Word-finding impairment in veterans of the 1991 Persian Gulf War

Kristin Moffett a, Bruce Crosson b,c,d,f,* , Jeffrey S. Spence g , Kimberly Case a , Ilana Levy a ,
Kaundinya Gopinath g , Parina Shah g , Aman Goyal g , Yan Fang g , Richard W. Briggs g,h , John Hart Jr. g ,
Anna Moore c,1 , Robert W. Haley g

a Department of Clinical and Health Psychology, University of Florida, 1225 Center Drive, Room 3151, Gainesville, FL 32611, USA
b Department of Veterans Affairs Rehabilitation Research and Development, Brain Rehabilitation Research Center of Excellence, Malcolm Randall VA Medical Center, 1601 S.W. Archer Road, Gainesville, FL 32608-1197, USA
c Department of Veterans Affairs Rehabilitation Research and Development, Center of Excellence for Visual and Neurocognitive Rehabilitation, 1670 Clairmont Rd., Decatur, GA 30033, USA
d Departments of Neurology and Radiology, Emory University, 101 Woodruff Circle, Suite 6000, Atlanta, GA 30322, USA
e Department of Rehabilitation Medicine, Emory University, PO Box 5010, Atlanta, GA 303205010, USA
f Department of Psychology, Georgia State University, PO Box 5010, Atlanta, GA 30302-5060, USA
g Departments of Internal Medicine (Epidemiology Division) and Radiology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390-8874, USA
h Department of Physics & Astronomy, Georgia State University, Atlanta, GA 30302-5060, USA

ABSTRACT

Approximately one quarter of 1991 Persian Gulf War Veterans experience cognitive and physiological sequelae that continue to be unexplained by known medical or psychological conditions. Difficulty coming up with words and names, familiar before the war, is a hallmark of the illness. Three Gulf War Syndrome subtypes have been identified and linked to specific war-time chemical exposures. The most functionally impaired veterans belong to the Gulf War Syndrome 2 (Syndrome 2) group, for which subcortical damage due to toxic nerve gas exposure is the suspected cause. Subcortical damage is often associated with specific complex language impairments, and Syndrome 2 veterans have demonstrated poorer vocabulary relative to controls. 11 Syndrome 1, 16 Syndrome 2, 9 Syndrome 3, and 14 age-matched veteran controls from the Seabees Naval Construction Battalion were compared across three measures of complex language. Additionally, functional magnetic resonance imaging (fMRI) was collected during a covert category generation task, and whole-brain functional activity was compared between groups. Results demonstrated that Syndrome 2 veterans performed significantly worse on letter and category fluency relative to controls. 11 Syndrome 1, 16 Syndrome 2, 9 Syndrome 3, and 14 age-matched veteran controls from the Seabees Naval Construction Battalion were compared across three measures of complex language. Additionally, functional magnetic resonance imaging (fMRI) was collected during a covert category generation task, and whole-brain functional activity was compared between groups. Results demonstrated that Syndrome 2 veterans performed significantly worse on letter and category fluency relative to Syndrome 1 veterans and controls. They also exhibited reduced activity in the thalamus, putamen, and amygdala, and increased activity in the right hippocampus relative to controls. Syndrome 1 and Syndrome 3 groups tended to show similar, although smaller, differences than the Syndrome 2 group. Hence, these results further demonstrate specific impairments in complex language as well as subcortical and hippocampal involvement in Syndrome 2 veterans. Further research is required to determine the extent of language impairments in this population and the significance of altered neurologic activity in the aforementioned brain regions with the purpose of better characterizing the Gulf War Syndromes.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

An estimated one quarter of the 700,000 veterans of the Persian Gulf War experience chronic physiological and psychological symptoms of unknown etiology (Lange et al., 2001). Affecting veterans largely report distressing multisymptom complaints that significantly impair daily functioning (Anger et al., 1999; Coker, Bhattacharyya, & Graham, 1999; Fukuda et al., 1998; Gray, Reed, Kaiser, Smith, & Gastanaga, 2002; Haley & Kurt, 1997; Lange et al., 2001). Through extensive epidemiological
investigation of hundreds of Gulf War veterans, Haley and colleagues used factor analysis of symptoms to develop a working research definition of Gulf War illness that includes three syndrome subtypes, or variants, referred to as the Gulf War syndromes (Haley, Kurt, & Hom, 1997), which they cross-validated with structural equation modeling, first in a clinical sample (Haley, Luk, & Petty, 2001) and later in a large national random sample (Iannacchione et al., 2011). They identified unique war-time chemical exposures that are associated with each syndrome (Haley & Kurt, 1997). Syndrome 1 is referred to as “impaired cognition” and is associated with exposure to pesticides in flea and tick collars used by soldiers to deter insects. It is characterized by sleep disturbance, attention and memory difficulties, and migraine headaches. Syndrome 2 is referred to as “confusion-ataxia” and is strongly associated with low-level sarin exposure and unusually severe side effects of the pyridostigmine anti-nerve gas medication soldiers were ordered to take. It is characterized by more profound cognitive difficulties involving confusion and difficulty processing information, as well as motor abnormalities and psychological symptoms. Syndrome 3, referred to as “central neuropathic pain,” is associated with use of high-concentration DEET and more severe side effects of pyridostigmine. It is characterized by pain and weakness in the joints and muscles of the body as well as milder cognitive complaints. Although veterans in all three syndrome groups show some degree of functional impairment, those with Syndrome 2 are the most cognitively and functionally impaired, and they are 12.5 times more likely to be unemployed than veterans from other syndrome groups (Haley, Maddrey, & Gershfield, 2002; Haley et al., 1997; Iannacchione et al., 2011).

