

This is the peer reviewed version of the following article: Pajkert A, Finke C, Shing YL, et al. Memory integration in humans with hippocampal lesions. *Hippocampus*. 2017;27:1230–1238, which has been published in final form at <https://doi.org/10.1002/hipo.22766>. This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for self-archiving.



Memory integration in humans with hippocampal lesions

Journal:	<i>Hippocampus</i>
Manuscript ID	HIPO-17-049.R2
Wiley - Manuscript type:	Research Article
Keywords:	hippocampus, encoding, memory integration, relational inference, decision making

SCHOLARONE™
Manuscripts

Review

Memory integration in humans with hippocampal lesions

Running title: Memory integration and hippocampal lesions

Authors: Anna Pajkert¹, Carsten Finke^{1,2}, Yee-Lee Shing^{3,4}, Martina Hoffmann¹, Werner Sommer⁵, Hauke R. Heekeren⁶, Christoph J. Ploner¹

¹ Department of Neurology, Charité – Universitätsmedizin Berlin, D-10117 Berlin, Germany

² Berlin School of Mind & Brain, Humboldt-Universität zu Berlin, D-10117 Berlin, Germany

³ Center for Lifespan Psychology, Max Planck Institute for Human Development, D-14195 Berlin, Germany

⁴ Division of Psychology, Faculty of Natural Sciences, University of Stirling, FK94LA Scotland, United Kingdom

⁵ Institute of Psychology, Humboldt-Universität zu Berlin, D-10099 Berlin, Germany

⁶ Department of Education and Psychology, Freie Universität Berlin, D-14195 Berlin, Germany

Number of text pages: 23 figures: 4 tables: 2

Corresponding author: Christoph J. Ploner, Department of Neurology,
Charité – Universitätsmedizin Berlin, Charitéplatz 1, D-10117 Berlin, Germany;

mail: christoph.ploner@charite.de

Grant sponsor: Deutsche Forschungsgemeinschaft, Grant number: DFG PI 248/4-1

Key words: hippocampus, encoding, memory integration, relational inference, decision making

Abstract

Adaptive behavior frequently depends on inference from past experience. Recent studies suggest that the underlying process of integrating related memories may depend on interaction between hippocampus and prefrontal cortex. Here, we investigated how hippocampal damage affects memory integration. Subjects with mediotemporal lesions and healthy controls learned a set of overlapping AB- and BC-associations (object-face- and face-object pairs) and were then tested for memory of these associations ('direct' trials) and of inferential AC-associations ('indirect' trials). The experiment consisted of four encoding/retrieval cycles. In direct trials, performance of patients and controls was similar and stable across cycles. By contrast, in indirect trials, patients and controls showed distinct patterns of behavior. Whereas patients and controls initially showed only minor differences, controls increased performance across subsequent cycles, while patient performance decreased to chance level. Further analysis suggested that this deficit was not merely a consequence of impaired associative memory but rather resulted from an additional hippocampal contribution to memory integration. Our findings further suggest that contextual factors modulate this contribution. Patient deficits in more complex memory-guided behavior may depend on the flexible interaction of hippocampus-dependent and – independent mechanisms of memory integration.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

An important prerequisite for adaptive behavior is the ability to form relational memory networks out of multiple distinct experiences and to draw on information from these networks in novel behavioral contexts. This process has been termed memory integration and is thought to underlie flexible decision making and multiple other cognitive faculties such as spatial navigation, imagination of the future and creativity (Zeithamova et al. 2012a; Preston and Eichenbaum 2013; Palombo et al., 2014; Schlichting and Preston, 2015; Shohamy and Daw, 2015). Following pioneering research in rats (Dusek and Eichenbaum, 1996; Bunsey and Eichenbaum, 1997), recent studies in healthy humans suggests a pivotal role of the hippocampus within networks for memory integration (Zeithamova et al. 2012a; Preston and Eichenbaum, 2013; Palombo et al., 2014; Schlichting and Preston, 2015; Shohamy and Daw, 2015). One hypothesis posits that the hippocampus provides integrated representations that shape individual behavioral preferences, even in conditions that involve implicit and reward-related learning, i.e. forms of learning and memory that are commonly thought to be less dependent on the hippocampus (e.g. Reber et al. 2012; Barron et al., 2013; Gluth, et.al., 2015; Wimmer and Shohamy, 2012; Zeithamova et al., 2012b). A mechanism supporting integration is supposed to be the flexible creation of overlapping neural representations between distinct but related memories via pattern completion and novelty detection (Zeithamova et al. 2012a; Preston and Eichenbaum, 2013; Schlichting and Preston, 2015; Shohamy and Daw, 2015). Behavioral and Imaging studies further suggest that memory integration may occur both during encoding of new information and during retrieval of previously formed memories immediately before behavior is initiated (e.g. Zeithamova et al. 2010, 2012b; Schlichting and Preston, 2015; Schlichting et al. 2014; Horner and Burgess 2014).

Although these findings allow for the prediction that hippocampal dysfunction should yield behavioral deficits beyond mere memory impairments (Shohamy and Turk-Browne, 2013; Schlichting and Preston, 2015), few human studies have addressed the interrelationship of hippocampal dysfunction, memory integration and other cognitive functions such as decision making so far (Gleichgerrcht, et al., 2010; Shadlen and Shohamy, 2016). This is surprising, as significant impairments in complex behaviors and real-world decision making concerning finance, healthcare etc. have been demonstrated in patients with disorders affecting the hippocampus, e.g. in mild cognitive impairment, Alzheimer's disease and schizophrenia (e.g. Lee, 2013; Martin, et al., 2013; Han et al. 2015).

Here, we investigated effects of hippocampal damage on memory integration. Adult human subjects with damage to the medial temporal lobe performed a visual associative inference task that required memory of face-object pairs and integration of these memories for decision between choice options (Zeithamova and Preston 2010; Zeithamova et al., 2012). We tested whether behavioral deficits in patients merely result from impaired memory or from dysfunction of a hippocampus-dependent integration mechanism.

