Electrophysiological correlates of semantic memory retrieval in Gulf War Syndrome 2 patients

Gail D. Tillman a, Clifford S. Calley a, Virginia I. Buhl a, Hsueh-Sheng Chiang a, Robert W. Haley b, John Hart Jr a,⁎, Michael A. Kraut c

a School of Behavioral and Brain Sciences, Center for BrainHealth, The University of Texas at Dallas, Dallas, TX, United States
b Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, United States
c Department of Radiology, Division of Neuroradiology, The Johns Hopkins University School of Medicine, Baltimore, MD, United States

ABSTRACT

Gulf War veterans meeting criteria for Haley Syndrome 2 of Gulf War illness endorse a particular constellation of symptoms that include difficulty with processing information, word-finding, and confusion. To explore the neural basis of their word-finding difficulty, we assessed event-related potentials (ERPs) associated with semantic memory retrieval in 22 veterans classified as Syndrome 2 and 28 veterans who served as controls. We recorded ERGs while subjects judged whether pairs of words that represented object features combined to elicit a retrieval of an object memory or no retrieval. Syndrome 2 subjects’ responses were significantly slower, and those participants were less accurate than controls on the retrieval trials, but they performed similarly on the nonretrieval trials. Analysis of the ERPs revealed a difference between retrievals and nonretrievals that has previously been detected around 750 ms at the left temporal region was present in both the Syndrome 2 patients and controls. However, the Syndrome 2 patients also showed an ERP difference between retrievals and nonretrievals at the midline parietal region that had a scalp voltage polarity opposite from that recorded at the left temporal area. We hypothesize that the similarities between task performance and ERP patterns in Syndrome 2 veterans and in patients with amnestic mild cognitive impairment reflect disordered thalamic cholinergic neural activity, possibly in the dorsomedial nucleus.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

It has been suggested that approximately 25–30% of those deployed in the 1991 Persian Gulf War have developed persistent cognitive deficits [1]. A common symptom reported in these individuals is difficulty with finding words [2–5]. The prominence of this dysfunction is such that it has been captured in symptom-derived definitions characterizing patients with Gulf War-related illnesses. Haley and colleagues developed a classification for those suffering symptoms following being deployed in the Persian Gulf [6–8]. Haley Syndrome 2 patients exhibit confusion that is characterized by difficulty with processing information, word finding, emotional lability, confusion, and balance problems [6].

We previously used functional MRI (fMRI) to study a group of US Naval Construction Forces personnel ("Seabees") as they performed a semantic memory retrieval task in order to localize the brain regions associated with performance of that task [4]. In that study, subjects were presented with two words that represent features of objects and were asked to indicate whether the words together resulted in retrieval of a specific object from memory. Neural correlates of normal subjects performing this task have been studied using behavioral [9], fMRI [10–12], event related potential (ERP) [13], and electroencephalographic time-frequency analysis [14,15] techniques. The task has also been used to probe dysfunction in patients with mild cognitive impairment and/or Alzheimer’s Disease [9,16,17], schizophrenia [18], stroke [19, 20], and concussion and aging [21]. In a study of normal controls performing the task during fMRI, significant BOLD signal changes were detected for the correct retrievals in bilateral medial Brodmann Area 6 (pre-SMA region), dorsomedial and pulvinar thalamic nuclei, caudate nuclei, and bilateral temporo-occipital regions [10,11,21]. There is also an ERP difference between retrievals and nonretrievals at approximately 750 ms with a maximum at the left fronto-temporal region that has been proposed to signify co-activation of common feature representations of the object being retrieved [13].

In our previous Seabee study [4], subjects with Haley Syndrome 2 made significantly more errors than did study subjects in the other groups (i.e., controls, Syndromes 1 and 3), consistent with their subjective complaints of word finding and memory difficulties. In addition, the Syndrome 2 patients had patterns of signal changes in the caudate and...
thalamus that were noticeably different from the other Haley Syndromes and normal controls during correctly performed trials. In these regions, we found increased BOLD signal changes with longer reaction times on the task, in contrast with the subjects in the other groups including the controls, who showed the opposite pattern. This atypical BOLD–reaction-time correlation in correctly performed trials was proposed to represent an increased effort in an attempt to maintain performance in the setting of dysfunctional underlying neural resources. We also administered word generation tasks to the same groups while we recorded fMRI [5]. The task required the subject to recall the names of as many members of a category of objects or of words that begin with a specific letter as he or she could. Syndrome 2 patients performed significantly worse behaviorally on letter and category fluency compared to Syndrome 1 subjects and controls. The Syndrome 2 subjects also showed reduced BOLD signal in the thalamus and putamen compared to controls, consistent with the proposal that the thalamus is involved in word generation when semantic input is used for word finding [22]. ERP studies were not obtained in conjunction with either of these two Seabees’ studies.

