Impaired processing of response conflicts in mesial temporal lobe epilepsy

Markus Ramm1,2, Gabriel Möddel3, Benedikt Sundermann4, Annegret Last1, Lisa Langenbruch3, Johannes Jungilligens5, Jörg Wellmer5, Peter Young2 and Nikolai Axmacher1*

1Department of Neuropsychology, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr University Bochum, Germany
2Institute of Sleep Medicine and Neuromuscular Disorders, University Hospital Muenster, Germany
3Department of Neurology with Institute of Translational Neurology, University Hospital Muenster, Germany
4Institute of Clinical Radiology, University Hospital Muenster, Germany
5Ruhr-Epileptology, Department of Neurology, University Hospital Knappschaftskrankenhaus, Ruhr University Bochum, Germany

Increasing evidence from neuroimaging studies points towards a hippocampal role in resolving approach-avoidance goal conflicts. Furthermore, previous neuroimaging findings suggest that the hippocampus (HC) contributes to successful conflict resolution as it is measured, for example, in a Stroop paradigm. However, it is still an open question whether the hippocampus is indeed causally relevant for resolving cognitive conflicts. Here, we investigated whether conflict resolution performance is affected by hippocampal pathology. N = 30 patients with mesial temporal lobe epilepsy (MTLE), almost exclusively showing MRI signs of hippocampal sclerosis, and an equal number of age-matched healthy controls performed an auditory Stroop paradigm. Participants listened to the words ‘high’ and ‘low’, spoken in either a high or a low pitch. Subjects’ response time and accuracy to the phonetic information in the presence of incongruent (conflict trials) or congruent (non-conflict trials) semantic information were assessed. In addition, patients’ regional grey matter (GM) brain volumes were analysed. We observed an increased effect of conflict on accuracy in patients with MTLE compared to healthy controls. This effect was negatively correlated with right HC volume. The results suggest that the impairment in the resolution of a response conflict is related to hippocampal structural integrity and thus add further support to the notion that the HC is not only involved but even causally relevant for successful cognitive conflict processing.

The hippocampus (HC) is primarily known for its crucial role in episodic memory and spatial navigation (Buzsaki & Moser, 2013; Scoville & Milner, 1957). In addition, the last years have witnessed an increasing interest in the contribution of the HC to processes beyond these domains (Bach et al., 2014; Chan, Morell, Jarrard, & Davidson, 2001; 

*Correspondence should be addressed to Nikolai Axmacher, Department of Neuropsychology, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr University Bochum, Universitätsstr. 150, D-44801 Bochum, Germany (email: nikolai.axmacher@ruhr-uni-bochum.de).

DOI:10.1111/jnp.12186
Davidson & Jarrard, 2004; Ito & Lee, 2016; Oehrn et al., 2015; Sakimoto, Okada, Hattori, Takeda, & Sakata, 2013; Schumacher, Vlassov, & Ito, 2016). Specifically, empirical evidence from a variety of different research fields points towards a hippocampal role in processing conflicts, that is, in resolving competition between incompatible stimulus representations, goals, or responses. First, several studies indicate that the HC supports the detection of item-context mismatches (Kumaran & Maguire, 2006, 2007; Thakral, Yu, & Rugg, 2015). Second, the HC is crucially involved in pattern separation, that is, in the orthogonalization of similar representations in order to reduce interference in episodic memory (Bakker, Kirwan, Miller, & Stark, 2008; Lee, Yoganarasimha, Rao, & Knierim, 2004; Leutgeb, Leutgeb, Treves, Moser, & Moser, 2004; Vazdarjanova & Guzowski, 2004). These findings suggest a hippocampal role in processing perceptual conflicts. In addition, conflicts may also arise at the level of goals and responses. Early research has demonstrated that hippocampal lesions in rats induce behavioural rigidity as evidenced by impairments of extinction learning (Isaacson & Wickelgren, 1962; Jarrard & Lewis, 1967). By now, the rodent literature offers convincing evidence that the HC is crucially involved in inhibitory response control, particularly under circumstances of an approach-avoidance conflict (Abela, Dougherty, Fagen, Hill, & Chudasama, 2013; Chudasama, Doobay, & Liu, 2012; Schumacher et al., 2016; Schumacher et al., 2018). Fewer studies investigated whether the HC is also relevant for conflict resolution in humans.

One recent functional magnetic resonance imaging (fMRI) study with healthy subjects and patients with mesial temporal lobe epilepsy (TLE) due to hippocampal sclerosis showed that adaptive behaviour in a computerized approach-avoidance conflict paradigm was related to left HC BOLD responses (Bach et al., 2014). Moreover, patients’ behaviour was less influenced by an aversive stimulus, providing first causal evidence for a role for the human HC in approach-avoidance conflict processing.

In later fMRI and magnetoencephalographic studies, the hippocampal role in approach-avoidance conflicts was confirmed and elucidated in more detail (Khemka, Barnes, Dolan, & Bach, 2017; Loh et al., 2017; O’Neil et al., 2015).