Sarin nerve gas, associated with Syndrome 2, is an organophosphate that produces immediate poisoning by binding to and thereby inactivating acetylcholinesterase, allowing build-up of acetylcholine at cholinergic synapses. It produces long-term neuropsychiatric symptoms and deficits by mechanisms not yet understood. Henderson et al. (2002) exposed rats to low levels of sarin nerve gas over an extended period of time, causing no immediate overt clinical symptoms of poisoning. One month post-exposure, the rat brains exhibited significant reductions of acetylcholinesterase in the basal ganglia, and additionally in the hippocampus when sarin exposure was coupled with heat-induced stress. In a similar study, rats that were treated with pyridostigmine bromide (PB) in combination with shock-induced stress demonstrated lower acetylcholinesterase activity in the basal ganglia and basal forebrain (Beck et al., 2001). These animal data provide a working model of the long-term effects of sarin nerve gas on the brain.

Veterans ill with Gulf War syndromes also exhibit changes in the basal ganglia and hippocampus, most consistent and severe in those with Syndrome 2. Magnetic resonance spectroscopy (MRS) data identified biochemical changes in basal ganglia bilaterally (Haley et al., 2000), with Syndrome 2 veterans showing the greatest reduction in the N-acetylaspartate-to-creatine ratio (NAA/Cr), indicating reduced functional neuronal mass in this region. Decreased NAA/Cr ratios in the left basal ganglia of this sample of veterans are also associated with altered central dopaminergic production, signifying additional neurotransmitter changes present in Syndrome 2 veterans (Haley et al., 2000). Reduction of NAA/Cr was also noted in hippocampus of another sample of ill Gulf War veterans (Menon, Nasrallah, Reeves, & Ali, 2004). Cholinergic challenge by intravenous infusion of physostigmine in a 1998 study produced abnormal changes in regional cerebral blood flow (rCBF), measured by single-photon emission computed tomography (SPECT), compared with controls (Haley et al., 2009). The most marked difference from controls was a paradoxical increase in rCBF in Syndrome 2 in the hippocampus, caudate and amygdala; less marked differences were seen in the right putamen in Syndrome 1, and in the left thalamus in Syndromes 1 and 3. In a repeat of the cholinergic challenge study of the same veterans a decade later, rCBF measured by MRI arterial spin labeling (ASL) showed the same paradoxical reversal in the hippocampus in Syndrome 2, but now also in the Syndrome 3 group (Li et al., 2011). One theory regarding these paradoxical changes posits that the increases in activity noted in Syndromes 2 and 3 veterans reflect hyperactivity of the cholinergic system as a downstream effect of sarin-induced cellular damage (Allon et al., 2011). Rats exposed to low-levels of sarin nerve gas exhibit chronic changes to the inhibitory muscarinic-2 cholinergic receptors of the hippocampus, thereby reducing their affinity. In this way, one could similarly interpret the observed activity increases in Syndromes 2 and 3 veterans as chronic disinhibition of hippocampal processes resulting from cholinergic receptor damage.