Materials & Methods

Subjects

A group of five subjects with damage to the right medial temporal lobe (MTL) was recruited from the Department of Neurology at the Charité – Universitätsmedizin Berlin, Germany (2 females; mean age 45.8 ± 7.2 ys, table 1). All patients acquired MTL damage in adulthood (4 post-surgical lesions following resection of a benign brain tumor, 1 post-encephalitic lesion). Four patients had already participated in previous studies of our group (patients 1, 4, 5 and

21 of Esfahani-Bayerl et al. (2016). The fifth patient is a 57 year-old female with a MTL lesion following resection of a low-grade glioma 11 months before testing. All patients were right-handed and normal on neurological examination. Mean delay between MTL lesion and testing was 92 ± 26 months. Lesions were reconstructed from routine coronal MRIs according to the method proposed in Esfahani-Bayerl et al. (2016). All five patients sustained damage to the right anterior hippocampus, right entorhinal cortex and parts of right perirhinal cortex (table 2). The right parahippocampal cortex was additionally involved in two patients. The left hippocampus, left entorhinal cortex, left perirhinal cortex and left parahippocampal cortex was involved in one patient. The control group consisted of 17 normal subjects without any history of neurological disorders (8 females; mean age 43.6 ± 2.5 ys). In all subjects, verbal and non-verbal intelligence were assessed using the MWT-B, a German equivalent to the National Adult Reading Test (Lehrl, 2005) and the sub-test no. 3 of LPS, a German equivalent to Raven's Progressive Matrices (Horn, 1983). Spatial working memory was tested with Corsi Block Tapping test forward and backward. Visuo-spatial abilities were tested with the Rey-Osterrieth Complex Figure test. No significant group differences were found for any of these measures (verbal IQ; patients, 111 ± 6.1 ; controls, 114 ± 2.7 ; $p = 0.68$, Mann-Whitney test; LPS-scores (t-values); patients, 57.5 ± 3.9 ; controls, 58.1 ± 1.6 ; $p = 0.69$; Corsi Blocks forward; patients, 8.3 ± 0.5 ; controls, 8.6 ± 0.6 ; $p = 0.50$; Corsi Blocks backward; patients, 9 ± 0.7 ; controls, 7.9 ± 0.2 ; $p = 0.1$; Rey-Osterrieth copy; patients, 35.3 ± 0.3 ; controls, 35.4 ± 0.2 ; $p = 0.45$; Rey-Osterrieth immediate; patients, 22.8 ± 3 ; controls 22.5 ± 1.6 ; $p = 1$; Rey-Osterrieth delayed; patients, 19.8 ± 1.4 ; controls 23.1 ± 1.4 ; $p = 0.28$). Informed consent was obtained from all subjects before participation in the study, which was approved by the local Ethical Committee and conducted in conformity with the Declaration of Helsinki.

Experimental Paradigm

Subjects were tested with a variant of an associative inference task originally used in fMRI studies of memory integration (Zeithamova and Preston, 2010; Zeithamova et al., 2012b). In this task, subjects learned a set of overlapping associations (A-B- and B-C-associations, i.e. object-face- and face-object pairs). After a delay, subjects were tested for memory either of these associations or of inferential A-C-associations (figure 1).

Stimuli consisted of 96 images of objects (e.g. fruits, tools, clothes etc.) and 64 color images of human faces (32 females). In AB-trials, objects and faces were presented pair-wise in pseudo-random and trial-unique combinations that overlapped with corresponding BC-trials. The overlapping B-stimulus always consisted of a face. As a control condition, subjects were presented DE-stimuli that consisted of pseudo-random and trial-unique combinations of faces and objects.

The structure of the task was as follows: During encoding blocks, subjects learned a set of AB-, BC- and DE-associations. During subsequent retrieval blocks, subjects were tested for inferential AC-associations ('indirect trials') and for memory of AB-, BC- and DE-associations ('direct trials'). The experiment consisted of a pseudorandom sequence of four cycles, each consisting of an encoding block followed by a delay and a retrieval block (figure1).

Encoding blocks consisted of 24 trials (8 AB-, 8 BC- and 8 DE-trials). Trials were presented in pseudorandom order. AB-trials always preceded the corresponding BC-trials. Overlapping AB- and BC-trials were presented between three and eight trials apart in order to avoid direct repetition of B-stimuli. DE-trials were interleaved with AB- and BC-trials. In each trial, stimulus pairs were presented for 5 s followed by a response period, during which subjects were asked to indicate by keypress whether two of the stimuli are living objects or only one

(e.g. a vehicle and a face, a plant and a face). The inter-trial interval was unpredictably varied between 1 and 3 s.

Retrieval blocks consisted of 32 trials (8 AC-, 8 AB-, 8 BC- and 8 DE-trials) and were started 5 min after termination of the corresponding encoding blocks ('delay' in figure 1). AC-trials were always presented at the beginning of a block in order to avoid re-learning of AB- or BC- associations before testing of AC-associations. In each trial, an A-, B-, or D-stimulus was presented in the center of the top half of the screen and two choice stimuli (B-, C-, or E-stimuli) were presented in the lower half of the screen (see figure). In indirect (AC-) trials, subjects were requested to indicate by keypress which of the two (C-) stimuli in the lower half of the screen shared an indirect association with the (A-) stimulus in the top half of the screen. That is, which choice stimulus in the lower half of the screen was presented together with the same (B-) face as the stimulus in the top half of the screen. In direct trials, subjects were requested to indicate by keypress which of the two (B-, C- or E-stimuli) stimuli in the lower half of the screen was presented together with the (A-) stimulus in the top half of the screen. Both during direct and indirect trials, foil stimuli were always part of other associations during encoding blocks. Thus, there was no difference in familiarity between target and foil stimuli. Stimuli were presented until a subject's manual response terminated the trial. All stimuli were only used within a given cycle. Neither retrieval of indirect associations nor of direct associations related to stimuli that were encoded in preceding cycles.

Before testing, all subjects received written standardized instructions with figures showing example stimuli. All subjects were well informed about the necessity to make inferential AC-judgments in retrieval blocks. All subjects received training of the paradigm with a fixed number of training trials. During training, encoding blocks consisted of three AB-, three BC- and three DE-stimuli. Retrieval blocks consisted of three indirect (AC-) trials and nine direct (AB-, BC- and DE-) trials. Stimuli from

training trials were not part of the final experiment. All subjects reported full comprehension of instructions.

Data Analysis

Behavioral measures consisted of percentage of correct responses for a given trial type and reaction times (RTs) of correct trials of a given trial type. Medians were used to describe individual average RTs. Since RTs and percentages are usually not normally distributed and since the sample sizes permitted no meaningful conclusions about data distribution, analysis was conducted by using non-parametric statistical tests throughout (Altman, 1991; Altman and Bland 2009).

