g. governmental offices), the MTL was not implicated at any time point. For categories for which both prototypical and episodically-related information exists (e.g. kitchen utensils), there was more robust MTL activity for the open-ended, late generation periods compared with the more well-defined, early item generation time periods. We interpret these results as suggesting that early in the generation phase, responses are based on well-rehearsed prototypical knowledge whereas later performance relies more on open-ended strategies, such as deriving exemplars from personally relevant contextual information (e.g. imagining one's own kitchen). These findings and interpretation were consistent with the results of an initial, separate behavioral study (Expt 1), that used the distinctiveness of responses as a measure of open-endedness across the generation phase: Response distinctiveness corresponded to the predicted open-endedness of the various tasks at early and late phases. Overall, this is consistent with the view that as generation of semantic information becomes open-ended, it recruits processes from other domains, such as episodic memory, to support performance. © 2011 Wiley Periodicals, Inc.

KEY WORDS: episodic memory; fluency; hippocampus; semantic memory; task characteristics

INTRODUCTION

Regions of the medial temporal lobe (MTL), particularly the hippocampus and the parahippocampal cortex, are widely considered episodic memory structures (Eichenbaum, 1997; Moscovitch et al., 2006; Eichenbaum et al., 2007). Recent studies, however, have extended the influence of these structures to domains outside of episodic memory (for examples see, Lee et al., 2005; Karlsgodt et al., 2005; Addis et al., 2007), including semantic retrieval (e.g. Kapur et al., 1995; Whatmough and Chertkow, 2007). Whilst such studies suggest that MTL-dependent processes can contribute to nonepisodic tasks, it is not clear what are the precise task conditions or characteristics that invoke these mnemonic processes. The purpose of the present article was to examine the nature of MTL contributions to semantic retrieval tasks. Specifically, we hypothesized that the MTL will be particularly involved in semantic retrieval for tasks that are open-ended. Open-ended tasks are those for which the to-be-retrieved information is not well specified. In this sense, they are likely to rely on the application of idiosyncratic strategies or information that can be based on episodic memories, rather than on application of algorithms or prototypical exemplars, which are common across individuals and upon which close-ended tasks rely. Usually, the set size is smaller for close-ended tasks given that prototypical responses are generated and larger for open-ended tasks given that these tasks are amenable to the implementation of various processes to generate alternative responses. We also suggest that tasks may be open-ended, and perhaps rely on episodic memory from the beginning, or they may become open-ended as nonepisodic strategies become more demanding or less useful.

To test this hypothesis, we used fMRI to track the activation of MTL structures during verbal fluency, a traditional semantic retrieval task, with the prediction that the MTL will be recruited when such fluency tasks (a) depend on episodic information from the beginning or (b) contain both semantic and episodic components, with the former driving performance at the beginning, closed-ended or well-defined portion of the task, and the latter gaining in importance as the task progresses and becomes more open-ended. Before describing the present study, we will first review the literature that supports a role for the MTL in semantic retrieval.

The Medial Temporal Lobes and Semantic Memory

The idea that MTL regions are involved in semantic retrieval has a history dating at least to Newcombe's (1969) observation that patients with temporal lobe lesions were impaired on tests of semantic, but not phonemic, fluency. Many investigators thought that the deficit arose from damage to lateral and anterior aspects of the temporal lobes, rather than to medial ones, since the former structures have long been implicated in semantic memory (for more recent contributions, see Thompson-Schill, 2002; Dronkers et al., 2004; Moscovitch et al., 2006; Lambon Ralph and Patterson, 2008). Current reports have implicated

[1] **Department of Psychology, University of Toronto, Toronto, ON, Canada;**
[2] **Rotman Research Institute, Toronto, ON, Canada**
Grant sponsors: NSERC (to S.S.) and CIHR (to M. M.)
*Correspondence to: Signy Sheldon, Department of Psychology, Sidney Smith, 4th floor,100 St. George Street, University of Toronto, Toronto, Ontario, M5S 3G3, Canada. E-mail: signy.sheldon@utoronto.ca
Accepted for publication 1 September 2011
DOI 10.1002/hipo.20985
Published online 18 November 2011 in Wiley Online Library (wileyonlinelibrary.com).

the medial aspects of the temporal lobe to performance on tests traditionally defined as semantic, such as category fluency which we discuss in detail in the next section. Neuroimaging studies have also implicated the MTL in tests of semantic memory (Kapur et al., 1995; Maguire and Mummery, 1999; Whatmough and Chertkow, 2007; Whitney et al., 2009). Together these studies imply that semantic and episodic memory are more interactive than they are distinct.

Consistent with this interactionist account of episodic and semantic memory, Barsalou proposed that, in some cases, semantic information can be contextually bound to autobiographical or episodic memories so that the latter can be used to help generate or retrieve semantic information (Barsalou, 1988, 2009; also see Westmacott and Moscovitch, 2002, 2003). It follows that episodic memory will be implicated more on semantic tasks that can benefit from recovering related autobiographical or spatial/contextual details than those that do not. As such, one objective of the present study is to explore the influence of this task characteristic on verbal fluency.

Verbal Fluency and Task Characteristics

Verbal fluency encompasses both category and phonemic fluency tasks (Lezak, 1995). Although both of these types of verbal fluency tasks have conventionally been thought to reflect semantic memory retrieval and frontal and lateral temporal lobe functioning, there is mounting evidence that category fluency also involves episodic memory retrieval processes (Vallee-Tourangeau et al., 1998) mediated by the MTL and the hippocampus, in particular (for lesion evidence see Gleissner and Elger, 2001, and for neuroimaging evidence see Pihlajamaki et al., 2000; Baldo et al., 2006; but see Schmolck et al., 2002 for a study that showed spared semantic retrieval in patients with hippocampal lesion). In a recent study, Ryan et al. (2008) compared activity in the hippocampus proper and other MTL structures across three category types that differed in their relation to episodic memory (nonspatial/autobiographical categories such as governmental positions were low episodic-relevance categories, and autobiographical nonspatial categories, such as names of friends and autobiographical spatial categories, such as kitchen utensils were high relevance categories). They found similar levels of activity in the hippocampus for all three category types and greater activity in the parahippocampal cortex for those categories that evoked a spatial or autobiographical context. Following this study, Greenberg et al. (2009) found that patients with bilateral or unilateral MTL lesions were more impaired for categories that required strategies based upon episodic memory processes (autobiographical and spatial categories) than those that did not (nonepisodic categories).

What has yet to be examined is precisely *how* and *when* episodic memory processes mediated by the MTL contribute to performance on tests of category fluency. The main objective of the present study is to investigate when episodic memory contributes to fluency tasks. We suggest that this contribution will vary based on the nature of the task, as discussed above, but also on the time-course of item generation, as fluency task parameters change. In the early stages of some fluency tasks, category fluency is defined as a closed-ended task in that responses are determined by a set of prototypical exemplars evoked by the category cue (e.g. generating knife, fork, spoon for the category kitchen utensils). These responses are common to most individuals in a cohort given that they can be derived from a network of items that are strongly associated with the cue or are highly representative of that semantic category (Crowe, 1998). For such categories, as the task progresses, the responses will become more idiosyncratic and variable. As a result, item generation will reflect the individual's particular experiences and associations, leading to responses that consist of less representative members of the category (e.g. scissors, turkey baster, apple corer for kitchen utensils). In other words, at these later stages, the task of generating items can be considered to be more "open-ended" in that the exemplars are now not as tightly determined by their prototypicality or associations in semantic memory. In such cases, if the task allows, responses are determined by other factors, such as one's experience with the items (e.g. "let me think of what is in my kitchen"), lending themselves to an episodic rather than semantic strategy.

Under our hypothesis, we suggest that the likelihood of drawing on episodic memory, and how soon one does so during performance, will vary with the type of category and the task characteristics that change during fluency. Taking this further, we suggest that open-ended tasks, those for which standard responses or set procedures for arriving at the needed information do not exist, will invoke MTL-based episodic memory processes.

Some categories lend themselves to an episodic or autobiographical strategy from the beginning and throughout the task given the open-ended and episodic nature of the task (e.g. names of friends), some categories only after a semantic strategy becomes less useful and the task becomes progressively more open-ended (e.g. kitchen utensils in which both prototypical and personal information exists). For other categories, an episodic strategy is difficult to implement or is unlikely to provide worthwhile information (e.g. government offices), and thus will not be implemented at all. In other words, if a category is open-ended and based on autobiographical information from the beginning, then episodic memory will be engaged at all time points. If a category can benefit from recovering related autobiographical or contextual details (given that recovering contextual information is a crucial aspect of episodic memory processes) as it becomes more open-ended as it progresses, then episodic memory will be recruited for the task after initial semantic search strategies have been exhausted or are too inefficient to prove useful. If strategies based on episodic memory are not useful (e.g. naming government offices or phonemic fluency), then even when the task is open-ended, the MTLs are unlikely to be engaged at any point. Put simply, the interaction of open-endedness and episodic relevance, both in terms of autobiographical and spatial content (as it defines the context of episodic memory), determine the time course and extent to which the MTL will be engaged in the semantic task.

The Current Study

To test the hypothesis that episodic memory processes supported by the MTL are engaged more when semantic tasks are open-ended, the degree of MTL activation was examined while participants generated items for categories that drew upon episodic memory to varying degrees (Ryan et al., 2008). The present study differed from a previous study that investigated regions involved in category fluency (Ryan et al., 2008) in two important respects. First, participants were given a longer generation time than in the previous study to enable examination of the time-course of performance, during both closed and open-ended task periods. If category generation can benefit from episodic memory, then episodic memory processes supported by the MTL will contribute to the extent that the task is open-ended. Second, phonemic fluency was included in this study not only for its own inherent interest, but also to serve as a baseline task to control for general characteristics of verbal fluency, such as response generation and difficulty across different time intervals.

