IMPAIRED CATEGORY FLUENCY IN MEDIAL TEMPORAL LOBE AMNESIA: THE ROLE OF EPISODIC MEMORY

Daniel L. Greenberg[a,*], **Margaret M. Keane**[a,b], **Lee Ryan**[c], and **Mieke Verfaellie**[a]

[a] Memory Disorders Research Center, 150 S. Huntington Ave (151-A), VA Boston Healthcare System, Boston University, Boston, MA, 02130

[b] Psychology, Wellesley College, Wellesley, MA

[c] Cognition & Neuroimaging Laboratories, Psychology Department, University of Arizona, Tuscon, AZ

Abstract

Memory tasks are often classified as semantic or episodic, but recent research shows that these types of memory are highly interactive. Category fluency, for example, is generally considered to reflect retrieval from semantic memory, but behavioral evidence suggests that episodic memory is also involved: Participants frequently draw on autobiographical experiences while generating exemplars of certain categories. Neuroimaging studies accordingly have reported increased medial temporal lobe (MTL) activation during exemplar generation. Studies of fluency in MTL amnesics have yielded mixed results but were not designed to determine the precise contributions of episodic memory. We addressed this issue by asking MTL amnesics and controls to generate exemplars of three types of categories. One type tended to elicit autobiographical and spatial retrieval strategies (AS). Another type elicited strategies that were autobiographical but nonspatial (AN). The third type elicited neither autobiographical nor spatial strategies (N). Amnesic patients and control participants generated exemplars for 8 categories of each type. Patients were impaired on all category types but were more impaired on AS and AN categories. After covarying for phonemic fluency (total FAS score), the N category impairment was not significant, but the impairment on AS and AN categories remained. The same results were obtained for patients with lesions restricted to the MTL and those with more extensive lesions. We conclude that patients' episodic memory impairment hindered their performance on this putatively semantic task. This interaction between episodic and semantic memory might partially account for fluency deficits seen in aging, mild cognitive impairment, and Alzheimer's disease.

Keywords

AMNESIA; EPISODIC; SEMANTIC; MEDIAL TEMPORAL; FLUENCY; MEMORY

Introduction

Theories of memory distinguish between episodic and semantic memory (Tulving, 1972). Episodic memory refers to memory for personally experienced events that occurred at a particular time and place, while semantic memory refers to general world knowledge that is not linked to specific spatiotemporal contexts. Most research examines each type of memory in isolation, but other work shows that they frequently interact (e. g. Westmacott et al., 2004;

*Corresponding author; dlg@bu.edu.

Kan et al., in press). This study investigated the manner in which episodic memory facilitates retrieval from the semantic knowledge base.

Category fluency tasks are usually considered semantic (Strauss et al., 2006), but behavioral studies show that episodic memory is also involved. Participants frequently use episodic memory to search for category exemplars: When naming makes of cars, for instance, they may imagine a familiar parking lot or list cars they themselves have owned (Walker & Kintsch, 1985). Many categories can evoke episodic memory, though the degree of episodic involvement can vary from category to category (Vallée-Tourangeau et al., 1998).

If category fluency is aided by episodic memory, then it might draw on the medial temporal lobe (MTL) structures on which episodic memory relies. Neuroimaging studies have indeed reported MTL activation during fluency tasks (Frith et al., 1991; Mummery et al., 1996; Gourovitch et al., 2000; Pihlajamäki et al., 2000; Greenberg et al., 2005; Hirshorn & Thompson-Schill, 2006; Ryan et al., 2008; but see Paulesu et al., 1997). Recently, one of us combined behavioral and neuroimaging evidence by comparing activation across three types of categories that drew on different kinds of memory (Ryan et al., 2008). All category types elicited equivalent activity in hippocampus proper. They also elicited activity in parahippocampal gyrus, but the activity was higher for categories that evoked familiar spatial contexts. The degree and location of MTL involvement thus depended on the type of category that was presented.

Patients with MTL damage have a severe episodic memory impairment and relatively preserved premorbid semantic memory (Schmolck et al., 2002), but evidence for a fluency deficit is mixed. Gleissner and Elger (2001) examined patients with unilateral damage limited to either the hippocampus or lateral temporal regions. They found that patients with hippocampal damage were impaired on fluency tasks regardless of lesion side, while patients with lateral damage were only impaired when they had left-hemisphere lesions. Schmolck and colleagues (2002), however, reported that patients with bilateral damage limited to the hippocampal formation performed normally on fluency tasks, though patients whose damage also involved lateral temporal regions were impaired.

Although these studies provided further evidence for the contribution of lateral temporal lobes to fluency tasks, they did not resolve the role of the MTL in category fluency and were not designed to compare categories that drew on episodic memory to different degrees. We therefore tested MTL amnesics and controls on three kinds of categories that varied in their reliance on episodic memory. By comparing performance across these categories, we sought to identify the circumstances in which episodic memory contributes to performance on a semantic task.

Materials and Methods

Pilot Study

We conducted a pilot study to identify categories that tended to elicit different types of retrieval strategies. We selected a broad range of candidate categories from a variety of sources (Battig & Montague, 1969; Barsalou, 1983; Vallée-Tourangeau et al., 1998; Van Overschelde et al., 2004; Ryan et al., 2008). Thirty-one participants (mean age=60.7, mean education=15.1, mean VIQ=103.6) each generated exemplars for 18–22 categories. They read the category name on a computer screen and were given 60 seconds to generate as many exemplars as they could; the experimenter recorded their answers using a separate sheet for each category. After participants had generated exemplars for all of the categories, they were presented with the answer sheets and asked to describe the thought processes they had used for each category. (An alternative approach would have involved asking participants about their strategy for a

category immediately after they had generated exemplars for it; we decided against this design because we did not want participants' response to one category to influence the strategy they used on the next.)

Participants' descriptions of their thought processes were classified into the three types that were developed by Ryan and colleagues (2008). If the participant described a strategy with a self-referential and spatial component (for example, "I imagined my bedroom, looked around, and described what I saw"), then it was classified as *autobiographical spatial* (AS). If the participant described a self-referential strategy without a spatial component ("I thought of what my sisters and I are giving our father for his 90th birthday"), then it was classified as *autobiographical nonspatial* (AN). If the participant described a nonautobiographical strategy ("I tried to go in alphabetical order") or no strategy ("I just said whatever came to mind"), then it was classified as *neither* (N). If necessary, a description could be given more than one classification (for example, "The first few items just popped into my head, and then I imagined my own living room" would be classified as N and AS).

We used these data to calculate the frequency with which each category elicited each of the three strategies. Based on these frequencies, we assigned categories to one of three types (AS, AN, or N). Each type consisted of 8 categories that usually elicited that strategy. Specifically, a category was assigned to a category type if it elicited that strategy in more than 50% of participants, but elicited other strategies in fewer than 50% of participants. Thus, *things in a bedroom* was classified as AS because 95% of pilot participants used that strategy on that category, while only 5% used other strategies. These category assignments are presented in Table 1.

Experiment

Participants—Participants consisted of 9 patients with amnesia resulting from anoxia or encephalitis and 10 controls matched on age, verbal IQ, and education. Controls were neurologically intact and had no history of medical conditions that could affect the MTL (e. g. stroke, transient ischemic attack, head injury with loss of consciousness, cardiac arrest, seizure disorder, psychiatric conditions, and uncontrolled diabetes or hypertension). Patients were considered amnesic if they had a General Memory score that was 20 or more points below their verbal IQ score. The patients had a mean age of 60.1 (SD = 12.6), an average of 15.0 years of education (SD = 2.3), and an average verbal IQ of 105.6 (SD = 19.6). The controls had an average age of 63.5 (SD = 9.3), an average of 14.5 years of education (SD = 2.5), and a mean verbal IQ of 107.0 (SD = 18.7). There were no differences between groups in terms of age, years of education, or verbal IQ (all ps > 0.5).