Does the role of the HC in conflict processing also extend to more general response conflicts? The Stroop task is a classic paradigm to investigate response conflicts (Stroop, 1935). In the most common version of this task, subjects are asked to name the colour of an incongruent colour word (e.g., the word ‘red’ written in a green font) or a congruent colour word (e.g., the word ‘red’ written in a red font). An incongruent colour word elicits a response conflict because it requires participants to inhibit a dominant response tendency (reading the colour word) that interferes with a less automatic but goal-relevant behaviour (naming the colour). The resolution of response conflicts is accompanied by less accurate responses and increased reaction times. On a neural level, the ‘conflict monitoring and cognitive control theory’ (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Carter & van Veen, 2007) has arguably become the most influential framework of conflict processing. This model proposes that the anterior cingulate cortex (ACC) acts as a conflict monitoring system which continuously evaluates levels of conflict, then passes this information to prefrontal brain regions responsible for cognitive control (mainly the dorsolateral prefrontal cortex, dPFC) which eventually adjusts information processing. Once additional resources have been recruited, the performance costs due to conflict are reduced, a phenomenon known as conflict adaptation. Imaging studies support a role of the ACC in detecting conflicts and a role for the dPFC in mediating cognitive control (Barch, Braver, Sabb, & Noll, 2000; MacDonald, Cohen, Stenger, & Carter, 2000; Oehrn et al., 2014).
The HC is not among the core areas that support conflict processing in this framework, and is typically not found to be activated in neuroimaging studies that use the Stroop paradigm (e.g. Haupt, Axmacher, Cohen, Elger, & Fell, 2009). However, conventional whole-brain fMRI analyses may not be sufficiently sensitive to detect a contribution of the HC. In order to specifically test a role of the HC in resolution of response conflict, Oehrn et al. (2015) combined two more sensitive methods: First, they investigated HC activity during an auditory Stroop task by applying a region of interest analysis to fMRI data in healthy subjects. Second, they recorded intracranial electroencephalography (iEEG) in presurgical TLE patients implanted with hippocampal depth electrodes. In brief, they found greater fMRI BOLD responses and iEEG theta (3–8 Hz) oscillations in the left HC during inconsistent as compared to consistent trials. Importantly, iEEG effects occurred only during correct trials and were related to reaction times, suggesting that the HC indeed plays a functional role in resolving response conflicts in the Stroop paradigm. However, while neuropsychological data suggest a crucial role of the HC for resolving approach-avoidance conflicts as described above (Bach et al., 2014), it remains unclear whether the HC is also causally relevant for conflict processing in the Stroop task.

Patients with mesial TLE due to circumscribed HC lesions allow further investigating this question. Episodic memory impairment is the main neuropsychological finding in mesial TLE patients (Helmstaedter, Grunwald, Lehnerz, Gleissner, & Elger, 1997; Tramoni-Negre, Lambert, Bartolomei, & Felician, 2017). Moreover, previous neuropsychological studies consistently revealed deficits in a range of executive functions in patients with mesial TLE (Zamarian et al., 2011; Zhao et al., 2014). Few neuropsychological studies used paradigms specifically targeting conflict resolution. An impairment in the colour word Stroop task was indicated in two studies (Pinto et al., 2017; Wang et al., 2007), whereas another study found no differences between mesial TLE and frontal epilepsy (Corcoran & Upton, 1993). In contrast to the fMRI findings in healthy subjects showing left lateralized BOLD responses during conflict resolution (Bach et al., 2014; Oehrn et al., 2015), the neuropsychological studies did not report a lateralization effect for unilateral mesial TLE.

The aim of the present study was to systematically explore the role of the HC for response conflict resolutions in patients with mesial TLE. Since different TLE-related pathologies may differ in their temporal evolution and the amount of possible spread of epileptic activity (and possibly neuropsychological impairment), we focused on patients with magnetic resonance imaging (MRI) signs of a hippocampal sclerosis. We investigated whether these patients showed deficits in the resolution of, or adaptation to, response conflicts during an auditory Stroop task. Moreover, we hypothesized patients’ performance to be related to the structural integrity of their HC formation as reflected by the residual grey matter volume. As a previous fMRI study using a similar auditory Stroop paradigm found conflict-related BOLD responses in the left HC (Oehrn et al., 2015), we expected mesial TLE patients with a left seizure focus to be more severely impaired than patients with right mesial TLE.

Methods

Subjects

N = 30 patients with mesial temporal lobe epilepsy (MTLE) according to International League Against Epilepsy criteria (Scheffer et al., 2017) were included in the study (age: 46.7 ± 15.2 years; 13 females; age at seizure onset: 26.9 ± 16.0 years; disease duration:
Patients were asked for their handedness, most of them (83%) were right-handed. Patients underwent a detailed neurological examination, EEG and high-resolution MRI. Based on EEG and seizure semiology, 26 patients were diagnosed as unilateral MTLE (left MTLE: \( n = 13 \); right MTLE: \( n = 13 \)), four patients had a bilateral MTLE. Twelve patients received antiepileptic drug (AED) monotherapy, 18 patients were on polytherapy (AED doses are reported in Table S1). Most patients had a therapy-refractory epilepsy according to International League Against Epilepsy criteria (Kwan et al., 2010). Patients with a multifocal epilepsy, comorbid neurological disorders or severe psychiatric disorders were excluded. When applicable, the presence of a comorbid psychiatric disorder was documented.

The clinical MRI scans (when available) were visually analysed by two independent radiologists following a standardized protocol (Dekeyzer et al., 2017). All findings were in consensus between the two radiologists. Classic signs of HS were as follows: volume reduction, abnormal internal layer and increased signal in T2-weighted (T2w) images. Detailed findings of individual MRI scans are listed in the Patient characteristics (Table S1). Signs of HS were detected in most of the MTLE patients. More specifically, unilateral signs of HS were found in the left HC in 12 of 13 (92.3%) patients with left MTLE. In one patient, MRI visual analysis of the latest scans did not confirm reliable signs of HS. A clear unilateral abnormality in the right HC was present in 11 of 13 (84.6%) patients with right MTLE. In one patient with right MTLE, bilateral signs of HS were documented, in the other patient, no reliable signs of HS were found in the latest clinical scans. All four patients with bilateral MTLE revealed congruent bilateral signs of HS in the MRI.