The categories used were similar to those of Ryan et al. (2008). Ryan et al. asked over 40 undergraduates to state the strategies they used to generate items to 60 categories. From these categories, 15 were assigned to each of three category types based on normative data on how generating exemplars proceeded: (1) open-ended throughout, which we called *autobiographical* categories, as they relied crucially on autobiographical information (e.g. *names of friends*). Exemplar generation from this category is likely to activate structures that comprise an autobiographical memory network that in addition to the MTL, includes the medial frontal and posterior parietal cortex (Maguire, 2001; Gilboa, 2004; Addis et al., 2007). (2) Closed-ended throughout, which we called *nonepisodic* (e.g. *governmental positions*) for which contextual episodic information or autobiographical memories are not likely to help in exemplar generation. For these categories, we predict that little MTL involvement will be found at any stage during the task. It is also predicted that these categories will activate a network of brain regions that is more representative of semantic retrieval, such as the inferior and lateral aspects of the temporal lobe (Binney et al., 2010). (3) *Spatial/context* categories (e.g. *kitchen utensils*) which change from closed-ended to open-ended with task progression. For such categories, one is likely to rely on highly frequent prototypical exemplars derived from semantic memory in the early stages, but then turn to processes associated with episodic memory once the semantic strategy loses its usefulness or items from it are exhausted. Increases in MTL activation are expected to accompany the shift from a semantic strategy in early stages to an episodic one in later stages. Moreover, given that these categories draw on scenes and on contextual information, we predict activation of a network of regions involved in inspection of spatial imagery, namely the right parahippocampal cortex and the precuneus (e.g. Hassabis et al., 2007).

As stated earlier, participants were also given a phonemic fluency task (e.g. words that start with *F*), which also has no episodic component and has been shown to rely upon executive functions mediated by the prefrontal cortex (Milner and Petrides, 1984; Baldo et al., 2006). This task served as a control measure for nonspecific factors associated with item generation in the other verbal fluency tasks.

MATERIALS AND METHODS

Experiment 1: Behavioral Study

Before turning to the neuroimaging study, we begin with a behavioral study to test our conjecture regarding the extent to which different category fluency tasks and task periods are open-ended or close-ended. As we noted earlier, we define a closed-ended task as one in which solutions are based on the application of algorithm and rules, or on the use of prototypical responses. With respect to category fluency, this implies a uniformity of responses across individuals. By contrast, an open-ended task is one in which solutions rely on the application of idiosyncratic strategies or information, based on episodic memories which are peculiar to each individual. As a result, as category fluency becomes more open-ended, the variability in responses across individuals will increase. In other words, exemplar generation will be more distinct across individuals (individual differences will be greater) for tasks that are open-ended than closed-ended.

To obtain an objective measure of endedness, we measured the distinctiveness/variability of items by comparing responses among participants in early and late item generation periods for the three assigned categories: autobiographical, spatial, and nonepisodic categories. According to our hypothesis, we predicted that autobiographical categories would show high levels of distinctiveness throughout the task, with no difference between early-generated and late-generated items; there will be many distinct items across participants throughout the generation time. By comparison, we predicted low-levels of distinctiveness for the nonepisodic categories throughout the task, as exemplar generation in this task is based more on semantic memory and are less likely to benefit from idiosyncratic responses based on episodic memory. For the spatial/context categories, we predicted a shift in distinctiveness between early and late stages of exemplar generation, reflecting the shift from a closed-ended task dependent on semantic memory (prototypicality) to a more open-ended task dependent on episodic memory (idiosyncratic).

Participants

There were 20 participants in this experiment [eight females; mean age = 23 years (SD = 5.3); mean years of education = 16 years (SD = 2.2)]. None of these participants participated in the subsequent neuroimaging experiment. All had normal or corrected to normal vision, were free from neurological or psychiatric illness, and English was their primary language. The majority of the participants were undergraduate students at the

TABLE 1.

Categories and Category Types From Which Exemplars Were Generated in Experiments 1 and 2

Autobiographical	Spatial	Nonepisodic	Phonemic
Books you've read	Things in a garage	Pollutants	F
Movies you've seen	Items in dresser drawers	Famous people	A
Jobs you've worked	Furniture in a living room	Cars	S
Brand name of clothes you own	Utensils in a kitchen	Random numbers	C
CD's you own	Things in a bedroom	Superheroes	L
People you work with	Things in a house	Things in a jail cell	B
Things you keep in your pocket	Attractions in Toronto	Animals from smallest to largest	M
Names of family members	Things on a desk	Crimes	R
Classes you took in High School	Things that go on feet	Modes of transportation	T
Sports that you've played	Food in a fridge	Sharp objects	D
Cities you've visited	Places to eat at	Things that are red	P
Your favorite foods	Things in a closet	Foreign countries	N
Names of your friends	Things in a bathroom	Governmental positions	G
Games on your computer	Office items	Types of sports	H
Shops in your local mall	Things on a wall	Things that cost less than $1	O

University of Toronto and all participants received an honorarium for their participation.

Materials

Categories from Ryan et al. (2008) were used in this experiment (Table 1). Of note, the category "random numbers" was simply "numbers" for this experiment.

Procedure

On a computer screen, participants were presented with instructions that described the task: they were to think of 12 items that belonged to various categories as quickly as they could. Twelve items were chosen because, based on previously collected data, the average time to generate 12 items for these categories was close to the time that subsequent participants would be given to generate items in the scanner in the following study. The categories were presented on the center of the computer screen, signaling the start of the task and remained on the computer screen for the entire duration of the task. Responses were recorded via an electronic recorder as well as by hand by the experimenter.

Results and discussion

Distinctiveness of the items generated was examined using the experimenter-assigned categories and category types used by Ryan et al. (2008; please refer to this article for a detailed description of category assignment). To obtain a measure of distinctiveness in early and late item generation periods, the number of different items across all 20 participants was counted for the first and final three items. Words that were close synonyms (e.g. *couch* and *sofa*, *rocket ship* and *spaceship*, or *notepad* and *paper*) were not counted as distinct because these items refer to the same semantic item using different linguistic labels. For some categories, albeit very few, some participants misunderstood the types of items to generate (e.g. for the category *famous people*, generating the items "greed" and "beautiful"). In these cases, those responses were eliminated. Thus to calculate an unbiased measure of distinctiveness, we divided the total number of distinct items by the total number of items generated during the generation periods across all participants, resulting in a percent distinctiveness rating for early and late item generation. The distinctiveness rating for these two periods was averaged for categories belonging to the autobiographical, spatial, and nonepisodic categories and is presented in Figure 1.

As illustrated in Figure 1, the results confirmed our predictions regarding distinctiveness in the three categories across time. The percent distinctiveness was high for both time periods for the autobiographical categories, mid to low for both time periods for the nonepisodic categories, and increased for the spatial categories from early to late generation periods. A χ^2 analysis examining the counts in early and late periods

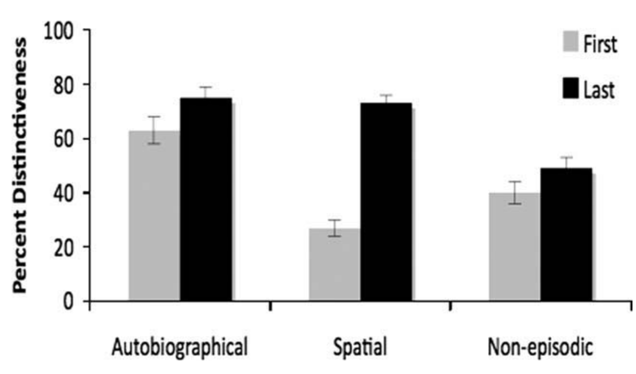

FIGURE 1. The average percent distinct items generated in the initial (first three items) and final (last three items) time periods across 20 participants to autobiographical, spatial, and nonepisodic categories for Experiment 1. Standard error bars are shown.

across the three category types revealed that early and late frequency or counts differed between the category types (χ^2 = 9.80, $P = 0.007$). A subsequent χ^2 analyses revealed that only the spatial categories showed such a significant difference in distinctiveness rating from what would expected if no change occurred ($\chi^2 = 21.16$, $P < 0.001$). There was no difference for the nonepisodic and autobiographical categories between the two time periods ($\chi^2 = 1.04$, $P = 0.31$; $\chi^2 = 0.72$, $P = 0.39$, respectively). Overall, these results provide evidence that it is only the spatial categories that switch from being a closed-ended task to an open-ended task. The autobiographical and the nonepisodic task remained, respectively, open-ended and closed-ended throughout.

Experiment 2: Functional Neuroimaging

Participants were scanned while generating exemplars during each of the three category fluency tasks, and during a phonemic fluency task. We first examined the overall activity in these four tasks to determine their effect on brain activation. Here we expect that for tasks that are nonepisodic, regions associated with semantic memory, such as the left lateral temporal lobe, and inferior prefrontal cortex will be activated preferentially, whereas for categories that are episodic, regions of the episodic memory networks, such as the MTL, medial prefrontal cortex, and the precuneus, will be activated preferentially. Using phonemic fluency as a baseline to control for nonspecific factors that may influence activation across time, we then tested our main hypothesis that the MTL would contribute most to verbal fluency in those conditions in which the task is open-ended. Based on the results of Experiment 1, we predict that activation of the MTL and other regions that comprise the episodic/autobiographical memory network will be greatest in the autobiographical task and the later stages of the spatial task, but least or absent for the nonepisodic task and the early stages of the spatial task.