Table 2 presents the demographics, neuropsychological information, etiology, and extent of brain damage of the patients. Two anoxic patients (MTL01 and MTL05) were unable to undergo MRI scanning because of medical contraindications (a cardiac defibrillator and a pacemaker, respectively). For one of these patients (MTL01), CT scan showed an incidental lesion of the putamen but no visible damage to any regions associated with memory, including the MTL; for the other patient (MTL05), no scan was available. In both of these cases, damage was inferred based on the anoxic etiology. In an autopsy study, Rempel-Clower and colleagues (1996) showed that patients with amnesia secondary to anoxia had MTL damage, but did not have significant damage to any other memory-associated regions. We therefore concluded that MTL damage secondary to anoxia was the most likely source of these patients' memory deficits (see Smith & Squire, 2005, for a similar rationale). For the remaining seven patients (MTL02, MTL03, MTL04, MTL06, MTL07, MTL08, MTL09), damage was determined based on MRI or CT scans that were interpreted by a board-certified neurologist. Reconstructions of these patients' lesions are shown in Figure 1. In addition, volumetric measurements of the temporal lobe were available for 5 patients (MTL03, MTL04, MTL06, MTL07, and MTL08), each with

reference to 9 age-and gender-matched controls. Temporal lobe structure volumes were measured using established procedures (see Kan et al., 2007, for details), and adjusted for total intracranial volume. The patients' corrected volumes were compared to the control means using Z-scores, and regions with $Z < -2$ were considered abnormal. The results for the MTL are summarized in Table 3, and a fuller description of the participants' lesions is given below.

MTL02 (anoxia): CT scan revealed significant damage to the right hippocampal formation and possible involvement of the right amygdala and perirhinal regions. No other brain regions appeared to be damaged.

MTL03 (anoxia): Volumetric analysis and examination of the MRI scan revealed bilateral damage limited to the hippocampal formation. No other damage was evident.

MTL04 (anoxia): Volumetric analysis revealed significant damage to the right hippocampal formation, and examination of the MRI scan indicated some atrophy of the left hippocampal formation as well. The scan also revealed a small incidental lesion to the posterior left putamen.

MTL06 (encephalitis): Volumetric analysis revealed bilateral damage to the hippocampal formation, the amygdala, the entorhinal cortex, and the perirhinal cortex; the left temporal pole was also severely affected, as was the posterior portion of the left parahippocampal cortex. The left insula was significantly reduced in size, as was the left cingulate. Inspection of the MRI scan also suggested bilateral damage to the fusiform gyrus and some involvement of the right temporal pole.

MTL07 (encephalitis): Volumetric analysis indicated bilateral damage to the hippocampus, amygdala, entorhinal cortex, and perirhinal cortex. The right temporal pole was significantly damaged, as was the posterior portion of the right parahippocampal cortex. In addition, the left and right insula were significantly reduced in size along with the right planum polare. The anterior portions of the middle temporal gyrus, inferior temporal gyrus, and fusiform gyrus were damaged in both hemispheres. Inspection of the MRI scan also indicated damage to the septal region and the anterior cingulate.

MTL08 (anoxia; left temporal lobectomy secondary to epilepsy): Volumetric analysis indicated significant bilateral atrophy of the hippocampal formation. In the left hemisphere, the temporal pole, the amygdala, the perirhinal cortex, and the entorhinal cortex were also significantly affected. In addition, in the lateral portions of the left temporal lobe, the anterior aspects of the fusiform and the superior temporal gyrus are reduced in size, as are the entire middle and inferior temporal gyri.

MTL09 (encephalitis): Examination of the CT scan revealed extensive bilateral damage to the temporal pole, the perirhinal and entorhinal cortex, the hippocampal formation, and the amygdala. In the left hemisphere, the insula, the inferior parahippocampal gyrus, the basal forebrain, the septum, and the deep white matter of the frontal lobes were also affected.

Experimental Design—The experiment was divided into three sessions that were spaced about two weeks apart. The first two sessions provided the main data for the experiment. In each of these sessions, participants generated exemplars for 14 categories. The first two trials were practice trials; they used categories that had often been given in other studies (specifically *animals, fruits and vegetables, tools,* and *jobs*). The next 12 trials used the categories that had been selected in the pilot experiment; participants were given 4 categories per category type per session. The order of presentation was randomized.

Participants were allowed 60 seconds to generate exemplars for each category. The experimenter recorded their answers and marked off 15-second intervals. If a participant appeared to lose set or run out of exemplars, the experimenter prompted them by asking "Can you think of any more examples of [*category name*]?" Intrusions and repetitions, which were rare, were not counted; including them did not change the results. As in the pilot experiment, participants were asked about their thought processes only after they had generated exemplars for all categories (i.e., at the end of each session).

This design might seem problematic in the case of the amnesic patients, who might be expected to forget what their thought processes had been. The third session of the experiment was designed to examine this possibility. Participants were retested on six categories (one AS, one AN, and one N category from each of the first two sessions). This time, however, they described their strategies immediately after completing each category. Their reports were quite consistent with those they gave at the end of the previous sessions: Controls endorsed the same strategy 83% of the time, while patients did so 78% of the time; there was no evidence for a difference between groups (Fisher's exact test, $p = 0.48$). Thus, participants tended to endorse the same strategies whether they were questioned immediately after each category (when their reports are as accurate as they can be) or at the end of the entire session.

We examined the control data to reassess the assignment of categories to category types. We found that two categories, *weather phenomena* and *insects*, had tended to elicit an AN strategy in the pilot study but more often elicited an N strategy in the actual experiment. We excluded these categories from the analysis; including them did not change the overall pattern of results.

Results

Experimenter-assigned strategies

The first set of analyses used the category assignments described above. Figure 2 presents the mean number of exemplars that the patients and controls generated for each category type (for comparison purposes, we also plot the results for the 3 MTL-only patients for whom imaging data were available.) Different category types yielded different numbers of exemplars, as shown by a main effect of category type ($F(2,17) = 19.53$, $p < 0.0001$). The patients were generally impaired at the task, as shown by a main effect of participant group ($F(1,17) = 13.84$, $p = 0.002$). Pairwise analyses showed that the patients were significantly impaired relative to the controls on all category types. For the AS categories, patients generated an average of 4.9 fewer items (95% CI = 2.3 – 7.4; $t(17) = -3.99$, $p = 0.0009$). For the AN categories, they generated 5.1 fewer items (95% CI = 2.3 – 7.9; $t(17) = -3.78$, $p = 0.002$). For the N categories, they generated 3.2 fewer items (95% CI = 0.7 – 5.7; $t(17) = -2.66$, $p = 0.02$). The significant participant group × category type interaction ($F(2,17) = 4.83$, $p = 0.02$), however, demonstrated that the extent of the impairment was not the same for all three category types. To examine this difference further, we conducted three 2×2 ANOVAs with participant group as one factor and one pair of categories (i.e., AS vs. N, AN vs. N, or AS vs. AN) as the other factor. As in the overall model, a significant interaction was taken as evidence for a differential impairment by category type. The interaction was significant for the AS vs. N comparison ($F(1,17) = 6.80$, $p = 0.02$) and the AN vs. N comparison ($F(1,17) = 6.17$, $p = 0.02$), in both cases showing that the patients were less impaired on the N categories; however, the interaction was not significant for the comparison of the AS and AN types ($F(1,17) = 0.08$, $p = 0.79$). Thus, the patients were most impaired with the categories that tended to evoke autobiographical strategies, but there was no evidence that one autobiographical category was associated with a greater impairment than the other. We then examined the interaction in a different way by testing for an effect of category type within each participant group. The effect of category type was significant for the controls ($F(2,17) = 22.65$, $p < 0.0001$). Pairwise comparisons showed that controls generated more exemplars for the AS and AN categories compared to the N categories: For the AS vs.