In addition to patients, \( n = 30 \) age-matched healthy controls participated in the study (age: 46.3 ± 14.8 years; 17 females). Subjects with a neurological or psychiatric disorder, a CNS medication or known cerebral lesions were excluded. All participants gave written informed consent. The study was approved by the local ethics committee.

**Auditory Stroop task**

Participants performed an auditory Stroop paradigm (Figure 1) identical to the one used in previous fMRI and iEEG studies (Haupt et al., 2009; Oehrn et al., 2015; Oehrn et al., 2014). These studies had consistently reported conflict effects in the auditory domain.

During the task, participants listened to the German equivalents of the English words ‘high’ and ‘low’, spoken in either a high or a low pitch. This results in congruent trials in which semantic and phonetic information are consistent and incongruent trials with inconsistent stimulus characteristics. As a control condition, the German word for ‘good’ was presented in either high or low pitch. The paradigm consisted of two blocks that differed in their task instruction: in the first block (semantic task), subjects were asked to indicate the word meaning (‘low’ vs. ‘high’, irrespective of pitch), while in the second block (phonetic task), participants had to identify whether the word was spoken in high or low pitch (regardless of the word meaning). The participants responded by left and right button presses with their dominant hand, counterbalanced across subjects. Responding to control trials was required only during the phonetic task. Behavioural effects of conflict were only expected in trials in which the subject was required to inhibit the more automatic response to the word meaning, that is, incongruent trials in the phonetic task.

The paradigm consisted of two blocks with 240 trials in total. Each block comprised 40 inconsistent, 40 consistent and 40 control trials, presented in a randomized order. All participants performed at least 10 practice trials to familiarize themselves with the
During this period, the experimenter verified that participants understood the instructions correctly. The auditory stimuli were presented for 0.5 s, followed by a constant 2s interval during which the task instructions remained on the screen. Afterwards, a fixation cross was presented for a variable duration of 1.5s – 3.3s. The words were spoken by a male person, transposed to high or low pitch and aligned to an equal length of 500 ms. For stimulus presentation, we used Presentation Software (Version 19.0; Neurobehavioral Systems, San Francisco, CA, USA).

**Behavioural data analysis**
In the auditory Stroop task, mean response time (RT; including only correct responses) and response accuracy were measured. Accuracy scores reflect the ratio between the number of correct responses and the total number of trials in a specific condition. In a control analysis of the accuracy measure, we excluded all miss trials. Moreover, the inverse efficiency score (IES) was estimated as a performance parameter combining both RT and accuracy (Bruyer & Brysbaert, 2011). For each participant, the IES was calculated by mean RT of correct responses divided by the proportion of correct responses. The IES is expressed in milliseconds.

Moreover, we performed an analysis based on signal detection theory (SDT). SDT is an analytical method for measuring behavioural performance independent of inter-individual differences in response tendencies for specific trial categories (Green & Swets, 1966; Macmillan & Creelman, 2005). Here, we defined the ratio of correct identifications of ‘high’ trials as hit rate (H) and the ratio of incorrect estimations of ‘low’ trials (as high) as false alarm rate (FA; note that the same results are obtained when focusing on hits and false alarms for low trials). Signal detection theory separates performance into sensitivity ($d'_{\text{prime}}$; i.e., how well correct and incorrect stimuli are discriminated) and response criterion ($c$; i.e., a participant’s bias to select one response over the other). We calculated the sensory index of discriminability ($d'_{\text{prime}} = z(H) - z(FA)$) and the response criterion

![Figure 1](image-url). Auditory Stroop paradigm. Subjects responded to the words ‘high’, ‘low’ and ‘good’ spoken in either a high or low pitch. In the semantic task indication of word meaning was required, in the phonetic task subjects had to name the pitch. Incongruent semantic information during the phonetic task was expected to result in a behavioural conflict.
In each participant, we contrasted the responses in incongruent and congruent trials of the phonetic task in order to compare conflict effects on the dependent variables between MTLE and controls. In addition, we assessed effects of facilitation (phonetic-congruent trials vs. phonetic-control trials) and of interference (phonetic-incongruent trials vs. phonetic-control trials) and compared them between the two groups. In order to measure effects of conflict adaptation, we compared responses in conflict trials that directly followed a correct conflict trial with responses in conflict trials that followed a correct non-conflict trial.

Moreover, in the MTLE group we investigated the effects of lesion lateralization by comparing patients with a clear unilateral (left vs. right) pathology.

Statistical analysis was carried out with IBM SPSS® Statistics Software (version 25.0, IBM, Armonk, NY, USA) and R Software (version 3.5.0, The R foundation for Statistical Computing, 2018, Vienna, Austria). The SDT measures $d'$ and $c$, the IES and RT measures were analysed using two-way mixed analyses of variances (ANOVAs) with ‘consistency’ (inconsistent vs. consistent) as a within subject factor and ‘group’ (MTLE vs. healthy controls) as a between subjects factor. None of the accuracy measures was normally distributed within groups and homogeneity of variance between groups was not given. Thus, we performed robust mixed ANOVAs using a 10% trimmed mean accuracy based on Wilcox’ WRS2 functions (Mair & Wilcox, 2018). Moreover, follow-up non-parametric Mann–Whitney U-tests were conducted in order to compare accuracy measures between groups. A multiple regression analysis was conducted to explore the influence of age at seizure onset, disease duration, mono- vs. polytherapy and depression on behavioural measures of conflict. Pearson’s correlation coefficients were used to analyse relationships between regional brain volume and RT data, Kendall’s tau was used for the relationship between brain volume and accuracy scores.

