The neuroanatomic correlates of semantic memory deficits in patients with Gulf War illnesses: a pilot study


Abstract We used functional magnetic resonance imaging (fMRI) to study semantic memory processing in 38 Gulf War veterans in 3 affected groups (Syndromes 1, 2, and 3) and normal-deployed controls. Subjects were given the Semantic Object Retrieval Test (SORT), which requires participants to decide whether two features combine and result in the retrieval of a specific object (e.g., “desert” and “humps” → “camel”). Differences between groups were calculated using a repeated measures analysis of variance (ANOVA). Then, regions of interest were constructed and correlations assessed between the percent signal change (PSC) within these regions, followed by correlations between behavioral measures and PSC. We found affected groups performed less well on the SORT than the controls did, and behavioral differences were correlated to PSC within the caudate and thalamus. The combination of performance deficits and functional neuro-imaging differences between affected Gulf War veterans and deployed normal controls begins to establish a neurobiological basis for their word-finding deficits.

Keywords Gulf War illness · Semantic memory · Word finding · fMRI

Introduction

Individuals deployed to the Persian Gulf during 1991–92 returned reporting a variety of symptoms, including some related to cognition (see Barrett et al. 2002). Two of the most salient symptoms are word-finding and memory difficulties (see Calley et al. 2010). Several studies of the symptoms reported by veterans with Gulf War Illnesses corroborate reports of difficulties with word finding (Fukuda et al. 1998; Barrett et al. 2002) and/or memory (Pierce 1997; Fukuda et al. 1998; Gray et al. 2002; Ishøy et al. 1999); however, the incidence and prevalence of deficits related to semantic memory with these symptoms, the etiology of the impairments, and the brain regions involved are unknown.

Previous investigations of retrieving an object memory after viewing either written or pictorial representations of its component features (e.g., “desert” and “humps” “camel”) have consistently demonstrated significant fMRI signal changes in bilateral medial Brodmann Area 6 (pre-SMA region), dorsomedial and pulvinar thalamic nuclei, caudate, and in temporo-occipital regions (Kraut et al. 2002; Slotnick et al. 2002; Assaf et al. 2006; Hart et al. 2007). This probe of object memory, known as the Semantic Object Retrieval
Test (SORT), has also been applied to clinical populations with semantic memory deficits and has demonstrated correlations between poor behavioral performance and degree of fMRI signal change in specific anatomical regions (Assaf et al. 2005). Patients with Formal Thought Disorder in that study showed positive correlations between signal change in the rostral anterior cingulate cortex and disorder severity during both correctly and incorrectly recalled trials. Also, that study showed greater-than-usual signal change in bilateral pre-SMA (BA6), temporo-occipital junctions, temporal poles, para-hippocampal gyri, right inferior frontal gyrus, and dorsolateral prefrontal cortex, and reduced signal change in the inferior parietal lobules in the patients compared to controls.

Investigations of Gulf War Illnesses (GWI) have used symptom report clustering to identify common groupings of patients. Haley et al. (1997) created a classification schema of GWI consisting of three syndromes based on symptom clusters. Syndrome 1 consists primarily of mild cognitive symptoms including problems with attention, memory, and reasoning as well as insomnia, depression, daytime sleepiness and headaches. Syndrome 2 patients are more debilitated than those with the other two syndromes, and report significant difficulties with thinking, disorientation, balance disturbances, vertigo, and impotence. Syndrome 3 patients report the least in terms of cognitive symptoms, with the major complaints being joint and muscle pains, muscle fatigue, difficulty lifting, and extremity paresthesias. None of these syndromes capture clearly the subjective complaints of word finding and memory difficulties indicated in general symptom reports.

We sought to characterize the disruptions in semantic memory retrieval, and the neural correlates of those disruptions, by using the SORT while acquiring fMRI data in veterans classified with Syndromes 1, 2, or 3, as well as in matched controls who were deployed to the region during the conflict but did not return with symptomatic complaints.

Methods

Subjects

53 right-handed male veterans of the 1991 Persian Gulf War (11 subjects with Syndrome 1; 16 with Syndrome 2; 11 with Syndrome 3; 16 normal deployed) were selected from a Construction Battalion in the United States Naval Reserve. All subjects had participated in prior studies of Gulf War Syndrome (Haley et al. 1997, 2000) and those placed into one of the symptomatic groups met the Haley criteria for that syndrome. None of the groups, including controls, differed significantly in age, sex, education, or ethnicity. Subjects were housed and monitored at The University of Texas Southwestern Medical Center’s Clinical and Translational Research Center (CTRC) in 2008 and 2009 and underwent a week-long multi-modal neuroimaging and biomarker study.