N comparison, the mean difference was 2.8 exemplars (95% CI = 1.9 – 3.8; $t(17) = 6.26$, $p < 0.0001$), and for the AN vs. N comparison, the mean difference was 2.4 exemplars (95% CI = 1.3 – 3.5; $t(17) = 4.47$, $p < 0.0001$). There was no statistically significant difference between the AS and AN categories in the control group (on average, the controls generated 0.43 more exemplars for the AS categories; 95% CI = −0.7 – 1.6; $t(17) = −0.75$, $p = 0.46$). There was a trend towards an effect of category type for the patients ($F(2,17) = 2.76$, $p = 0.09$). Further analysis revealed no significant differences between the AS and AN categories (with patients generating an estimated 0.7 more exemplars for the AS categories; 95% CI = −0.5 – 1.8; $t(17) = −1.19$, $p = 0.24$) and no significant differences between the AN and N categories (with an 0.5 exemplar advantage for the AN categories; 95% CI = −0.7 – 1.6; $t(17) = 0.84$; $p = 0.41$). There was a trend towards better performance on the AS categories than the N categories (by 1.1 exemplars; 95% CI = 0.0 – 2.2; $t(17) = 2.03$; $p = 0.0504$). As indicated above, however, this advantage was smaller in patients than it was in controls. These results are consistent with the idea that the controls were able to use episodic memory to facilitate performance on the autobiographical categories while the patients had little or no ability to do so.

To further characterize the nature of the deficit, we examined the time course of responses for each participant group and category type. As noted above, the experimenter marked off 15-second intervals during generation; therefore, the data are plotted and analyzed as the cumulative total for each 15-second time period (Figure 3). As expected, these analyses yielded a significant main effect of time on the number of exemplars that were generated ($F(3,51) = 320.93$, $p < 0.0001$). They also revealed a trend towards an interaction between time and category type ($F(6,102) = 1.96$, $p = 0.08$). Further analysis of this trend suggested that the generation rate was fastest for the AS categories (mean ± SE = 2.2 ± 0.3 exemplars/15 sec), followed by the AN categories (2.0 ± 0.3 exemplars/15 sec) and then the N categories (1.7 ± 0.2 exemplars/15 sec), though these differences were not statistically significant. We did not find a time × participant type × category type interaction ($F(6,102) = 0.35$, $p = 0.91$). We did, however, find a time × participant type interaction ($F(3,51) = 23.00$, $p < 0.0001$), showing that the rate of generation differed for the patients and the controls. Patients generated exemplars at 1.4 items per 15 sec period (SE = 0.2), while the control rate was 2.5 items per 15 sec (SE = 0.2).

As before, these results are consistent with the idea that the patients' episodic memory deficit impaired their ability to generate category exemplars. At the same time, however, cognitive slowing is a common consequent of brain injury (Lezak, Howieson, & Loring, 2004); thus, slowing could have an effect on the patients' performance as well. To examine a possible effect of response speed, we compared performance on categories which tended to elicit a higher rate of exemplar generation ("high-fluency" categories, as defined by number of exemplars generated in 1 minute) to categories which tended to elicit a slower rate of exemplar generation ("low-fluency" categories). We separated each category type (AS, AN, N) into high-fluency and low-fluency subgroups by conducting a median split on the average number of exemplars generated by the controls. If the patients' impairment were a consequence of slower speed, then their impairment should have been greater on the high-fluency categories. To the contrary, we did not find a participant group × fluency level interaction ($F(1,13) = 2.57$, $p = 0.13$). Thus, differences in speed do not account for the patients' impairment on the autobiographical categories.

In addition, we tested for a general fluency impairment by giving all participants the FAS test (Benton & Hamsher, 1976), a measure of phonemic fluency in which participants are given 60 seconds to generate as many words as they can that begin with a particular letter. The FAS test draws on many of the same cognitive processes as semantic fluency does (such as effortful retrieval, self-monitoring, and inhibition), but places fewer demands on semantic memory, particularly semantic category knowledge (see Henry et al., 2004, for a discussion). As Table

2 shows, the majority of patients were at or below the 50th percentile, though none was significantly impaired. We therefore reanalyzed the data with the total FAS score as a covariate (Figure 4). We again found a main effect of category type (F(2,16) = 19.53, $p < 0.0001$), a main effect of participant group (F(1,16) = 5.06, $p = 0.04$), and a participant group $\times$ category type interaction (F(2,16) = 4.83, $p = 0.02$). Pairwise comparisons on adjusted means again showed that the patients generated fewer exemplars for each category type. For the AS categories, patients generated an average of 3.4 fewer items (95% CI = 0.5–6.3; $t(16) = -2.49$, $p = 0.004$). For the AN categories, the mean difference was 3.7 items (95% CI = 0.6–6.7, $t(16) = -2.57$, $p = 0.02$). The difference in score for the N categories, however, was not significant in this analysis (the mean difference was 1.7 items, $-1.1 - 4.5$; $t(16) = -1.30$, p = 0.21). Thus, once verbal fluency was taken into account, we found a selective impairment for the autobiographical categories.

In addition, we selected a subset of categories half from each type such that control performance was equated across category types. Analyses revealed that the patients generated significantly fewer exemplars for the autobiographical categories than for the N categories. Similarly, when we compared their performance to that of controls, the patients were more impaired on the autobiographical categories than on the N categories. Therefore, our observation of a selective impairment of the autobiographical categories is not an artifact of differential performance across category types on the part of the controls; it was observed even when performance in the control group was equated.

The patients are generating fewer exemplars than are controls for the autobiographical categories, but what kind of exemplars are they leaving out? By definition, episodic memory draws on particular spatiotemporal contexts that can be unique to each individual, so the patients and controls might therefore differ in the prototypicality of the exemplars they generate. Specifically, the patients might still be able to generate prototypical exemplars, such as utensils that are common to most kitchens (pots, pans, knives), but they might have more difficulty generating exemplars that are less prototypical and particular to their own kitchen (zesters, corkscrews, ice picks). To examine this possibility, we returned to the pilot data and generated prototypicality measures for each exemplar (defined as the frequency with which an exemplar had been given for a particular category). We then applied these measures to the exemplars that were given in the experiment itself (if an exemplar was given in the experiment but not the pilot, it was assigned a prototypicality rating of 0). We then conducted a median split for each category type; exemplars above the median were considered high-prototypical while those below the median were considered low-prototypical. For the AS categories, chi-squared analyses showed that the patients tended to give a greater proportion of high-prototypicality exemplars than did control participants ($\chi^2 = 4.53$, $p = 0.03$), but this was not the case for the AN categories ($\chi^2 = 0.98$, $p = 0.32$) or the N categories ($\chi^2 = 1.60$, $p = 0.21$). Thus, at least for the AS categories, the patients' episodic memory impairment was associated with greater difficulty generating exemplars that were less prototypical.

Self-reported strategies

We recoded and reanalyzed the data according to the participants' self-reported strategies in order to ensure that between-group and interindividual differences in strategy use were not affecting the results. As an example, consider a participant who endorsed an N strategy for *things in a garage*. In the previous analysis, the participant's result for this category would have been included in his or her AS score (based on classification from the pilot study); in this analysis, it was included in his or her N score. Trials that elicited multiple strategies were excluded. The results of this analysis are presented in Figure 5. As before, analyses showed a main effect of participant group (F = 12.82, $p = 0.002$), a main effect of category type (F = 6.59, $p = 0.008$), and a participant group $\times$ category type interaction (F = 6.26, $p = 0.009$).

Pairwise comparisons again showed that the patients scored lower than the controls on all category types. For the AS categories, the mean difference was 5.4 exemplars (95% CI = 2.7 – 8.0, $t(17) = -4.21$, $p = 0.0006$). For the AN categories, the mean difference was 4.4 exemplars (95% CI = 1.3 – 7.6, $t(17) = -2.97$, $p = 0.009$). For the N categories, the difference was 2.8 exemplars (95% CI = 0.2 – 5.4, $t(17) = -2.25$, $p = 0.04$).