For all statistical analyses, the alpha level was set to 5%. Pearson’s correlation coefficient $r$ was used as an effect size for significant effects.

**MRI data acquisition and volumetric analysis**

A subgroup of patients ($n = 25$) who had clinical MRI data of sufficient quality were used for retrospective volumetric analysis. $N = 21$ patients underwent MRI using a 3T Achieva Philips MR scanner (Philips Medical Systems, Best, NL, USA) equipped with a six-channel head coil. MRI scanning included a 3D T1-weighted (T1w) turbo field echo (TFE) sequence (repetition time [TR] = 7.0ms, echo time [TE] = 3.4ms; 160 slices; matrix = 256 x 255; Field of view [FOV] = 256 x 256mm, slice thickness (reconstructed) = 1mm; flip angle = 9°). Moreover, paracoronal T2-weighted (T2w) images perpendicular to the long axis of the temporal lobes were acquired using one of two turbo spin echo (TSE) sequences (sequence 1: TR = 3,000 ms, TE = 80 ms, parallel imaging factor [SENSE] = 1.5, turbo factor = 15, 72 slices, FOV = 180 x 180 mm, acquisition matrix = 288 x 225, slice thickness = 2 mm, slice gap = 0.2 mm; sequence 2: TR = 5,169 ms, TE = 118 ms, parallel imaging factor [SENSE] = 1.5, turbo factor = 26, 38 slices, FOV = 200 x 200mm, acquisition matrix = 372 x 312, slice thickness = 2 mm, slice gap = 0.2 mm).

$N = 4$ patients underwent MRI investigation using a 3T Siemens Magnetom Prisma MR scanner (Siemens Medical Solutions, Erlangen, Germany). Scanning comprised a 3D T1w magnetization prepared gradient echo (MP-RAGE) sequence (TR = 2,200 ms, TE = 2.29 ms;
parallel imaging factor [GRAPPA] = 2; 176 slices; matrix = 256 × 256; FOV = 256 × 256 mm, slice thickness = 1 mm; flip angle = 8°) and a paracoronal T2w TSE sequence (TR = 3,600 ms, TE = 103 ms, turbo factor = 20, 31 slices, FOV = 220 × 150, acquisition matrix = 384 × 261, slice thickness = 2 mm, slice gap = 0.2 mm).

Processing of the T1w MRI scans and region of interest (ROI) volumetric analysis of grey matter volumes were performed using the FreeSurfer (version 6.0) ‘recon-all’ pipeline which includes a surface-based stream as well as a volume-based stream (https://surfer.nmr.mgh.harvard.edu/). Technical details of the procedure are described in previous publications (Dale, Fischl, & Sereno, 1999; Fischl et al., 2002; Fischl et al., 2004; Fischl, Sereno, & Dale, 1999). In brief, the fully automated procedure includes removal of non-brain tissue, registration to a common stereotactic space, image correction for magnetic field inhomogeneity, segmentation of subcortical white matter (WM) and grey matter (GM), tessellation of the WM-GM boundary, correction for topological errors and aligning of surface models to a spherical atlas based on individual folding patterns. Given our a priori hypotheses, we restricted our volumetric analysis to the HC formation as well as to the ACC and dlPFC, two brain regions that have been consistently related to conflict processing. GM volumes of ACC and dlPFC were extracted via the cortical parcellation procedure and automatically labelled based on the Destrieux Atlas dividing each hemisphere into 74 regions (Destrieux, Fischl, Dale, & Halgren, 2010). The dlPFC comprises the regions 15 (middle frontal gyrus), 16 (superior frontal gyrus), 52 (inferior frontal sulcus), 53 (middle frontal sulcus) and 54 (superior frontal sulcus). Total left and right HC volumes were derived from segmentation statistics (FreeSurfer’s ‘volume-based stream’) based on T1w MRI scans (Fischl et al., 2002).

Automatic tissue segmentations were visually inspected and volume measures were checked for outliers. Regional brain volumes were normalized by using FreeSurfer’s estimation of the total intracranial volume (ICV). We calculated the volume of interest to intracranial volume (ICV) fraction which has been shown to be a valid method to compensate for head size (O’Brien et al., 2011).

To investigate the role of structural alterations of the HC for conflict resolution, we assessed Pearson’s correlation coefficients and Kendall’s tau correlations between adjusted volumetric measures and behavioural parameters.

Results

Subjects

Patients with MTLE did not differ from healthy controls with regard to age \((t_{58}) = 0.1; \ p = .97\) and gender \((\chi^2_{1(1)} = 11; \ p = .30)\).

Table 1 summarizes the demographic, clinical and volumetric data of \(n = 23\) patients in whom seizure semiology, EEG and MRI indicated a clear unilateral pathology. \(N = 9\) patients (4 with left MTLE, 2 with right MTLE and 3 with bilateral MTLE) suffered from a comorbid depressive disorder. When controlling for seizure lateralization, depressive patients showed decreased left HC volume, \(F(1, 25) = 7.5, \ p = .012, \ r = .48\), but similar right HC volume, \(F(1, 25) = 0.1, \ p = .93\), compared to non-depressive patients.

Behavioural data

Effects of response conflict on accuracy

First, we analysed conflict resolution performance, that is, the effect of conflict on response accuracy (Figure 2). We expected greater response conflicts in incongruent
trials of the phonetic task in which subjects are required to respond to the pitch in the presence of conflicting semantic information (Haupt et al., 2009; Oehrn et al., 2015).