Results

Accuracy

Similar to previous studies on associative inference (e.g. Zeithamova and Preston, 2010; Armstrong et al., 2012), we first analyzed overall group performance differences across the two major categories of trials, i.e. direct trials and indirect trials. In direct trials, we found no significant differences between the three trial types (AB-, BC- and DE-trials), neither in patients nor in controls (patients: $p = 0.091$; controls: $p = 0.092$ Friedman-ANOVA). Direct trials were thus pooled for further analysis. On average, patient performance in direct trials was inferior to controls (patients: 75.6 ± 5.8 %, controls: 85.2 ± 2.8 % correct, figure 2). These differences did however not achieve significance ($p = 0.071$; Mann-Whitney test). In indirect trials (AC-trials), patients showed a slightly more pronounced and significant deficit (patients: 58.1 ± 5.8 , controls: 76.8 ± 3.7 % correct; $p = 0.038$, Mann-Whitney test). Although these differences may suggest that performance in indirect trials may have involved

computational processes that were less required in direct trials, defective memory of direct associations may have contributed to impaired overall performance in indirect trials. We therefore conducted an additional analysis, where we calculated accuracy of those indirect trials, for which both AB- and BC-stimuli were correctly remembered in the corresponding direct trials. With this analysis, overall group differences for indirect trials were still present, but smaller and no longer significant (patients: 67.2 ± 7.2 , controls: 77.7 ± 4.8 % correct; $p = 0.19$, Mann-Whitney test).

Next, we took a closer look at possible performance changes across cycles. We reasoned that repeated presentation of encoding and retrieval blocks across the four cycles of the experiment may have allowed for adjustment of behavior following individual perception of performance in initial retrieval blocks. Indeed, performance and performance differences between groups were not constant across cycles. In cycles 1 and 2, patients and controls showed only minor performance differences in indirect trials (cycle 1: patients: 65.0 ± 7.3 %, controls: 71.3 ± 5.3 % correct, $p = 0.49$ group difference; cycle 2: patients 62.5 ± 6.9 %, controls: 75.0 ± 5.5 % correct, $p = 0.36$ group difference, Mann-Whitney test, figure 3). This suggests that – at least in certain conditions - patients are not necessarily inferior to controls in tasks that require integration of associations in memory. However, during the course of the experiment, controls increased their performance in indirect trials up to 80.1 ± 3.2 % in cycle 3 and 80.9 ± 4.6 % in cycle 4, while patient performance decreased to chance level (50.0 ± 10.5 % in cycle 3 and 55.0 ± 6.4 % in cycle 4, respectively, $p = 0.005$ and 0.015 group differences, Mann-Whitney test, figure 3). In parallel, performance in direct trials remained fairly constant across cycles, both in patients and controls. In particular, patient performance was far above chance level in all cycles and did not drop during the course of the experiment (cycle 1: 72.5 ± 6.8 %; cycle 4: 75.8 ± 7.1 %, figure 3).

In order to investigate whether the observed group differences for indirect trials in the last two cycles of the experiment mainly result from distinct changes in memory integration across cycles rather than from a deficit in associative memory, we again calculated accuracy of those indirect trials, for which both AB- and BC-stimuli were correctly remembered in the corresponding direct trials. With this analysis, in cycles 1 and 2, performance in indirect trials was almost identical in both groups (patients: 73.8 ± 10.7 , controls: 75.0 ± 6.0 %; $p = 0.82$, Mann-Whitney test). By contrast, in cycles 3 and 4, performance differed significantly (patients: 60.8 ± 6.0 , controls: 80.8 ± 4.3 %; $p = 0.015$, Mann-Whitney test). Furthermore, we quantified the contribution of memory integration to impaired performance in indirect trials by calculating an additional measure ('Integration index'). We subtracted accuracy of direct trials from accuracy in indirect trials in each subject, separately for cycles 1 and 2 and cycles 3 and 4. We reasoned that this measure controls for inter-individual differences in general associative memory ability and more directly reflects a subjects' integration performance. Thus, subjects with low memory integration capabilities should show low integration indices. We found that the Integration index did not differ between groups in cycles 1 and 2 (patients: -10.0 ± 6.6 , controls: -13.7 ± 2.4 , $p = 0.669$, Mann-Whitney test) but significantly in cycles 3 and 4 (patients: -25.0 ± 7.8 , controls: -3.2 ± 2.7 , $p = 0.008$, Mann-Whitney test), with a decrease of the index in patients and an increase in controls.

Reaction Times

We observed no significant differences in RTs between groups in direct and indirect trials (direct trials; patients: $3399 \text{ ms} \pm 264$, controls: $2786 \text{ ms} \pm 242$, $p = 0.12$; Mann-Whitney test; indirect trials; patients: $5283 \text{ ms} \pm 594$, controls: $4212 \text{ ms} \pm 386$, $p = 0.14$; Mann-Whitney test; figure 2 B). In direct trials, RTs remained fairly constant and decreased only slightly across cycles in both groups (Δ RT cycles 1-4, patients: $346 \pm 327 \text{ ms}$; controls: $303 \pm 242 \text{ ms}$;

p = 1.0 Mann-Whitney test; figure 3C). Similarly, both groups decreased their RTs in indirect trials, without significant differences between groups (Δ RT cycles 1-4, patients: 2325 ms \pm 1173, controls: 821 ms \pm 563; p = 0.54 Mann-Whitney test; figure 3D). In controls, the RT decrease may reflect practice effects and increased familiarity with the temporal structure of the paradigm. In patients, the RT decrease in indirect trials is difficult to interpret, as performance in cycles 3 and 4 of indirect trials was at chance level. Theoretically, patients may have become increasingly tired and simply tended to give priority to fast RTs over correct choices in cycles 3 and 4 in indirect trials. However, since direct and indirect trials were tested in the same retrieval blocks and since RTs and accuracy of direct trials did not deteriorate in patients, this hypothesis is not supported by our data.

