Treatment of Memory Disorders in Gulf War Illness with high Definition transcranial Direct Cortical Stimulation (GWI HDtDCS)

NCT number: pending

Secondary ID: CDMRP-GW150060

October 6, 2017

BACKGROUND

An estimated 25-32% (175,000 to 210,000) of US military personnel deployed to the theater of operations in Southwest Asia during the 1990-1991 Gulf War (GW) continue to experience a wide array of adverse health problems, including headache, chronic and widespread pain, fatigue, gastrointestinal problems, respiratory problems, sleep problems, and cognitive difficulties, including memory problems.[1,2,3] Word retrieval deficits, or trouble with finding or "remembering" words one wants to say when speaking, are among the cognitive difficulties reported by deployed GW veterans with possible Gulf War Illness (GWI) and are included in the symptom-report case definitions of GWI.[1-7] About 24% of deployed GW veterans (GWV), compared to 6-8% of non-deployed controls (NDCs), have reported word retrieval problems, with odds ratio estimates of OR = 2.76-6.39.[1, 2] Word retrieval is important in verbal fluency (e.g., smoothness or flow with which sounds, syllables, words and phrases are joined together when speaking quickly) and verbal episodic memory (e.g., memory of autobiographical events – times, places, associated emotions, and other contextual who, what, when, where, why knowledge – that can be explicitly *stated*), and may contribute to other reported GWI symptoms, such as deficits in thinking and reasoning.[5, 7]

The cognitive tests typically used to probe word retrieval via verbal fluency require the production of words that begin with a given letter, that fit a category, or that mediate the association between words. These verbal fluency tasks have been administered to GWV suspected of or identified as having Gulf War Illness (GWI) and have been found useful in identifying those with word retrieval deficits.[8, 9] Retrieval processes also are required for successful retrieval of episodic memories, and those with GWI and memory complaints also show impairments on tests of verbal episodic memory.[1, 3, 6] Lesions associated with impairments on these word retrieval tasks have also been reported in those with GWI, including the frontal lobes,[22-24] temporal lobes,[24] the thalamus,[25] and the caudate nuclei.[26]

Studies reporting symptom prevalence in deployed GWV have led to the proposal of several case definitions.[1-3, 5, 10-12] The Center for Disease Control and Prevention Chronic Multisymptom Illness case definition requires the presence, for 6 months or longer, of at least one or more symptoms from at least two symptom clusters, which include general fatigue, mood and cognitive abnormalities (including word retrieval/word finding), and musculoskeletal pain.[1, 3] The Kansas-Steele case definition requires patients to have multiple and/or moderate to severe chronic symptoms in at least three of six defined symptom domains, and qualifying symptoms must have first been a problem during or after the Gulf War and must have persisted over a 6-month period.[2, 4] Symptoms include a) fatigue/sleep problems, b) somatic pain, c) neuro-logic/cognitive/mood symptoms (including word retrieval/word finding), d) gastrointestinal symptoms, e) respiratory symptoms, and f) skin abnormalities. A third major classification schema, comprising the Haley Syndromes[5-7, 12] (detailed below), has also been utilized to group patients based on symptom report. None of these classification schemas have clearly delineated specific mechanisms that underlie the observed phenotype.

Our group has published a specific study on verbal fluency (see Preliminary Studies below) in GWV with possible GWI based on the Haley case definitions: impaired cognition (Haley Syndrome 1: HS1), confusion-ataxia (Haley Syndrome 2: HS2), and central pain (Haley Syndrome 3: HS3).[5-7] The study showed that on standardized neuropsychological measures of verbal fluency, HS2s and HS3s exhibited impaired verbal fluency compared to case-controls.[9] Additionally, HS2s showed verbal fluency-related functional and anatomical deficits within the left putamen, in both thalamic hemispheres, and in the right amygdala.

All of these tasks require semantic memory retrieval and associated semantic processes, including selection and inhibitory control.[13, 14] We have also developed the Neural Hybrid Model of semantic memory[15-19] and subsequent word retrieval. In this model, we have proposed roles for pre-supplementary motor area (preSMA), caudate, and thalamus in semantic memory and word retrieval.[19] In our model, signals from cortical systems coding for features, categories, and other semantic properties of objects feed into the retrieval circuit via the thalamus. PreSMA mediates comparator/switching processes and provides input into caudate, which facilitates selection, sending its signal to the cortex via the thalamus functioning as a relay. In our model, retrieval emerges from processing of widely distributed cortical signals sent via the thalamic relay to preSMA, which in turn facilitates transmission to the caudate for selection, which then signals the thalamus to synchronize widely distributed cortical regions that store information about the target object. This results in sharpening the focus on the target object, and in the retrieval of an integrated memory representation, including retrieval of verbal components associated with the target object.

One of the studies that were used to establish this model was based on the Semantic Object Retrieval Test (SORT: judging whether pairs of words led to the retrieval of an associated object or not). We administered this test while we gathered fMRI data in a group of GWI patients with word finding difficulties. We found that HS2s and HS3s were impaired on the SORT compared to case-controls, and that performance for the HS2s was differentially related to thalamic and caudate function compared to case-controls and the other two syndrome groups. This demonstrates how semantic memory deficits, due to disruption of regions in the semantic retrieval circuit, can contribute to deficiencies in word finding/retrieval.

Guided by the model and our findings in GWI patients, we propose to examine the effects of transcranial direct-current stimulation (tDCS) on preSMA to improve semantic retrieval processes that support word finding/retrieval. TDCS appears to be a viable alternative or adjunct treatment option for cognitive dysfunction.[20] The application of tDCS has resulted in improved language abilities, including verbal fluency[21, 22], in aphasic patients and healthy controls after delivery of stimulation to frontal and temporal lobe regions.[23]. TDCS has been has been shown to lead to relatively localized metabolic change[24] and change in neuronal resting membrane potentials mediated by changes in sodium and calcium dependent channel function.[25] TDCS has been proposed to increase the neuronal firing rate leading to changes in long-term potentiation and depression[20] and purportedly leads to long-term enhancement of signal transmission between neurons that can strengthen neuronal connections, and subsequently strengthening the retrieval circuit. Thus, we will use tDCS to improve function in a dysfunctional semantic retrieval circuit in order to treat word finding deficits in patients with GWI.