Based on poorer-than-chance performance and/or excessive movement determined by deviations larger than the image voxel size, 15 subjects were excluded from the study (2 with Syndrome 1, 6 with Syndrome 2, 4 with syndrome 3, and 4 controls). These subjects were excluded from all further analyses, bringing the respective group totals to nine in Syndrome 1, ten in Syndrome 2, seven in Syndrome 3, and 12 controls.

Stimuli and task

The stimuli consisted of pairs of written words which represent features of common objects. The subjects were to determine whether the two features elicited the retrieval of a memory of a specific object (e.g., “desert” and “humps” ➤ “camel”) or a nonretrieval (e.g., “sleeve” and “jungle”). Subjects were instructed that the target needed to be a specific object, not merely an association between the two words. Fifty trials comprised stimulus pairs that have been shown in previous work (Assaf et al. 2006) to elicit retrieval of a specific object, and 50 were nonretrieval trials. The same feature words that were used in the object-retrieval pairs were re-paired to form the nonretrieval pairs.

Procedures

Cognitive task

Subjects were presented with word-pair stimuli, one word above the other, for 2,700 ms followed by a fixation point, which was presented for between 4,150 ms to 7,550 ms. The stimuli were presented as an event-related design in which the order of word pair presentation was randomized across experimental runs. Subjects used a button box positioned under the fingers of their right hand to make their responses, and the stimuli were displayed on a projector screen positioned at the subject’s head. Instructions for the task require subjects to respond as fast as possible by pressing one button for retrievals and another button for nonretrievals.

Scanning procedures

Functional MRI data were acquired on a 3.0 T Siemens Trio TIM MRI scanner, using a 12-channel receiver array head coil. The data were acquired in the axial plane, using an EPI sequence with a TR of 2 s, a TE of 25 ms, and a flip angle of 90°. Slice thickness was 3.2 mm, with no interslice gap, and slices collected in an interleaved fashion. Field of view was 24 cm, with a 64×64 acquisition matrix.
Image analysis

Image analysis was performed using SPM5. Each individual’s data were slice-time corrected for an interleaved bottom-up acquisition, adjusted for motion, spatially normalized into a standardized MNI template and spatially smoothed (8×8×8 mm Gaussian kernel). Signal changes were modeled as delta functions temporally coincident with the onset of each stimulus and convolved with a canonical hemodynamic response function. Initial event-related analyses were conducted using a within subject fixed-effects model, excluding all false-positive and false-negative responses from the model. We then established two regressors in the model to represent retrievals and nonretrievals, and contrast files were created for each regressor individually as well as for the difference between them.

Results from these within-subject analyses were then used for subsequent random-effects models. An ANOVA was run using SPM5 (2nd-level full-factorial design with repeated measures) with two factors and represented the primary analyses of the study. One factor represented group and consisted of four levels (one for each group) and the other represented condition and consisted of two levels (retrievals and nonretrievals).

For secondary analyses, we chose a priori regions of interest (ROI) that have been consistently associated with performance of this semantic retrieval task and shown to be damaged in previous experiments of GWI (Abdel-Rahman et al. 2002; Assaf et al. 2006; Haley et al. 2000). These included the bilateral medial pre-SMA, thalamus and caudate head. For each subject and condition, average percent signal change within these ROIs was calculated and entered as dependent variables in a two-factor repeated measures ANOVA, similar to the whole brain model. Our primary interest centered on differences between the syndrome variants and the control group for retrievals, non-retrievals and the contrast between these two levels. We considered pairwise group contrasts only following a significant F-test for either the group effect or the group-by-condition interaction. Since these group comparisons were planned and measured strictly in cases of omnibus Type I error control, we did not impose further restrictions on the observed significance probabilities.

Anatomical masks were defined by the WFU Pickatlas (Maldjian et al. 2003), and values of percent signal change were calculated for each regressor (retrievals, nonretrievals, and fixations) within these regions using MARSBAR (Brett et al. 2002). Subsequent correlation analyses were performed using the percent signal change values both between anatomical regions and between behavioral data and anatomical regions.

Results

Behavioral data results

A significant difference was found between the deployed control group and the Syndrome 3 group in correct responses to retrievals ($p<.05$) using a two-sample t-test, with the deployed control group having more correct responses (see Fig. 1). Also, trends of better performance
were observed in the deployed control group compared to the Syndrome 1 group for number of correct nonretrieval responses ($p < .06$), and in correct retrieval reaction times ($p < .07$).