Correlation analyses

We hypothesized that the pattern of results may reflect an increased reliance on integrated representations for AC-decisions as the experiment proceeded - at least in controls. With repeated necessity of AC-decisions, controls may have increasingly tended to form integrated ABC triplets out of AB- and BC-associations during or after the encoding blocks of the experiment in anticipation of future AC-decisions. Alternatively, subjects may have become more efficient in recombining AB- and BC-stimuli by the time an AC-decision is required. Imaging studies show that accuracy of AC-decisions in conditions promoting memory integration particularly correlates with hippocampal activation during encoding of BC-associations, presumably leading to reactivation of AB-associations and formation of integrated representations (Zeithamova et al., 2010, 2012b; Schlichting et al. 2014). Building on these results, we conducted exploratory correlation analyses between individual performance in indirect trials (AC-trials) and the three direct trial types (AB-, BC- and DE-trials) in the control group. The number of subjects in the patient group was too small for

these analyses. We expected a correlation between accuracy of AC-decisions and direct BC-trials in the case of memory integration. We further reasoned that if it is indeed mainly the efficient recombination of previously encoded AB-, BC- and DE-associations that determines AC-decisions, AC-performance may show a relationship with all of these direct trial types. This latter pattern was found for the beginning of the experiment. In cycle 1, AC-performance correlated significantly with performance in all direct trial types (AC-AB: $r = 0.75$, $p = 0.001$; AC-BC: $r = 0.76$, $p = 0.0001$; AC-DE: $r = 0.51$, $p = 0.04$; Spearman rank correlation, figure 4). At the end of the experiment, in cycle 4, AC-performance correlated significantly with performance in BC-trials only (AC-AB: $r = 0.44$, $p = 0.08$; AC-BC: $r = 0.57$, $p = 0.002$; AC-DE: $r = 0.39$, $p = 0.12$; Spearman rank correlation, figure 4).

Discussion

The results of our experiment show that patients with medial temporal lobe damage perform inferior to controls in a visual associative inference task that requires memory integration for correct inferential decisions. This deficit is not readily explainable by a memory deficit following hippocampal damage but appears to result from dysfunction of an additional integration mechanism. Our findings further suggest that this contribution is modulated by contextual factors.

The paradigm that we have used here is a modification of a visual associative inference task that has repeatedly been used to study how the brain constructs memory representations that can flexibly be used for future behavior (Preston et al., 2004; Zeithamova and Preston, 2010; Zeithamova et al., 2012b). Similar paradigms have incorporated information about value (Wimmer and Shohamy, 2012; Gerraty et al., 2014, Murty et al. 2016) or used verbal

material that was presented subliminally (Reber et al., 2012; Duss et al., 2014). Despite differences in stimulus material and the degree of consciousness of processing of to-be-remembered items, the results of these and other fMRI studies converge on a central role of the hippocampus within networks for memory integration (Zeithamova et al., 2012a; Palombo et al., 2015; Shohamy and Daw, 2015; Schlichting and Preston, 2015). In healthy controls, the individual intrinsic connectivity of the hippocampus to ventromedial prefrontal cortex has been shown to predict a subjects' memory integration performance (Gerraty et al., 2014). This interaction appears also to be critical for simulation and evaluation of outcomes by integration of activated representations of past experiences (Barron et al., 2013). Moreover, coupling between both regions may be involved in valuation and biasing of choice options during memory-guided decisions (Gluth et al., 2015).

The central role of hippocampal-neocortical interactions for memory integration conversely suggests a multitude of behavioral deficits, including decision making, in patients with hippocampal dysfunction that transcend the well-investigated memory deficits for 'direct' associations. Previous studies in patients with various neuropsychiatric disorders such as amnesic mild cognitive impairment or schizophrenia suggest that this might be the case (see Gleichgerrcht et al., 2010; Lee, 2013; Caceda et al., 2014 for reviews). The precise neurobiological underpinnings of these deficits are however unclear and probably manifold. In particular, whether and how deficient hippocampus-dependent memory integration may contribute to more complex behavioral impairments has only rarely been investigated in neuropsychiatric patients. In studies on subjects with schizophrenia with an associative inference paradigm that uses house-face associations, patients performed inferior to healthy controls in indirect trials, while performance in direct trials was relatively preserved (Armstrong et al., 2012). Both the overall pattern of results and the absolute accuracy values

show remarkable similarity to the performance levels in our study. Although it has been speculated that associative inference in schizophrenia may result from hippocampal dysfunction, it is principally difficult to establish unequivocal brain-behavior relationships in a complex neurodevelopmental disorder like schizophrenia or in a neurodegenerative disorder like Alzheimer's disease. So far, there have been no lesion studies of associative inference in human patients and only a handful of studies with related paradigms. In a study on aged human subjects with hippocampal atrophy, performance was deficient in an acquired equivalence task, i.e. a task that assesses a subjects' ability to generalize from previously trained associations on novel associations (Myers et al., 2003). Similar results were found in patients with amnesic syndromes following global cerebral hypoxia, a classic lesion model for hippocampal dysfunction (Myers et al., 2008). Despite uncertainties about the extent and selectivity of hippocampal damage in these disorders, neuro-computational simulation of hippocampal dysfunction during acquired equivalence yields behavioral impairments that match those observed in patients (Moustafa et al., 2010). Our data thus provide a link between imaging studies of memory integration / inferential reasoning in healthy controls and behavioral studies in neuropsychiatric patients, by showing that damage to the hippocampus is sufficient to create a memory integration deficit that manifests in inappropriate decisions, even when memory of individual associations is relatively preserved.

Our results complement recent imaging evidence that hippocampal and prefrontal contributions to memory integration are not constant, but rather modulated by contextual factors, e.g. the number of stimulus repetitions or blocked vs. mixed presentation of AB- and BC-stimuli during encoding blocks of associative inference tasks (Zeithamova et al. 2012b; Schlichting et al. 2015). Memory integration may occur by the time an AC-decision is