We undertook the current study to determine whether the findings in the initially characterized Seabee sample are also detectable in Haley Syndrome 2 patients more generally, and whether the ERP correlates of semantic memory retrieval that we found in normal controls are also present in patients with Haley Syndrome 2.

2. Method

2.1. Participants

All the participants had been in the military during the 1991 Persian Gulf War. The exacting measures taken to identify, contact, and recruit a representative sample of veterans have been described fully in previous reports [8,23] and supplementary materials [23]. For this report, data from 3 of the 31 veterans in the control groups and from 2 of the 24 veterans in the Syndrome 2 group were excluded from the analysis due to there being too few artifact-free epochs to create reliable ERP averages. Thus, we analyzed data from 50 participants (11 female). Twenty-two (6 female) of these met the Haley et al. [6,7] criteria for Syndrome 2 of GW Illness. Syndrome 2 is associated with more debilitating neurocognitive issues—confusion, word-finding and reasoning difficulties, emotional lability—and balance problems such as frequent stumbling and vertigo. The remaining 28 (5 female) veterans who did not meet the criteria for any of the six GW Illness Syndromes [6,8,23] served as controls. Chi square analysis indicated an expected distribution of male and female across the two groups studied here, $\chi^2 = 0.636, p = 0.425$. Additional medical information within each group is listed in Table 1. The subjects were housed and monitored at The University of Texas Southwestern Medical Center’s Clinical and Translation Research Center in 2009 and 2010, and underwent a week-long multi-modal neuropsychological, neuroimaging, and biomarker study. All subjects gave written informed consent according to a protocol approved by the university’s institutional review board.

2.2. Task and stimuli

Participants performed a task based on the Semantic Object Retrieval Test (SORT) [9,16]. We presented one hundred pairs of printed words that represent features of common objects, with one word above the other in black letters on a white screen. Fifty of the trials were made up of word pairs that have been shown to elicit retrieval of a specific object (e.g., “desert” paired with “hump” elicits the object “camel”) [11]; the remaining 50 word pairs were nonretrieval trials (e.g., “sleeve” paired with “jungle”). Each word pair was presented on a computer monitor positioned approximately 1 m in front of the participant for 3000 ms, and was followed by a 3000-ms fixation point. Participants were instructed to press the response pad button under their index finger when the word pair called to mind a specific object, rather than merely an association between the words. When the word pair did not call to mind a specific object, they were to press the response pad button under their middle finger. Six versions of the word pair presentation order were randomized across subjects.

2.3. Procedure

After the participants were fitted with the electrode cap, they were shown the instructions as they were read aloud to them. Participants were allowed to ask questions to assure that they understood the task. At the beginning of each task, the first image repeated the instructions.

2.4. EEG acquisition

We recorded EEG using a 128-electrode array mounted within an elastic cap. We positioned electrodes at the superior and inferior orbital margins to monitor blinks and vertical eye movements. The reference electrode was located near the vertex, and the APZ electrode served as the ground electrode. Before we started recording EEG data, we assured that the impedance for each electrode was below 10 kΩ.

We used Stim2 (Compumedics Neuroscan, Charlotte, NC, USA) software to record the accuracy and reaction time of the responses and to mark each stimulus onset and response in the electronic EEG record. The EEG was recorded using a Neuroscan Synamps2 (Compumedics Neuroscan) amplifier at a 500-Hz sampling rate. The continuous EEG data were high-pass filtered at 0.15 Hz and re-referenced to the global mean amplitude. Blink artifacts were filtered from the continuous EEG file by using a spatial filter process included in the Scan 4.5 Edit (Compumedics Neuroscan, Charlotte, NC, USA) software. Data from 200 ms before the onset to 1800 ms after the onset of each stimulus were included in each epoch. From each subject’s task data, retrieval and nonretrieval conditions were averaged. Each average consisted of epochs that had been baseline-corrected based on the 200-ms prestimulus data.