Participants

Sixteen participants (ten females; mean age 24.8 years, SD = 4.5; mean education 17.0 years, SD = 2.3; all right handed) with normal or corrected to normal vision participated in this study. All participants were free of psychiatric and neurological disorders, serious head injury, substance abuse, hypertension, or other conditions that are unsuitable for an MRI environment. Informed consent was obtained from all participants and all experimental procedures were in accordance with the Rotman Research Institute/Baycrest Hospital ethical guidelines. Participants received compensation upon completion of the study.

Stimuli

Forty-five different categories (Table 1) were presented to participants. The categories used drew on one of three types as determined and used in a previous study by Ryan et al. (2008; also see Vallée-Tourangeau et al., 1998). Fifteen of these categories draw upon autobiographical information (*autobiographical categories*, e.g. movies you've seen), 15 categories that are likely not to evoke episodic processes to generate items or they draw upon semantic information (*nonepisodic categories,* e.g. governmental positions), and 15 categories draw upon potentially autobiographical and/or spatial contextual information (*spatial/context categories*, e.g. kitchen items). An additional 15 categories that are phonemic categories (think of words that begin with a particular letter) and 15 trials of a visual-motor baseline task (press a button when you see an "X") were also given to the participants.

Procedure

Before scanning, the participants received experimental instructions and several examples. They were told that the name of a category would appear and they were to generate (silently) as many items that belonged to that category as they could, pressing a button with their right index finger every time they generated an item. For the phonemic fluency category, they were told not to generate proper names (Lezak, 1995). While in the scanner, participants were presented visually with each category in a pseudo-random order. Based on pilot testing to establish an appropriate amount of generation time, participants were given 36 s to generate as many items as they could which belong to that category, again, pressing the button each time an item was generated. The name of the category remained on the screen for the entire generation time. Each category run was separated by a 12 s interstimulus interval in which the participants were instructed to rest. For the visual-motor baseline condition, the presentation of the X was randomly jittered across the 36 s so that each X is presented with an interstimulus interval of 1,000 ms to 7,000 ms. Participants completed all 60 categories and 15 baseline tasks in five functional imaging runs. Each run lasted approximately 12 min.

Postscan interview

After scanning, participants completed a postscan interview. They were taken to a testing room outside of the MRI suite. On a computer, they were again presented with each category (in random order) and for each category they were asked to rate how difficult it was to think of items for each category on a scale of 1 to 7. Participants were also asked to recall items that they generated while in the scanner. This procedure was included as a means to determine the number of items recollected from each category, but given that there was no objective way of determining if the items the participants said they recalled were, in fact, recalled (verbal responses were not permitted inside the scanner), these data should be interpreted cautiously.

Image acquisition

All imaging was performed on a 3T Siemens full-body MRI machine with a standard 12-channel array head coil located at the Rotman Research Institute/Baycrest Hospital. Anatomical

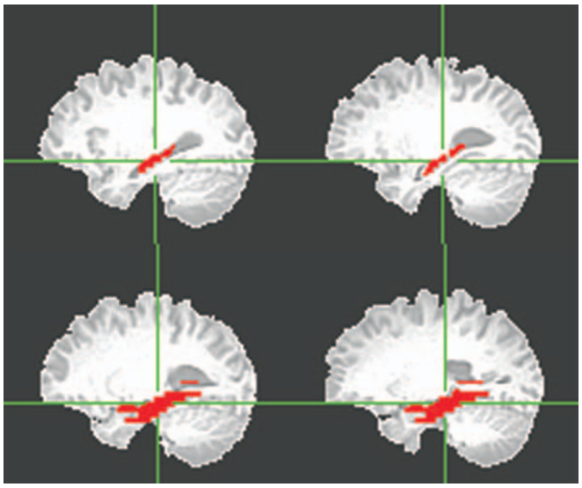

FIGURE 2. Regions that are include in the AFNI ROI templates for the left, right hippocampus, and the left and right parahippocampal gyrus (voxels included in each ROI are defined in the San Antonio Talairach Daemon created by tracing Talairach and Tournoux brain illustrations). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

scans were acquired via T1-weighted volumetric MRI (TR = 2,000 ms, TE = 2.63 ms, 160 axial slices, 1.0 mm thick, FOV = 256 mm). For the functional images, 30.5 mm thick axial slices with T2*-weighted EPI pulse sequence were obtained (TR = 2,000 ms, TE = 30 ms, flip angle = 70°, FOV = 200 mm) with no spacing. The spatial resolution (voxel size) was 1 × 1 × 1 mm.

Imaging analysis

All imaging processing and analyses was done using AFNI software package, version 2.0 (Cox, 1996; Cox and Hyde, 1997). The first 10 images that were acquired before the task in each functional run were excluded from all analyses because it is likely that brain magnetization had not yet reached a steady state. Images were first reconstructed, then corrected for physiological movement due to heart rate and respiration, slice-timing corrected to the first slice and motion corrected using a three-dimensional Fourier transform interpolation with a functional volume that minimized the amount of motion to approximately 1.5 mm. The images were normalized to a standard Talairach space using linear transformation, smoothed with an isotropic 6 mm FWHM Gaussian filter and resliced to 2 mm × 2 mm × 2 mm voxels. Each participant's anatomical image was transformed to standard Talairach atlas space (Talairach and Tournoux, 1988).

The five functional runs were concatenated and the BOLD signal activation for each participant in each condition was established with respect to activity during the visual-motor baseline task for examining overall category activation. Functional images were matched to the reference brain with alignment parameters from the structural scans.

Statistical analyses were performed in two phases. A two-factor analysis of variance (ANOVA) with category type as a fixed factor and participant as a random factor was used to examine the BOLD response associated with each category type. The statistical threshold was set to $P < 0.001$ when looking at activations within the whole brain. Based on the AFNI AlphaSim program (available at: http://afni.nimh.nih.gov/afni/docpdf/alphasim.pdf); a minimum cluster size of 62 was used to correct for multiple comparisons.

Given the outlined specific hypotheses concerning brain regions (i.e. the MTL; hippocampus and parahippocampal cortex), region of interest (ROI) analysis was used. The hippocampus, bilaterally, and the parahippocampal gyri, bilaterally, were anatomically defined at a group level, using the AFNI brain region templates (see Fig. 2 for the areas that are included in the AFNI standard hippocampal and parahippocampal templates). Because of these small volumes and because MTL regions have lower amplitude hemodynamic responses than other brain regions (e.g. Addis et al., 2007; Ryan et al., 2008), we used a less stringent statistical threshold of $P < 0.005$ with a cluster size of 11 for the hippocampal ROI and a cluster size of 32 for the parahippocampal cortex ROI, both corrected for multiple comparisons based upon the results from AlphaSim (available at: http://afni.nimh.nih.gov/afni/docpdf/alphasim.pdf).

RESULTS

Behavioral Results

Figure 3 plots the average number of items that were generated for each category type while the participants were in the scanner. There was a difference in the mean number of items generated among the four category types ($F_{(3,45)} = 8.72$, $P < 0.001$). Post hoc pair-wise comparisons revealed that there was no difference in the mean number of items generated between the phonemic and spatial/context categories ($P = 0.87$) nor between the phonemic and nonepisodic categories ($P = 1.00$),

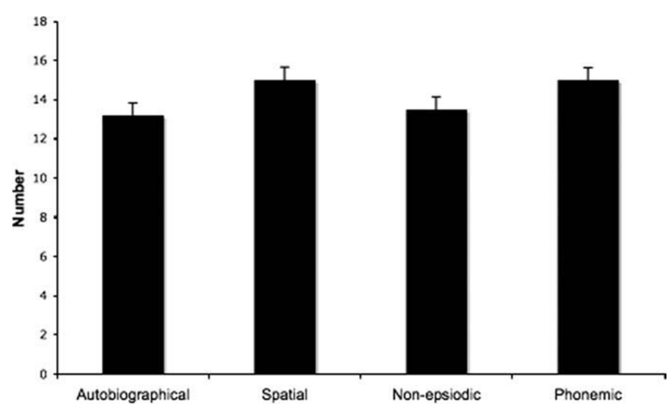

FIGURE 3. The average number of generated items for autobiographical, spatial, nonepisodic, and phonemic categories across 16 participants while in the scanner (Experiment 2). Standard error bars are shown.

TABLE 2.

Brain Regions and Their Co-ordinates (Talairach) Within the Medial Temporal Lobes Activated for Each Category Type Compared With Visual Motor Baseline (P < 0.005)

Region	L/R	x	y	z	Volume	Max int
Hippocampus	L					
Autobiographical		−30	−30	−6	272	1.22
Spatial		−32	−30	−10	328	1.87
Nonepisodic*		−32	−32	−4	224	0.74
Phonemic						
Hippocampus	R					
Autobiographical						
Spatial		30	−38	−2	136	1.21
Nonepisodic						
Phonemic						
Parahippocampal gyrus	L					
Autobiographical		−10	−48	2	2,304	7.20
Spatial		−26	−44	−10	3,816	5.14
		−10	−48	2	720	6.62
Nonepisodic		−12	−50	4	384	5.38
		−22	−44	−10	376	2.18
Phonemic						
Parahippocampal gyrus	R					
Autobiographical		12	−50	4	552	4.96
Spatial		26	−32	−18	2,544	2.48
		12	−50	4	576	4.30
Nonepisodic						
Phonemic						

xyz values represent the maximum intensity value coordinates. Volume is that which makes the cluster, in microliters, and max int represents the maximum intensity value for that volume or cluster.
*$P < 0.05$.

but there were significant differences between the nonepisodic and spatial/context categories ($P = 0.027$), and between the autobiographical categories and the phonemic ($P = 0.007$) and the spatial/context ($P < 0.001$) categories.