required and may be achieved by retrieving and integrating individual AB- and BC-associations ('recombination at retrieval', Zeithamova et al., 2012a or 'prospective integration', Shohamy and Daw, 2015). This option is highly flexible but associated with processing delays by the time a behavioral response is required (Schlichting et al., 2014; Shohamy and Daw, 2015). Integration may also occur at encoding or offline during the delay, with encoding of BC-associations leading to re-activation of previous AB-associations and subsequent integration to ABC-representations before an AC-decision is required ('integrative encoding', Zeithamova et al., 2012a or 'retrospective integration', Shohamy and Daw 2015). Previous imaging studies suggest that the hippocampus may be involved in integration at retrieval (Heckers et al., 2004; Preston et al., 2004). More recent studies have shown that the hippocampus may equally be important for integration at encoding (Shohamy and Wagner, 2008; Zeithamova and Preston, 2010; Zeithamova et al., 2012b; Wimmer and Shohamy, 2012). This function appears to be particularly dependent on posterior ventromedial prefrontal cortex and the anterior hippocampus on the right side – at least in visual integration tasks (Schlichting et al., 2015), i.e. the brain region that was affected in all our patients. By using a visual associative inference paradigm that inspired the design of our study, it has further been shown that activity in hippocampus and parahippocampal cortex during encoding of individual associations relates to later AC-inference (Zeithamova et al. 2010). Notably, it was activity in BC-trials that showed this effect, suggesting that stimuli in BC-trials triggered hippocampal pattern completion mechanisms that reactivated preceding overlapping AB-representations (Zeithamova et al., 2010). In a subsequent study, the degree of reactivation of AB-representations in anterior MTL-regions during encoding in BC-trials directly predicted AC-inference (Zeithamova et al., 2012b). Consistent with the idea that integration at encoding reduces processing delays at

retrieval, it has been shown that neural activity in hippocampal CA1 is more similar between BC- and AC-trials when the reaction time in AC-trials is short (Schlichting et al., 2014). The results from our explorative correlation analyses are compatible with these findings and suggest that memory integration does not occur automatically, but may also be modulated by the repetition of encoding-retrieval cycles as contextual factor. The differential correlations between indirect and direct trials at the beginning and at the end of our experiment must however be interpreted with caution, as the sample size of our control group was small. Weaker but relevant correlations between memory of direct associations and inferential decisions may have remained undetected. It would thus be premature to suggest a major strategy change for memory integration across the cycles of our experiment - in addition to likely practice effects. Furthermore, recombination at retrieval and integrative encoding are not mutually exclusive and can both be recruited during a single task (Zeithamova et al., 2012a; Palombo et al., 2015; Shohamy and Daw, 2015). Inferential decisions may thus be influenced by both mechanisms, with only relative predominance of one or the other. Together with the changes in integration index, our results may thus indicate an increasing tendency of control subjects to form integrated ABC-representations during or following encoding in anticipation of future AC-decisions in indirect trials, but they do not rule out a significant contribution of recombination at retrieval.

While average performance of patients in indirect trials was – as expected - inferior to controls, the pattern of accuracy across cycles is surprising. Despite increasing familiarity with the task, patients showed less efficient behavior with deterioration of accuracy across cycles. It may be speculated that the underlying mechanism was the same as in controls, i.e. the formation of integrated representations, presumably reflecting a still intact cognitive routine for new behavioral contexts that is however prone to fail, since it depends on an

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

intact hippocampus. It will thus be important to investigate whether it is possible to revert the increasing impairment by creating conditions that discourage reliance on integrated representations, e.g. by providing feedback about deficient performance during the experiment. The dynamics of our patients' deficits further suggest that the role of the hippocampus within networks for memory integration is dependent on contextual factors. Therefore, in some conditions, inferential decisions may sufficiently be supported by networks outside the damaged medial temporal lobe. Whether these networks involve intact residual hippocampus or neocortex may be clarified by combining lesion with functional imaging studies.

Taken together, the findings reported here suggest that lesions to the medial temporal lobe can impair performance in tasks requiring memory integration. The magnitude of behavioral deficits may depend on whether a particular behavioral context promotes the formation of integrated representations in anticipation of future decisions. It remains to be determined how this deficit translates into real-world decision making in neuropsychiatric patients and whether strategies that encourage less hippocampus-dependent mechanisms of memory integration optimize behavior in patients with disorders affecting the medial temporal lobe.

References

- Altman DG. 1991. Practical Statistics for Medical Research. London: Chapman & Hall.
- Altman DG, Bland JM. 2009. Parametric v non-parametric methods for data analysis. *BMJ*. 338:a3167.
- Armstrong K, Williams LE, Heckers S. 2012. Revised associative inference paradigm confirms relational memory impairment in schizophrenia. *Neuropsychology* 26:451-8.
- Barron HC, Dolan RJ, Behrens TEJ. 2013. Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat Neurosci* 16:1492–1498.
- Bunsey M, Eichenbaum H. 1996. Conservation of hippocampal memory function in rats and humans. *Nature*. 379:255-7.
- Cáceda R, Nemeroff CB, Harvey PD. 2014. Toward an Understanding of Decision Making in Severe Mental Illness. *J Neuropsychiatry Clin Neurosci* 26: 196–213.
- Dusek JA, Eichenbaum H. 1997. The hippocampus and memory for orderly stimulus relations. *Proc Natl Acad Sci U S A*. 94:7109-14.
- Duss SB, Reber TP, Hänggi J, Schwab S, Wiest R, Müri RM, Brugger P, Gutbrod K, Henke K. 2014. Unconscious relational encoding depends on hippocampus. *Brain* 137: 3355–3370.
- Esfahani-Bayerl NE, Finke C, Braun M, Düzel E, Heekeren HR, Holtkamp M, Hasper D, Storm C, Ploner CJ. 2016. Visuo-spatial memory deficits following medial temporal lobe damage: A comparison of three patient groups. *Neuropsychologia* 81:168–179.
- Gerraty RT, Davidow JY, Wimmer GE, Kahn I, Shohamy D. 2014. Transfer of Learning Relates to Intrinsic Connectivity between Hippocampus, Ventromedial Prefrontal Cortex, and

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Large-Scale Networks. *J Neurosci* 34: 11297–303.

Gleichgerricht E, Ibanez A, Roca M, Torralva T, Manes F. 2010. Decision-making cognition in neurodegenerative diseases. *Nat Rev Neurol* 6:611–623.

Gluth S, Sommer T, Rieskamp J, Büchel C. 2015. Effective Connectivity between Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices from Memory. *Neuron* 86:1078–1090.

Han SD, Boyle PA, James BD, Yu L, Bennett DA. 2015. Mild cognitive impairment is associated with poorer decision-making in community-based older persons. *J Am Geriatr Soc.* 63: 676–683.

Heckers S, Zalesak M, Weiss AP, Ditman T, Titone D. 2004. Hippocampal activation during transitive inference in humans. *Hippocampus* 14:153-62.