A two-way mixed ANOVA showed significant effects of ‘conflict’, incongruent vs. congruent; \( F(1, 28.0) = 11.3, p = .002, r = .54 \), and ‘group’, \( F(1, 26.4) = 13.3, p = .001, r = .57 \). Importantly, we also found a significant interaction, \( F(1, 28.0) = 5.0, p = .03, r = .39 \), indicating an impairment of conflict resolution in patients with MTLE (Figure 2b). This was confirmed by follow-up non-parametric Mann–Whitney \( U \)-tests, that is, patients’ accuracy difference (congruent minus incongruent) as well as the accuracy ratio (congruent divided by incongruent) were significantly higher than in controls (accuracy difference: \( U = 286.5, p = .014, r = .31 \); accuracy ratio: \( U = 274.5; p = .009; r = .34 \)).

Next, we assessed whether this effect was due to increased levels of facilitation (i.e., a higher difference between congruent and control trials) or enhanced interference (i.e., a more pronounced difference between incongruent and control trials) in the patient group. Regarding facilitation, we found significant main effects of ‘consistency’, congruent vs. control; \( F(1, 29.3) = 5.1, p = .03, r = .39 \), and ‘group’, \( F(1, 27.2) = 13.4, p = .001, r = .57 \), and a trend for an interaction, \( F(1, 27.2) = 4.0, p = .057, r = .36 \), indicating higher levels of facilitation in patients than in controls. This was confirmed by follow-up non-parametric Mann–Whitney \( U \)-tests on accuracy ratio (\( U = 250.5, p = .003; r = .38 \)) and accuracy difference (\( U = 255, p = .003; r = .38 \); Figure 2b). Notably, this difference was partly due to reduced accuracy of patients as compared to control subjects in the control trials (\( U = 708.5, p < .001, r = .50 \); Figure 2a), suggesting that patients experienced a substantial degree of interference even in control trials. This interpretation is supported by our analysis on interference, where we found a main effect of ‘group’, \( F(1, 25.4) = 11.2, p = .003, r = .55 \), but no effect of ‘consistency’, incongruent vs. control; \( F(1, 27.7) = 0.1, p = .7 \), and no interaction, \( F(1, 27.7) = 0.2, p = .7 \). Non-parametric Mann–Whitney \( U \)-tests confirmed that the difference between control and incongruent phonetic trials (\( U = 443.5, p = .9 \)) as well as the ratio between the two trial types (\( U = 249.1, p = .9 \)) did not differ between groups. Patients showed a lower accuracy than controls in incongruent trials (\( U = 701, p < .001; r = .48 \)) but also to a lesser extent in congruent trials (\( U = 610, p = .014; r = .32 \)).

Table 1. Demographic, clinical and volumetric data of patients with a clear unilateral pathology

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Left MTLE (n = 12)</th>
<th>Right MTLE (n = 11)</th>
<th>Statistics</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>45.7 ± 18.4</td>
<td>47.4 ± 14.0</td>
<td>( t_{(21)} = -0.25 )</td>
<td>.81</td>
</tr>
<tr>
<td>Gender (female/male)</td>
<td>5/7</td>
<td>7/4</td>
<td>( \chi^2_{(1)} = 1.1 )</td>
<td>.3</td>
</tr>
<tr>
<td>Clinical data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at seizure onset (years)</td>
<td>25.5 ± 18.7</td>
<td>23.7 ± 13.7</td>
<td>( t_{(21)} = 0.26 )</td>
<td>.80</td>
</tr>
<tr>
<td>Duration of disease (years)</td>
<td>20.2 ± 15.1</td>
<td>23.7 ± 13.7</td>
<td>( t_{(21)} = -0.48 )</td>
<td>.63</td>
</tr>
<tr>
<td>Depression (yes/no)</td>
<td>3/9</td>
<td>2/9</td>
<td>( \chi^2_{(1)} = 0.16 )</td>
<td>.69</td>
</tr>
<tr>
<td>Medication (Mono-/Polytherapy)</td>
<td>3/9</td>
<td>5/6</td>
<td>( \chi^2_{(1)} = 1.1 )</td>
<td>.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regional brain volumes(^a)</th>
<th>( n = 11 )</th>
<th>( n = 9 )</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Left HC volume</td>
<td>0.24 ± 0.05</td>
<td>0.28 ± 0.04</td>
<td>( t_{(18)} = -1.9 )</td>
<td>.07</td>
</tr>
<tr>
<td>Right HC volume</td>
<td>0.29 ± 0.03</td>
<td>0.24 ± 0.05</td>
<td>( t_{(18)} = 2.5 )</td>
<td>.02</td>
</tr>
</tbody>
</table>

Notes. HC = hippocampus; MTLE = mesial temporal lobe epilepsy.
Data are expressed as mean ± standard deviation.
\(^a\)Volume measures represent the percentage of regional volume of total intracranial brain volume.
Conflict effects on RT

For the RT data (Figure 2), a corresponding two-way mixed ANOVA yielded significant main effects of ‘conflict’, incongruent vs. congruent; $F(1, 58) = 56.6, p < .001; r = .70$, and ‘group’, $F(1, 58) = 6.0, p = .02, r = .31$, but no interaction, $F(1, 58) = 0.4, p = .55$, indicating that the effect of conflict on RT was similar between groups.

Next, we also analysed effects of facilitation and interference. Regarding facilitation, a two-way mixed ANOVA revealed a main effect of consistency, congruent vs. control; $F(1, 58) = 33.3, p < .001; r = .60$, and group, $F(1, 58) = 7.9, p = .007, r = .35$, but no interaction, $F(1, 58) = 2.0, p = .16$. Regarding interference, a two-way mixed ANOVA yielded main effects of ‘consistency’, incongruent vs. control; $F(1, 58) = 11.3, p = .001, r = .40$, and ‘group’, $F(1, 58) = 6.8, p = .012, r = .32$, but again no interaction, $F(1, 58) = 0.82, p = .37$. Together, these analyses did not reveal any conflict or interference-related RT differences between patients and control subjects.