Three participants' data were not recorded during the postscan interview due to computer error (note this error was only for the postscan interview, not for collecting neuroimaging data for which all 16 participants' data were used). For the remaining 13 participants there was no difference in the average difficulty rating across category type ($F_{(3,36)} = 1.102$, $P > 0.05$). Important to the analysis of early versus late item generation, there was no significant difference in terms of perceived difficulty between the phonemic categories and any of the other category types. Not surprisingly, there was a significant negative correlation between mean difficulty scores and the mean number of items generated for each category type across participants (autobiographical $r = -0.469$, $P < 0.001$; spatial/context $r = -0.320$, $P < 0.001$; nonepisodic $r = -0.629$, $P < 0.001$; phonemic $r = -0.519$, $P < 0.001$). That is, the higher the difficulty rating (the more difficult), the fewer the number of items that were generated.

Because the number of items generated differed among the categories, we examined postscanning recall by considering the ratio of the number recalled/the number generated in the scanner. As noted previously, these data should be interpreted with some caution because we cannot ascertain if the item recalled actually was generated in the scanner, though the likelihood of recalling new items is very low. An ANOVA showed there was a significant difference in the recall ratio across category type ($F_{(3,36)} = 19.579$, $P < 0.001$) with the ratio being smaller for the phonemic category (0.22, SD = 0.19) than for the three others (autobiographical = 0.49, SD = 0.08; spatial = 0.37, SD = 0.14; nonepisodic = 0.45, SD = 0.09). Subsequent simple comparisons among the three other categories showed that the recall ratio was significantly greater for autobiographical categories compared to the spatial categories ($P = 0.005$), with no significant difference between these two categories and the nonepisodic categories.

fMRI Results: Overall Category Activity

ROI analysis of the medial temporal lobe regions

Table 2 displays the coordinates and cluster sizes for areas within the left and right hippocampus and left and right parahippocampal cortex in response to each category type using the visual motor condition as a baseline over the generation period (also see Fig. 4). The generation period was defined individually for each participant and each category. This period began when the initial item was generated and ended when the last item was generated so as to ensure that the participant was on task. Within the hippocampus, there were no regions of activity that survived the correction for multiple comparisons for the phonemic category and the nonepisodic categories; but, at a higher threshold of $P < 0.05$, a cluster within the left hippocampus was significantly active for the nonepisodic categories (corrected for multiple comparisons). Autobiographical categories showed a significant cluster of activity in the left hippocampus that was similar in extent to the significant cluster in the left hippocampus found for the spatial categories. The spatial categories were the only category type to elicit a significant cluster within the right hippocampus. Autobiographical, nonepisodic, and spatial category types showed significant regions of activation within the left parahippocampal cortex whereas the phonemic categories did not. The spatial and autobiographical categories activated the areas within the right parahippocampal cortex significantly, but the nonepisodic and phonemic categories did not.

Whole brain analysis

Activity for each category was contrasted to the visual motor baseline task across the whole brain. Table 3 lists the regions with the peak coordinates and clusters that were active for each category type. A large cluster that covered the frontal and temporal regions, including the medial temporal lobe regions, was found for the autobiographical categories (with peak coordinates in the lingual gyrus) and for the spatial and nonepisodic categories (with peak coordinates in the middle frontal regions for the latter two categories). Of note was the finding of significant activity within the left precuneus region for all three category types, but not for phonemic fluency.

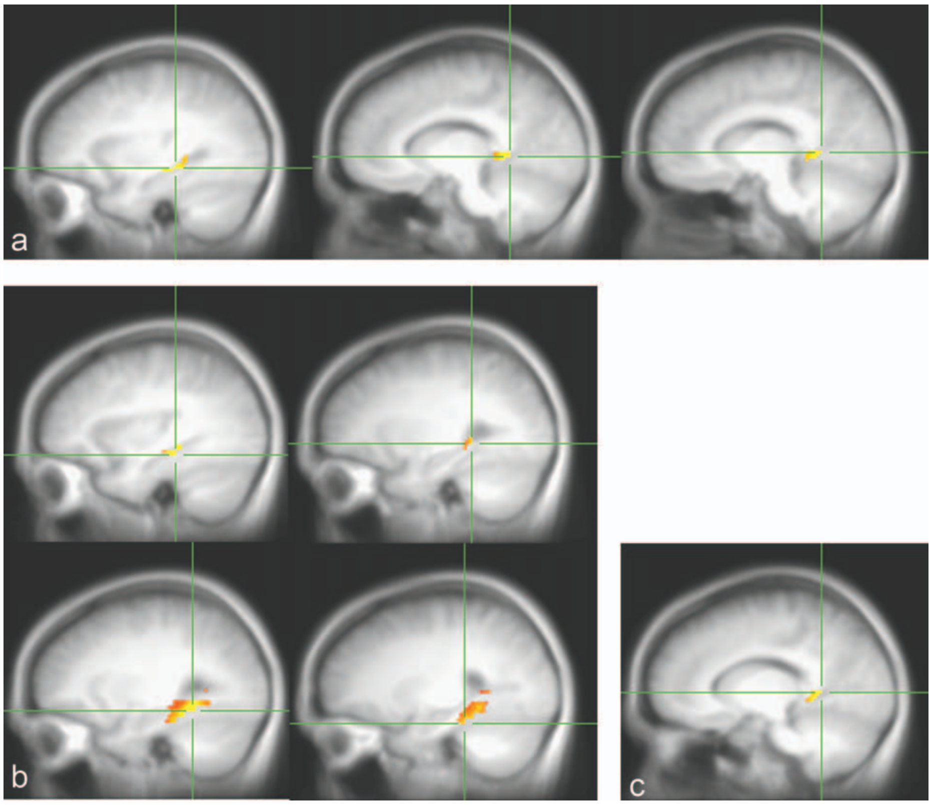

FIGURE 4. a. A ROI group analysis of hippocampal and parahippocampal regions using AFNI templates for the autobiographical categories when compared with a visual motor baseline. *X-Y-Z* coordinates of the intersection is the voxel of maximum intensity. Left panel is the left hippocampal ROI (−30 −30 −6), middle panel is the left parahippocampal ROI (−10 −48 2) and right panel is the right parahippocampal ROI (12 −50 4). The clusters shown are significant at $P < 0.005$, SVC. b. A ROI group analysis of hippocampal and parahippocampal regions using AFNI templates for the spatial categories when compared to a visual motor baseline. *X-Y-Z* coordinates of the intersection is the voxel of maximum intensity. Top left panel is the left hippocampal ROI (−32 −30 −10), top right panel is the right hippocampal ROI (30 −38 −2), bottom left panel is the left parahippocampal ROI (−26 −44 −10) and bottom right panel is the right parahippocampal ROI (26 −32 −18). The clusters shown are significant at $P < 0.005$, SVC. c. A ROI group analysis of hippocampal and parahippocampal regions using AFNI templates for the nonepisodic categories when compared with a visual motor baseline. *X-Y-Z* coordinates of the intersection is the voxel of maximum intensity. Only the left parahippocampal gyrus ROI (−12 −50 4) had a significant cluster at $P < 0.005$, SVC. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

"Episodically relevant" versus "episodically irrelevant" categories

To examine brain regions that were active for the categories thought to depend on episodic information and those regions involved in the categories thought to be episodically irrelevant, we contrasted autobiographical + spatial to the nonepisodic + phonemic categories. Areas that were more active for the "episodically relevant" categories included the MTL bilaterally as well as bilateral posterior parietal areas, medial frontal, and temporal regions. Areas that were more active for the "episodically irrelevant" categories were bilateral frontal regions, particularly the inferior frontal regions, and the left fusiform gyrus (Table 4).

Autobiographical versus spatial categories

To compare brains regions that were active for the two categories predicted to involve episodic memory, the autobiographical and spatial categories were then contrasted with one another (Table 5) across the whole brain. Areas that were more active for the autobiographical categories included regions in the medial frontal, left cuneus, right angular gyrus, left fusiform gyrus, bilateral middle temporal gyrus, and an anterior region in the left parahippocampal cortex. Areas that were more active for the spatial categories were the left and right parahippocampal cortex (posterior regions), the left precuneus, and an area of the medial frontal lobe that is more posterior than the area found for the autobiographical categories.

fMRI Results: Early Versus Late Activity

Analysis

To investigate possible changes in activation over time, generation time was divided into early and late for each category type (with a middle generation period to allow for signal drop off). We then examined the magnitudes of neural activity that

TABLE 3.

Brain Regions Outside of the Medial Temporal Lobes More Activated by Each of the Four Category Types When Compared With a Visual Motor Baseline Task (P < 0.001)

Category type	x	y	z	Volume	Max int
Autobiographical					
Left lingual gyrus	−10	−52	2	178,144	8.86
Right cerebellum	4	−54	−42	3,944	2.67
Left precuneus	−30	−73	46	3,264	5.17
Left middle temporal gyrus	−58	−42	0	496	2.60
Spatial					
Left middle frontal gyrus	−48	12	34	166,568	8.97
Left precuneus/superior parietal lobule	−28	−74	46	8,040	5.89
Right cerebellum	24	−34	−36	1,392	2.46
Right thalamus	14	−4	10	672	1.33
Nonepisodic					
Left middle frontal gyrus	−48	12	34	88,704	8.89
Right cerebellum	32	−64	−20	25,704	5.53
Left precuneus/superior parietal lobule	−30	−74	44	2,664	4.66
Phonemic					
Left inferior frontal gyrus	−48	10	34	88,056	9.56
Right cerebellum	32	−64	−20	30,464	7.15
Left fusiform gyrus	−48	−60	−16	8,200	6.09
Left superior parietal lobule	−30	−68	44	4,176	4.66
Left culmen	−10	−52	−2	584	4.39
Right caudate	16	−22	22	584	1.37

xyz values represent the maximum intensity value coordinates. Volume is that which makes the cluster, in microliters and max int. represents the maximum intensity value for that volume or cluster.

corresponded to the initial 12 s and final 12 s of interest (with a middle generation time of at least 6 s). We chose to examine time bins rather than item bins because we were interested in the generation process at different time points rather than the number of items generated.