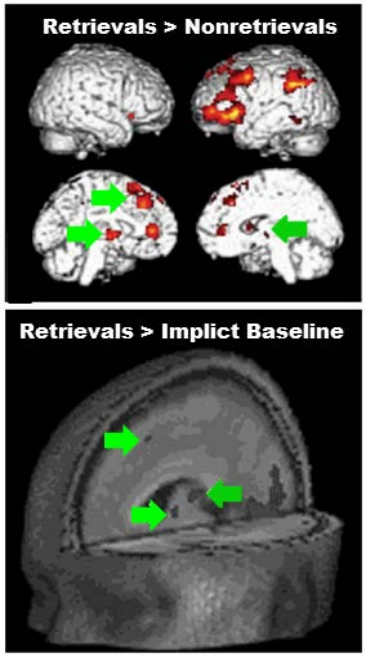

Figure 1. Regions showing fMRI BOLD change as a function of successful memory retrieval on the SORT. Green arrows indicate preSMA, caudate, and thalamus regions showing retrieval-related BOLD signal change. Upper image shows regions where retrievals led to greater BOLD signal change than non-retrievals, and lower images shows regions, in another study, where retrievals led to elevated BOLD signal change from the implicit baseline (i.e., "rest").

While we plan to stimulate the preSMA, we predict that the stimulation will affect the entire retrieval circuit. Based on our Neural Hybrid Model in which the preSMA mediates comparator/switching processes in word retrieval, applying tDCS at preSMA has the potential to strengthen neuronal connections disrupted by pathology within the preSMA, caudate and thalamus, and to improve the functioning of the retrieval circuit[19]. In turn, we predict improvements in verbal fluency, word finding problems, and potentially verbal episodic memory (see Preliminary Studies).

Preliminary Studies

Our group has delineated a network of brain regions that is engaged in verbal retrieval for multiple memory operations. The retrieval circuit consists of a thalamo-cortical (preSMA)-caudate network as part of the Neural Hybrid Model of semantic memory.[19] Under the model that we have advanced, preSMA-caudate-thalamic interactions govern processes fundamental to semantic retrieval of an integrated object memory, including its verbal components. Aspects of this model were derived from studies of the Semantic Object Retrieval Test (SORT) which consists of deciding whether two words that are features of an object (e.g., "desert" and "humps") combine to elicit the retrieval of a single object (e.g., "camel").[26, 27; also see supporting verbal fluency findings by 28, 29]

Preliminary studies on SORT in normal controls: Participants in these studies viewed word-pairs on each trial and judged whether the pairs triggered the retrieval of a target object. Across a series of studies, successful retrievals have been found to elicit retrieval-related BOLD signal within preSMA, caudate nuclei, and thalamus, in addition to activation in ventral visual-object areas (Figure 1).[15, 16, 30]

We also have administered the SORT to a patient with thalamic depth electrodes and a limited number of scalp electrodes. For this patient, showing word pairs that triggered an object retrieval led to a significant increase in EEG 20-30 Hz power in the thalamic depth electrodes and the occipital scalp electrodes.[31] This finding led to our proposing that beta frequency rhythms signify the selection of the correct memory in the retrieval process. That study was followed with recording EEG from scalp electrodes in a group of normal controls while they performed the SORT. First we noted an increase in global theta-band EEG power within the preSMA that was present from the beginning to the end of the task. This was proposed to reflect mediation

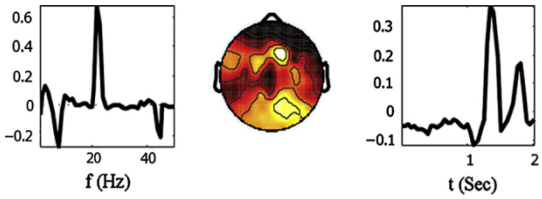

Figure 2. SORT EEG Power Spectra. Middle figure shows where (bright region over preSMA) 20-30 Hz beta-power (left) increases approximately1200 ms into (right figure) into the retrieval trials compared to non-retrieval trials (average RT ~1500 msec).

of the initiation and maintenance of the memory search. We also found that when a memory is successfully retrieved, there is a significant increase in 20-30 Hz beta power signal overlying the preSMA (Figure 2).[19] This suggests that the increased beta power interaction (~1200 msec) between the preSMA and thalamus late in the trial (avg. RT ~1500 msec) marks the selection of the correct memory.[32]

We also found that before the beta power increase there is a clear separation between retrieval and nonretrieval event-related potentials (ERP) from the left fronto-temporal area at 750 msec (Figure 3),[33] which we hypothesized represents the strengthening of verbal semantic associations for correct retrievals. This separation between retrievals and nonretrievals is evident at the individual subject level.[see Figure 4 in 33]

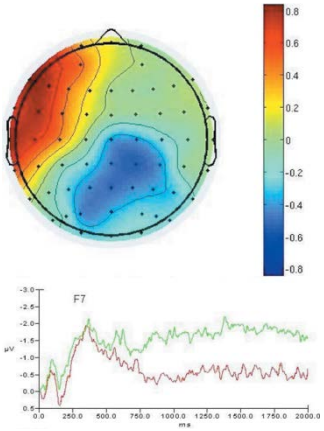

Figure 3. SORT EEG. Red area designates the ERP maximal intensity in the left fronto-temporal region (upper panel). The green line designates retrieval and the red line nonretrievals of the ERP in the lower panel. The separation is at approx. 750 msec.

To investigate the role of the caudate in this retrieval circuit, we administered a series of scalp EEGs during Semantic Selection Go/NoGo Tasks that vary in semantic complexity.[34, 35] In the simplest condition, the stimuli are a line drawing of one car (80% of the stimuli) and another drawing of one dog (20%); the task requires the subject to push a response button upon seeing the drawing of the car, but to withhold button-push responses when presented with the depiction of the dog. The middle level of difficulty is the same task and proportion of stimulus types, but the stimuli include multiple exemplars of cars and dogs. The semantically most complex task consists of 20% of the stimuli being drawings of various animals and 80% drawings of a variety of objects, with the task being to push the response button upon seeing a depiction of an inanimate object, but to withhold button pushes for animals. We found that the amplitude of the NoGo P3 (a positive polarity ERP component with a peak at close to 300 msec after stimulus onset) was attenuated with increased semantic complexity across all three tasks (see Figure 5, upper plot), where the blue wave is the highest amplitude for the semantically simplest task, the green wave in the middle of the group for the middle complexity task, and the red wave is the lowest amplitude for the most semantically complex task (Figure 4, upper plot).[34] Also, for the most semantically complex task (e.g., "don't push for animals,"), the P3 component is significantly delayed compared to the other two conditions. These ERPs were all maximal in amplitude over the preSMA region. In addition, we detected variations in theta power with selection or inhibition of correct or incorrect memory retrievals, respectively, and these theta power signals varied with semantic complexity— the more semantically complex tasks yielded lower theta power than the more simple tasks for both the retrievals and nonretrievals (Figure 4, lower graph).[35] These findings suggest that the caudate facilitates activation of correct and suppresses activation of incorrect target memories based on input from preSMA and from signaling sent to thalamus.

