8-19-2013 12:00 AM

# A P300 Based Cognitive Assessment Battery for Severely Motor-impaired and Overtly Non-responsive Patients

Aaron M. Kirschner
*The University of Western Ontario*

Supervisor
Adrian Owen
*The University of Western Ontario*

Graduate Program in Neuroscience
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

# A P300 Based Cognitive Assessment Battery for Severely Motor-impaired and Overtly Non-responsive Patients

(Thesis format: Monograph)

By

Aaron Kirschner

Graduate Program in Neuroscience

A thesis submitted in partial fulfillment
Of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

**Abstract**

Diagnosing disorders of consciousness (DOC) is notoriously difficult, with estimates of misdiagnosis rates as high as 40%. Moreover, recent studies have demonstrated that patients who do not show signs of volitional motor responses can exhibit preserved command following detected by functional magnetic resonance imaging (fMRI) and electroencephalography (EEG). Although these patients clearly retain some cognitive abilities, lack of consistent motor responses makes administration of standard neuropsychological tests impossible. Consequently, the extent of their cognitive function is unknown. In the current study, we developed and validated a P300b event related potential (ERP) neuropsychological battery in healthy participants to assess components of executive function without requiring motor output. First, participants were instructed to attend to a target auditory stimulus. P300b responses to attended relative to unattended stimuli were used as a neural proxy for detecting command following. To assess working memory capacity we adapted a digit span test to use a similar P300b response mechanism. Finally, reasoning was assessed by adapting a verbal reasoning task in the same manner. At the group level, and in a large majority of participants at the single-participant level, accurate performance could be detected using the P300b ERP, validating the potential utility of the battery. Additionally, the normalized magnitude of the P300b predicted individual differences in performance, but only when a suitable level of variability between participants was present. A post hoc Monte Carlo analysis was conducted to examine the necessary time required to conduct the battery as well as the interaction between time and performance in determining statistically significant performance. At 100% accuracy, a mean time of five minutes was required to achieve a significant result, with time increasing as a function of decreasing performance. These results demonstrate that covert control of attention, as measured by

the P300b ERP, can be used to assess command following, working memory and reasoning abilities with a high degree of reliability

**Keywords:** disorders of consciousness, neuropsychological assessment, executive function, P300, brain-computer interface, EEG

**Acknowledgements**

There are a number of people without whom this thesis would not have been possible, and to whom I am greatly indebted. I offer my sincere gratitude to my supervisor, Dr. Adrian Owen, for his continued intellectual support and for providing an excellent example of scientific skill and thoroughness. I also offer my appreciation to Dr. Adam Hampshire for his guidance through all stages of this thesis and wish him all the best in his new position. My advisory committee, comprised of Dr. Rhodri Cusack and Dr. Paul Gribble, was instrumental in helping develop many of the statistical analyses used in this project. Likewise, I would like to thank Dr. Damian Cruse for helping me with EEG analysis and taking the time to patiently explain difficult methodological concepts. Lastly, I would like to thank all of the members of the Brain and Mind Institute for contributing to an extremely enjoyable and rewarding experience during my Master's degree.

## Table of Contents

## Table of Figures

# 1. Introduction

## Disorders of Consciousness

Improvements in intensive care have resulted in a surge of survivors of severe brain injury (Owen, 2008a). Consciousness is lost following many neurological traumas, including stroke, drug overdose, hypoxia and traumatic brain injury (Royal College of Physicians Working Group, 1996; The Multi-Society Task Force on PVS, 1994). The majority of patients experience a significant recovery in the first few days, following a typical progression through several states of consciousness before recovering awareness (Figure 1).