One of the symptoms most characteristic of Gulf War illness is difficulty finding previously familiar words, e.g., “coming up with the right word” (Haley et al., 1997, 2001) or “trouble finding words” (Fukuda et al., 1998). Previous research has documented that word generation increases activity in the basal ganglia and that basal ganglia disturbances are associated with impairment of complex verbal functions (Henry & Crawford, 2004). Crosson et al. (2003) demonstrated that covert word generation tasks evoked basal ganglia activity bilaterally in normal subjects. Further, Copland and colleagues demonstrated that patients with dominant non-thalamic basal ganglia lesions perform significantly worse than normal controls on complex language tasks including word fluency, sentence generation, interpretation of ambiguous sentences, and definition of words (Copland, Cheryn, & Murdoch, 2000). Veterans with Gulf War syndromes also demonstrate impaired performance on the Vocabulary subtest of the Wechsler Adult Intelligence Scale in comparison to controls (Hom, Haley, & Kurt, 1997). These findings led us to question whether (1) the verbal impairments noted in ill Gulf War veterans extend beyond the reduced ability to correctly define words, (2) vary in presentation or severity across syndrome subtypes, and (3) are associated with basal ganglia pathology.

Hence, the current study compared veterans with Gulf War syndromes to control veterans on a short neuropsychological battery of word fluency tasks and on a word fluency task during functional magnetic resonance imaging (fMRI). Given the previous evidence of basal ganglia abnormalities in this sample of Gulf War veterans, most severe in the Syndrome 2 group, and the prominence of their word-finding problems, we expected veterans with Gulf War illness to perform worse than the control veterans, with Syndrome 2 veterans demonstrating the lowest performance across all tasks. Similarly, given the reduction in NAA/Cr in the basal ganglia as an indicator of reduced basal ganglia function, we expected veterans with Gulf War illness to exhibit reductions in brain activity primarily in the basal ganglia during a covert word generation task, with Syndrome 2 veterans exhibiting the greatest difference in activity when compared to control veterans. We also included the left and right thalamus in our analyses due to their significant interconnections with the structures of the basal ganglia (Alexander, DeLong, & Strick, 1986; Middleton & Strick, 2000) as well as the hippocampus because of its apparent involvement in Gulf War syndromes (Haley et al., 2009; Li et al., 2011; Liu et al., 2011; Menon et al., 2004).

2. Methods
2.1. Participants

The subjects included 50 right-handed male Gulf War-era veterans, all members of the Twenty-Fourth Reserve Naval Mobile Construction Battalion (Seabees), selected on the basis of a
validated case definition (Haley et al., 1997, 2001; Iannacchione et al., 2011). Those with Gulf War illness included 11 with Syndrome 1, 16 with Syndrome 2, and 9 with Syndrome 3. The control group was composed of 14 members of the battalion, age–sex–education-matched to the Syndrome 2 group, who did not meet the case definitions for any of the syndromes. There are a greater number of participants in the Syndrome 2 group than the control group because we were able to run additional Syndrome 2 veterans through the study after the initial data-collection phase. We did not feel it was necessary to increase the size of the control group, however, because there was already sufficient power to detect group differences. All participants were previously enrolled in studies of Gulf War illness (Haley et al., 1997, 2000, 2009). There were no significant group differences in education or in age, except that the mean age for the Syndrome 1 group was significantly younger than the other groups (Table 1). For this reason, we investigated the effect of age on cognitive performance variables and BOLD change for the functional imaging analyses.

Structured Clinical Interview for DSM-IV Axis I Disorders (SCID; First, Spitzer, Gibbon, & Williams, 2007) was used to assess psychiatric comorbidities including major depressive disorder (MDD) and alcohol and drug (cannabis, amphetamine) abuse or dependence. Biochemical assays were also used to detect the presence of drugs (cannabis, cocaine) in urine. Post-traumatic stress disorder (PTSD) was determined by a score \( \geq 40 \) on the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1995; Weathers, Keane, & Davidson, 2001; Weathers, Ruscio, & Keane, 1999).

This research was approved by the Institutional Review Board at the University of Texas Southwestern Medical Center, and written informed consent was obtained before the study from all participants in accordance with these guidelines.

2.2. Experimental design and procedure

As part of a week-long battery of testing, participants completed neuropsychological testing which included the verbal fluency subtests of the Delis–Kaplan Executive Function System (D-KEFS) (Delis, Kaplan, & Kramer, 2001), administered outside the scanner, and a word-generation task performed during fMRI scanning.

2.2.1. Behavioral testing

Participants completed three behavioral measures of verbal fluency from the D-KEFS battery. The Letter Fluency task requires participants to generate as many words as possible beginning with a specific letter of the alphabet for three consecutive 60 s trials, the Category Fluency task requires participants to generate as many members from two semantic categories (e.g., fruits, furniture) for 60 s. Responses were recorded verbatim and each subtest was scored for the total number of correct responses. Four Syndrome 2 subjects did not complete the Category Fluency and Category Switching tasks.