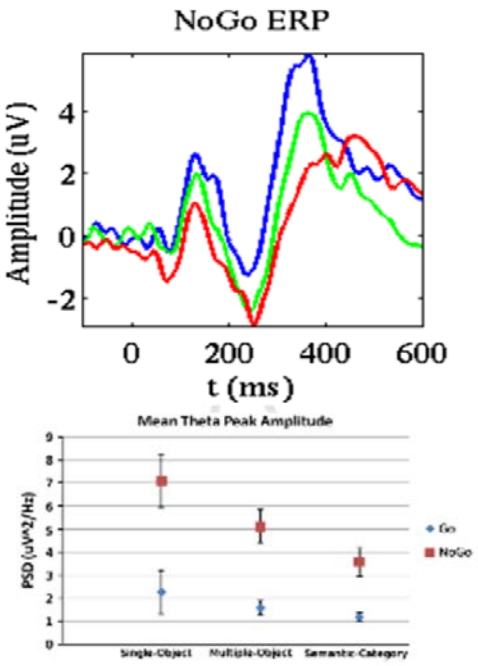

Figure 4. EEG signal changes related to inhibiting responses under varying semantic processing demands. Upper panel shows ERP amplitude change on successful no-go trials under simple (blue), moderate (green), and complex (red) semantic processing conditions. Lower paned shows EEG theta power changes under simple (left), moderate (center), and complex (right) semantic processing conditions for no-go (red) and go (blue) trials.

The scalp EEG findings from the Semantic Selection Tasks yielded clear P3 Go/NoGo amplitude differences at preSMA, where the amplitude is greater with the semantically shallow (single dog, single car) processing condition. We posit that when the prepotent response is strong (semantically simpler stimuli that rely on perceptual features), a stronger inhibitory signal is required to actually stop the response (i.e., larger P3 with shallower semantic processing requirements).

Our data support the idea that when the correct memory is

retrieved, high beta-band power increases between the preSMA and thalamus, and subsequently more widely distributed cortical regions, and reflect the retrieval of an integrated object representation from multiple semantic memory subsystems, including the verbal information about that object. The caudate engages in the selection of correct and inhibition of incorrect retrievals by modulating the circuit via theta power increases. These changes were detected by fMRI during SORT task performance to localize the regions involved in the circuit, and the sources of EEG power changes over time were derived from EEG collected during the performance of semantic retrieval tasks (SORT and Semantic Selection Tasks).[19] These markers are only measurable as a group effect. However, the SORT and Semantic Selection Tasks ERP measures demonstrate robust effects for both group and individual subjects.

Preliminary studies on retrieval deficits in GWV: As described above, behavioral and fMRI data were collected on members of the 24th Reserve Naval Mobile Battalion (Seabees) deployed during the Gulf War. Subgroups of the GWVs from the Seabees (Syndrome 1 n = 11; Syndrome 2 n = 16; Syndrome 3 n = 9; case-controls n = 14) were compared on standardized measures of verbal fluency from the DKEFS battery.[36] The HS2s generated fewer words in both the letter and category fluency tasks than did the case-controls, and the HS3s generated fewer words in the letter fluency task than did the case-controls.[9]

Functional MRI data also were collected on the subgroups of Seabees while they completed SORT (Syndrome 1 n = 11; Syndrome 2 n = 16; Syndrome 3 n = 16; deployed controls n = 16).[8] We detected significant performance impairment in the HS2 group on the SORT test compared to the controls and other HSs in terms of having decreased accuracy and longer reaction times (RT). Functional MRI using the SORT test revealed a significant increase in BOLD signal in the thalamus for correct responses in HS2 compared to the other groups and controls. In both the thalamus and caudate, the HS2 group had increased signal change with increased RTs, which was significantly different than that of the controls and other syndrome groups as these groups showed decreased signal change with increased RT. These findings show that the retrieval circuit for HS2s was dys-functional, given the behavioral performance and BOLD imaging parameters for the group of participants.

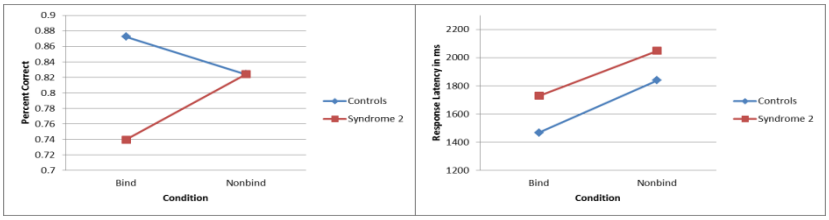

Figure 5. Behavioral data from the SORT task. Left panel: An interaction between condition and group (p = .023) on percent correct. The main effect of group was due to Syndrome 2's poorer performance in the bind condition. Right panel: Main effects of group and condition in response times.

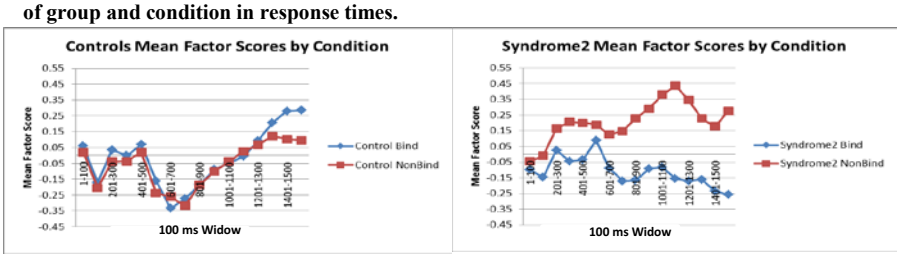

Figure 6. Interaction between condition (retrieval, nonretrieval) and group (Controls, HS2). Whereas controls' scores for retrievals and nonretrievals are similar, HS2 factor scores for the nonretrieval condition are strongly positive, especially after 600 ms, and scores for the bind condition are negative.

Preliminary Study #2. In addition to studies of the Seabee group, we conducted an ERP study using the SORT in a national sample[12] of selected individuals with possible GWI. In this study, the HS2 group (n=22) patients made significantly more errors on retrievals (Figure 5, left panel) and had longer reaction times than did 28 deployed normal controls (Figure 5, right panel). In addition, the typical ERP difference between the retrieval and nonretrievals at 750 msec (see Figure 4 for typical ERP) was present for the control group and HS groups in the left frontotemporal area. However, in the right centroparietal region in the HS2 group showed a separation of the retrieval and nonretrieval signals at 750 ms (Figure 6, left panel) that was not present in the control (Figure 6, right panel). Analysis of the ERPs for the controls and HS2s using a spatial principal components analysis yielded four components, with one related to a group x stimulus type x time point interaction. The variance explained by this factor showed a significant stimulus type x group interaction, $F(1, 48) = 4.244$, $MS_e = 3.11$, $p = .045$. 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 Figure 7. Whereas controls' scores for retrievals and nonretrievals are similar (Figure 6, left panel), HS2 factor scores for the nonretrieval condition are strongly divergent, especially

after 500 msec (Figure 6, right panel). This provides a clear set of behavioral and ERP markers for a targeted, objective treatment response.