Horn W. 1983. Leistungsprüfsystem. Göttingen, Germany: Hogrefe Verlag.

Horner AJ, Burgess N. 2014. Pattern completion in multielement event engrams. *Curr Biol.* 24:988-92.

Lee D. 2013. Decision making: from neuroscience to psychiatry. *Neuron* 78:233-48.

Lehrl S. 2005. Mehrfachwahl-Wortschatz-Intelligenztest. Göttingen, Germany: Hogrefe Verlag.

Martin RC, Triebel KL, Kennedy RE, Nicholas AP, Watts RL, Stover NP, Brandon M, Marson DC. 2013. Impaired financial abilities in Parkinson's disease patients with mild cognitive impairment and dementia. *Parkinsonism Relat Disord* 19:986-90.

Moustafa AA, Keri S, Herzallah MM, Myers CE, Gluck MA. 2010. A neural model of hippocampal-striatal interactions in associative learning and transfer generalization in

- various neurological and psychiatric patients. *Brain Cogn* 74:132-44.
- Murty VP, FeldmanHall O, Hunter LE, Phelps EA, Davachi L. 2016. Episodic memories predict adaptive value-based decision-making. *J Exp Psychol Gen.* 145:548-58.
- Myers CE, Shohamy D, Gluck MA, Grossman S, Kluger A, Ferris S, Golomb J, Schnirman G, Schwartz R. 2003. Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J Cogn Neurosci* 15:185–193.
- Myers CE, Hopkins RO, DeLuca J, Moore NB, Wolansky LJ, Sumner JM, Gluck MA. 2008. Learning and generalization deficits in patients with memory impairments due to anterior communicating artery aneurysm rupture or hypoxic brain injury. *Neuropsychology* 22:681-6.
- Palombo DJ, Keane MM, Verfaellie M. 2014. The medial temporal lobes are critical for reward-based decision making under conditions that promote episodic future thinking. *Hippocampus* 24: 345–353.
- Palombo DJ, Keane MM, Verfaellie M. 2015. How does the hippocampus shape decisions? *Neurobiol Learn Mem* 125:93-7.
- Preston AR, Shrager Y, Dudukovic NM, Gabrieli JDE. 2004. Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus* 14:148–52.
- Preston AR, Eichenbaum H. 2013. Interplay of hippocampus and prefrontal cortex in memory. *Curr Biol* 23:R764– R773.
- Reber TP, Luechinger R, Boesiger P, Henke K. 2012. Unconscious Relational Inference Recruits the Hippocampus. *J Neurosci* 32: 6138–6148.
- Schlichting ML, Zeithamova D, Preston AR. 2014. CA1 subfield contributions to memory

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

integration and inference. *Hippocampus* 1260:1248–1260.

Schlichting ML, Mumford JA, Preston AR. 2015. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Comm* 6: 8151.

Schlichting ML, Preston AR. 2015. Memory integration: neural mechanisms and implications for behavior. *Curr. Opin. Behav. Sci* 1, 1–8.

Shadlen MN, Shohamy D. 2016. Decision Making and Sequential Sampling from Memory. *Neuron* 90:927–939.

Shohamy D, Daw ND. 2015. Integrating memories to guide decisions. *Curr. Opin. Behav. Sci.* 5:85–90.

Shohamy D, Turk-Browne NB. 2013. Mechanisms for widespread hippocampal involvement in cognition. *Exp Psychol Gen* 142:1159–1170.

Shohamy D, Wagner AD. 2008. Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* 60:378-89.

Wimmer GE, Shohamy D. 2012. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* 338:270–273.

Zeithamova D, Schlichting ML, Preston AR. 2012a. The hippocampus and inferential reasoning: building memories to navigate future decisions. *Front Hum Neurosci* 6:1-14.

Zeithamova D, Dominick AL, Preston AR. 2012b. Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron* 75: 168–179.

Zeithamova D, Preston AR. 2010. Flexible memories: differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *J Neurosci* 30:14676–14684.