2.2.2. fMRI task and parameters

The fMRI portion of the study was completed using a block word-generation task wherein participants were visually presented with category cues (e.g., ocean animals) while in the scanner and asked to silently generate as many members from that category as possible (e.g., whale, shark, dolphin, etc.) until a picture of a stop sign appeared (Fig. 1). Periods of word generation alternated with periods of visual fixation wherein participants were instructed to relax, clear their minds, and fixate on a “plus sign.” In total, each participant completed 4 fMRI runs. Each run consisted of eight 15-s blocks of word generation, alternated with pseudo-random visual fixation intervals of 9 s, 12 s, or 15 s. Intervals of visual fixation were varied for the purpose of minimizing the effects of periodic physiological noise and were at least 9 s in length to allow the hemodynamic response to return to baseline levels.

We selected a silent generation version of this task because the target of our inquiry was the basal ganglia and thalamus. In our experience, the silent version of the task with continuous fMRI sampling (Crosson et al., 2003) yields more consistent and robust basal ganglia (caudate and putamen) activity bilaterally and left thalamic activity across multiple word generation tasks (two versions of a category member generation paradigm and generating rhyming words) than overt category member generation with sparse sampling (Meinzer et al., 2009, 2012). For example, in groups and tasks similar to those of Crosson et al. (2003), Meinzer et al. (2009) found basal ganglia activity increases only in the right caudate nucleus and Meinzer et al. (2012) showed activity only in the caudate nucleus bilaterally, and in the left thalamus. Similarly, Shuster and Lemieux (2005) reported basal ganglia activity increases only in the right caudate nucleus during overt single-word repetition, and Riecker et al. (2005) showed activity only in the putamen/pallidum bilaterally during overt syllable repetition. Hence, although the right caudate activity increases show up consistently during overt word generation, left basal ganglia activity increases show up consistently only during covert generation, and the left basal ganglia were a specific target of this investigation.

2.2.3. Image acquisition

Data were collected on a 3T Siemens Trio scanner with a body coil to transmit and a 12-channel head array coil to receive. Whole-brain high-resolution fMRI data were acquired with an echo planar

| Table 1 Demographic characteristics and comorbidities of study subjects. |
|-----------------------|-----------------|-----------------|-----------------|-----------------|
| Characteristics       | Gulf War syndrome groups | Controls (N = 14) |
|                       | Syndrome 1 (N = 11) | Syndrome 2 (N = 16) | Syndrome 3 (N = 9) |
| Age, years (SD)       | 51.4 (6.1) | 63.4 (6.9) | 57.1 (7.2) | 60.1 (6.3) |
| Education, years (SD) | 13.6 (9.0) | 13.5 (3.3) | 13.0 (2.1) | 13.6 (2.6) |
| Active PTSD (%)       | 2 (18) | 4 (27)* | 1 (11) | 0 (0) |
| Active MDD (%)        | 2 (18) | 3 (20)* | 0 (0) | 0 (0) |
| Active alcohol abuse or dependence (%) | 3 (27) | 2 (13)* | 2 (22) | 3 (21) |
| Active drug abuse or dependence (%) | 1 (9) | 1 (7)* | 1 (11) | 1 (7) |

PTSD = post-traumatic stress disorder; MDD = major depressive disorder.

* One Syndrome 2 subject declined the SCID interview.
imaging (EPI) sequence and parallel imaging with generalized auto-calibrating partially parallel acquisitions (GRAPPA) acceleration factor of 2, TR/TE/Flip Angle (FA) = 3000 ms/30 ms/90°. Forty-six 3 mm thick sagittal slices were acquired with 1.8 mm × 1.8 mm in-plane resolution. In each run, 82 whole-brain volumes were acquired in 4.1 min. A high-resolution T1-weighted 3-dimensional structural MRI scan (3D MP-RAGE, TR/TE/FA = 2250 ms/900 ms/3°/9°, 0.9 mm × 0.9 mm × 1.0 mm, 4 min 52 s acquisition time) was also obtained to provide an anatomical reference. Head motion was minimized by using foam padding.

2.3. Data analysis

2.3.1. Behavioral analysis

Each verbal fluency measure was scored for total correct responses and tested for normality. The right skewed distribution of the Category Fluency task was log transformed for the statistical hypothesis testing. We tested the overall group effect across all 3 behavioral measures with a multivariate analysis of variance model using the repeated-measures option of the SAS Mixed procedure (Wright, 1998). The model included contrasts testing the group effect for each behavioral measure and the difference between each syndrome and the control group on each behavioral measure. The contrasts were adjusted for multiple testing by the Benjamini–Hochberg method allowing for the correlation among measures. The mixed procedure includes non-normal values of all patients at baseline on some measures and provides a chi square test appropriate for data with missing values. Due to the lower mean age of the Syndrome 1 group, we examined the effect of age on verbal fluency. However, finding that age was not significantly associated with any of the 3 outcome measures and that introducing it into the multivariate model only increased the standard errors without altering the least squares estimates of the group means, we did not include age in the final models.