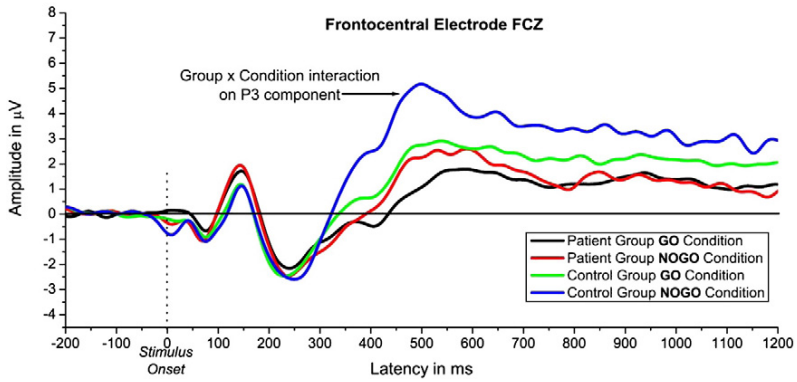

Figure 7. The amplitude of the P3 component, whose maxima were at frontocentral midline electrode showed interaction (p=.0137) wherein the difference between the Go and NoGo P3 amplitudes from the control group with greater difference between the Go and NoGo P3 amplitudes from the patient group compared to controls.

Preliminary Study #3. A group of 25 Gulf War veterans who complained of cognitive difficulties and 23 matched controls, who were deployed but not symptomatic, performed the Semantic Selection Tasks. In the more semantically complex task, the cognitive complaint group had significantly greater false-alarm rates in the complex semantic Go/NoGo task and this was accompanied by ERPs whose NoGo P3 was significantly reduced in amplitude compared to the control group (Figure 7). This inability to inhibit an incorrect retrieval, marked by increased false positive responses and a decrease in P3 NoGo amplitude, will be used in this proposed study as one marker of a dysfunctional retrieval circuit.[37]

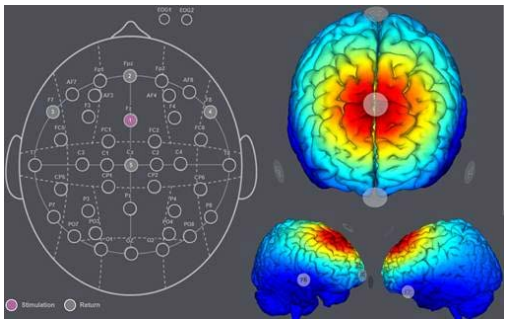

Figure 8. Electrode locations on the HD tDCS cap (left) and the projected field if the electrical stimulation mapped onto the brain (right). In this study, the midline frontal electrode (purple) is the stimulating electrode and the solid gray ones are the return electrodes. The maps on the right show the electrical stimulation intensity mapped to the brain with red being most intense and blue least.

Preliminary Studies with HD tDCS administration for verbal retrieval deficits: Our group also has experience using electromagnetic modulation therapies to improve cognitive and behavioral function in Post-Traumatic Stress Disorder (DOD W81XWH-09-DMRDP-ARATDA) and with Multiple Sclerosis (MS). Preliminary data from our current study of MS have been collected on participants with relapsing-remitting MS. MS patients and healthy controls have undergone neuropsychological baseline testing and then received ten sessions of high-definition (HD) tDCS delivered to the preSMA region at 1 mA for 20 min each session, as we propose to use in this project (Figure 8). Following completion of the treatment sessions, the patients have shown improvement in verbal fluency (with significant improvement in SORT RT, $t[3]=6.47$, $p=.003$), verbal memory, selection and inhibition (with significant improvement in Go RT, $t[3]=4.19$, $p=.01$), and executive function. One of the main contributors to neurological deficits in MS is decreased effectiveness of inter-regional connections due to white matter pathology, but it appears that tDCS-mediated changes in primarily neuronal, as opposed to axonal, activity, has been effective in overcoming the effects of white matter damage. Similarly, while white matter abnormalities have been hypothesized to be an important mediator of the neurological deficits in GWI[see 38], we believe that our therapeutic approach, will also be effective in mitigating the deleterious neurological effects of GWI.

We hypothesize that increased activation of this circuit via HD tDCS stimulation of the preSMA can lead to improved verbal retrieval. We also hypothesize that open-loop up-regulation of this circuit via stimulation of the preSMA will improve verbal retrieval that will be reflected in verbal retrieval-related tasks. We also predict that we will also see changes in SORT and/or Semantic Selection Tasks ERP measures that we will be able to use to assess outcomes of the treatment.

OBJECTIVES

Objective #1: Screen 120 Gulf War Illness patients (e.g., Haley Syndrome 2, but also including all others with subjective word-retrieval complaints) who report having word retrieval difficulties. These patients will be recruited from a) Gulf War veterans referred to Dr. Robert Haley at UTSW for assessment but have not been in the previous Seabee or National Sample studies, b) Dallas VA (Stuve), and c) local advertising.

Objective #2: A subject that passes the phone screen for having a word-finding issue on the Controlled Oral Word Association Test (COWAT) will be enrolled, consented, and screened in-person for having a) a Gulf War-related Illness, b) no significant concurrent medical illness that could account for a retrieval deficit (e. g., CVA, dementia, Parkinson's disease), and c) objective evidence of a word-retrieval impairment on neuropsychological testing. Subjects will then be evaluated with the SORT clinical test, COWAT, Boston Naming Test (BNT), California Verbal Learning Test (CVLT), and focused neuropsychological battery. These measures will serve as outcome measures for the treatment intervention.

Objective #3: If the subject demonstrates a clinically significant deficit (> 2 sd below the mean for age and education) on any of the 4 neuropsychological measures engaging word retrieval, they will be enrolled in the treatment phase of the study. They will then be classified using the CDC criteria for typing of GWI.

Objective #4: Each subject will undergo a detailed neuropsychological battery to characterize his or her cognitive and behavioral status—including targeted tests that should be affected by improved verbal retrieval, tests of cognitive functions to which the treatment may generalize, and tests consisting of purportedly unrelated functions to account for global, generalized cognitive improvement. They will also undergo pre-treatment EEG (SORT task with EEG, Semantic Selection task with EEG). These will be used as outcome measures for the treatment trial. The two primary outcome measures will be the COWAT and CVLT. Secondary outcome measures will be SORT and Semantic Selection task, performance and ERP for both sets of tasks.