As before, we tested for an effect of category type within each participant group. These analyses revealed a significant effect of category type within the control group ($F(3,17) = 10.31$, $p = 0.0004$). Pairwise comparisons showed that controls generated more exemplars for the AS categories than for the N categories (an estimated difference of 2.6; 95% CI = 1.5 – 3.6; $t(17) = 5.22$; $p < 0.0001$). There was a trend for an advantage of the AN categories over the N categories (by 1.8 exemplars; 95% CI = −0.2 – 3.7; $t(17) = 1.92$; $p = 0.07$). There was no significant difference between the AS and AN categories (an estimated advantage of 0.8 exemplars for the AS categories; 95% CI = −1.0 – 2.7; $t(17) = -0.93$; $p = 0.36$). There was no effect of category type for the patients ($F(2,17) = 0.01$, $p = 0.99$). Pairwise comparisons showed no significant difference between the AS categories and the AN categories (the estimated advantage for the AS categories was 0.1 exemplars; 95% CI = −1.7 – 2.0; $t(17) = 0.14$, $p = 0.89$) and no significant difference between the AN and the N categories (0.2; 95% CI = −1.8 – 2.1; $t(17) = 0.16$, $p = 0.87$). Furthermore, we found no significant difference between the AS categories and the N categories (an estimated advantage of 0.03 exemplars for the AS categories; 95% CI = −1.04 – 1.17; $t(17) = 0.06$, $p = 0.95$). While an earlier analysis had indicated a trend towards an advantage on the AS categories over the N categories in amnesic participants, it was no longer evident when the data were analyzed using the strategies that the patients themselves had endorsed.

We reanalyzed the data with FAS score as a covariate and found a trend towards a main effect of participant group ($F = 4.14$, $p = 0.06$), a main effect of category type ($F = 6.66$, $p = 0.008$), and a participant group × category type interaction ($F = 6.17$, $p = 0.01$). Pairwise comparisons showed that the patients scored significantly lower than the controls on the AS categories ($t(16) = -2.68$, $p = 0.02$), and that there was a trend towards lower performance on the AN categories ($t(16) = -1.88$, $p = 0.08$). Performance on the N categories did not differ ($t(16) = -0.90$, $p = 0.38$). We then collapsed across autobiographical categories to obtain a more nearly equal number of observations per subject; the pattern of results did not change. These results indicate that performance on autobiographical categories, particularly the AS categories, was more impaired in amnesia even when the assignment of categories to category types was based on participants' own reports.

Effect of damage outside the MTL

As noted above, Schmolck and colleagues (2002) found a fluency impairment only in patients with damage that extended outside the MTL. We had four such patients in our patient group (MTL06, MTL07, MTL08, and MTL09); we compared their performance to that of the five patients with damage limited to the MTL. When we used the experimenter-assigned category types, we found no effect of patient subgroup ($F(1,7) = 1.14$, $p = 0.33$), no effect of category type ($F(2,14) = 2.28$, $p = 0.14$), and no subgroup × category type interaction ($F(2,14) = 1.21$, $p = 0.33$). When we conducted the same analysis using the self-described strategies, we found no effect of patient subgroup ($F(1,7) = 1.17$, $p = 0.32$), no effect of category type ($F(2,13) = 0.09$, $p = 0.92$), and no subgroup × category type interaction ($F(2,13) = 0.60$, $p = 0.57$). Thus, there was no evidence for a greater or different pattern of impairment in patients with damage outside the MTL.

Discussion

We administered a set of category fluency tasks to patients with MTL amnesia and matched controls. The tasks consisted of three sets of categories that differed in their tendency to evoke episodic memories. One set of categories tended to evoke autobiographical and spatial retrieval strategies, a second set evoked autobiographical but nonspatial strategies, and a third set evoked neither autobiographical nor spatial strategies. The patients were impaired on the task, but the degree of impairment varied across category types: Their impairment was larger for the categories that tended to elicit autobiographical strategies.

This pattern of impairment was not attributable to differences in response speed: the rate of exemplar generation (defined by total number of exemplars generated in one minute) was not associated with different levels of impairment in amnesia. Furthermore, the results are not attributable to a broad impairment of generative abilities: After covarying for performance on the FAS test, which places similar demands on executive processes but much less of a demand on semantic category knowledge, the impairment was selective to categories that elicited autobiographical strategies. These results are not consistent with a specific deficit for spatial information, since patients were equally impaired on the spatial and nonspatial autobiographical categories. Nor are the results attributable to damage outside the MTL, as the pattern of results did not differ based on the presence or absence of such damage. Instead, our findings suggest that the patients' episodic memory deficit hindered their performance on this putatively semantic task.

Previous studies have provided evidence for several kinds of interactions between episodic and semantic memory. Indeed, one of the formulations of the episodic/semantic distinction explicitly posits such an interaction: Newly-acquired information is thought to be mostly episodic in nature; repeated presentation "converts" it to semantic memory by dissociating it from its original spatiotemporal context (Baddeley, 1988). In accordance with this view, an episodic memory impairment impedes the acquisition of new semantic knowledge (e. g. Verfaellie, Croce, & Milberg, 1995), and relatively unimpaired episodic memory can help preserve or re-establish degenerating semantic knowledge to some extent, as seen in semantic dementia (Westmacott et al., 2004). Other work has focused on interactions between episodic memory and premorbid semantic memory. Premorbid semantic knowledge can facilitate episodic memory for novel material in amnesic patients with MTL damage (Kan et al., in press). In the present study, we provide evidence for the converse interaction by showing that poor episodic memory impedes the retrieval of information from the semantic knowledge base.

Episodic memory is known to depend on the MTL (Rosenbaum et al., 2008). Could the episodic involvement in category fluency therefore account for the MTL activation seen in neuroimaging investigations of this task? While it may play a role, the results of a previous neuroimaging study show that a more complex explanation is required (Ryan et al., 2008). In that experiment, participants were presented with the same three types of categories used here. All categories activated the hippocampal formation equally, including those that made no apparent use of episodic memory, suggesting that the hippocampal activity may reflect the encoding of new information as well as the retrieval of old information. The present results suggest that while category exemplar generation is often associated with MTL activity, only "autobiographical" category exemplar generation depends on the integrity of the MTL.

When combined with research on the role of the lateral temporal lobes, our results provide a potential explanation for the pattern of impairment that Gleissner and Elger (2001) reported. These authors administered fluency tasks to two groups of patients with temporal-lobe epilepsy. One group of patients had damage limited to the hippocampus; the other group had temporal lobe damage that spared the hippocampus entirely. Within each group, half the

patients had left-lateralized brain damage; the other half had right-lateralized damage. Patients with hippocampal damage were impaired on fluency tasks, regardless of the side of the lesion; patients with extrahippocampal damage were only impaired when the damage was left-lateralized. The effect of the left lateral temporal lobe damage is presumably attributable to its well-established role in semantic memory (Levy et al., 2004); the effect of the medial temporal lobe damage may reflect the episodic contribution to this task. Accordingly, we would predict that the patients with hippocampal damage would be impaired on autobiographical categories, while the patients who had left lateral temporal damage would have a broader semantic impairment that affects performance on all categories. The conflicting results of Schmolck and colleagues (2002), who found no fluency impairment in two patients with damage limited to the hippocampal formation, are harder to account for, though a partial explanation may lie in the choice of categories. In Schmolck and colleagues' experiment, the responses for several categories (e.g., animals, birds, water creatures) could overlap. While it is unclear what precise effect this design might have, it is possible that the responses to one category primed exemplars to subsequent categories, thus facilitating generation. Further investigation of episodic-semantic interactions may clarify these issues.

The demonstration of an episodic contribution to category fluency suggests that the usual interpretation of a fluency deficit warrants re-examination. For instance, previous research has shown that category fluency declines with age (Kempler et al., 1998; Gladsjo et al., 1999; Tombaugh et al., 1999; Acevedo et al., 2000; Brickman et al., 2005), a finding that has been attributed to an impairment of semantic memory (Brickman et al., 2005). Other research, however, has shown that aging has little effect on the sort of well-established semantic knowledge that is tapped in fluency tasks (Prull et al., 2000). Episodic memory, by contrast, does decline with age; for example, the autobiographical memories of older adults tend to contain fewer specific details and a greater number of general statements (Piolino et al., 2006). We therefore speculate that the age-related decline in category fluency may arise from an impairment not in semantic memory, but in episodic memory instead. Along related lines, patients with mild cognitive impairment (MCI) or Alzheimer's disease also have an impairment of category fluency (Jones et al., 2006; Nutter-Upham et al., 2008; see Henry et al., 2004, for review), but this deficit may arise from the impairment of multiple processes. Patients with these conditions have an episodic memory impairment; unlike older adults, however, they perform poorly on semantic memory tests as well (Petersen et al., 1999). Researchers have debated whether the impairment on these tests results from degraded semantic memory or impaired access to an intact semantic store (e. g. Rogers et al., 2006). Our data cannot resolve this question directly, and a possible contribution of a general fluency impairment would need to be assessed, but our data nonetheless suggest that impaired access to the semantic store could arise from the episodic memory impairment.