Because many more affected veterans compared to controls were removed from the analyses due to poor performance on the SORT, we ran a permutation test in order to determine whether subjects with a specific

Table 1  BOLD signal changes: location and size of signal maxima

<table>
<thead>
<tr>
<th>Anatomic location of maximum activation</th>
<th>Talairach coordinates</th>
<th>Size (mm3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L/R pre-SMA (BA 6, R&gt;L)</td>
<td>7 16 43</td>
<td>16425</td>
</tr>
<tr>
<td>R thalamus</td>
<td>14 −20 −1</td>
<td>156</td>
</tr>
<tr>
<td>L/R precentral gyrus (BA 4, 6, 9, L&gt;R)</td>
<td>−48 −1 47</td>
<td>21050</td>
</tr>
<tr>
<td>L/R superior parietal lobule (BA 7, 40, R&gt;L)</td>
<td>34 −48 59</td>
<td>8350</td>
</tr>
<tr>
<td>L/R inferior parietal lobule (R&gt;L)</td>
<td>34 −48 47</td>
<td>8389</td>
</tr>
<tr>
<td>L/R superior frontal gyrus (BA 32, R&gt;L)</td>
<td>7 16 43</td>
<td>7487</td>
</tr>
<tr>
<td>L/R inferior frontal gyrus (BA 9, R&gt;L)</td>
<td>52 3 38</td>
<td>18698</td>
</tr>
<tr>
<td>R superior temporal gyrus</td>
<td>51 −39 11</td>
<td>1529</td>
</tr>
<tr>
<td>L/R medial temporal (R&gt;L)</td>
<td>34 −66 27</td>
<td>1137</td>
</tr>
<tr>
<td>L/R inferior temporal (BA 37, R&gt;L)</td>
<td>48 −41 −16</td>
<td>4586</td>
</tr>
<tr>
<td>L/R inferior occipital gyrus (R&gt;L)</td>
<td>31 −91 −4</td>
<td>14817</td>
</tr>
<tr>
<td>L/R middle occipital gyrus (R&gt;L)</td>
<td>31 −90 2</td>
<td>21325</td>
</tr>
<tr>
<td>L/R superior occipital gyrus (R&gt;L)</td>
<td>34 −85 22</td>
<td>4079</td>
</tr>
<tr>
<td>L/R middle cingulate (BA 24, 32, R&gt;L)</td>
<td>7 19 41</td>
<td>6350</td>
</tr>
<tr>
<td>L/R anterior cingulate (BA24)</td>
<td>0 20 −9</td>
<td>392</td>
</tr>
<tr>
<td>R parahippocampal Gyrus</td>
<td>34 −22 −23</td>
<td>784</td>
</tr>
</tbody>
</table>

ANOVA: Main effect of condition $P<0.001$

<table>
<thead>
<tr>
<th>Talairach coordinates</th>
<th>Size (mm3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>7</td>
<td>−27</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>45</td>
<td>8</td>
</tr>
<tr>
<td>45</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
</tr>
</tbody>
</table>
syndrome were more likely to be performance outliers. We found that relative to controls and subjects with Syndromes 1 or 3, subjects in the Syndrome 2 group were significantly more likely ($p < .05$) to exhibit poor performance that resulted in removal from the study cohort.

Primary imaging results

The fMRI data were entered into a two-factor ANOVA and assessed for effects of condition and group. For the effect of condition (correct retrievals vs. correct nonretrievals), whole-brain analysis of the fMRI data demonstrated areas of significant signal change ($p < .001$, uncorrected) that corresponded to previously published data using the SORT task, including bilateral pre-supplementary motor area (pre-SMA), superior and inferior parietal lobules, middle occipital gyri, and anterior cingulate gyri (see Fig. 2 and Table 1). Whole-brain analysis of fMRI to detect group-wise effects demonstrated significant signal change ($p < .001$, uncorrected) again in regions previously noted with this task including right thalamus, right caudate head and anterior cingulate, as well as bilateral superior temporal gyrus, right medial temporal region, and right superior frontal gyrus (Table 1).

Comparisons between all affected veterans as a group and controls were also assessed in order to discover any regions affecting semantic memory that are common to all syndromes. No significant findings were present.