2.5. Data analysis

Only the ERP averages that comprised 20 or more artifact-free sweeps were used in the analysis. In order to reduce the dimensionality of the ERP data, 25 regions based on equivalent scalp areas were designated. Average amplitude for each 100-ms time window from stimulus onset to 1200 ms post-stimulus for each electrode within a region was calculated. This yielded 25 (space) × 16 (time) data points for each participant. A principal components analysis (PCA) was performed on this matrix, followed by Varimax rotation. Four orthogonal spatial factors

---

Table 1: Demographic and comorbidity data.

<table>
<thead>
<tr>
<th></th>
<th>Control M (SD)</th>
<th>Syndrome 2 M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age M (SD)</td>
<td>49.39 (7.65)</td>
<td>49.41 (7.43)</td>
</tr>
<tr>
<td>Age range</td>
<td>38–65</td>
<td>37–65</td>
</tr>
<tr>
<td>Number of females (%)</td>
<td>5 (18%)</td>
<td>6 (27%)</td>
</tr>
<tr>
<td>PTSD</td>
<td>0</td>
<td>9 (41%)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>1 (4%)</td>
<td>16 (73%)</td>
</tr>
<tr>
<td>Depression NOS active</td>
<td>0</td>
<td>16 (73%)</td>
</tr>
<tr>
<td>Major depressive disorder</td>
<td>0</td>
<td>1 (56%)</td>
</tr>
<tr>
<td>Alcohol abuse or dependence</td>
<td>3 (11%)</td>
<td>10 (43%)</td>
</tr>
<tr>
<td>Drug abuse</td>
<td>4 (14%)</td>
<td>4 (18%)</td>
</tr>
<tr>
<td>Smoking</td>
<td>0</td>
<td>5 (23%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>5 (18%)</td>
<td>8 (36%)</td>
</tr>
<tr>
<td>Cholesterol-reducing medication</td>
<td>3 (11%)</td>
<td>7 (32%)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1 (4%)</td>
<td>1 (5%)</td>
</tr>
</tbody>
</table>

* Indicates a significant difference between the groups.
Fig. 1. Behavioral data from the SORT task. Left panel: An interaction between condition and group \((p = 0.023)\) on percent correct. The main effect of group was due to Syndrome 2’s poorer performance in the retrieval condition. Right panel: Main effects of group and condition on response times.

Fig. 2. Interaction between condition (retrieval, nonretrieval) and group (Controls, Syndrome 2): While controls’ scores for retrieval and nonretrieval are similar (bottom left panel), Syndrome 2 factor scores (bottom right panel) for the nonretrieval condition are strongly positive, especially after 600 ms, and scores for the retrieval condition are negative.
explaining 76% of the variance were extracted using scree criteria. Factor scores were used as the dependent variables in subsequent mixed-measure analyses of variance (ANOVA) where condition (retrievals, nonretrievals) and time point were the repeated measures, and group (Control, Syndrome 2) was the between-subjects factor.

3. Results

3.1. Behavioral results

We found a main effect of group on percent correct, $F(1, 48) = 4.216, MSe = 0.026, p = 0.046, \eta^2_p = 0.081$. This effect was driven by the interaction between condition and group ($F(1, 48) = 5.539, MSe = 0.020, p = 0.023, \eta^2_p = 0.103$), depicted in the left panel of Fig. 1. While accuracy in the nonretrieval condition was quite similar for the control group and the Syndrome 2 group, the control group’s accuracy in the retrieval condition was considerably better. There was no main effect of condition, $F(1, 48) = 0.405, MSe = 0.020, p = 0.528$. The $d’$ metric of sensitivity was not different between the two groups ($p = 0.353$), but response bias (using $c’$) was ($p = 0.002$). Syndrome 2 showed a greater bias toward the nonretrieval responses.

A similar analysis of variance using reaction time as the dependent variable indicated a main effect of condition, $F(1, 48) = 107.165, MSe = 27,443.081, p < 0.0005, \eta^2_p = 0.691$. As shown in the right panel of Fig. 1, responses in both groups to the retrieval condition were faster than those to the nonretrieval condition. There was also a main effect of group, $F(1, 48) = 7.571, MSe = 180,681.879, p = 0.008, \eta^2_p = 0.136$. Response times from the control group were faster than those from the Syndrome 2 group. There was no interaction between condition and group, $F(1, 48) = 0.592, MSe = 27,443.081, p = 0.445$.