Throughout this paper, we have argued that amnesic patients' impairment in exemplar generation arises from their episodic impairment. The autobiographical categories do tend to elicit familiar contexts and memories, but are these really "episodic"? In some cases, they clearly are: When a participant generates *buildings on the main street of a town* by imagining "what I saw on my drive over here," then he or she is using episodic memory by any common definition of the term. But what about a participant who generates *kitchen utensils* by "thinking of what I use in my own kitchen every day"? This kind of memory is not quite episodic; it is bounded in space but not in time, and is thus not really an "episode." Nor does it meet the usual definition of "semantic," because the specific spatial information indicates that it is not decontextualized. Instead, it seems to fall somewhere between the two (see Ryan et al., 2008, for a similar argument). The idea of an intermediate form of memory, or of a continuum between episodic and semantic memory, roughly comports with Neisser's (1981) concept of "repisodic memory," Conway's (2001) "general events" level of autobiographical knowledge, or Barsalou's (1988) conception of "summarized" or "extended events". Other evidence

suggests that MTL amnesics have difficulty constructing the sort of rich, detailed spatial and/ or self-referential contexts that these memories require, whether the patients are trying to remember the past (e. g. Moscovitch et al., 2005) or imagine the future (Hassabis et al., 2007). The results of this study demonstrate that an inability to retrieve such contextual memories leads to poor performance on tasks that are usually considered semantic.

Acknowledgments

References

Acevedo A, Loewenstein DA, Barker WW, Harwood DG, Luis C, Bravo M, Hurwitz DA, Aguero H, Greenfield L, Duara R. Category fluency test: normative data for English- and Spanish-speaking elderly. J Int Neuropsych Soc 2000;6:760–769.

Baddeley AD. Cognitive psychology and human memory. Trends Neurosci 1998;11:176–181. [PubMed: 2469187]

Barsalou LW. Ad hoc categories. Mem Cognition 1983;11:211–227.

Barsalou, LW. The content and organization of autobiographical memories. In: Neisser, U.; Winograd, E., editors. Remembering reconsidered: ecological and traditional approaches to the study of memory. Cambridge: Cambridge University Press; 1988. p. 193-243.

Battig WF, Montague WE. Category norms for verbal items in 56 categories. J Exp Psychol 1969;80:1– 42.

Benton, AL.; Hamsher, K. Multilingual Aphasia Examination. Iowa City, IA: University of Iowa; 1976.

Brickman AM, Paul RH, Cohen RA, Williams LM, MacGregor KL, Jefferson AL, Tate DF, Gunstad J, Gordon E. Category and letter verbal fluency across the adult lifespan: relationship to EEG theta power. Arch Clin Neuropsych 2005;20:561–573.

Conway MA. Sensory-perceptual episodic memory and its context: autobiographical memory. Philos T Roy Soc B 2001;356:1375–84.

Frith CD, Friston KJ, Liddle PF, Frackowiak RSJ. A PET study of word finding. Neuropsychologia 1991;29:1137–1148. [PubMed: 1791928]

Gladsjo JA, Schuman CC, Evans JD, Peavy GM, Miller SW, Heaton RK. Norms for letter and category fluency: demographic corrections for age, education, and ethnicity. Assessment 1999;6:147–178. [PubMed: 10335019]

Gleissner U, Elger CE. The hippocampal contribution to verbal fluency in patients with temporal lobe epilepsy. Cortex 2001;37:55–63. [PubMed: 11292161]

Gourovitch ML, Kirkby BS, Goldberg TE, Weinberger DR, Gold JM, Esposito G, Van Horn JD, Berman KF. A comparison of rCBF patterns during letter and semantic fluency. Neuropsychology 2000;14:353–360. [PubMed: 10928738]

Greenberg DL, Rice HJ, Cooper JJ, Cabeza R, Rubin DC, Labar KS. Coactivation of the amygdala, hippocampus, and inferior frontal gyrus during autobiographical memory retrieval. Neuropsychologia 2005;43:659–674. [PubMed: 15721179]

Hassabis D, Kumaran D, Van SD, Maguire EA. Patients with hippocampal amnesia cannot imagine new experiences. PNAS 2007;104:1726–1731. [PubMed: 17229836]

Henry JD, Crawford JR, Phillips LH. Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. Neuropsychologia 2004;42:1212–1222. [PubMed: 15178173]

Hirshorn EA, Thompson-Schill SL. Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. Neuropsychologia 2006;44:2547–2557. [PubMed: 16725162]

Jones S, Laukka EJ, Bäckman L. Differential verbal fluency deficits in the preclinical stages of Alzheimer's disease and vascular dementia. Cortex 2006;42:347–355. [PubMed: 16771040]

Kan IP, Alexander MP, Verfaellie M. Contribution of prior semantic knowledge to new episodic learning in amnesia. J Cognitive Neurosci. (in press)

Kan IP, Giovanello KS, Schnyer DM, Makris N, Verfaellie M. Role of the medial temporal lobes in relational memory: Neuropsychological evidence from a cued recognition paradigm. Neuropsychologia 2007;45:2589–2597. [PubMed: 17433382]

Kempler D, Teng EL, Dick M, Taussig IM, Davis DS. The effects of age, education, and ethnicity on verbal fluency. J Int Neuropsych Soc 1998;4:531–538.

Levy DA, Bayley PJ, Squire LR. The anatomy of semantic knowledge: medial vs. lateral temporal lobe. PNAS 2004;101:6710–6715. [PubMed: 15090653]

Lezak, MD.; Howieson, DB.; Loring, DW. Neuropsychological assessment. Oxford University Press; 2004.

Moscovitch M, Rosenbaum RS, Gilboa A, Addis DR, Westmacott R, Grady CL, McAndrews MP, Winocur G, Nadel L. Functional neuroanatomy of remote episodic (autobiographical), semantic and spatial memory in humans as determined by lesion and functional neuroimaging studies: A unified account based on Multiple Trace Theory. J Anat 2005;207:35–66. [PubMed: 16011544]

Mummery CJ, Patterson K, Hodges JR, Wise RJS. Generating 'tiger' as an animal name or a word beginning with T: differences in brain activation. Proc Roy Soc Lond B Bio 1996:263, 989–995.

Neisser U. John Dean's memory: A case study. Cognition 1981;9:1–22. [PubMed: 7196816]

Nutter-Upham KE, Saykin AJ, Rabin LA, Roth RM, Wishart HA, Pare N, Flashman LA. Verbal fluency performance in amnestic MCI and older adults with cognitive complaints. Arch Clin Neuropsych 2008;23:229–241.

Paulesu E, Goldacre B, Scifo P, Cappa SF, Gilardi MC, Castiglioni I, Perani D, Faziio F. Functional heterogeneity of left inferior frontal cortex as revealed by fMRI. NeuroReport 1997;8:2011–2016. [PubMed: 9223094]

Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999;56:303–308. [PubMed: 10190820]

Pihlajamäki M, Tanila H, Hänninen T, Könönen M, Laakso M, Partanen K, Soininen H, Aronen HJ. Verbal fluency activates the left medial temporal lobe: a functional magnetic resonance imaging study. Ann Neurol 2000;47:470–476. [PubMed: 10762158]

Piolino P, Desgranges B, Clarys D, Guillery-Girard B, Taconat L, Isingiri M, Eustache F. Autobiographical memory, autonoetic consciousness and self-perspective in aging. Psychol Aging 2006;3:510–525. [PubMed: 16953713]

Prull, MW.; Gabrieli, JDE.; Bunge, SA. Age-related changes in memory: a cognitive neuroscience perspective. In: Craik, FIM.; Salthouse, TA., editors. The handbook of aging and cognition. Lawrence Erlbaum; 2000. p. 91-154.