These time points were determined for each category and for each participant individually. Those categories for which few items were generated (fewer than six items) were excluded from this analysis because the generation time was too short to permit examination of any process change that occurred in the interval. To control for increasing drift associated with stimulus presentation time as well as for possible increasing levels of difficulty associated with later compared with earlier generated items, phonemic fluency was used a baseline in all the subsequent analyses. A three-factor analysis of variance (ANOVA) with category type and time period (early vs. late) as fixed factors and participant as a random factor was used to examine the BOLD response associated with each time period for each category type. A more liberal threshold of $P < 0.05$ was used for this set of analysis because of the fewer data points involved in each condition.

TABLE 4.

Contrast Between Episodically Relevant (Autobiographical and Spatial) and Episodically Irrelevant (Nonepisodic and Phonemic) Categories (P < 0.001)

Category type	x	y	z	Volume	Max int
Episodically relevant (autobiographical and spatial)					
Right posterior cingulate gyrus	6	−54	12	62,376	7.65
Left medial frontal gyrus	0	52	0	31,288	8.54
Left angular gyrus	−40	−78	32	12,600	5.87
Right precuneus	38	−78	36	9,736	4.51
Right middle frontal gyrus	28	22	54	9,280	2.93
Right cerebellum	10	−44	−42	5,792	3.26
	40	−66	−46	1,056	2.71
	12	−90	−28	584	2.56
Left middle temporal gyrus	−60	−16	−8	2,760	2.83
Left cerebellum	−36	−54	−32	688	1.47
Right middle temporal gyrus	58	−4	−16	568	2.12
Episodically irrelevant (nonepisodic and phonemic)					
Left inferior frontal gyrus	−52	4	26	7,016	4.41
Left inferior parietal lobule	−50	−42	54	4,440	2.91
Right inferior parietal lobule	52	−40	54	3,520	2.64
Right inferior frontal gyrus	50	6	28	2,136	3.23
Left fusiform gyrus	−44	−56	−10	992	3.13

xyz values represent the maximum intensity value coordinates. Volume is that which makes the cluster, in microliters, and max int represents the maximum intensity value for that volume or cluster.

ROI of the medial temporal lobes

Spatial categories showed significant clusters of activation within the MTL when later generated items were contrasted with earlier generated items. Specifically, clusters within the left

TABLE 5.

Contrast Between Autobiographical and Spatial Categories (P < 0.001)

Category type	*x*	*y*	*z*	Volume	Max int
Spatial					
Left parahippocampal gyrus	−28	−46	−6	7,728	2.60
Left precuneus	−12	−76	48	5,936	3.18
Right parahippocampal gyrus	30	−42	−6	3,184	2.10
Right middle temporal gyrus	42	−80	24	784	1.50
Right precuneus	18	−58	20	712	1.80
Left medial frontal	−22	2	54	600	1.01
Autobiographical					
Left medial frontal	−2	52	8	33,584	5.88
Left cuneus	0	−68	34	13,576	4.25
Right angular gyrus/TPJ	48	−66	38	8,384	2.66
Left angular gyrus	−46	−70	36	6,768	2.24
Left cerebellum	−6	−82	−14	4,920	1.79
Left fusiform gyrus	−26	−84	−18	3,376	1.90
Left middle temporal gyrus	−60	−18	−8	2,568	2.43
Right middle temporal gyrus	60	−6	−12	2,048	1.67
Right cerebellum	4	−52	−42	1,176	1.32
Left parahippocampal gyrus (anterior)	−20	−18	−19	944	2.33
Right parahippocampal gyrus (anterior)	18	−16	−10	504	1.98

xyz values represent the maximum intensity value coordinates. Volume is that which makes the cluster, in microliters, and max int represents the maximum intensity value for that volume or cluster.

and right parahippocampal gyri and the right hippocampus at $P < 0.05$ (corrected for multiple comparisons) were more active for the later items (Table 6; also see Fig. 5). There were no regions of higher activation for the earlier items. No regions within the MTL survived the chosen statistical threshold for the autobiographical categories suggesting equivalent activation at both points. No differences between early and late items were found for the nonepisodic categories, as there was an absence of activation in each.

Effect size

Percent signal change was also calculated for each time period as a numerical measure of the effect size by anatomically-defined masks of the left hippocampus and the right hippocampus and using a finite impulse response model, dividing the generation time into miniblocks and then binning these IRFs into early, late, and middle time bins. The signal was averaged for the time bins of interest (i.e. early and late time periods) for each category type and for each participant. These two time periods were then compared for each category.

Figure 6 shows that there was a significant difference in the signal for the later items compared with the early items only for the spatial category, in both the left and right hippocampus ($t_{(15)} = 2.31$, $P < 0.05$; $t_{(15)} = 2.16$, $P < 0.05$, for the left and right, respectively) but not for the autobiographical or nonepisodic categories in the left ($t_{(15)} = 0.436$, $P > 0.05$; $t_{(15)} = 1.074$, $P > 0.05$, respectively) or the right hippocampus ($t_{(15)} = 0.695$, $P > 0.05$; $t_{(15)} = 0.003$, $P > 0.05$, respectively). Also notable is the finding that the activity level for both early and late generation times for the autobiographical category was similar to that of the later generated items for the spatial categories. Conversely, the activity for the earlier generated items for the spatial categories was similar to the activity level to that of the nonepisodic categories at both the early and late generation periods.

Brain regions outside of the MTL

A whole-brain analysis was done using the same 12-s epochs (two analyses: one for early and one for late generation periods) for each category type with the threshold set at a more conservative 0.001, given that we are looking at the whole brain. As shown in Table 7, an area within the left precuneus was more active for late than early items for the spatial categories. Areas that were more involved in early item generation compared with later generated items for the spatial categories included the left precentral gyrus, left inferior parietal lobule, right cerebellum, left superior frontal gyrus, and left fusiform gyrus. While there were no areas that were specifically involved in late item generation for the autobiographical categories, there were regions that were specifically involved in early item generation. These included regions in the right superior frontal lobe, left middle frontal gyrus, left fusiform gyrus, and the right superior temporal gyrus. Nonepisodic categories also did not show areas that were more active for later compared with earlier generated items, but did show areas that were more involved in early item generation such as regions in the left precentral gyrus, left inferior frontal gyrus, and left inferior parietal lobule.

TABLE 6.

Brain Regions Within the Medial Temporal Lobes Activated for Late > Early Time Periods by Each of the Four Category Types (P < 0.05)

Region	L/R	x	y	z	Volume	Max int
Hippocampus	R					
Autobiographical						
Spatial		32	−26	−10	168	0.88
Nonepisodic						
Parahippocampal gyrus	L					
Autobiographical						
Spatial		−26	−44	−10	3,056	0.83
Nonepisodic						
Parahippocampal gyrus	R					
Autobiographical						
Spatial		22	−40	−10	1,400	0.48
Nonepisodic						

xyz values represent the maximum intensity value coordinates. Volume is that which makes the cluster, in microliters, and max int represents the maximum intensity value for that volume or cluster.

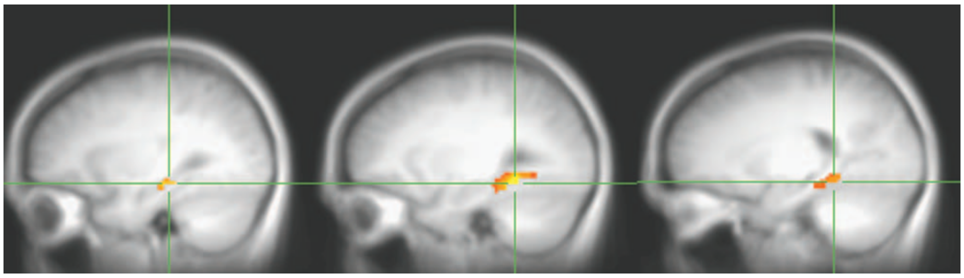

FIGURE 5. A ROI group analysis of the right hippocampus (leftmost pane, 32 −26 −10) and left (middle pane, −26 −44 −10) and right parahippocampal gyrus (rightmost pane, 22 −40 −10) using AFNI template comparing late > early generated items for the spatial categories. *X-Y-Z* coordinates of the intersection is the voxel of maximum intensity. The cluster shown is significant at $P < 0.05$, SVC. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

DISCUSSION

The present results suggest that even on tests that are ostensibly semantic, there are contributions from both semantic and episodic memory processes. As predicted, and based on the results of Experiment 1, the extent to which the MTL and the hippocampus in particular was activated during performance on verbal fluency tests varied with type of category and the time-course of item generation. Involvement of the MTL and related structures in the autobiographical network was evident throughout the autobiographical category tasks which, Experiment 1 showed, drew on distinct information from the beginning. For the spatial category tests, similar levels of activation in those regions were evident only at later stages when responses became progressively more distinctive, indicative of more open-endedness. Activation of the MTL and related structures was minimal or absent on category fluency tests, such as the nonepisodic tasks and the early stages of the spatial task, that drew more on information from semantic, rather than episodic, memory, reflecting the closed-ended nature of the tasks. The results will be discussed in more detail, first in terms of overall activity summed over the entire generation time for each type of category, and then in terms of early and late item generation periods within each category type.