Secondary imaging results

1. A group x condition interaction ANOVA targeted to the anatomical ROIs delineated in both previous investigations and in the above-described primary analysis (pre-SMA, caudate heads, and thalami bilaterally) revealed significant signal change in the caudate heads and thalamic hemispheres (see Fig. 3). Post-hoc testing of the thalamus showed a significant signal change increase for correct nonretrieval epochs in Syndrome 2 compared to both Syndrome 1 and controls. Similar testing of the caudate heads showed a significant signal change decrease for correct nonretrievals in Syndrome 1 compared to both Syndrome 2 and controls.

2. To evaluate the relationships between performance and signal change in specific areas, we measured correlations between reaction time and percent signal change (PSC) and between percent correct responses and PSC for the bilateral thalami, caudate heads, and pre-SMA. An inverse correlation ($p < .05$) was found between PSC in the thalami and reaction time across all groups (see Fig. 4). Also, a direct correlation was found between PSC in the caudate heads and percent correct responses across all groups.

3. To assess whether reaction times affected percent signal change similarly for all groups, an ANCOVA was performed within the caudate heads bilaterally. An increase in reaction times was associated with a decrease in percent signal change for normal controls, Syndrome 1 and Syndrome 3 groups. In contrast, a positive association between reaction time and percent signal change was found in Syndrome 2, significantly different than the other Syndrome groups and normal controls ($p = 0.02$, see Fig. 5). A similar ANCOVA was performed within the thalamic hemispheres and found a trend towards a statistically significant interaction, at $p < .08$, again with Syndrome 2 showing an increase in PSC associated with an increase in reaction times which was distinctly different from all other groups.

4. PSC inter-region correlations between thalamus and pre-SMA, thalamus and caudate head, and pre-SMA and caudate head were also assessed within each
Discussion

This study demonstrated both behavioral deficits and fMRI differences in veterans with Gulf War Illness compared to normal controls as the participants performed a semantic memory task. There were differences in both accuracy and reaction time between affected veteran groups and the deployed control group, with specific affected groups making more errors and having longer reaction times. The anatomic locations of signal change differences between semantic retrievals and nonretrievals were consistent with previous findings using SORT (Kraut et al. 2002, 2002; Assaf et al. 2006), and correlations between these regions and performance measures showed significant differences between affected veteran groups and controls. Brain regions shown to work in conjunction to perform semantic retrieval exhibited covarying activity only in the deployed control group, and not in the affected groups. These significant differences between clinical groups relative to controls demonstrate the sensitivity of the SORT task in distinguishing clinical populations, similar to what was found in previous studies using the same task (Assaf et al. 2005; Kraut et al. 2006, 2007).

Comparisons of task performance showed both significant differences and trends toward poorer performance when comparing Syndromes 1 and 3 with the deployed control group, and there were no clear differences in specific performance parameters that distinguished between

![Fig. 4](image1)

Significant correlations between reaction time (RT) and % signal change within the thalamus (top) and % correct responses and % signal change within the caudate heads (bottom) across all groups. These analyses revealed significant correlations between each region-pair in the deployed control group, and one correlation between the thalamus and caudate in Syndrome 3. No significant region-pair correlations were found within Syndromes 1 and 2.

![Fig. 5](image2)

Analysis of covariance (ANCOVA) showing the effect of group membership on the correlation between % signal change (SC) and reaction time (RT) within the thalamus and caudate heads.

Table 2: Strength of inter-region % signal change correlations organized by group membership

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>Synd. 1</th>
<th>Synd. 2</th>
<th>Synd. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thalamus &amp; Pre-SMA</td>
<td>0.04*</td>
<td>0.21</td>
<td>0.88</td>
<td>0.54</td>
</tr>
<tr>
<td>Thalamus &amp; Caudate</td>
<td>0.01*</td>
<td>0.58</td>
<td>0.95</td>
<td>0.01*</td>
</tr>
<tr>
<td>Caudate &amp; Pre-SMA</td>
<td>0.02*</td>
<td>0.58</td>
<td>0.29</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Values followed by asterisk (*) represent a significant correlation ($p<.05$) between regions.
the affected veteran groups. This latter finding suggests a common etiology for impairment across these two Syndromes. It should be noted that Syndrome 2 subjects showed no significant performance differences compared to controls, but a significant number of subjects from this group were so impaired on task performance that they could not participate in the study. This exclusion of the most impaired subjects belonging to the Syndrome 2 group was thought to be a major contributing factor to the lack of significant behavioral differences between this group and the normal controls.