Rempel-Clower NL, Zola SM, Squire LR, Amaral DG. Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. J Neurosci 1996;16:5233–5255. [PubMed: 8756452]

Rogers TT, Ivanoiu A, Patterson K, Hodges JR. Semantic memory in Alzheimer's disease and the frontotemporal dementias: a longitudinal study of 236 patients. Neuropsychology 2006;20:319–335. [PubMed: 16719625]

Rosenbaum RS, Moscovitch M, Foster JK, Schnyer DM, Gao F, Kovacevic N, Verfaellie M, Black SE, Levine B. Patterns of autobiographical memory loss in medial-temporal lobe amnesic patients. J Cogn Neurosci 2008;20:1490–1506. [PubMed: 18303977]

Ryan LR, Cox C, Hayes SM, Nadel L. Hippocampal activation during episodic and semantic memory retrieval: comparing category production and category cued recall. Neuropsychologia 2008;46:2109–2121. [PubMed: 18420234]

Schmolck H, Kensinger EA, Corkin S, Squire LR. Semantic knowledge in patient HM and other patients with bilateral medial and lateral temporal lobe lesions. Hippocampus 2002;12:520–533. [PubMed: 12201637]

Smith C, Squire LR. Declarative memory, awareness, and transitive inference. J Neurosci 2005;25:10138–10146. [PubMed: 16267221]

Strauss, E.; Sherman, EMS.; Spreen, O. A compendium of neuropsychological tests. Oxford University Press; 2006.

Tombaugh TN, Kozak J, Rees L. Normative data stratified for age and education for two measures of verbal fluency: FAS and animal naming. Arch Clin Neuropsych 1999;14(14):167–177.

Tulving, E. Episodic and semantic memory. In: Tulving, E.; Donaldson, W., editors. Organization of memory. New York: Academic Press; 1972. p. 381-403.

Vallée-Tourangeau F, Anthony SH, Neville GA. Strategies for generating multiple instances of common and ad hoc categories. Memory 1998;6:555–592. [PubMed: 10197163]

Van Overschelde JP, Rawson KA, Dunlosky J. Category norms: an updated and expanded version of the Battig and Montague (1969) norms. J Mem Lang 2004;50:289–335.

Verfaellie M, Croce P, Milberg WP. The role of episodic memory in semantic learning: an examination of vocabulary acquisition in a patient with amnesia due to encephalitis. Neurocase 1995;1:291–304.

Walker WH, Kintsch W. Automatic and strategic aspects of knowledge retrieval. Cognitive Sci 1985;9:261–283.