Category Type

When examining the overall MTL activity involved in generating exemplars for various types of categories, those that drew upon information from episodic memory (*autobiographical* and *spatial/context* categories) activated both the hippocampus and the parahippocampal cortex. Looking at the activation across the entire generation time, the spatial/context categories (e.g. kitchen utensils) activated the hippocampus bilaterally and more extensive regions of the left and right posterior parahippocampus cortex than did the autobiographical categories, which activated only the left hippocampus and small areas of the left parahippocampal cortex.

Aside from the hippocampus, the different category types also activated other brain regions. Categories that required the retrieval of autobiographical information (autobiographical categories) activated areas that are implicated in self-processing and self-perspective taking, such as the medial frontal regions and an area close to the temporoparietal junction (TPJ) (Saxe et al., 2006; Addis et al., 2007: Muscatell et al., 2010). Generating items for spatial/context categories activated regions in the posterior right and left parahippocampal gyri and a region in the left precuneus. Activation of these areas has been found during spatial navigation and retrieval of episodic spatial relations (Fletcher et al., 1995; Parslow et al., 2004; Rosenbaum et al., 2004) as well as during scene construction (Hassabis and

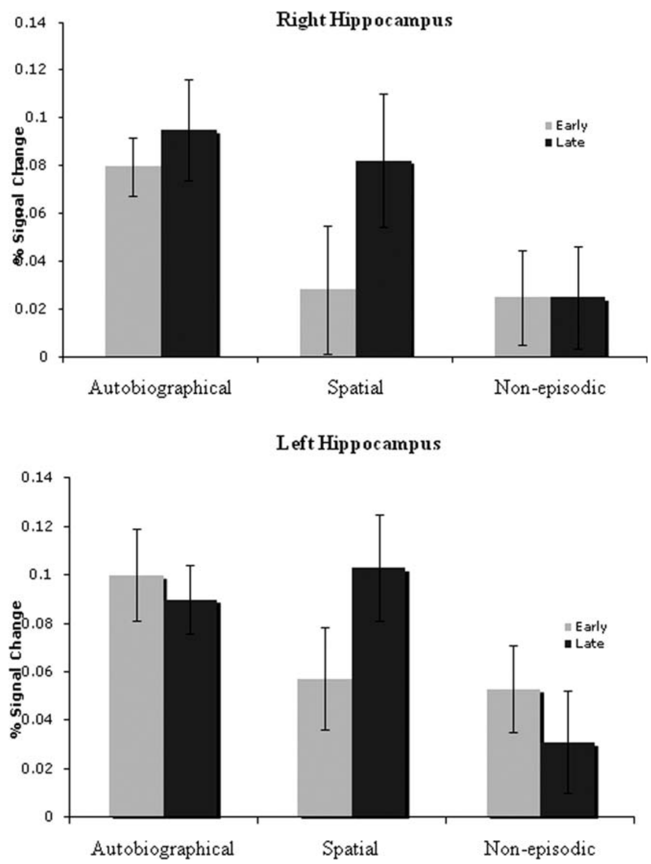

FIGURE 6. The average percent signal change across all 16 participants in the left and the right hippocampus for initial and final items. This percent change is shown for autobiographical, spatial, and nonepisodic categories. Standard error bars are shown.

TABLE 7.

Brain Regions Outside of the MTL That are Specifically Activated by Each of the Four Category Types in Early and Late Item Generation Time Periods (P < 0.001)

Category type	Time period	*x*	*y*	*z*	Volume	Max int
Region						
Autobiographical	Early					
Left declive/fusiform gyrus		−48	−56	−18	20,416	2.81
Left precentral gyrus		−56	10	4	16,216	3.08
Right superior temporal gyrus		62	44	10	5,600	1.75
Left superior frontal gyrus		−2	−4	66	3,680	3.15
Right inferior frontal gyrus		38	20	−6	2,960	1.48
		48	6	30	528	0.77
Right cerebellum		20	−72	−44	2,920	1.96
		30	−70	−16	2,048	2.87
Right inferior occipital gyrus		40	−86	0	2,384	1.08
Left middle frontal gyrus		−42	46	24	1,440	2.07
Right precentral gyrus		48	2	48	696	1.56
Right lentiform nucleus		18	8	4	656	1.18
Autobiographical	Late					
		–	–	–	–	–
Spatial	Early					
Left precentral gyrus		−56	6	8	6,944	2.55
		−54	0	44	3,416	2.49
Left inferior parietal lobule		−58	−32	48	4,008	1.76
Right cerebellum		26	−64	−48	2,536	1.48
		22	−70	−18	880	1.98
Left superior frontal gyrus		−2	−4	66	2,032	2.93
Left fusiform gyrus		−44	−56	−12	1,112	1.56
Right middle frontal gyrus		44	46	16	872	1.69
		2	34	40	808	1.00
Right inferior parietal lobule		46	−40	58	864	2.10
Spatial	Late					
Left precuneus		−12	−56	24	1,232	1.07
Nonepisodic	Early					
Left precentral gyrus		−56	−2	44	1,784	1.78
Left inferior frontal gyrus		−56	8	22	984	1.28
Left inferior parietal lobule		−54	−32	34	704	1.03
Left superior frontal gyrus		−4	−4	64	608	1.70
Left superior parietal lobule		−32	−58	50	496	1.11
Nonepisodic	Late					
		–	–	–	–	–

xyz values represent the maximum intensity value coordinates. Volume is that which makes the cluster, in microliters, and max int represents the maximum intensity value for that volume or cluster.

Maguire, 2007). In fact, the distinct roles of autobiographical and spatial retrieval were reflected even within the hippocampus: as already noted the spatial/context categories involved both the left and right hippocampi and posterior regions of the parahippocampal cortex bilaterally whereas the autobiographical categories activated the left hippocampus and regions in the parahippocampal cortex, namely the left. In fact, the laterality of activation between the autobiographical and spatial categories fit nicely with theories of left and right MTL function (Milner, 1971), with the right parahippocampus being more active during the spatial category task (Burgess et al., 2002), and the left hippocampus more active on the autobiographical task (Maguire, 2001; Gilboa, 2004).

There was no significant involvement of the hippocampus for the phonemic and nonepisodic categories at a comparable threshold to the other categories. At a lower threshold, however, regions in the left hippocampus and posterior parahippocampal cortex were active for the nonepisodic categories. There are two plausible explanations for this MTL activity: first, it may be that the MTL activity found for the nonepisodic categories reflect a base level of incidental encoding. Second, the MTL activity may indicate that even for some of the categories that were ostensibly nonepisodic, participants, nonetheless, occasionally resorted to retrieving information from episodic memory. For example, if a participant was taking a criminology course, then he or she may recall what was learnt in class that day to generate items for the category *Crimes.* Even so, less MTL activity was found for these nonepisodic categories compared with the autobiographical and spatial categories. The nonepisodic categories also activated regions in the middle temporal gyrus (Thompson-Schill, 2002; Lambon-Ralph and Patterson, 2008) and areas in the angular gyrus, which have been implicated in previous semantic retrieval tasks (e.g. Binder et al., 1999).

The absence of MTL involvement in phonemic fluency was not surprising given that other investigators have found that phonemic fluency relies more on frontal, executive control regions than category fluency (Milner and Petrides, 1984; Baldo et al., 2006). There was significant activity in left lateralized frontal regions, particularly the inferior frontal gyrus, and posterior cortical regions, such as the left superior parietal lobule, for the phonemic categories. These regions have been implicated in phonemic fluency tasks in past investigations (Phelps et al., 1997; Baldo et al., 2006; Whitney et al., 2009) and are thought to be implicated in strategic retrieval and shifting (see Troyer et al., 1997).

These particular findings are consistent with those of previous studies that have found a distinction between category fluency and phonemic fluency. Baldo et al. (2006) found that while category fluency activated temporal cortices, phonemic fluency relied on frontal cortices (with the parietal cortices implicated in both). More recently, Whitney et al. (2009) found that category, associative (e.g. list associated words to *tree)*, and phonemic fluency all activated medial frontal cortices, but there was significant activity in the MTL (the hippocampus in particular) only for the category fluency and associative task, which may reflect participants accessing related past episodes. Our results are consistent with findings from patients with selective frontal and temporal lesions or degeneration (e.g. Newcombe, 1969; Troyer et al., 1998).

In fact, neurodegenerative conditions, such as Alzheimer's disease and semantic dementia both result in a deficit in category fluency (Rogers and Friedman, 2008), with semantic dementia being associated with a more severe impairment and

Alzheimer's disease with a more progressive decline (Rogers et al., 2006a). In light of the current results, perhaps the decline associated with Alzheimer's dementia (more subtle deficit early on in the disease) could be due to hippocampal and anterior temporal lobe contributions to category fluency (Troyer et al., 1998; Gleissner and Elger, 2001). Hippocampal and entorhinal degeneration is associated with the early stages of Alzheimer's disease, and neocortical and specifically anterior temporal degeneration is associated with later stages of Alzheimer's disease (e.g. Fox and Schott, 2004) and with semantic dementia (Rogers et al., 2006). Early deficits in Alzheimer's disease, due to hippocampal contributions to semantic tasks, should result in less impairment; and later deficits, due to loss of the integrity of the semantic representations, should lead to more severe deficits similar to those observed in semantic dementia (Westmacott et al., 2004).