Objective #5: These subjects will then be enrolled in the active treatment aspect of the study. Enrollment of 80 participants is the target for enrollment in the active phase of the treatment trial. Half the subjects will first receive HD tDCS at 1 mA for 20 minutes over the preSMA region for 10 daily sessions (1 session 5 days per week for a 2 week period). The remaining subjects will initially receive sham HD tDCS; after all follow-up evaluations are completed, they will receive the active HD tDCS treatment.

Objective #6: Each subject will receive follow-up neuropsychological and EEG testing at 1 week, 3 months, and 6 months after treatment. This last time point for evaluation extends from our preliminary work showing that later, further improvement in electromodulation treatments can occur at 6 months post-treatment. HD tDCS active treatment subjects enrolled early in the study will be followed longer than 6 months if agreeable to assess length of effect.

Following completion of this clinical trial, we will be able to determine whether HD tDCS is an effective treatment for the verbal retrieval deficits associated with GWI. An HD tDCS cap is a relatively inexpensive piece of equipment that can be re-used many times for multiple patient treatments. These caps are readily available, and the application of the treatment regimen is straightforward to administer. Thus, this treatment can be utilized in any VA medical center or clinic around the nation. In addition, the neuro-psychological tests and ERP/time frequency findings may function as clinical markers for diagnosis and treatment efficacy in subsequent treatment trials to measure treatment response.

SPECIFIC AIMS

The goal of this proposal is to entrain the verbal retrieval circuit using HD tDCS to target the preSMA in order to improve retrieval in impaired GWI patients, using behavioral and electrophysiological markers to assess outcomes. After identifying GWI patients with retrieval deficits using standard neuropsychological measures, we will address the following aims:

Aim 1. <u>Administer 10 20-minute sessions of either 1 mA high definition tDCS or sham to the preSMA region.</u> Subtypes of Gulf War Illness patients have been shown to have dysfunction that disrupts the above-described retrieval circuit.[8] Transcranial DCS targeted to the preSMA has been shown in preliminary studies in other diseased populations to result in improved performance in tasks that engage this circuit by purportedly strengthening potentially weak connections within the preSMA-caudate-thalamus circuit for verbal retrieval.

Hypothesis #1:HD tDCS to the preSMA will produce long-term modulation of the basal state of the preSMA-caudate-thalamic-cortical retrieval circuit.

Aim 2. <u>Assess neuropsychological and electrophysiological (ERP) markers of verbal retrieval following tDCS therapy.</u> Long-term modulation of the preSMA-caudate-thalamic-cortical retrieval circuit, via tDCS to the

preSMA region, will alter the neurophysiological properties of the circuit due to modulation of the readiness of the retrieval circuit. This will result in changes in neural markers as follows:

Hypothesis #2a:Neuropsychological measures of verbal retrieval that were impaired at baseline will be improved to expected premorbid levels of functioning for the treated individuals.

Hypothesis #2b: Patients with impairments in making a correct retrieval will exhibit both high false negative errors and false positive errors compared to norms[27] and a decreased 750-ms ERP amplitude difference between retrievals and nonretrievals on the SORT task. *We hypothesize that after HD tDCS treatment, these patients will improve performance and/or have an increase in amplitude difference in the ERP between retrievals and nonretrievals.*

Hypothesis #2c: An increase in false negative errors and a decreased "Go"P3 ERP response on the Semantic Selection Tasks is indicative of an impaired ability to select a correct memory for retrieval.[33] Conversely, there will be increased false positive errors and a decreased "NoGo" P3 ERP on the Semantic Selection Tasks if a patient cannot inhibit an incorrect memory. *We hypothesize that following HD tDCS therapy, there will be a normalization of the patient's behavior (reduced errors, improved RTs) and/or ERP responses (increased P3 amplitudes for Go and/or NoGo trials) due to modulation of the retrieval circuit, with the changes based on the pre-treatment impaired retrieval function (selection and/or inhibition).*

CLINICAL TRIAL

A. General Experimental Approach.

1. We will screen 120 GWI patients who report subjective word-retrieval deficits over the phone with the COWAT.

2. A subject who on the phone screen has a subjective complaint and decrement in COWAT performance will then be consented and enrolled in person at CBH, and then screened to determine that they have GWI and objective evidence of a verbal retrieval impairment on neuropsychological testing.

3. If a subject scores > 2 SD below the mean on any of the following – the SORT clinical test, COWAT, BNT, and/or CAVLT, the subject will be enrolled in the treatment phase of the study and classified using the CDC criteria for subtype of Gulf War Illness.

4. Subjects will undergo a pre-treatment battery of detailed neuropsychological tests and EEG (SORT and Semantic Selection tasks with ERP). These will be used as outcome measures for the treatment trial.

5. We will enroll 80 subjects in treatment with one-half of the subjects receiving HD tDCS at 1 mA for 20 minutes over the preSMA region for 10 daily sessions; the other half of the subjects will initially receive sham HD tDCS and, after all follow-up evaluations are completed, they will receive active HD tDCS treatment. Participants will be randomly assigned to stimulation and sham conditions via algorithm.

6. Determine at 1 week, 3 months, and 6 months post-last-treatment whether the therapy has been effective by repeating the outcome measures (parallel versions when available) of the SORT, COWAT, BNT, CVLT, and ERP tasks.

B. Subject Recruitment

Stage 1. Interview. We will phone screen approximately 120 Gulf War veterans (to obtain the targeted 80 subjects) with symptoms they attribute to service in the 1991 Persian Gulf War who have sought care at UT Southwestern Medical Center (Dr. R. Haley) or the Dallas VA Medical Center Neurology Clinic (Dr. O. Stuve) for GWI. In addition, we have enrolled two large (>100 subjects enrolled after screening) veterans studies at the Center for BrainHealth using local advertisement and particularly our social networking public relations team. This team includes a combat veteran with extensive experience in using social media outlets to effectively recruit subjects for clinical trials.

The subjects will be screened initially by short telephone interview to determine whether they appear to have a Gulf War Illness, and will be briefly screened for exclusion criteria and the co-existing conditions listed below. The subjects will include men and women, ages 43 to 75 (ages 18 to 50 during the 1991 Gulf War) ofany race/ethnicity, and both enlisted and officer ranks. The potential subjects will then perform over the phone the Controlled Oral Word Association Test (COWAT) for three letters. Performance below 1 sd of the mean is the criterion, in our experience, to warrant further screening with more extensive testing for retrieval deficits (COWAT is a sensitive, brief measure of verbal retrieval). The telephone interview will be conducted

by a trained telephone interviewer. Those meeting the inclusion criteria will then be invited to participate in the second stage of the evaluation and be given an appointment for it.