2.3.2. fMRI image processing

fMRI data were spatially normalized and overlaid onto structural images using the Analysis of Functional Neuroimages (AFNI) program from the National Institutes of Health (http://afni.nimh.nih.gov/afni). To minimize head movement, functional runs were registered to the image immediately following the acquisition of the anatomical T1-weighted scan with an intensity-based linear least squares regression algorithm. Locally estimated scatterplot smoothing (LOESS) was used to detrend the time series images with 40% span to remove linear trends and low frequency drift (Cleveland & Devlin, 1988). The four blocks of fMRI data were then concatenated. BOLD amplitude maps for each participant were obtained by regressing the detrended time series maps on the convolution of a canonical hemodynamic response function (two-gamma mixture with varying amplitude coefficient) with a boxcar stimulus vector, which represented the duration of each word generation epoch. BOLD amplitude maps were transformed to percent signal change from baseline and smoothed by kriging interpolation (Cressie, 1993), which removed small-scale measurement error variance but did not alter the range of spatial correlation inherent in the images themselves. The data analyses were completed using the R statistical computing language (http://www.R-project.org).

2.3.3. Hypothesis tests of group differences in BOLD signal response

Statistical inference at the group level followed methodology described by Spence et al. (2007), derived from standard methods of spatial modeling common to geostatistical applications (Chilès & Delfiner, 1999). In summary, BOLD amplitude maps, along with their respective variances, from each subject were analyzed in a general linear ANOVA model testing the effects of group classification on mean BOLD signal response to the word-finding task. F-statistic maps were generated from the full 4-group model, and t-statistic maps were generated following the planned contrasts between each syndrome group and the control group.

2.3.4. Mapping of significant group differences in mean BOLD signal of voxel spatial clusters

For each statistical map, contiguous spatially correlated voxels were grouped into clusters based on the estimated range of spatial correlation obtained from the geostatistical model, and local signal was calculated as a weighted average of the voxels within each cluster. Statistical inference was then drawn from the cluster-level averages. The rate of false positive errors due to the multiple tests over the entire set of cluster-level statistics was controlled with the Benjamini–Hochberg method (Benjamini & Hochberg, 1995), setting the expected false discovery rate (FDR) at 0.05. For plotting the group-level means and standard errors within anatomical regions of interest (ROIs), we averaged the significant cluster-level BOLD amplitude means that were spatially contiguous within the anatomical ROIs.

2.3.5. Covariate analysis

Effects of age on BOLD amplitude were also investigated due to the significantly lower mean age for the Syndrome 1 group (Table 1). In all regions of the brain that were examined in this study, however, age did not significantly affect analyses of BOLD signal response. We, therefore, did not include age in the final linear models.

3. Results

3.1. Behavioral results

The overall multivariate analysis of variance across all 3 fluency measures found a significant group effect ($\chi^2 = 17.99$, $df = 9$, $p = 0.035$). Contrasts in the model showed that the number of correct words generated varied significantly among the 4 comparison groups for the Letter Fluency test and the Category Fluency test, but not for the Category Switching test (Table 2). In the contrasts testing individual syndrome group differences from the control group on the Letter Fluency test, the subjects with Syndrome 1, the least ill group, generated on average the same number of correct words as those in the control group; whereas, those in the Syndrome 2 group and the Syndrome 3 group generated fewer correct words than those in the control group (Table 2). The two-group differences on the Category Fluency test were similar but smaller (Table 2).

3.2. fMRI results

3.2.1. Full 4-group model and group contrasts

Mean BOLD change in response to the fMRI word-finding task—particularly in large areas of the posterior thalamus (pulvinar) bilaterally, left anterior putamen, right amygdala, and right posterior hippocampus—expressed differentially across the groups, as indicated by significant F statistics (multiple-testing corrected, FDR = 0.05, Fig. 2).

In the thalamus bilaterally the right amygdala, and the left anterior putamen, the BOLD response to the word-generation cues was generally lower in the three syndrome groups than in the controls (Fig. 2). This difference between the Syndrome 2 group and the controls was large and significant in all 4 areas; whereas, the difference between the controls and the Syndromes 1 and 3 groups was smaller and of borderline significance in the thalamus bilaterally and the right amygdala, but uniformly large and highly
significant in the left anterior putamen (Fig. 2). In the right amygdala, the Syndrome 2 group exhibited negative activity relative to baseline, while the activity differences in all other groups was positive. At a FDR of .10, group differences in BOLD response were evident in the left head of the caudate nucleus as well, particularly reflecting lower Syndrome 2 response compared to controls.