For Peer Review

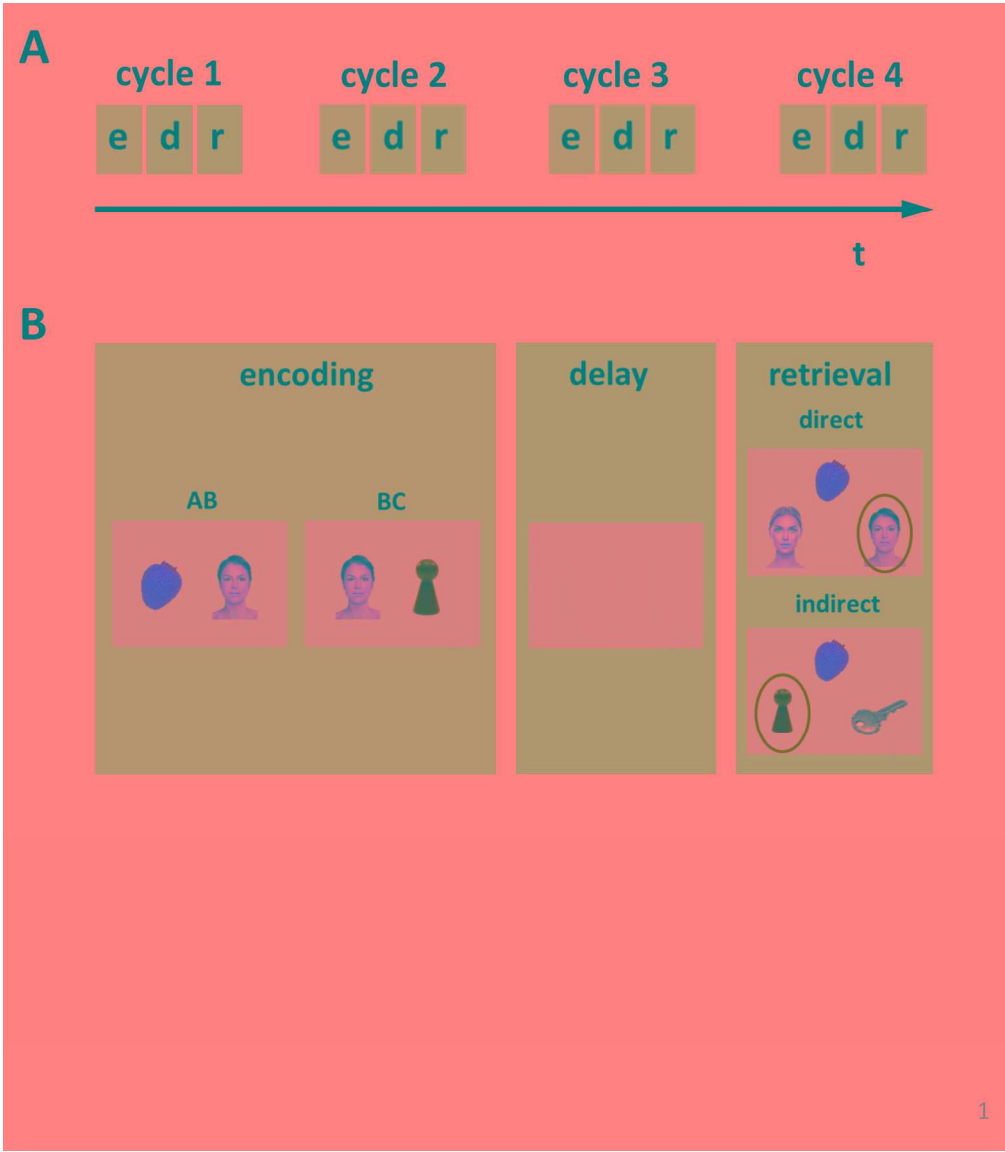
Legends

Figure 1. Task and stimuli. A, structure of the experiment; four consecutive cycles consisting of encoding- and retrieval-blocks ('e' and 'r'), separated by a delay ('d'). B, example stimuli during encoding and retrieval blocks. Note overlap of AB- with BC-stimuli during encoding and distinct stimulus configurations of direct and indirect retrieval trials. Correct choices are encircled in green.

Figure 2. Group results in patients (black bars) and controls (white bars). A, accuracy in direct and indirect trials; B, reaction time in direct and indirect trials; error bars, standard error of the mean; * $p = 0.038$ significant difference between patients and controls.

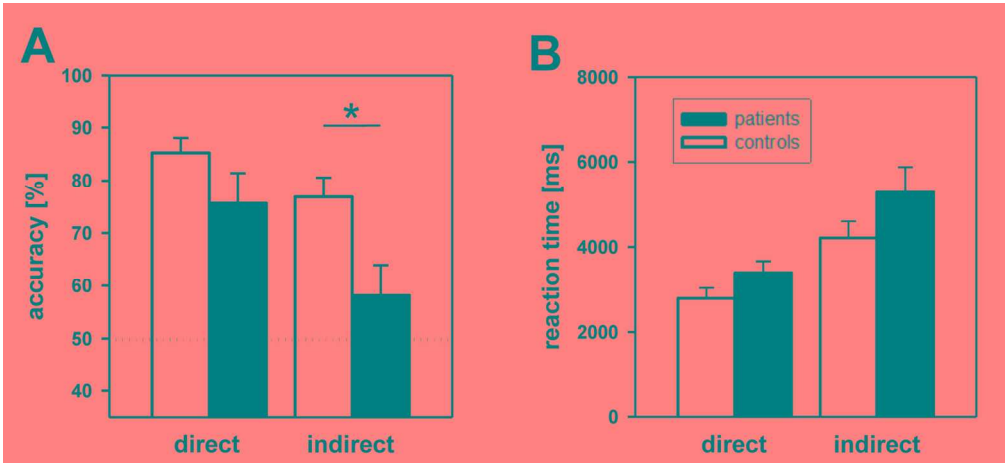
Figure 3. Results in patients (black dots) and controls (white dots) across the four cycles of the experiment. A, accuracy in direct trials; B, accuracy in indirect trials; C, reaction time in direct trials; D, reaction time in indirect trials; error bars, standard error of the mean; * $p = 0.015$, ** $p = 0.005$ significant difference between patients and controls.

Figure 4. Correlation analyses in controls ($n = 17$ subjects). Correlation of accuracy in indirect trials (AC-accuracy, x-axis) with accuracy in direct trials (AB-, BC-, DE-accuracy, y-axis). A, B, C, correlations for cycle 1 of the experiment; D, E, F, correlations for cycle 4 of the experiment. r , correlation coefficient from Spearman rank correlation. Dot size indicates number of identical values.



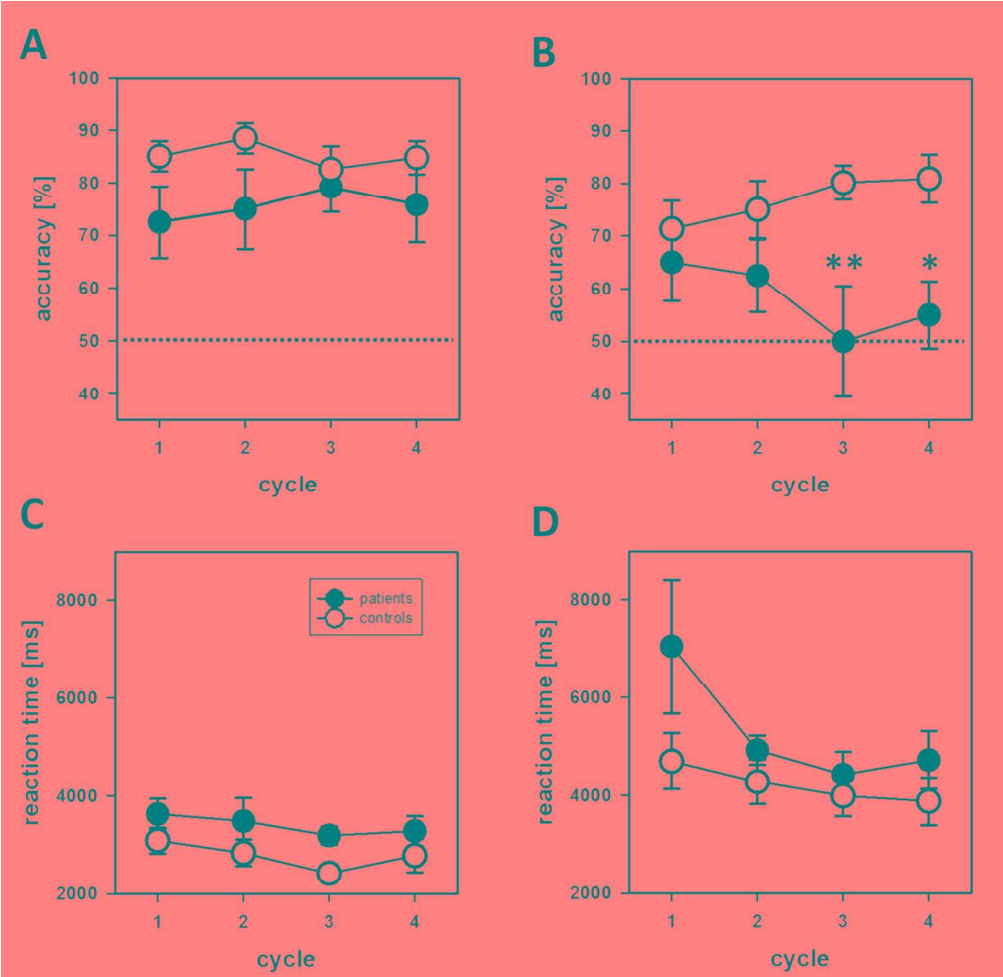
Task and stimuli. A, structure of the experiment; four consecutive cycles consisting of encoding- and retrieval-blocks ('e' and 'r'), separated by a delay ('d'). B, example stimuli during encoding and retrieval blocks. Note overlap of AB- with BC-stimuli during encoding and distinct stimulus configurations of direct and indirect retrieval trials. Correct choices are encircled in green.