The present findings differ slightly from those of Ryan et al. (2008) on which this study was based. Although they found similar hippocampal activity across the three types of categories (autobiographical, spatial/context, nonepisodic), we found hippocampal involvement for the autobiographical and spatial/context, but not for the nonepisodic categories when using a statistical threshold of $P < 0.005$. Furthermore, right hippocampal involvement was found only for the spatial/context categories and not the autobiographical categories. Differences between our study and that of Ryan et al. is that they used fewer categories for each type and the spatial categories names that they used were often defined by personal pronouns such *you* and *your* (e.g. *things in your kitchen)*, making the task more autobiographical. Our rationale for not including such pronouns was based on the presumption that it necessarily evoked an autobiographical strategy for generating items. It may be that using such pronouns throughout the task, as Ryan et al. did, put participants in an autobiographical strategy mode throughout the experiment. This would account for the similar levels of MTL activation for all of their category types in their study and the difference between their findings and ours.

Early versus Late Item Generation

The main purpose of the present study was to look more closely at how MTL activity during category fluency may be affected by changes in processing strategies across different time periods during the task. To do so, we first reported a behavioral study that showed that distinctiveness an indicator of openendedness, was high for both early and late generation periods in the autobiographical task, but changed from low to high as the spatial task progressed. In our fMRI experiment, we predicted that MTL activity would be related to distinctiveness as the latter reflects the open-endedness of the task that is determined by idiosyncratic responses based on information in episodic memory. Activity during early and late item generation was contrasted for the three category types (*autobiographical categories, nonepisodic categories*, and *spatial/context categories*) using phonemic fluency as a baseline to control for changing levels in difficulty and item generation across time.

As predicted, the MTL (right hippocampus and left and right parahippocampal cortex) was significantly more active for later generated items compared with earlier generated items for the spatial/context categories. The degree of activation during later generated spatial/context items was comparable to the BOLD signal during autobiographical item generation at both early and late time points. The signal change during early spatial/context item generation was, in contrast, similar to that of nonepisodic items at both time points. This suggests that initially generated spatial/context items made little processing demands on the hippocampus as did the nonepisodic items, whereas later generated spatial/context items made similar processing demands on the hippocampus as the autobiographical items. The region that was active for the later spatial categories was located posterior consistent with observations that these regions typically are active during spatial navigation tasks. Such findings are in line with observations about the division of labor within the hippocampus, with spatial relations during nonmnemonic tasks activating posterior regions of the MTL (hippocampus and parahippocampal cortex) compared with nonspatial semantic relations that are thought to activate more anterior regions of the MTL, and the hippocampus specifically (see Ryan et al., 2010). The increase in MTL activation for later spatial/context category items was also accompanied by greater activity in the left precuneus. As noted above, the precuneus has been thought to be involved in tasks that require the inspection of mental imagery (Cavanna and Trimble, 2008) and scene construction (Hassabis and Maguire, 2007). Thus, the concurrent activation of these regions suggests that some assessment of spatial environments or construction of a scene was involved for later items compared to early items. No regions were more active for later generated items compared with earlier generated items for the autobiographical and the nonepisodic categories, likely because changes in strategy with time were not useful in these situations; in the former case, participants stuck with an episodic strategy throughout, and in the latter case, they never adopted one.

In terms of the early > late contrast, all category types did show greater activation in left precentral gyrus and in a frontoparietal network for early compared with late item generation. Left precentral gyrus activation likely occurred because more items were generated for initial periods than later periods (i.e. more button pressing) and this region is associated with motor movement. Also, the frontoparietal network activation is likely related to task initiation (see Dosenbach et al., 2007). Greater activation for early items was found in bilateral frontal and temporal regions for the autobiographical categories but in left frontal and parietal regions for nonepisodic categories.

One may speculate that the greater MTL activity found for late spatial/context items represents up-regulation of the default mode network; as generation times slow down for later items, participants' mind may tend to wander more. A related possibility is that when generation times slow down there is more time for elaborative processing of an item, which is supported by the episodic memory network, and specifically, the MTL. We do not think these are likely interpretations. If either were the case, we should also expect to see greater late-onset activity in the same networks for the nonepisodic and phonemic categories, but we did not. Nonetheless, to control for these default mode effects, and to also control for changing levels of difficulty between early and late time periods, we used phonemic fluency as a baseline. Participant's rated phonemic fluency equally as difficult compared with the three other category types. If there was a late onset default network involved when item generation slows down, comparison with phonemic fluency should eliminate the late-onset activation observed in the spatial/context condition, especially considering the same number of items was generated for these two category types. The late onset MTL activation, however, survived such a comparison. Furthermore, the use of phonemic fluency as a baseline rules out the explanation that the greater activation in the MTL and related structures for later items is due to greater proactive interference during item generation.

Another possible interpretation is that the MTL activity reflects incidental encoding of the items. However, if this were the case, then we should see also increased MTL activity for the early > late generation period contrast across all category conditions given that more items were typically generated in the earlier periods. Alternatively, it is possible that incidental encoding applies primarily to those items which were personally meaningful. If this were so, we would expect subsequent memories of the generated items to reflect the extent to which they drew on autobiographical or episodic memory. Thus, memory for generated exemplars should be greatest in the autobiographical category > spatial category > nonepisodic category > phonemic category. While the recall ratio for the phonemic category was the lowest, memory for the remaining categories did not follow the predicted order. In particular the position of the spatial and nonepisodic categories was reversed, with the recall ratio for the latter being no worse than that for the autobiographical category, but also not significantly better than the spatial category. It is unlikely, therefore, that the pattern of activation reflects incidental encoding of generated items that were personally meaningful.

The results also argue against a general "spatial" interpretation of hippocampal activation. According to such an interpretation, hippocampal activation should be evident throughout the spatial task and should be greater than in the autobiographical category, which often does not carry a spatial component (e.g. things you keep in your pocket, names of your friends). In fact, the reverse was found. Nonetheless, MTL activity for later, but not early spatial/context items, may reflect participants' use of episodic information to retrieve old memories/familiar environments or to reconstruct novel scenes or scenarios to help generate items (e.g. Schacter et al., 2007; Hassabis et al., 2007).

After considering these alternatives, we believe our hypothesis provides the best explanation of the results: generating items for ostensibly semantic categories implicates the hippocampus and related structures to the extent that the categories are open-ended and can draw on information from episodic memory to benefit performance when retrieval from semantic memory is exhausted or becomes too demanding. It is important to note with regard to our hypothesis that MTL-based processes will only be useful in open-ended retrieval situations if episodically relevant information is available or useful. Thus, for situations that are seemingly open-ended, such as phonemic fluency, we would not expect, and did not observe, any MTL involvement because such a task is not autobiographically or episodically relevant. Also, one can conceive of tasks that are quasi-episodically relevant, such as thinking of items found on a space station. While one could not use strictly autobiographical information to help complete this task, one could do so by constructing imaginary scenes or scenarios, an activity that recent research has shown also depends on the hippocampus and related structures (Hassabis et al., 2007; Schacter et al., 2007).

Study Limitations

One methodological limitation of the present experiment was that the use of an episodic strategy or autobiographical/episodic relevance was based on interviews and normative data from another study and a previous behavioral study we conducted (Sheldon and Moscovitch, in preparation). In terms of our fMRI analysis, the statistical threshold was varied according to our expected level of power for the contrast that we were conducting (i.e. lower thresholds for contrasts with fewer data points in smaller volumes and higher thresholds for contrasts with a large number of data points and greater volumes). We believe that adopting this procedure, which is not uncommon in the literature, ensured that we were revealing meaningful clusters of activations which otherwise may have been lost if an arbitrarily high threshold were applied uniformly.

CONCLUSION

Our experiments suggest that episodic memory processes mediated by the MTL contribute to semantic retrieval particularly when tasks are open-ended and can draw on episodic or spatial information to support performance. When the retrieval of semantic information is sufficiently open-ended in the sense that searching through semantic networks has been exhausted or if such networks do not exist, the MTL will be involved since it mediates processes related to recalling or constructing an environment or scenario from which the needed semantic information can be derived. Overall, our results and conclusions reinforce views advanced by Barsalou (1988, 2009) and Ryan et al. (2008) that, in some cases, semantic information can be derived from information in episodic memory as well as from semantic memory. The present results also support Pillemer (2003) who noted that autobiographical memory has a directive function in everyday semantic retrieval. Following a tradition that dates back at least as far as Bartlett (1932), Pillemer (2003) has also emphasized the role that episodic memory can play in a variety of social, decision making, and problem solving tasks. This position is beginning to garner support from behavioral, lesion, and functional neuroimaging studies on the contribution of episodic memory, mediated by the MTL and related structures, to such tasks (Kumaran et al., 2009; Spreng et al., 2010, also see Mosovitch, 2008). Our study on category fluency adds to that growing body of work.

Acknowledgments

The authors thank Lee Ryan and her colleagues (Christine Cox, Scott Hayes, and Lynn Nadel) for graciously providing them with the testing materials. The authors also thank Kathy Li for her help in data collection, Marilyne Ziegler for help with programming, and the members of the thesis committee for S.S., Mary Pat McAndrews and Lynn Hasher, for advice during all stages of the experiment. This work was supported by grants from the National Sciences and Engineering Research Council of Canada to Signy Sheldon and from the Canadian Institutes of Health Research to Morris Moscovitch.