3.2. Electrophysiology results

Only the fourth spatial PCA component represented variance related to our factors—group (Controls and Syndrome 2), condition (retrieval, nonretrieval), and time point (100-ms windows from stimulus onset to 1600 ms). The fourth component accounted for 5.2% of the variance and had strong positive loadings in the temporal areas, especially on...

Fig. 3. Mean amplitudes in 100-ms latency windows from stimulus onset to 1600 ms following stimulus onset. The left temporal area showed the most positive factor loading with the fourth spatial component. The midline parieto-occipital area showed the most negative factor loading. At left temporal electrodes, both controls and Syndrome 2 mean amplitudes show more negative amplitudes in the retrieval condition. At midline parieto-occipital, controls retrieval and nonretrieval conditions are very similar, but Syndrome 2 nonretrieval amplitudes are more negative.

Fig. 4. Mean amplitude differences computed from the electrode areas with the most positive and the most negative factor loadings (Left temporal mean amplitude minus that of midline parieto-occipital). This pattern is reflected in the pattern of the factor scores (Fig. 2).
the left. Weaker negative loadings were indicated in the midline parieto-occipital area. The variance explained by this factor showed an interaction between condition and group, $F(1, 48) = 4.244, MS_e = 3.11, p = 0.045, \eta^2_p = 0.081$. No main effect of group, condition, or time point was indicated ($p > 0.12$), and there were no other interactions ($p > 0.21$).

As shown in Fig. 2, while the mean factor scores for the retrieval and nonretrieval conditions were similar for the control group, the mean factor scores from the Syndrome 2 data show considerable divergence, especially after 500 ms. Mean amplitudes from the left temporal area and midline parieto-occipital area, which showed the most positive and most negative factor loadings, respectively, are shown in Fig. 3. The ERPs from both the controls and the Syndrome 2 participants show more negative potentials for the retrieval condition in the left temporal area, but the midline parietal area is notably different. Controls’ amplitudes include a steep positivity in the 300–500-ms latency region and a subsequent drop toward baseline in both the retrieval and nonretrieval conditions. Syndrome 2

Fig. 5. Event-related potentials (ERPs) from areas whose mean amplitudes are depicted in Fig. 4. TL = Left Temporal Area; POZ = Midline Parieto-Occipital.

![Graphs showing ERPs from left temporal and midline parieto-occipital areas](image)

Fig. 6. The pattern seen in the difference waves computed from left temporal area (TL) and midline parieto-occipital (POZ) is similar to the patterns seen in the mean factor scores (Fig. 2) and the mean amplitude differences of 100-ms latency windows (Fig. 4).
waveforms lack the steep positive deflection and show an effect of condition in the last half of the epoch. The waveform that represents the difference between the amplitudes in these two areas is similar to the mean factor scores for the groups and conditions (Fig. 4). This pattern is also observable when examining the original ERP averages from the temporal and parieto-occipital electrodes whose variance is most strongly associated with this principal component (Figs. 5 & 6).

4. Discussion

We found that the Syndrome 2 patients made significantly more errors (in the retrieval trials) than did the controls, and Syndrome 2 patients were also significantly slower in reaction times than were controls. The ERP difference between retrievals and nonretrievals that has previously been detected at 750 ms, maximal at the left frontal temporal region, was present in both the Syndrome 2 patients and controls (retrievals negative; nonretrievals positive). However, the Syndrome 2 patients also have an ERP difference between retrievals and nonretrievals at the midline parietal region, and with a reversal of polarity from that of the left fronto-temporal ERP (retrievals positive; nonretrievals negative).

The semantic process at 750 ms that has been associated with this task has been attributed to the coactivation of feature representations common to the same object [13]. The timing of the separation between retrievals and nonretrievals was prior to memory retrieval but consistent temporally and spatially with correlating similar features. Disruptions of or alterations in this ERP has been correlated with a variety of behavioral outcomes, depending on the etiology of the neural dysfunction causing the ERP changes, the degree of disruption, and the brain regions that are dysfunctional.