### 3.2.2. Areas of the significant differences overlaid on an anatomical MRI image

Fig. 3 provides a map of significant voxel clusters overlaid on a structural MRI scan image, specifically showing the location and size of areas where the mean BOLD response to category-member generation differed from controls after multiple-testing correction (FDR = 0.05). All of the voxel clusters that survived the multiple-testing correction involved differences between Syndrome 2 and controls. The voxel clusters where BOLD response was lower in the Syndrome 2 group than the controls are shown on the heat scale; whereas, those where the BOLD response was higher in the Syndrome 2 are shown on the cool scale. Table 3 provides additional information about the spatial coordinates, volumes, and t-statistics for these clusters.

### 4. Discussion

The present study sought to assess whether the Gulf War Syndromes, especially Syndrome 2, were associated with cognitive and/or neurophysiological abnormalities related to verbal

<table>
<thead>
<tr>
<th></th>
<th>Letter fluency</th>
<th>Category fluency</th>
<th>Category switch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn1</td>
<td>11 43.00 11.86</td>
<td>11 39.00 7.48</td>
<td>11 14.00 3.35</td>
</tr>
<tr>
<td>Syn2</td>
<td>16 31.31*</td>
<td>12 31.75*</td>
<td>12 11.50 3.68</td>
</tr>
<tr>
<td>Syn3</td>
<td>9 33.78*</td>
<td>9 34.00c</td>
<td>9 12.78 3.23</td>
</tr>
<tr>
<td>Controls</td>
<td>14 43.57 12.09</td>
<td>14 38.07 6.44</td>
<td>14 12.50 2.53</td>
</tr>
<tr>
<td><strong>p value</strong></td>
<td>0.007</td>
<td>0.03</td>
<td>0.29</td>
</tr>
</tbody>
</table>

* On the Letter Fluency test, Syndrome 2 (p = 0.002) and Syndrome 3 (p = 0.029) differed significantly from controls.

b The Category Fluency and Category Switch tests were not completed by 4 subjects with Syndrome 2.

c On the Category Fluency test, Syndrome 2 differed significantly from controls (p = 0.013), but the difference between Syndrome 3 and controls was not significant (p = 0.19).

d The overall multivariate analysis of variance testing the group effect across all 3 behavioral measures was significant ($X^2 = 17.99, df = 9, p = 0.035$), and these p values are from the contrasts testing the group effect for each behavioral measure, adjusted for multiple testing.
executive functioning (as measured by verbal fluency). Prior cognitive testing on this sample of veterans revealed verbal impairment on a neuropsychological measure of vocabulary (Hom et al., 1997). Thus, the first aim of the current study was to compare verbal fluency performance across groups to determine if these previously noted language difficulties extended to verbal executive functioning. Although we expected to detect group differences across all the behavioral measures of verbal fluency, significant differences were found only for the Letter and Category Fluency measures. Smaller group differences for the Category than the Letter Fluency task may be related to a lower difficulty for the category than the letter cues. The Category Fluency task provides broad semantic categories such as “Animals” from which participants are instructed to generate exemplars. This breadth may enable participants to generate correct responses from a number of semantic subcategories (e.g., zoo animals, domesticated animals, etc.). Conversely, the lack of group differences for the Category Switching measure may have resulted from the high level of task difficulty.

Although our analyses included all syndrome groups, our strongest hypotheses for subcortical effects involved the Syndrome 2 group. The basal ganglia have been shown to play a role in the frontal-executive aspects of word generation through their participation in frontal–basal ganglia–thalamic networks. The dorsal caudate is the striatal component of the loop that connects pre-SMA (supplementary motor area) to subcortical structures (pre-SMA loop), and the pre-SMA is recognized to be reciprocally connected to the prefrontal cortex (Matsuzaka, Aizawa, & Tanji, 1992; Middleton & Strick, 2000; Picard & Strick, 2001). On the other hand, the anterior putamen is the striatal component connecting Broca’s area to its basal ganglia loop (Ford et al., 2013). In the left hemisphere these loops appear to play a role in lexical retrieval during verbal fluency tasks, with subcortical structures providing a response bias toward one lexical item among competing choices for each generated word (Bohsali & Crosson, in press; Crosson, 2013; Crosson et al., 2003; Crosson, Benjamin, & Levy, 2007). Indeed, previous research on patients with Huntington’s disease (Monsch et al., 1994), which is also characterized by basal ganglia pathology, suggested that damage to these structures affect the initiation and retrieval aspects of fluency, thereby negatively influencing performance across all verbal fluency subtests. Thus, we expected our sample of Syndrome 2 veterans with previously documented basal ganglia abnormalities to perform significantly worse than the other veteran groups on all verbal fluency tasks.