Stage 2. Screening for a Retrieval Deficit. Those meeting the inclusion criteria will visit the UTD Center for BrainHealth for the study to be fully explained and formal consent obtained by study personnel. Once consent is obtained, the subject will be screened for a retrieval deficit with the neuropsychological tests of the Semantic Object Retrieval Test (SORT), COWAT (both letter and category fluency), Boston Naming Test (BNT), and the California Verbal Learning Test (CVLT). If a subject scores > 2 SD below the mean on any of these tests, the subject will be deemed to have a retrieval deficit and will be enrolled in the treatment phase of the study. These subjects will then be classified under the CDC criteria for GWI subtype and screened for co-morbid conditions/exclusion criteria that could potentially obfuscate the results of a treatment response.

Stage 3. Pre-Treatment Testing. Those meeting criteria of having an established GWI and retrieval deficit will be scheduled for a neuropsychological test battery and EEG testing (SORT task with EEG, Semantic Selection task with EEG) as outlined under *Neuropsychological Measures and EEG Tasks* below. The two primary outcome measures will be the COWAT and CVLT. Secondary outcome measures will be SORT task behavioral and ERP measures, as well as Semantic Selection task performance and ERP measures.

Expected sample sizes. Our target sample of 80 Gulf War veterans who will complete the treatment phase of the study will afford us power to detect a statistically significant difference. Treatment-related change in verbal fluency and memory measures will be the primary outcome measures. Treatment-related change in verbal fluency and memory measures will be the primary outcome measures. A study of the effects of single-session tDCS applied to PFC on verbal fluency measures in normal controls showed a large effect-size, with standardized mean difference Cohen's $d = 1.18$.[20] Additionally, this treatment effect estimate on healthy controls is comparable to or exceeds the effect-size estimates for verbal fluency deficits in GWVs with GWI: Letter Fluency averaging over S1-3 Cohen's d = 0.69; S2 Cohen's $d = 1.12$, S3 Cohen's $d = 0.85$; Category Fluency averaging over S1-3 Cohen's d =.41; S2 Cohen's $d = 0.81$, S3 Cohen's $d = .61$.[11] For the treatment effect-size Cohen's $d = 1.18$[20], a sample size of 40 per group yields Power=0.92, and for the deficit effect-sizes for S2, Cohen's $d = 1.12$, and for S3 Cohen's $d = .61$, Power =0.94 and Power =0.89, respectively.

Exclusion Criteria. We will exclude non-English speakers because not all of the screening forms, questionnaires, and tests are available in any language except English. Other exclusion criteria are a history of a neurological disorder, including dementia of any type, moderate to severe traumatic brain injury (TBI), brain tumors, present or past drug abuse, stroke, blood vessel abnormalities in the brain, Parkinson's disease, Huntington's disease, or multiple sclerosis. Traumatic brain injury will be screened by history. No subjects will be enrolled who are cognitively or clinically incompetent to give informed consent. In addition, the patient cannot be taking medications that interact with the tDCS effect amphetamines, L-dopa, carbamazepine, sulpiride, pergolide, lorazepam, rivastigmine, dextromethorphan, D-cycloserine, flunarizine, ropinirole,or citalopram.[39]

C. Procedures

1. Neuropsychological Measures

a. The Semantic Object Retrieval Test (SORT; clinical paper and pencil test; Kraut et al., 2006, 2007) consists of 32 pairs of words that are features of objects. The task is to determine if the two feature words elicit retrieval of the memory of a unique object. The test comprises two components, one score for semantic memory retrieval and another for name production (word finding). The word pairs will be presented orally for the subject to provide a yes/no oral response based on whether the word pairs resulted in object memory retrieval and, if yes, to provide the name of the item retrieved. The nonretrieval word pairs have been constructed from the same words in the retrieval word pairs, but re-paired to not result in object retrieval. For example, the word pair "humps" and "desert" would elicit the memory of "camel" (retrieval pair). In contrast, "humps" and "monitor" would not elicit any object memory (nonretrieval pair). Previous studies have revealed SORT to be a sensitive measure of word-finding deficits in GWV[8], and SORT has been shown to involve preSMA, caudate, and

thalamus.[19] The SORT has been normed for age, test-retest reliability $r = .90$ for memory and $r = .93$ for correct names, $p < .001$, with test of independence for these measures significant at chi-squared (1) = 6.70, p = 0.01.

b. Controlled Oral Word Association Test (COWAT).[40] This is an assessment of word retrieval during which the patient is required to produce as many words in one minute that begin with either a specific letter (e.g., *F, A,* or *S*) or members of a specific category (e.g., *animals*, *fruit and vegetables*). Psychometrics: Cronbach's alpha = .83 for the number of items generated per condition; test-retest reliability $r = .74$; test-retest also has shown significant gains observed from first (M=39.7, SD=10.48) to second (M=42.5, SD=9.9) administration, $t(1, 119)=4.19$, $p<.001$.[41]

c. Boston Naming Test (BNT).[42] The test consists of 60 pictures of items where each is presented one at a time for the patient to provide the name of the item depicted. The task will be split into two equivalent 30 item tests (Cronbach alpha = .88; whole sample agreement between forms kappa = .93; Graves et al., 2004). The test assesses visuoperceptual processing, access to semantic knowledge and access to and production of the name of the object, which marked word retrieval from a visual input. Internal consistency ranges from $r = .78$ to .96 across studies. Test–retest stability in cognitively normal adults varies as a function of time interval and sample composition and ranges from $r = .59$ to .92.[43]

d. California Verbal Learning Test (CVLT-II).[44] The test is a list-learning task where the tester reads aloud a list, called "Monday's shopping list." The list contains sixteen common words, each of which belongs to one of four categories. The tester records how many items the patient remembers over several repeated trials and if the patient is making use of category information. There is also a delayed recall trial after 20 minutes. The test is an indirect measure of word retrieval in that successful word retrieval from semantic memory is needed to perform well in this episodic memory task. The CVLT-II has generally large test-retest correlation coefficients for the primary CVLT-II measures in both the standard/standard (range r = 0.80–0.84) and standard/alternate (range r = 0.61–0.73) cohorts, with reliability coefficients for primary measures ranging from 0.80 (e.g., total trials 1–5) to 0.84 (recognition discriminability).[45]

Parallel versions are available for all of the above neuropsychological tests. Additional neuropsychological measures included will be (organized by domain of cognitive function):

a) General Cognitive: American National Reading Test (AMNART) for estimated IQ.[46] This test is chosen to determine if changes from treatment could be associated with IQ and to compare groups that may have differential responses to treatment.

b) Attention and Cognitive Flexibility: Trail Making Test parts A and B[47]; Color Word Interference Test.[36]

1. For Trails A and B, Trails A measures processing speed and simple attention, and Trails B utilizes the same cognitive components as Trails A with the additional executive function of alternating sequencing. This test is not a direct measure of retrieval but is used to assess for generalization or indirect improvement with HD tDCS. Test-retest reliability is reported as r = 0.745 for Trails A, r = .849 for Trails B, and r = .971 for B-A.[48]