REFERENCES

Addis DR, Moscovitch M, McAndrews MP. 2007. Consequences of hippocampal damage across the autobiographical memory network in left temporal lobe epilepsy. Brain 130:2327–2342.
Baldo JV, Schwartz S, Wilkins D, Dronkers NF. 2006. Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. J Int Neuropsychol Soc 12:896–900.
Barsalou LW. 1988. The content and organization of autobiographical memories. In: Neisser U, Winograd E, editors. Remembering Reconsidered: Ecological and Traditional. New York, NY: Cambridge University Press. pp 193–243.
Barsalou LW. 2009. Simulation situated conceptualization and prediction. Philos Trans R Soc Lond B Biol Sci 364:1281–1289.
Bartlett FC. 1932. Remembering: A Study in Experimental and Social Psychology. Cambridge: University Press.
Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW. 1999. Conceptual processing during the conscious resting state: A functional MRI study. J Cogn Neurosci 11:80–93.
Binney RJ, Embleton KV, Jefferies E, Parker GJ, Lambon Ralph MA. 2010. The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: Evidence from a novel direct comparison of distortion-corrected fMRI rTMS and semantic dementia. Cereb Cortex 20:2728–2738.
Burgess N, Maguire EA, O'Keefe J. 2002. The human hippocampus and spatial and episodic memory. Neuron 35:625–641.
Cavanna AE, Trimble MR. 2008. The precuneus: A review of its functional anatomy and behavioural correlates. Brain 1293:564–583.
Cox RW. 1996. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:2–17.
Cox RW, Hyde JS. 1997. Software tools for analysis and visualization of FMRI data. NMR Biomed 10:171–178.
Crowe SF. 1998. Decrease in performance on the verbal fluency test as a function of time: Evaluation in a young healthy sample. J Clin Exp Neuropsychol 203:391–401.
Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Peterson SE. 2007. Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci USA 10426:11073–11078.
Dronkers NF, Wilkinsa DP, Van Valin RD Jr, Redferna BB, Jaegerc JJ. 2004. Lesion analysis of the brain areas involved in language comprehension. Cognition 92:145–177.
Eichenbaum H. 1997. Declarative memory: Insights from cognitive neurobiology. Annu Rev Psychol 48:547–572.
Eichenbaum H, Yonelinas AP, Ranganath C. 2007. The medial temporal lobe and recognition memory. Ann Rev Neurosci 30:123–152.
Fox NC, Schott JM. 2004. Imaging cerebral atrophy: Normal ageing and Alzheimer's disease. Lancet 363:392–394.
Fletcher PC, Frith CD, Baker SC, Shallice T, Frackowiak RS, Dolan RJ. 1995 The mind's eye–precuneus activation in memory-related imagery. Neuroimage 2:195–200.
Gilboa A. 2004. Autobiographical and episodic memory- One and the same? Evidence from prefrontal activation in neuroimaging studies. Neuropsychologia 42:226–1249.
Gleissner U, Elger C. 2001. The hippocampal contribution to verbal fluency in patients with temporal lobe epilepsy. Cortex 37:1155–1163.
Greenberg DL, Keane MM, Ryan L, Verfaellie M. 2009 Impaired category fluency in medial temporal lobe amnesia: The role of episodic memory. J Neurosci 2935:10900–10908.
Hassabis D, Maguire EA. 2007. Deconstructing episodic memory with construction. Trends Cogn Sci 11:299–306.
Hassabis D, Kumaran D, Maguire EA. 2007. Using imagination to understand the neural basis of episodic memory. J Neurosci 27:14365–14374.
Hassabis D, Kumaran D, Vann SD, Maguire EA. 2007. Patients with hippocampal amnesia cannot imagine new experiences. Proc Natl Acad Sci USA 104:1726–1731.
Kapur N, Friston KJ, Young A, Frith CD, Frackowiak RS. 1995. Activation of human hippocampal formation during memory for faces: A PET study. Cortex 31:99–108.
Karlsgodt KH, Shirinyan D, van Erp TG, Cohen MS, Cannon TD. 2005. Hippocampal activations during encoding and retrieval in a verbal working memory paradigm. Neuroimage 25:1224–1231.
Kumaran D, Summerfield JJ, Hassabis D, Maguire EA. 2009. Tracking the emergence of conceptual knowledge during human decision making. Neuron 636:889–901.
Lambon Ralph MA, Patterson K. 2008. Generalization and differentiation in semantic memory: Insights from semantic dementia. Ann NY Acad Sci 1124:61–76.
Lee AC, Bussey TJ, Murray EA, Saksida LM, Epstein RA, Kapur N, Hodges JR, Graham KS. 2005. Perceptual deficits in amnesia: Challenging the medial temporal lobe 'mnemonic' view. Neuropsychologia 431:1–11.
Lezak MD. 1995. Neuropsychological Assessment,3rd ed. New York, NY: Oxford University Press.
Maguire EA. 2001. Neuroimaging studies of autobiographical event memory. Philos Trans R Soc Lond B Biol Sci 356:1441–1451.
Maguire EA, Mummery CJ. 1999. Differential modulation of a common memory retrieval network revealed by positron emission tomography. Hippocampus 9:154–161.
Milner B. 1971. Interhemispheric differences in the localization of psychological processes in man. Br Med Bull 27:272–277.
Milner B, Petrides M. 1984. Behavioural effects of frontal-lobe lesions in man. Trends Neurosci 711:403–407.
Moscovitch M, Nadel L, Winocur G, Gilboa A, Rosenbaum RS. 2006. The cognitive neuroscience of remote episodic semantic and spatial memory. Curr Opin Neurobiol 16:2179–2190.
Moscovitch M. 2008. The hippocampus as a "stupid" domain-specific module: Implications for theories of recent and remote memory and of imagination. Can J Exp Psychol 621:62–79.
Muscatell KA, Addis DR, Kensinger EA. 2010. Self-involvement modulates the effective connectivity of the autobiographical memory network. Soc Cogn Affect Neurosci 5:68–76.

Newcombe F. 1969. Missile Wounds of the Brain: A Study of Psychological Deficits. London: Oxford University Press.

Parslow DM, Rose D, Brooks B, Fleminger S, Gray JA, Giampietro V, et al. 2004. Allocentric spatial memory activation of the hippocampal formation measured with fMRI. Neuropsychology 183:450–461.

Phelps EA, Hyder F, Blamire AM, Shulman RG. 1997. FMRI of the prefrontal cortex during overt verbal fluency. Neuroreport 82:561–565.

Pihlajamaki M, Tanila H, Hanninen T, Kononen M, Laakso M, Partanen K, Soininen H, Aronen HJ. 2000. Verbal fluency activates the left medial temporal lobe: A functional magnetic resonance imaging study. Ann Neurol 47:470–476.

Pillemer DB. 2003. Directive functions of autobiographical memory: The guiding power of the specific episode. Memory 11:193–202.

Rogers SL, Friedman RB. 2008. The underlying mechanism of semantic memory loss in Alzeheimer's disease and semantic dementia. Neuropsychologia 46:12–21.

Rogers TT, Ivanoiu A, Patterson K, Hodges JR. 2006. Semantic memory in Alzheimer's disease and the frontotemporal dementias: A longitudinal study of 236 patients. Neuropsychology 20:319–335.

Rosenbaum RS, Ziegler M, Winocur G, Grady CL, Moscovitch M. 2004. "I have often walked down this street before": FMRI studies on the hippocampus and other structures during mental navigation of an old environment. Hippocampus 14:826–835.

Ryan L, Cox C, Hayes SM, Nadel L. 2008. Hippocampal activation during episodic and semantic memory retrieval: Comparing category production and category cued recall. Neuropsychologia 46:2109–2121.

Ryan L, Lin CY, 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.

Saxe R, Moran JM, Scholz J, Gabrieli J. 2006. Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. Soc Cogn Affect Neurosci 1:229–234.

Schacter DL, Addis DR, Buckner RL. 2007. Remembering the past to imagine the future: The prospective brain. Nat Rev Neurosci 8:657–661.

Schmolck H, Kensinger EA, Corkin S, Squire LR. 2002. Semantic knowledge in patient H.M. and other patients with bilateral medial and lateral temporal lobe lesions. Hippocampus 124:520–533.

Spreng RN, Stevens WD, Chamberlain J, Gilmore AW, Schacter DL. 2010. Default network activitycoupled with the frontoparietal control network supports goal-directed cognition. Neuroimage 31:303–317.

Talairach J, Tournoux P. 1988 Co-planar Stereotaxic Atlas of the Human Brain. New York, NY: Thieme Medical Publishers Inc.

Thompson-Shill S. 2002. Neuroimaging studies of semantic memory: inferring "how" from "where". Neuropsychologia 413:280–292.

Troyer AK, Moscovitch M, Winocur G. 1997. Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. Neuropsychology 11:138–146.

Troyer AK, Moscovitch M, Winocur G, Leach L, Freedman M. 1998. Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. J Int Neuropsychol Soc 4:137–143.

Vallée-Tourangeau F, Anthony SH, Austin NG. 1998. Strategies for generating multiple instances of common and ad hoc categories. Memory 6:555–592.

Whatmough C, Chertkow H. 2007. rCBF to the hippocampal complex covaries with superior semantic memory retrieval. Behav Brain Res 181:262–269.

Whitney C, Weis S, Krings T, Huber W, Grossman M, Kircher T. 2009. Task-dependent modulations of prefrontal and hippocampal activity during intrinsic word production. J Cogn Neurosci 214:697–712.

Westmacott R, Moscovitch M. 2002. Temporally graded semantic memory loss in amnesia and semantic dementia: Further evidence for opposite gradients. Cogn Neuropsychol 192:135–163.

Westmacott R, Moscovitch M. 2003. The contribution of autobiographical significance to semantic memory. Mem Cognit 315:761–774.

Westmacott R, Black SE, Freedman M, Moscovitch M. 2004. The contribution of autobiographical significance to semantic memory: Evidence from Alzheimer's disease, semantic dementia and amnesia. Neuropsychologia 42:25–48.