Our analyses showed that GWS veterans in the Syndrome 2 and Syndrome 3 groups performed less well on measures of verbal fluency than their counterparts in the control group. As expected, Syndrome 2 veterans were the lowest performers across the verbal fluency tasks, confirming the presence of lexical-semantic retrieval problem in this group. The question remains as to whether these noted deficits reflect isolated impairments in verbal fluency or a broader dysfunction in complex language functions as a whole.

Utilizing fMRI, the present study also sought to localize and quantify the neural abnormalities associated with reduced verbal fluency in this GWS population. A word generation task known to elicit robust activity in the left basal ganglia and thalamus as well as the right basal ganglia (Crosson et al., 2003) was chosen as the task to be performed in the scanner. It should be noted that the category cues used in the scanner had much smaller numbers

<table>
<thead>
<tr>
<th>Cluster location</th>
<th>MNI coordinates</th>
<th>Volume (mm³)</th>
<th>t-Statistic (Syn 2 vs. controls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right Hippocampus</td>
<td>33 -24 -12</td>
<td>530</td>
<td>-7.10</td>
</tr>
<tr>
<td>Right Amygdala</td>
<td>21 -5 -24</td>
<td>174</td>
<td>6.35</td>
</tr>
<tr>
<td>Right Thalamus (pulvinar)</td>
<td>15 -17 7</td>
<td>4330</td>
<td>10.15</td>
</tr>
<tr>
<td>Left Thalamus (pulvinar)</td>
<td>-12 -21 9</td>
<td>2850</td>
<td>8.55</td>
</tr>
<tr>
<td>Left Putamen</td>
<td>-24 10 3</td>
<td>156</td>
<td>6.38</td>
</tr>
</tbody>
</table>
of items in the categories than the items used in the semantic fluency tests outside the scanner. For example, the category of “Ocean Animals” (used in the scanner) has many fewer members than “Animals” (used outside of the scanner), thereby increasing the difficulty of the task and increasing our ability to detect group differences during the scanning session. Analysis of the imaging data identified various discrete brain regions where GWS veterans’ activity differed from that of the control veteran group. As expected, many of our significant findings were in regions of the basal ganglia. In left anterior putamen, all the syndrome groups exhibited less activity than the control group. This area of the putamen has been shown to connect via white matter pathways to Broca’s area (Ford et al., 2013), suggesting that activity in this structure is associated with the production of verbal output. Group differences between Syndrome 2 veterans and controls were also evident in the left caudate head at a FDR of .10, with Syndrome 2 veterans exhibiting minimal activation while the other syndrome groups and controls exhibited a strong response. This finding did not survive more conservative statistical thresholding, but is important because of its clinical and cognitive implications for Syndrome 2 veterans. Most notably, it is suggestive of caudate dysfunction during verbal fluency tasks in Syndrome 2 veterans. In our experience, the amygdala is not commonly a site of increased activity during word fluency; thus, its decrease in activity relative to baseline is difficult to interpret relative to the semantic fluency task. However, given the role of the amygdala in the limbic system, the relative decrease in right amygdala activity for Syndrome 2 veterans may be related to the changes in the psychological functions they experience (Hom et al., 1997). In addition, the right and left thalami both exhibited less activation for all three syndrome groups than for control veterans.

For the left thalamus, the differences were significant for the control–Syndrome 2 comparison only; whereas, in the right thalamus, the differences were significant for the control–Syndrome 2 comparison and the control–Syndrome 3 comparison. Our left and right thalamic activity differences are generally confined to the pulvinar, which is known to play a role in modulating attention as part of its participation in dorsal and ventral attentional networks (Benarroch, 2015) but also receives robust connections from language eloquent cortices (Parent, 1996). These robust thalamic findings are consistent with the selective engagement theory proposed by Nadeau and Crosson (1997), and further developed by Crosson (2013), which states that the thalamus is involved in word generation through the selective engagement of cortical resources needed for lexical retrieval when lexical retrieval is based on semantic input.