2. Delis-Kaplan Color Word Interference Test[36] measures processing speed in the control conditions of color naming and word reading with color word names written in a different ink color than the name measures inhibitory control, an executive function. This test is not a direct measure of retrieval but is used to assess for generalization with HD tDCS, particularly to compare to performance on the inhibitory aspects of the Semantic Selection Tasks. Internal consistency is reported as Cohen's alpha = .62-.86 over age groups from 9 to 89 years of age (Manual), and within in the age GWV, Cohen's alpha = .72-.86. Test-retest reliability is reported as Color Naming r = .76, Word Reading, r = .62, Inhibition r = .75, and Inhibition/ Selection r = .65, with lower test-retest reliability coefficients for 50-89 year olds, r = .50-.57.[36]

3. Digit Span Forwards and Backwards[49] measures simple attention with digit span forwards and working memory with digit span backwards. This is chosen not because it is to be directly affected by the HD tDCS, but to determine that verbal episodic memory increases are from improvements in attention (digits forward) and to determine if there is generalization to working memory (digits backward). Test-retest reliability coefficients for Digits forwards ranges from .66-.89 and Digits Backwards approximating 0.70.[50]

c) Visual Memory: Rey-Osterrieth Complex Figure Test (Copy and Memory components)[51] The task requires a subject to copy a complex geometric figure and then later to draw what aspects of the figure that they

still remember. It is included in this study as a check to determine if retrieval improvement is not only in verbal retrieval but also generalizes to nonverbal material retrieval. Reliability across alternate forms is reported as Copy r = .50, Immediate Recall r = .76, and Delayed Recall r = .69, and reliablility over one year is Copy r = .18, Immediate Recall r = .47, and Delayed Recall r = .592.[52]

d) Psychomotor Speed: Grooved Pegboard Test. This test of psychomotor speed utilizes placement of pegs into slots of varying orientations. It is included to assess the specificity of tDCS to retrieval versus generalization to processing speed and the possibility that improved processing speed might mediate improved retrieval. The test has good test-retest reliability (0.91 and 0.85 for right and left hands, respectively).[53, 54]

e) Beck Depression Inventory.[55] (Beck et al., 1996) is a subjective questionnaire of depressive symptomatology. The test is not being used as a marker of retrieval, but to determine if mood improvement that may be secondary to HD tDCS application may be accounting for minor improvements in retrieval.[54] Cronbach's average alpha coefficient average around 0.9, ranging from 0.83to 0.96, test-retest reliability ranges between 0.73 to 0.96.[56, 57]

2. EEG Tasks

a. Semantic Object Retrieval Test (EEG version).[33] The stimuli in this EEG paradigm comprise 112 pairs of words. There are 56 retrieval and 56 nonretrieval word pairs. The "retrieval" and "non-retrieval" attributes of the word pairs have been validated in previous studies with different groups of subjects. Words will be presented simultaneously one pair at a time for three seconds, with one word above the other. Between trials, a + sign will be presented for three seconds as a visual fixation target. Participants will be instructed to respond "yes" or "no" depending on whether they could think of any particular object upon seeing each word pair. Word stimuli will be presented on an LCD screen placed about 46 inches from the eyes using Stim2 software (Compumedics Neuroscan, USA). Responses will be made by pressing a button under the right index finger for *yes* responses and a button under the right middle finger for *no*. The entire task lasts about 11 minutes. The retrieval trials index successful access to an object memory represent-ation including verbal information, with the correct nonretrievals indicating a memory search that leads to a conclusion that the features do not combine to activate an object memory.

b. Semantic Selection Tasks.[34, 35, 58] This paradigm comprises three variants of the Go/NoGo task (200 trials each task) that require varying levels of semantic categorization: a) Single Car task (SC): The stimuli are a line drawing of a car (Go) and a line drawing of a dog (NoGo). Participants are required to not push a button for the dog (20% of trials) and to push for the car (80% of trials). b) Multiple Cars task (MC): This task includes 40 different exemplars of cars (Go) and 10 different exemplars of dogs (NoGo). The instructions for this task are identical to those in the SC task. Each drawing of the 40 cars and of the 10 dogs is presented 4 times over the course of the task. Successful response and inhibition is dependent on the ability to retrieve the concept based on abstraction and integration of salient perceptual features associated with each object, followed by object level identification. c) Object Animal task (OA): This task includes 160 different exemplars of objects (Go) (40 food items, 40 cars, 20 clothing items, 20 kitchen items, 20 body parts, and 20 tools) and 40 different exemplars of animals of varying visual typicality (NoGo). Participants are to push a button for items that are not animals. The OA task has the greatest semantic complexity, as correct categorization between multiple categories necessitates retrieval of object knowledge that relies on more than just perceptual features and a determination of category membership.[34] Overall, each task comprises 200 trials presented in random order, 160 (80%) of which are Go trials that require a button press and 40 (20%) of which are NoGo trials that require inhibition/withholding of a response. Each stimulus is presented for 300 ms followed by a 1700-ms blank fixation period. The Go trials indicate selection and retrieval of the correct object in semantic memory and the NoGo trials index inhibition of a response when retrieval of a memory does not lead to the search target.

3. EEG Procedure

Continuous EEG will be recorded from a 64-electrode elastic cap (Neuroscan Quickcap) while the participants are preforming the task, through a Neuroscan SynAmps2 amplifier and using Scan 4.5 software (Compumedics Neuroscan, USA; sampling rate: 1kHz, DC-200Hz). Electrode impedances will typically be below 10 kΩ. The reference electrode located at midline between Cz and CPz, and vertical electroocculogram (VEOG) will be

recorded at sites above and below left eye. Data will be processed off-line using scripts developed in our lab that implement functions from EEGLAB version 13[59] (http://www.sccn.ucsd.edu/eeglab) running under Matlab 7.11.0 (The MathWorks, Inc.). Preprocessing will consist of down-sampling to 512 Hz, removing data recorded from poorly functioning electrodes, and correcting for stereotyped artifacts including eye blinks, lateral eye movements, muscle, line noise, and heart rate using the "Runica" algorithm (with the *extended, 1* option)[59], an implementation of the logistic infomax independent component analysis algorithm.[60] Stereotyped artifacts will be identified by visual inspection of the spatial and temporal representation of the independent components. Continuous data will be segmented into 2-second non-overlapping epochs spanning from 500 ms before to 1800 ms after the presentation of the visual stimuli. Epochs containing high amplitude, high frequency muscle noise, and other irregular artifacts will be removed. Only trials to which the subject responded correctly and those without artifacts will be subjected to further analysis, retaining on average 75 percent of all epochs. We will estimate data for missing electrodes by interpolation. The interpolated data will be re-referenced to the average reference.[61]

D. HD tDCS Treatment Procedures

High-definition (HD)-tDCS uses arrays of smaller electrodes, combinations of which can be optimized for targeting. So far, HD-tDCS has been shown to reliably target specific brain areas and proved to produce plastic changes that may outlast conventional tDCS. [62]. A study comparing conventional tDCS with a HD-tDCS design using a set of small electrodes approximating the conventional set-up (covering the large pad-electrodes) found that the HD-tDCS approach achieved electrical fields with greater focality (80% improvement) and higher target intensity (98% improvement) at cortical targets using the same total current applied.[63]

Cap. The EEG cap is gently secured on the head of a subject and positioned with Cz at the vertex, as measured using surface anatomical landmarks, and defined as the intersection of the nasion–inion and interaural lines.