We also explored the relationship between RTs and accuracy. In healthy subjects, RTs and accuracy in incongruent trials were positively related (Kendall’s tau = .52; $p < .001$). Longer reaction times in participants with better performance are indicative of a speed/accuracy trade-off. In patients, we found no significant correlation between RTs and accuracy (Kendall’s tau = -.17; $p = .2$).

Figure 2. Behavioural results. Response times (a) and accuracy (b) for each experimental condition. Difference scores of response times (c) and accuracy (d) reflecting effects of conflict (incongruent vs. congruent trials), interference (incongruent vs. control trials) and facilitation (congruent vs. control trials). Difference of accuracy and response time are calculated in opposite directions providing mainly positive differences. Errors bars represent standard error of mean.
Control analyses
First, we re-analysed the accuracy results by excluding all miss trials. In total, 2% of the trials were excluded in healthy controls and <7% in patients (averaged across conditions). Notably, patients missed most trials in the control condition (11%). Thus, excluding miss trials enhanced patients’ apparent performance in the control condition but their accuracy remained lower compared to healthy controls ($U = 601.5; p = .015; r = .31$), supporting the notion that patients perceived an increased interference already in control trials. Moreover, the analysis of interference now yielded a significant main effect of ‘consistency’, incongruent vs. control; $F(1, 26.8) = 9.7, p = .004, r = .52$, and also a significant interaction effect, $F(1, 26.8) = 4.3, p = .048, r = .37$, indicating an increased interference effect in patients when excluding miss trials.

Second, the data were analysed with respect to the sensitivity of pitch discrimination ($d$-prime) and the response criterion ($c$) using SDT. The analysis of $d$-prime scores replicated increased effects of conflict and interference in patients compared to controls, as reflected by significant interaction effects, conflict: $F(1, 58) = 12.8, p = .001; r = .43$; interference: $F(1, 58) = 5.0, p = .03; r = .28$. As a result, patients had significantly lower $d$-prime scores than controls in incongruent ($U = 737, p < .001; r = .56$) and control trials ($U = 602, p = .014; r = .31$) but not in congruent trials ($U = 505, p = .37$). Importantly, in each of the three conditions, the response criterion was not modulated by group, consistency or an interaction between both (lowest $p = .22$).

Third, we analysed the inverse efficiency score (IES), a measure combining both RT and accuracy. The results are consistent with the main accuracy measure. In contrast to the findings of control analyses, i.e. accuracy (excluding misses) and $d$-prime, we found a similar interference effect in patients and controls, interaction; $F(1, 58) = 1.7, p = .19$. Since RTs and accuracy were not related in patients (i.e., no speed/accuracy trade-off), the IES needs to be interpreted with caution (Bruyer & Brysbaert, 2011).

Effects of conflict adaptation
Next, we analysed effects of conflict adaptation, that is the influence of previous exposure to conflict on subsequent conflict resolution. Overall, conflict resolution was improved by previous conflict exposure, as shown by faster RTs, $F(1, 58) = 63.8, p < .001; r = .72$. Notably, however, we did not observe a significant interaction between conflict adaptation and group on mean RT, $F(1, 58) = 0.8, p = .38$. Regarding accuracy, a two-way mixed ANOVA revealed an effect of group, $F(1, 44.6) = 13.9, p < .001; r = .49$, but no effect of conflict adaptation, $F(1, 43.3) = 0.1, p = .77$, and no interaction, $F(1, 44.6) = 0.1, p = .84$. Also, non-parametric analyses showed that differences between pre-conflict and pre-non-conflict trials were similar between groups ($U = 422, p = .67$). This suggests that conflict adaptation is not impaired in patients with MTLE.

Effect of lesion lateralization
We investigated the effect of lesion lateralization on conflict resolution performance (RT and accuracy differences between incongruent and congruent trials). Patients with unilateral left and unilateral right MTLE showed similar conflict effects on RTs ($t_{(17)} = -0.35, p = .73$) and accuracy ($U = 85.5, p = .69$). Moreover, effects of facilitation, interference and conflict adaption did not differ between left and right MTLE (all $p > .05$).
Influence of clinical variables on measures of conflict resolution

Multiple regression analyses revealed that clinical variables (age at seizure onset, disease duration, mono- vs. polytherapy, depressive disorder, handedness) had no significant influence on conflict effects on RT, $R^2 = .02; F(5, 29) = 0.1, p = 1$, and accuracy, $R^2 = .17; F(5, 29) = 0.95, p = .47$.

Next, we compared accuracy in seven patients who were at the time of testing on GABAergic medication with 23 patients receiving non-GABAergic drugs. This analysis revealed a significant effect of ‘conflict’, incongruent vs. congruent; $F(1, 21.2) = 10.9, p = .003, r = .58$, no ‘group’ effect, GABAergic vs. non-GABAergic drugs; $F(1, 22.3) = 2.7, p = .12$, but a trend for an interaction, $F(1, 21.2) = 4.1, p = .056, r = .40$, on accuracy.

While this result might be interpreted as suggesting a relevant impact of GABAergic drugs on conflict resolution performance, we also found that the HC volume was significantly lower in patients on GABAergic drugs compared to patients receiving non-GABAergic drugs ($U = 19, p = .006; r = .34$). Moreover, we found a trend for a longer epilepsy duration in patients with as compared to without GABAergic drugs ($U = 41.5, p = .054; r = .24$). Thus, these confounding effects prevent a clear interpretation.