Westmacott R, Black SE, Freedman M, Moscovitch M. The contribution of autobiographical significance to semantic memory: evidence from Alzheimer's disease, semantic dementia, and amnesia. Neuropsychologia 2004;42:25–48. [PubMed: 14615074]
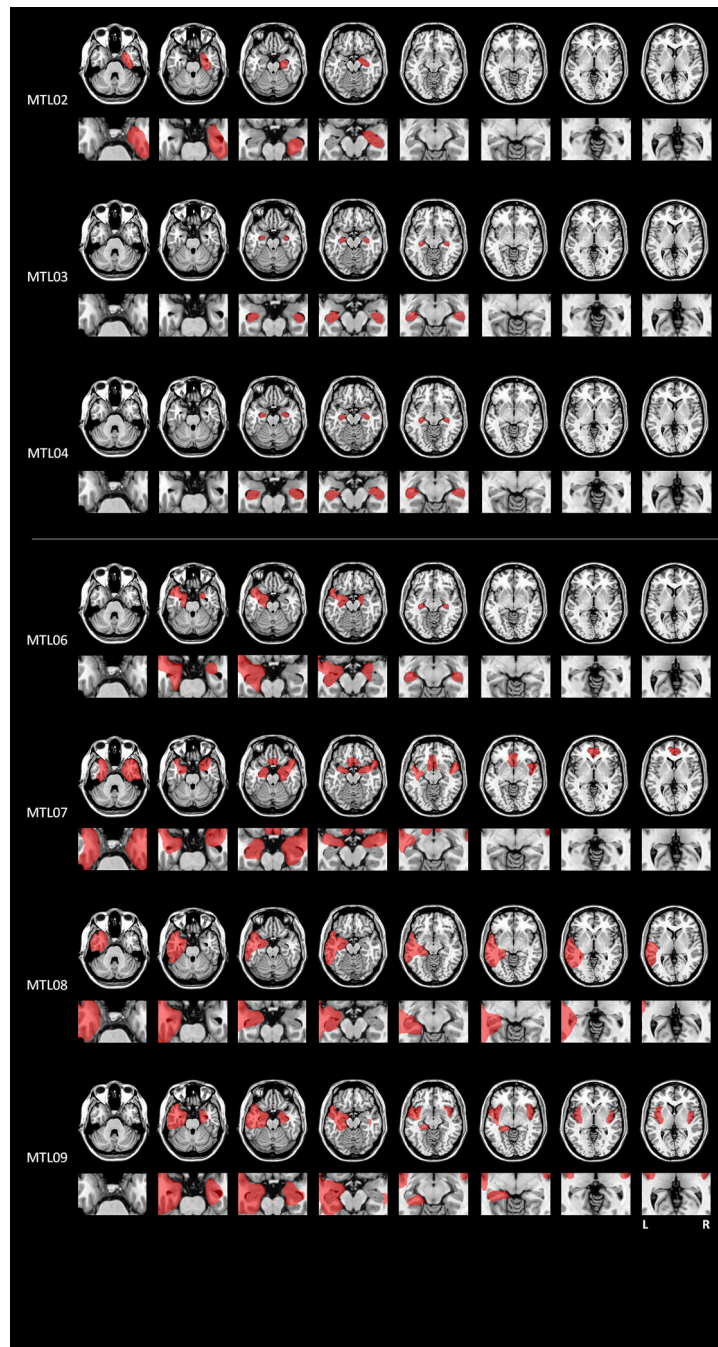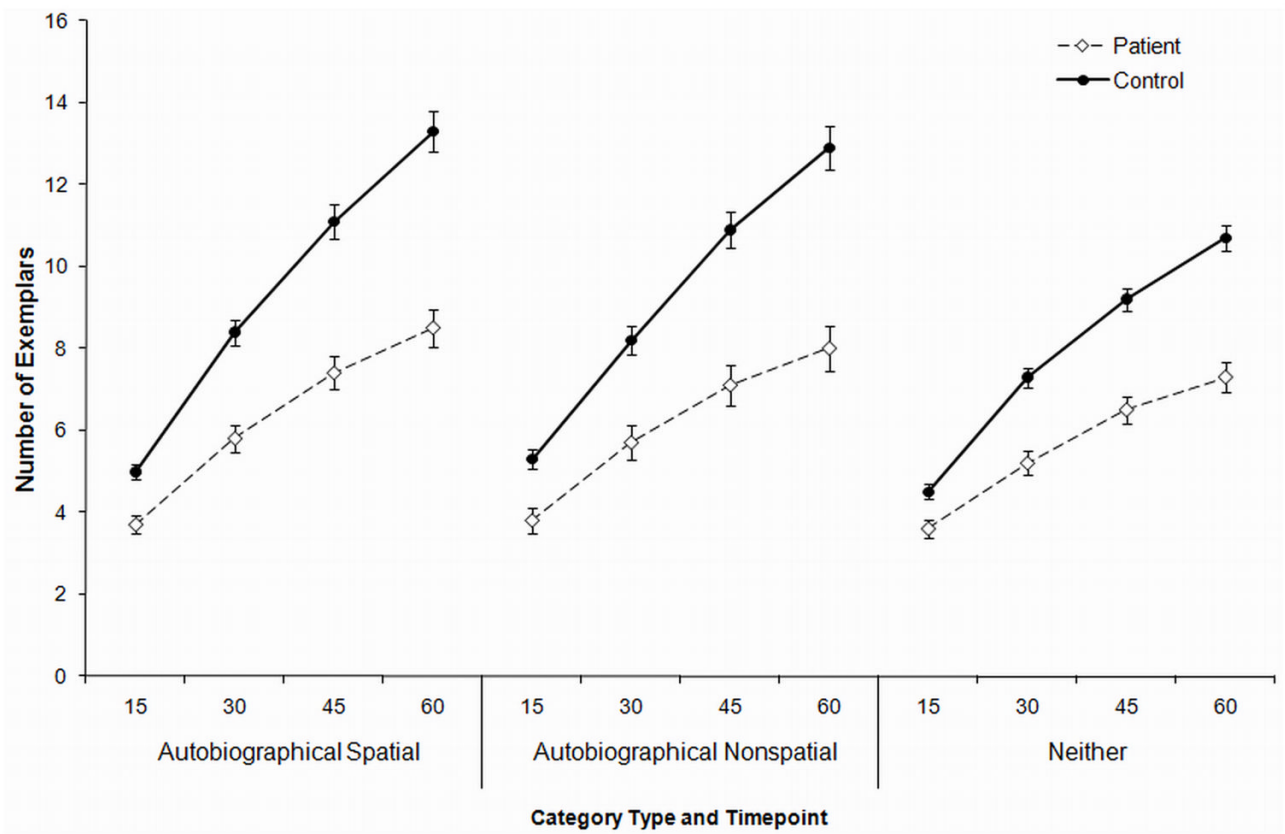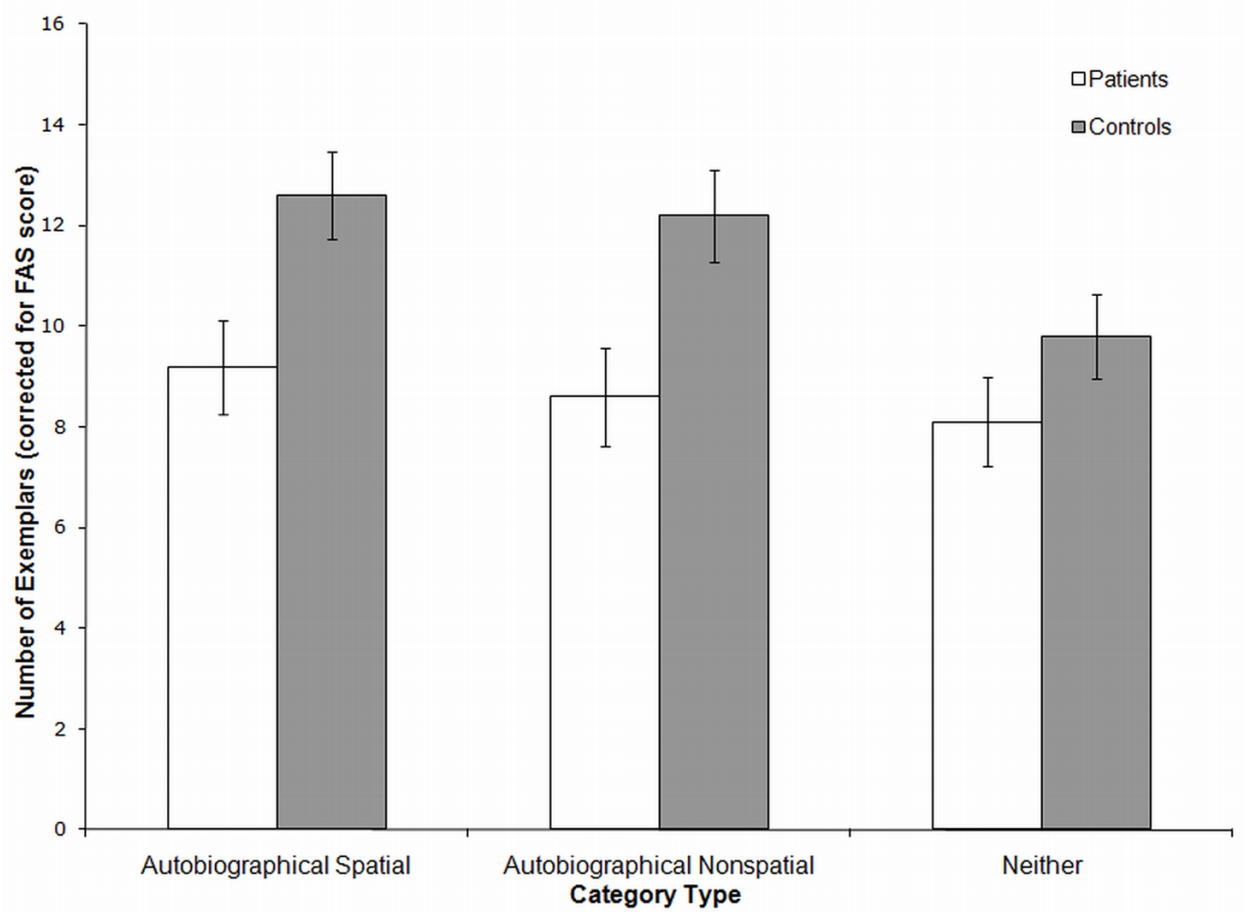
Figure 1.
Reconstruction of the lesions of 7 of the patients who participated in the experiment. The top panel shows the reconstructions for the patients whose damage was limited to the medial temporal lobe; the bottom panel shows the reconstructions for the patients whose damage extended to other areas of the temporal lobe. For each participant, the top row provides an overall view of the lesions while the bottom row focuses on medial temporal lobe structures. The left side of the image depicts the left side of the brain.
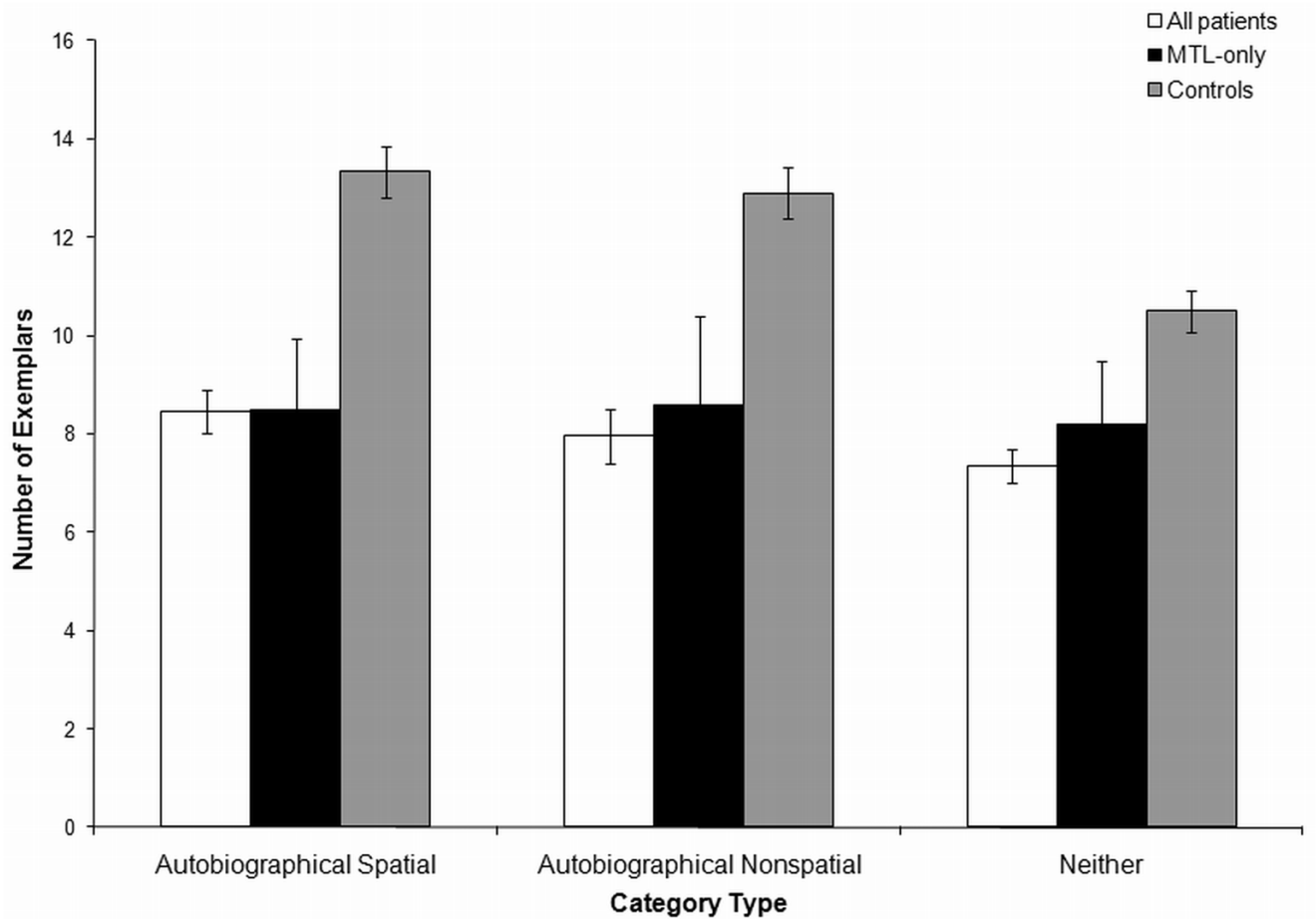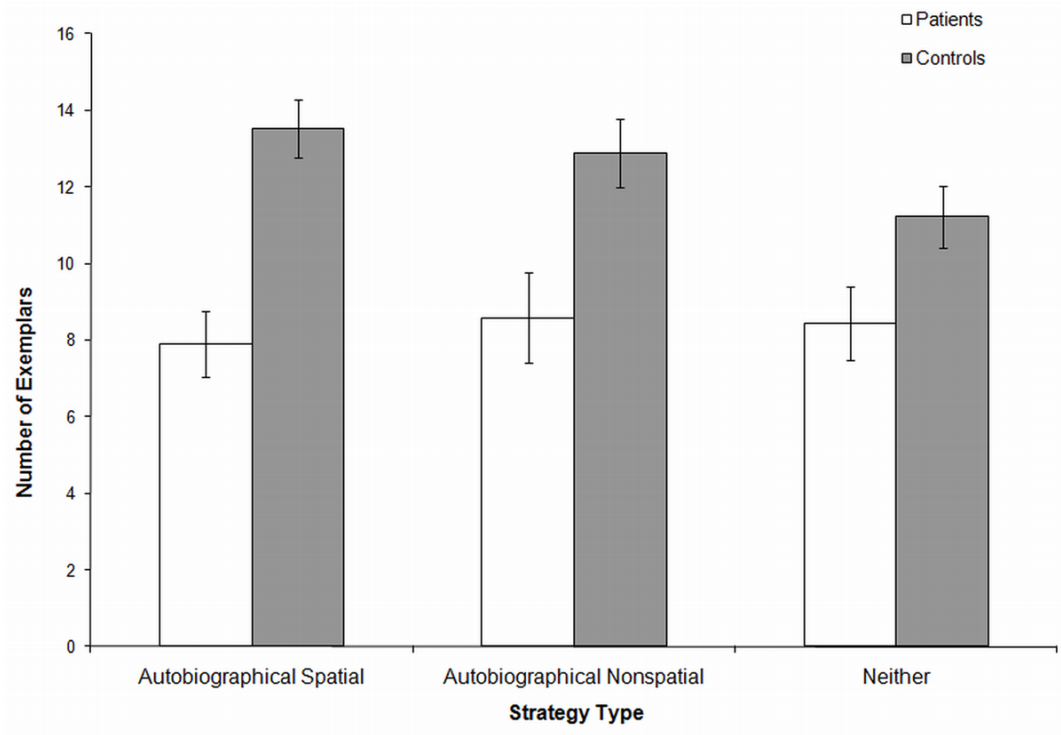
Figure 2.
Mean number of exemplars generated for each category type (using experimenters' assignment of categories to category types). Error bars indicate standard errors.