Given that only correct trials were used in the analyses and that the signal changes across all groups matched areas previously found to be active during the SORT task, the pattern of signal change found in the current study suggests that these subjects with semantic memory difficulties still engaged approximately the same anatomic regions when correctly performing the task as do normal individuals. However, even though the pattern of signal change showed some overlap across groups, there were significant interaction effects between specific groups and signal changes within the thalamus and caudate, indicative of dysfunction in the affected groups (Fig. 3).

Among the brain regions that demonstrated signal changes for correct retrievals were medial BA6, thalamus, and caudate. Previous studies have suggested that increased 20–35 Hz synchronized electrical activity links the activity between these and perhaps other brain regions during correct retrievals (Slotnick et al. 2002; Hart et al. 2007). Several of these brain regions demonstrated a correlation between degree of signal change and observable test performance, and it is likely that these regions or the connections between them are dysfunctional in Gulf War illnesses. Both the thalamus and basal ganglia regions, including the caudate, are thought to have been adversely affected by exposure to sarin and/or physostigmine bromide during the first Gulf War (Haley et al. 2000; Abdel-Rahman et al. 2002). White and colleagues have also suggested that these same agents, particularly sarin, can damage white matter connecting subcortical regions as well as cortical-subcortical connections (Heaton et al. 2007). Damaged white matter connections between medial BA6 and these subcortical structures could result in behavioral impairments due to impeded propagation of inter-regional synchronizing electrical rhythms (Slotnick et al. 2002).

Syndrome 2 appeared to have markedly different patterns of thalamic and caudate signal change as related to reaction time, even though the most impaired participants in the group were removed, suggesting a distinct type of dysfunction in these patients compared to the other two syndromes (Fig. 5). These differences between the three syndromes themselves, and the lack of significant differences across all affected veterans compared to controls, also support the clustering of affected veterans into syndromes. Analyzing each syndrome separately, as done in this study, can illuminate deficits that are more severely expressed in one syndrome relative to others.

The thalamic and caudate signal changes associated with decreased performance parameters in Syndrome 2 patients compared to the other groups could be similar to changes in early degenerative diseases. For example, increased BOLD signal in early Mild Cognitive Impairment (MCI) is proposed to indicate an increase in effort in an attempt to maintain performance in a dysfunctional system, similar to Syndrome 2. As degenerative changes progress to later MCI and/or early Alzheimer’s disease (AD), these findings change to decreased BOLD responses accompanied by impaired task performance (Kraut et al. 2007).

Disruption of a neural circuit comprising thalamus, caudate head, and medial BA6 is further supported by the significant pair-wise inter-regional correlations between percent fMRI signal change in these regions for the control group but not for the affected veteran groups. Given that fMRI signal changes in the thalamus and caudate were shown to correlate with accuracy and reaction time during task performance across all groups, the integrity of this circuit appears essential to effective semantic memory retrieval, and is disrupted to varying degrees in these patients. The pre-SMA area, which did not show significant differential signal change between groups in the primary analysis, nor correlations with performance measures, does not appear to be affected in the same manner as the subcortical nuclei. This is consistent with previous investigations of GW related illnesses which have shown thalamic and basal ganglia abnormalities without significant cortical disruptions.

The thalamic and caudate regions exhibited signal changes that correlated with performance measures across all subjects, demonstrating their involvement in performance of this task, which is consistent with previous studies. In addition, there was evidence of increased signal changes in the thalami and caudate heads for nonretrievals in Syndrome 2. Correlations between signal changes in these areas and reaction times also showed pronounced differences in Syndrome 2 compared to other groups, with Syndrome 2 being the only group to show reaction time increases relative to signal change. This syndrome group included the patients who have the most profound performance deficits on this task, the most impaired of whom could not perform the task in fMRI. These patients appear either to have the most profound impairment within a single disease entity, or may be manifesting a different pathophysiological process. Further characterization of the
neurochemical (MR Spectroscopy), ultrastructural (diffusion tensor indices) and electrophysiological (EEG, event related potentials) differences between these syndrome groups may help clarify these issues.

Acknowledgments This work was supported by the VA IDIQ contract number VA549-P-0027 awarded and administered by the Dallas, TX VA Medical Center and by the Department of Defense. The content of this paper does not necessarily reflect the position or the policy of the U.S. government, and no official endorsement should be inferred. The authors report no conflict of interest in the publication of this work, financial or otherwise. The authors would like to thank Tim Green, Virginia Buhl, and Jack Grinnan for their invaluable assistance, and D T’ib for insightful comments.

References