Previous fMRI studies of GW Syndrome 2 patients on the task used in this study showed thalamic and caudate dysfunction [4]. The correlation of reaction time for correct responses and signal change in BOLD in the caudate and thalamus for Syndrome 2 patients showed pronounced differences relative to controls: Syndrome 2’s reaction time increase was associated with greater signal change, whereas the controls’ reaction time increase was associated with decreased signal change [4]. Interactions between the pre-Supplementary Motor Area (preSMA) and the thalamus have been linked to the memory retrieval process in previous studies [10,24,25]. In addition, these brain structures participate in pre-retrieval semantic processes of semantic search and association, along with other regions (e.g., left inferior parietal-superior temporal gyrus, left inferior frontal gyrus, etc.) [26,27].

Another group of patients who have exhibited a similar ERP and behavioral performance pattern to GW Syndrome 2 patients is aging individuals and those who progress to degenerative neurological states. A study of younger and older normal adults [28] using the SORT [9,16] task with ERP showed that both groups demonstrated the left fronto-temporal ERP difference around 750 ms, but that the older adults had a later frontal positive ERP around 800–1000 ms; that is, the ERP associated with nonretrievals were more positive than the ERP associated with retrievals. Since behavioral performance of the older adults was comparable to that of the younger adults, it was posited that this frontal component represented activation of different and perhaps more extensive brain regions in response to nonretrieval trials. This also was supported by the observation that maintaining an active search for an extended time in order to link two features to an object in memory, and then terminating that search when not successful is cognitively more taxing than that required for retrieval trials. Given that there was no difference in the accuracy of response between younger and older adults, the authors hypothesized that these additional neural resources in older adults serve as a compensatory mechanism for maintaining and/or terminating the search when features result in no retrieval.

Chiang et al. [17] also studied amnestic Mild Cognitive Impairment (aMCI) patients and found lower accuracy in behavioral performance on the SORT compared to controls. ERPs in both the aMCI and control cohorts showed a left fronto-temporal component starting at around 750 ms post-stimulus. The aMCI subjects showed an increase in the frontal–parietal scalp potential that distinguished retrieval from nonretrieval trials between 950 and 1050 ms post-stimulus. There was also a reversal of polarity in the retrieval–nonretrieval relationship similar to that of GW Syndrome 2 s. At the neural level, the patterns of synaptic activity that contribute to scalp recorded EEG/ERP likely differ between the control and the disease populations. Those different patterns could be happening within the same regions of brain, resulting in distinct neural generators with different polarities. Alternatively the activity could be in different brain regions altogether, with consequent opposite-polarity summations at the scalp.

We propose that the altered neural activity in the aMCI and GW Syndrome 2 s compared to controls reflect a more sustained and effortful semantic search during object memory retrieval [17]. The common neuropathological findings in the GW Syndrome 2 patients and aMCI is the loss of acetylcholine neurons or cholinergic dysfunction [23,29–31]. In adult rats, low exposure to pyridostigmine bromide, DEET, and permethrin, combined with stress—similar to the exposures of Desert Storm veterans—was associated with blood–brain barrier disruption, neuronal death, decreased acetylcholine esterase activity, and increased acetylcholine receptor binding [32–34]. Haley et al. [23] found higher degrees of autonomic dysfunction in Syndrome 2 veterans, and that these deficits were more related to cholinergic autonomic systems than to adrenergic autonomic systems. Amnestic MCI patients have been shown to have significant gray matter reduction in the thalamus [35–39] and the caudate [37,38] compared to age-matched normal controls. In particular, there is a loss of cholinergic innervation of the dorsomedial nucleus of the thalamus as patients progress to AD [40]. Older adults using anticholinergic drugs showed lower verbal fluency and naming performance—but not poorer Mini-Mental State Examination performance—than matched controls who were not using in anticholinergic drugs [41]. Poorer performance in a word-generating task by Syndrome 2 Gulf War veterans as compared to controls was accompanied by a BOLD response in basal ganglia and thalamus that was lower in especially the Syndrome 2 veterans as compared to veterans in the control group [5]. Thus, cholinergic dysfunction in the thalamus present in both GW Syndrome 2 s and aMCI could account for the similarities in behavioral performance and ERP patterns that we have found.

Acknowledgement

This study was supported by IDIQ contract VA549-P-0027, awarded and administered by the Department of Veterans Affairs Medical Center, Dallas, TX; U.S. Army Medical Research and Materiel Command grant number DAMD17-01-1-0741; and Grant Number UL1RR024982, titled North and Central Texas Clinical and Translational Science Initiative (Milton Packer, M.D., Ph.D.), from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research. The content does not necessarily reflect the position or the policy of the Federal government or the sponsoring agencies, and no official endorsement should be inferred.

References