Correlations between measures of conflict resolution and regional brain volumes

Finally, we tested whether conflict effects on accuracy were significantly related to HC volume. Indeed, we found more pronounced effects of conflict on accuracy in patients with a smaller right HC (Kendall’s tau = −.32; $p < .03$, Figure 3). This correlation remained significant when using accuracy ratio instead of difference (Kendall’s tau = −.29; $p = .047$). We did not find this correlation with the left HC (Kendall’s tau = −.04; $p = .8$). Moreover, correlation values in patients with unilateral left MTLE (Kendall’s tau = −.28) were numerically but not significantly ($z = −0.1; p = .46$) higher than in right MTLE patients (Kendall’s tau = −.23). Conflict effects on accuracy did not correlate with GM volumes of ACC and dlPFC. Similarly, there were no correlations with neocortical areas of the right temporal lobe when corrected for multiple comparisons (10 areas resulting from FreeSurfer segmentation; all Kendall’s tau < .33, all $p_{corr} > .3$).

Related to RT, we did not find any significant correlations between effects of conflict and adjusted GM volumes of HC (left: $r = .1, p = .5$; right: $r = −.1, p = .7$) or ACC (left: $r = −.2, p = .3$; right: $r = −.1, p = .9$). However, greater effects of conflict on RT were accompanied by a lower adjusted left dlPFC volume ($r = −.43; p = .03$).

We found a trend towards a correlation between GM volumes of right HC and right dlPFC ($r = .37; p = .067$). This correlation was similar in patients with an epilepsy duration below ($r = .38$) and above ($r = .35$) the patients’ median epilepsy duration (Fisher Z = 0.125; $p = .45$).

Discussion

We investigated conflict resolution performance in an auditory Stroop task in patients with MTLE due to HS. Our results revealed increased effects of interference and conflict on accuracy, but not on RTs, in patients compared to healthy controls. Control analyses of accuracy (accuracy excluding misses, SDT $d$-prime scores) replicated the findings. Reduced accuracy was not only observed in incongruent trials, but also to a lesser extent in control trials, suggesting a high sensitivity of patients to even low degrees of interference. Conflict adaptation was not affected. Conflict effects on accuracy were
related to a reduced right hippocampal GM volume irrespective of seizure lateralization, but were not related to ACC and dIPFC GM volumes, suggesting a direct role of the hippocampus for response conflict resolution.

In Stroop paradigms, an increased RT and/or a reduced accuracy reflect higher executive control demands due to the recruitment of conflict resolution processes that are engaged in incongruent trials. In these trials, subjects need to inhibit a more automatic task behaviour in order to initiate a less automatic one. Patients showed generally increased RTs across all trial types, suggesting increased attempts to recruit executive control processes. Still, their response accuracy was significantly reduced during conflicts, suggesting that RT slowing was not sufficient. This is supported by our finding that RTs and accuracy were only related in healthy controls but not in patients, indicating lack of a beneficial speed/accuracy trade-off.

Interestingly, accuracy was not only reduced in incongruent but also to a lesser extent in control trials. One explanation refers to the fact that the semantic task was presented first, which requires subjects to ignore the word ‘good’, while in the subsequent phonetic task they need to respond to ‘good’ (control trials). It seems plausible that this leads to a proactive interference effect of the learned inhibition to the control word ‘good’, possibly indicating impaired extinction of the previously learned inhibitory response. This interpretation would be in line with the patients’ high rate of omissions (11%) in control trials. However, analysis of accuracy values when excluding misses and \( d \)-prime scores still confirmed the patients’ lower performance in control trials so that it cannot exclusively be explained by an increased miss rate. Presumably, this trial type required greater attentional demands in patients. In sum, both conflict and interference effects were consistently found across different accuracy measures, supporting the conclusion that mesial TLE due to HS is associated with a neuropsychological deficit in the resolution of (but not in the adaptation to) response conflicts.
Is impaired conflict resolution a direct consequence of hippocampal lesions? The previous neuropsychological studies remained inconclusive with regard to this point. Corcoran and Upton (1993) found preserved visual Stroop task performance in patients with mesial TLE compared to patients with frontal lobe epilepsy. However, their study did not include healthy control participants. Wang et al. (2007) found an impairment in a range of executive tasks including the Stroop task in patients with TLE but did not describe the underlying pathology and whether or not they restricted their analysis to MTLE patients. A recent study found behavioural effects of conflict on accuracy in a flanker task in individuals with amnestic mild cognitive impairment (Borsa et al., 2018), a condition with structural damage of the HC and the neocortical temporal lobe (Chetelat et al., 2005). Our results extend these findings by showing that reduced conflict resolution in the Stroop task relates to reduced GM volume of the HC but not the neocortical temporal lobe in patients with mesial TLE. These neuropsychological data fit to previous findings of a hippocampal involvement in the successful resolution of a Stroop response conflict (Oehrn et al., 2015). In sum, these results suggest that the HC is not only involved but may even be causally relevant for the resolution of response conflicts.

The neuropsychological data suggest that in humans, a HC lesion results not only in disinhibited behaviour during the processing of mixed-valence stimuli (Bach et al., 2014) but also in impaired behavioural control over established response patterns as in the Stroop task used in our current study. We assume that an intact ‘pattern completion’ which is provided by hippocampal CA3 subfield (Rolls, 2013) might be required to activate representations of incongruent motor responses leading to a response conflict. On the other hand, the dentate gyrus might increase dissimilarity between the two competing response representations, that is, a process known as ‘pattern separation’ (Leutgeb, Leutgeb, Moser, & Moser, 2007). Which of the two responses is executed may finally depend on PFC mediated motor selection based on the current task. Thus, impaired pattern separation would cause overlap between the automated and the task-relevant response representations ultimately reducing the accuracy of response selection. The interaction between hippocampal subfields, in particular dentate gyrus, CA3 and CA1 might therefore be of interest to investigate in future lesion studies (Loh et al., 2017).