Given that semantic fluency was the behavioral task used in scanning, another unexpected finding was that Syndrome 2 and Syndrome 3 veterans exhibited significantly greater task-related activation in the right hippocampus than control veterans. This finding is consistent with the aforementioned SPECT and ASL studies which showed that Syndrome 2 and Syndrome 3 veterans of this sample exhibited abnormalities within the cholinergic system of the right (as well as left) hippocampus (Haley et al., 2009; Li et al., 2011; Liu et al., 2011), with a pattern strikingly similar to that observed for the group fMRI activations in right hippocampus in Fig. 2. This might result from loss of the normal inhibitory muscarinic-2 cholinergic receptor function that has been shown to be impaired chronically after low-level sarin exposure at levels to which the subjects in our study were exposed in the Gulf War (Allon et al., 2011). It also is possible that the Syndrome 2 hippocampal activation is a downstream effect of inadequate attentional and inhibitory processes. Compared to control veterans who can appropriately engage in controlled, search-driven semantic retrieval, GWS veterans may have difficulty inhibiting prepotent and task-irrelevant responses elicited by the category cue. The hippocampal activation may reflect engagement of long-term memory during such automatic, bottom-up semantic retrieval processes. Given that four of our Syndrome 2 veterans carried diagnoses of PTSD at the time of the study, one might be inclined to interpret the increased hippocampal activation as being representative of active PTSD symptomatology. However, if this were true then we would have expected to detect hyperactivity of the amygdala of our Syndrome 2 group as well. Indeed, this is the opposite of what we found, and Syndrome 2 veterans had reduced amygdala activation compared to controls.

The regional activation differences observed in our present study are unrelated to the volumetric characteristics of our regions of interest, and reflect differential BOLD activation patterns among groups in the absence of gray-matter differences. Previous work with this sample of Gulf War Veterans confirmed this, as it revealed no significant regional volume differences in gray matter between syndrome groups and controls (Li et al., 2011) using whole-brain voxel-based morphometry.

Overall, the imaging results further confirm that Syndrome 2 veterans show evidence of brain changes in the left basal ganglia, right amygdala, and right hippocampus when compared to healthy control veterans. In addition, this analysis revealed that brain changes exist in the right and left thalamus of veterans with Gulf War syndromes. Hart and colleagues recently documented thalamic abnormalities in a Gulf War syndrome population as well (Calley et al., 2010), adding credence to this particular finding and to reports of thalamic blood flow abnormalities in ill Gulf War veterans (Haley et al., 2009; Liu et al., 2011). Considered together, these thalamic findings invoke important questions regarding the role of the thalamus in Gulf War syndromes. Although beyond the scope of this study, future investigations of the GWSSs should take a network approach to determine whether the brain changes noted in the basal ganglia and thalamus are related or distinct. Our current hypotheses regarding the interrelationship of basal ganglia loops and verbal fluency additionally lend themselves well to connectivity analyses. Given the network-basis of our interpretation, it would be wise to employ functional connectivity methods to explore the specific network-effects of our findings as a next step. To determine whether these basal ganglia findings reflect a unique pathology, future investigations should also compare Gulf War syndrome veterans to patients with basal ganglia diseases and damage of known mechanism.

While the brain changes observed in this study are most robust for the Syndrome 2 veteran group, it is important to note that the other syndrome groups also demonstrate changes in the basal ganglia, hippocampus, and thalamus. Most notably, Syndrome 3 veterans exhibit task-related activity changes that resemble those of Syndrome 2 veterans but of lesser magnitude. Interestingly, and perhaps not coincidentally, a progressive increase in relative risk of Syndrome 2 and Syndrome 3 Gulf War illness is associated with increased number of chemical agent alarms heard by veterans deployed in northern Saudi Arabia immediately following bombing raids on chemical warfare plants in central Iraq in January of 1991 (Haley & Tuite, 2013). Evidence demonstrating long-distance transit of fallout from the bombed chemical warfare plant sites to U.S. troop locations in northern Saudi Arabia has recently been presented (Tuite and Haley, 1991). From these two papers, the neurocognitive and fMRI data of this paper, and the recently reported abnormalities in subcortical blood flow, especially in hippocampus, for Syndromes 2 and 3 (Haley et al., 2009; Li, Spence, Buhner, Haley, & Briggs, 2012; Li et al., 2011; Liu et al., 2011), a preponderance of evidence suggests that anticholinesterases were the likely cause of these syndromes.

The mechanism of Gulf War Syndrome is still unknown, but this study’s findings help to emphasize the importance of basal ganglia, thalamus, and hippocampus in Gulf War Illness. Our ability to
understand the precise mechanism of Gulf War syndromes depends on clarifying which distinct cognitive functions are impaired in these veterans, what specific neural systems are involved, and how.

Acknowledgments

Study supported by United States Department of Veterans Affairs contract number VA549-P-0027 awarded and administered by the Dallas, TX VA Medical Center; United States DOD grant DAMD 17-01-1-0741; and United States NIH grants UI1RR024982 and P50NS23352. The content of this paper does not necessarily reflect the position or the policy of the U.S. government. No official endorsement should be inferred.

References