Electrodes. tDCS is delivered with Sintered Ag/AgCl disc electrodes using customized high-definition electrode with 8-mm diameter resulting in a ~85 $\pm$ 5 mm². The electrode were then filled with > 3 mL of conductive Gel into which the Ag/AgCl disc electrode was immersed. The cotton swab was used to adjust the gel and hair until the electrode resistance was less than 10 kΩ prior to stimulation. The anode was placed over Fz according to the International 10-20 EEG system, corresponding to the approximate location of the pre-SMA. Four return cathodal electrodes were placed approximately 5 cm radially from Fz, corresponding to locations FPz, F7, F8 and Cz (see Figure 8). The preparation procedure consisted of exposing the scalp by separating the hair underlying each electrode and adding approximately 1.5 mL of highly conductive gel. Contact quality and impedance levels were verified for each electrode before each stimulation session began.

Stimulation. Electrical stimulation is performed using a constant voltage Cortical Stimulator (Model D185, Digitimer Ltd, UK, maximal output 1000 V/1.5 A) with single square-wave 50-μs pulses (0.1 A/μs rise time). The adaptor measured resistance prior to stimulation, and divide current during stimulation. HD tDCS will be delivered in the active condition for 20 minutes of 1 milliamp high definition tDCS per session and administered for 10 total sessions on separate days (10 sessions over 10-14 days). This provides focal delivery of current maximally at the preSMA. There is no sensation associated with stimulation at this level, so that the sham condition will consist of placing the HD tDCS cap in place for the 10 sessions of 20 minutes each, but without delivering electrical stimulation. After each session, patients will be administered a standard post-treatment symptom questionnaire (e.g., assessing for skin irritation, fatigue).[64]

E. Potential Problems

1. There is the possibility that the HD tDCS will affect regions other than the preSMA. This will be assessed by reviewing changes in neuropsychological test measures.
2. There is also the possibility that stimulation will be ineffective because the target tissues have been too badly damaged. We will monitor for response patterns consistent with ineffective stimulation.
3. There could be difficulty in recruiting subjects, but we have not only identified the individuals referred to Dr. Haley's cohort as noted in the subjects section, but also have the availability to recruit subjects

through the Dallas VAMC, and we have been quite successful in recruiting for two ongoing veteran studies at our Center through local advertising.

4. Some individuals with Gulf-War–related complaints and impaired performance will have these difficulties from etiologies other than GWI. It is possible that individuals with complaints of symptoms not attributable to GWI will have abnormal findings based on some other underlying pathology (e.g., traumatic brain injury, depression, thyroid disease, etc.). We have performed studies in aging[65], concussion [66], dementia[67, 68], and depression[69] using several of the tests proposed for this study and the pattern of performance in these groups differs from those with GWI. We will try to exclude individuals with co-morbidities that result in retrieval deficits similar to GWI and will be mindful to assess for age and possible dementing-related factors affecting outcome.

STATISTICAL PLAN

The efficacy of HD tDCS will be assessed on each outcome. For the cognitive-behavioral measures (i.e., from SORT, COWA, BNT, CVLT, Trails A & B, Color Word Interference, Digits Forward & Backward, Rey-Osterrieth Complex Figures Test, Grooved Pegboard Test, and Beck Depression Inventory), a mixed-model MANOVA initially will be used to test for omnibus group (active vs sham) effects over time (i.e., Group X Time interaction effects). Given a significant multivariate effect, separate univariate mixed-model ANOVAs for each cognitive-behavioral measure will be evaluated for group (active vs sham) effects over time, and significant interaction effects will be followed up by separate pre- versus post-intervention evaluations for each tDCS group. Verbal fluency has been associated with sex, age, and education[70, 71 but see 72]; thus, these variables will be included in the models. Additionally, potential mediating effects of IQ (based on NAART) and pre-test to post-test change in executive function, simple attention, working memory, processing speed, and depression (measured via Trails A & B, Color Word Interference, Digits Forward & Backward, Rey-Osterrieth Complex Figures Test, Grooved Pegboard Test, and Beck Depression Inventory) will be explored through covariate mediation modeling. For all testing, data will be thresholded for False Discovery Rates (FDR).

After removing artifacts and segmenting EEG data into multiple trial-by-trial EEG epochs (-200 ms to 1800 ms for SORT and to 1500 ms for SIT), we will apply the STAT-PCA technique that we published, and with which we have longitudinal experience [32, 73] for ERP and time frequency data. Baseline corrected ERPs at each space-time (ST) point, and log normalized squared absolute values at each space, time, frequency (STF) point will be calculated and used as input for the inferential stage of STAT-PCA. The statistical model will include the additive effects of group (active vs sham) and time and higher-order additive effects of group, time, and condition (SORT: Bind vs Nonbind; Semantic Selection: Trial-Type [Go vs No-Go] X Categorization [Simple vs. Moderate vs. Complex) on ERP at each space-time point and log power at each STF. Both within- and between-subject variance components will be included. Interaction contrast *t*-statistics for each ERP ST and spectral STF point will be calculated to assess the influence of the interaction terms on log power, and planned t-contrasts will be calculated from condition differences between the groups. Mean differences that survive a threshold of 5% FDR will be stored as a ST and STF arrays and entered into the sequential PCA. For ERP data, the ST array will be put in a matrix in which columns index space and rows index time. For ERP data, PCA will be performed and will return a set of eigenvalues and eigenvectors, which will be subjected to Parallel Analysis [74] to determine the number of factors to retain. Only eigenvalues above the 95% confidence intervals of the null distribution will be retained. The retained eigenvectors will be varimax rotated and used to calculate the factor scores by projecting the original dataset onto the eigenvectors. The resulting factor scores will be reshaped such that the columns index electrodes and the rows index time points, and PCA will be performed on this matrix. Factor retention will be determined as in the first sequence. The resulting eigenvectors will represent topographies and their corresponding factor scores will represent the time courses for the interaction effects.