Figure 3.
Time courses of responses for the patients and the controls. Error bars indicate standard errors.

Figure 4.
Mean number of exemplars generated for each category type (using experimenters' assignment of categories to category types) adjusted for FAS score. Error bars indicate standard errors.

Figure 5.
Mean number of exemplars generated for each category type (using participants' own assignments of categories to category types). Error bars indicate standard errors.

Table 1

Classification of categories

Autobiographical Spatial	%	Autobiographical Nonspatial	%	Neither	%
Things in a bedroom[a]	95	Things people like when they're sick[a, c]	85	Units of measurement[a, c]	89
Living room furniture[a, c]	89	Reading materials[a]	84	Things that make noise[a, c]	85
Things in a park	84	Birthday presents	84	Things that are usually red	80
Buildings on the main street of a town[c]	83	Makes of cars[c]	79	Elected offices	83
Things people put on the walls	80	Diseases[c]	79	Precious stones	65
Kitchen utensils	78	Things people take on camping trips	71	Heavy objects[c]	74
Things in a garage[c]	63	Weather phenomena[b]	65	Things that can fall on your head	72
Things made of wood[c]	55	Insects[b]	58	Weapons[c]	65

Note.

% = percentage of pilot participants who endorsed this strategy for this category

[a] = categories that were reused in the third session of the experiment

[b] = categories that were excluded (see Method)

[c] = categories used to equate control participants across category types

Table 2

Demographic and neuropsychological characteristics of the amnesic patients

Patient	Etiology	Damage	Age	Educ	FAS	WAIS-III	WMS-III				
						VIQ	GM	VD	AD	WM
MTL01	Anoxia	MTL	79	18	−1.16	113	75	72	80	102
MTL02	Anoxia	MTL	58	12	−1.07	83	52	56	55	91
MTL03	Anoxia	MTL	50	14	−0.96	90	45	53	52	93
MTL04	Anoxia	MTL	52	14	−1.22	111	59	72	52	96
MTL05	Anoxia	MTL	56	17	0.03	134	70	75	67	126
MTL06	Encephalitis	MTL+	53	14	0.11	92	45	56	55	85
MTL07	Encephalitis	MTL+	65	12	−0.61	106	69	68	77	111
MTL08	Anoxia + L Temp Lobectomy	MTL+	45	16	−1.94	86	49	53	52	93
MTL09	Encephalitis	MTL+	80	18	0.66	135	45	53	58	141

Note. WAIS-III = Wechsler Adult Intelligence Test-III; WMS-III = Wechsler Memory Scale-III; VIQ = Verbal IQ; GM = General Memory; VD = Visual Delay; AD = Auditory Delay; WM = Working Memory; MTL = damage limited to the medial temporal lobes; MTL+ = damage extended outside medial temporal lobes.

Table 3

MTL brain volumes for 5 patients (expressed as z-scores).

Patient	L HPC	R HPC	L AMYG	R AMYG	L PHG(A)	R PHG(A)	L PHG(P)	R PHG(P)	L TP	R TP
MTL03	**-7.70**	**-6.82**	0.22	0.36	-1.65	0.24	0.85	0.52	-1.01	0.06
MTL04	-1.61	**-3.24**	0.17	0.97	-1.15	0.21	3.67	-0.66	-0.92	-0.46
MTL06	**-8.46**	**-6.22**	**-3.92**	**-2.53**	**-3.75**	**-3.53**	**-2.98**	-0.41	**-6.24**	-1.02
MTL07	**-5.34**	**-7.82**	**-2.53**	**-2.57**	**-3.73**	**-6.64**	-1.25	**-3.87**	-1.39	**-3.99**
MTL08	**-8.19**	**-4.40**	**-3.92**	1.00	**-2.48**	-0.17	4.40	2.71	**-3.92**	1.00

Note. L=left; R=right; HPC=hippocampus; AMYG=amygdala; PHG(A)=anterior portion of the parahippocampal cortex (includes the medial portion of the temporal pole, the medial and lateral perirhinal cortices, and the entorhinal cortex); PHG(P)=posterior parahippocampal cortex; TP=temporal pole.