Possible alternative explanations for our findings need to be mentioned. First, conflict resolution deficits might be the result of secondary structural or functional changes of other regions such as the prefrontal cortex (PFC). This would be congruent to the finding of a correlation between behavioural measures of conflict and ACC GM volume in aMCI patients (Borsa et al., 2018). It would also be in line with previous studies showing GM volume reductions in prefrontal brain areas of patients with MTLE (Keller, Baker, Downes, & Roberts, 2009). In our study, we found a trending positive correlation between dlPFC and HC GM volumes which was not modulated by epilepsy duration. Moreover, increased conflict effects on RTs in the Stroop task were related to a reduced GM volume of the dlPFC, suggesting that prefrontal cognitive control structures also reflect performance in the resolution of a response conflict in patients with mesial TLE due to HS. On the other hand, right HC but not dlPFC GM volume correlated with conflict effects on accuracy, the measure of conflict resolution that was impaired in our patient sample. This suggests that conflict resolution performance indeed depends on the integrity of brain structures beyond ACC and dlPFC, that is, the HC. Second, epileptiform EEG activity may have negatively affected performance in our patient sample with chronic therapy-refractory epilepsy. A recent study showed an impairment in the visual Stroop paradigm only in patients with unilateral HS who had concordant epileptiform EEG activity while patients with contralateral EEG epileptiform activity showed preserved Stroop test performance.
(Pinto et al., 2017). However, a previous review across various studies concluded that direct effects of epileptiform EEG discharges on cognition are rather mild (Aldenkamp & Arends, 2004). Third, antiepileptic drug treatment as well as a comorbid depressive disorder may affect cognition (Epp, Dobson, Dozois, & Frewen, 2012; Witt, Elger, & Helmstaedter, 2015). We found that the number of AEDs (mono- vs. polytherapy) and comorbid depression did not show a significant influence on the conflict effects. Patients on GABAergic drugs showed worse performance than patients receiving drugs without GABAergic action. However, these patients also had a lower HC volume and a trend for a longer epilepsy duration, precluding any clear interpretation of these results. We conclude that conflict resolution impairment cannot be fully explained by secondary PFC changes, epileptiform EEG activity, medication or a depressive disorder but that the HC itself causes a relevant impact on the conflict resolution performance.

Is the Stroop conflict resolution impairment lateralized? Although a recent fMRI-iEEG study revealed predominant BOLD responses in the left HC during an auditory Stroop conflict in healthy subjects (Oehrn et al., 2015), no effect of seizure lateralization on conflict resolution performance was observed in our study. This is in line with previous neuropsychological studies using a visual Stroop task (Corcoran & Upton, 1993; Wang et al., 2007). We found that conflict resolution performance was related to right but not left HC GM volume irrespective of seizure lateralization and assumed hemispheric dominance, a finding for which we do not have a conclusive explanation. As patients in our study had a long mean duration of disease, functional reorganization of the corresponding neural networks has likely occurred as described earlier for the long-term memory network in post-surgical patients with TLE (Helmstaedter, Kurthen, Lux, Reuber, & Elger, 2003; Salvato et al., 2016). Thus, compensatory processes could have masked a lateralization effect.

Our study has several limitations. First, we found a high inter-individual variation in the patients’ group especially for the accuracy measures. The accuracy variation can only in small parts ($R^2 = 9\%$) be explained by the structural integrity of the HC. We therefore assume that additional factors as described above contribute to performance. Second, using a standardized MRI visual analysis protocol, we detected structural brain lesions outside the hippocampal formation in some cases. Although patients with frontal brain lesions were excluded, we cannot rule out that these additional brain abnormalities might have affected the results. Another limitation is that FreeSurfer’s automatic segmentation may be of reduced reliability in case of a hippocampal lesion. Even though inspection of segmentation results did not reveal any outliers, future studies employing manual segmentation would be desirable. Moreover, in two patients MRI signs of HS were not replicated in the final radiological evaluation, thus, intact HC structure in these patients possibly might have attenuated the performance decline. Although visual inspection of MRI is an important part of the multimodal clinical assessment of suspected HS, its independent sensitivity is limited (Coan, Kubota, Bergo, Campos, & Cendes, 2014) even though it can significantly be improved by using a specific MRI protocol designed for the early detection of epileptogenic lesions (Wellmer et al., 2013). Apart from MRI, seizure semiology and results of neuropsychological testing were indicative of typical MTLE and thus suggestive of HS in these two patients.

Conclusions
We observed a neuropsychological deficit in Stroop conflict resolution which was related to structural integrity of the HC in mesial TLE patients. These findings add support to the
idea that the HC is crucially involved in higher-order cognitive functions like cognitive conflict processing. Our findings may help to explain everyday symptoms like affective impulsivity and distractibility in patients with a hippocampal pathology. Functional neuroimaging studies in patients are needed to explore the interplay between the cognitive control network and the HC.

Acknowledgements

We would like to thank Werner Weber from the Department of Radiology, University Hospital Knappschaftskrankenhaus, Ruhr University Bochum, for providing neuroimaging data.

References


**Received 18 December 2018; revised version received 10 May 2019**

**Supporting Information**

The following supporting information may be found in the online edition of the article:

**Table S1.** Patient characteristics.