165x190mm (300 x 300 DPI)



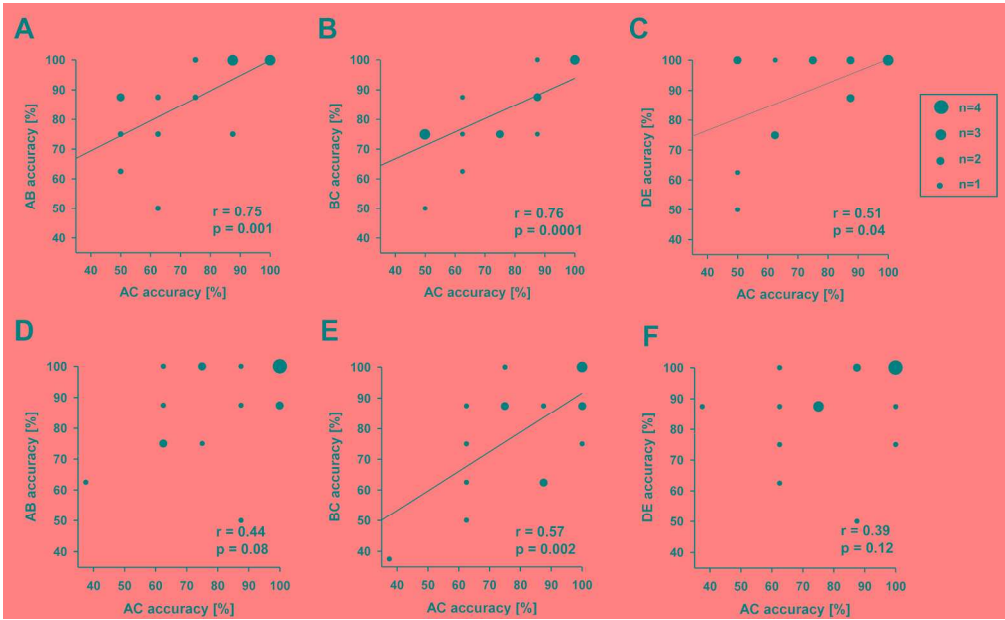
Group results in patients (black bars) and controls (white bars). A, accuracy in direct and indirect trials; B, reaction time in direct and indirect trials; error bars, standard error of the mean; * $p = 0.038$ significant difference between patients and controls.

146x66mm (300 x 300 DPI)



Results in patients (black dots) and controls (white dots) across the four cycles of the experiment. A, accuracy in direct trials; B, accuracy in indirect trials; C, reaction time in direct trials; D, reaction time in indirect trials; error bars, standard error of the mean; * $p = 0.015$, ** $p = 0.005$ significant difference between patients and controls.

148x144mm (300 x 300 DPI)



Correlation analyses in controls (n = 17 subjects). Correlation of accuracy in indirect trials (AC-accuracy, x-axis) with accuracy in direct trials (AB-, BC-, DE-accuracy, y-axis). A, B, C, correlations for cycle 1 of the experiment; D, E, F, correlations for cycle 4 of the experiment. r, correlation coefficient from Spearman rank-sum correlation. Dot size indicates number of identical values.

247x151mm (300 x 300 DPI)

Table 1: Individual patient characteristics. R-MTL, post-surgical damage to the right medial temporal lobe; B-MTL post-encephalitic damage to the right and left medial temporal lobe; F, female, M, male

Patient	Diagnosis	Sex	Age	Years of education	Time since lesion (months)	Clinical notes / Histopathology	Anticonvulsant medication/ centrally acting drugs
1	R-MTL	F	42	12	133	Pilomyxoid astrocytoma	gabapentin 900 mg/d
2	R-MTL	M	24	13	107	Neuroepithelial tumor	lamotrigine 200 mg/d
3	R-MTL	M	24	17	155	Pigmented astrocytoma	none
4	B-MTL	M	53	15	53	Herpes encephalitis, HSV1	citalopram 20mg/d; opipramol 150mg/d
5	R-MTL	F	57	15	11	Low-grade glioma	levetiracetam 2000 mg/d eslicarbazepine 1200 mg/d

Table 2: Lesion analysis. R, right; L, left; HIP, hippocampus; ERC, entorhinal cortex; PRC, perirhinal cortex; PHC, parahippocampal cortex; 0, region unaffected; +, rostrocaudal lesion extent ≤ 20 mm; ++, ≤ 40 mm; +++, > 40 mm

Subject	R-HIP	R-ERC	R-PRC	R-PHC	L-HIP	L-ERC	L-PRC	L-PHC
1	+	+	+	0	0	0	0	0
2	+	++	++	0	0	0	0	0
3	+++	++	++	+	0	0	0	0
4	++	++	++	+	++	++	++	+
5	++	+	+	0	0	0	0	0