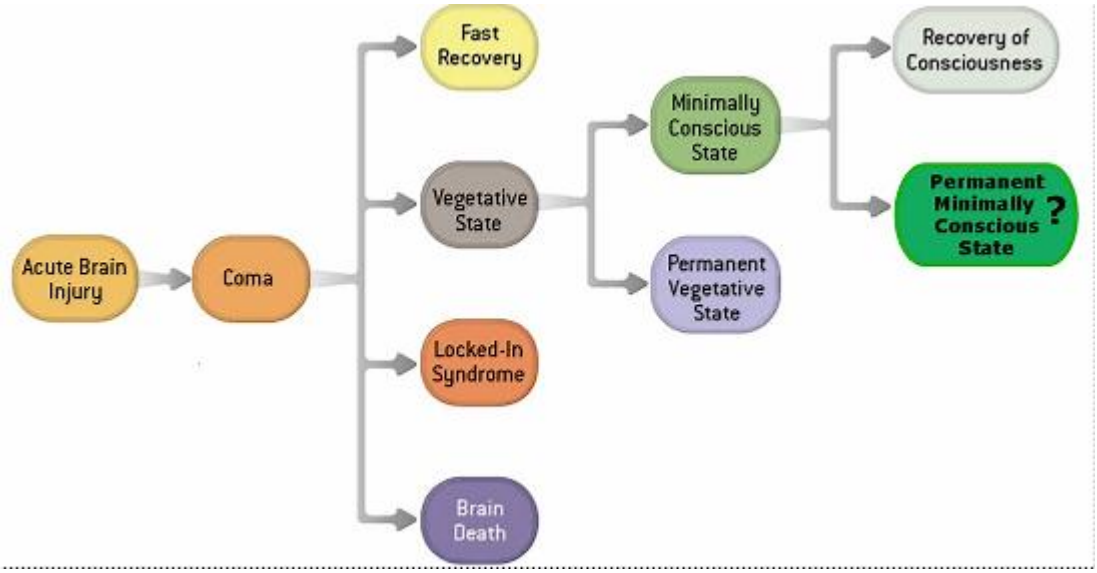



Figure 1. Typical clinical pathways after disruption of consciousness. (From Laureys, 2007)

In some cases, however, recovery is incomplete, resulting in a prolonged disruption of consciousness. Neurological disorders that involve a persistent impairment of the patient's awareness of their self and environment are collectively referred to as disorders of consciousness (DOC) and include coma, vegetative state (VS) and minimally conscious state (MCS) (Owen, 2008a).

Clinical characterizations of consciousness typically make a distinction between wakefulness and awareness, and DOC can be understood as pathologies varying within these two dimensions (Posner, Saper, Schiff, & Plum, 2007). Wakefulness refers to the level of general arousal. With the exception of certain stages of sleep, sufficient wakefulness is considered a prerequisite for conscious awareness (Posner, Schiff, & Plum, 2007). Neurologically, wakefulness is a function of the reticular activating system, a collection of excitatory neuronal circuits from the brainstem to the cerebral cortex relying upon acetylcholine, histamine, serotonin and dopamine neurotransmitter systems (Young, Ropper, & Bolton, 1997).

Awareness, often referred to as conscious awareness or simply consciousness, refers to the content of conscious perception (from here on these terms will be used interchangeably). Although an exact definition of awareness is elusive, it is usually defined as the subjective experience of the internal and external environment (Metzinger, 2010). The neural correlates of consciousness remain to be understood, but current theories emphasize that awareness cannot be localized to a single area of the brain and relies instead on the dynamic interaction between distributed brain regions throughout the cortex and thalamus (Crick & Koch, 1990; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Seth & Baars, 2005).

**Coma**

Following loss of consciousness, patients typically experience a period of coma (Posner, Schiff, & Plum, 2007). Although involuntary reflexes are generally intact, coma patients do not open their eyes, do not display arousal in the presence of external stimulation, and never initiate voluntary movements. In contrast to VS and MCS, coma lacks sleep-wake cycles and it is

generally assumed that comatose patients have no awareness of themselves or their environment. Coma can therefore be understood as the absence of both wakefulness and awareness (Figure 2).

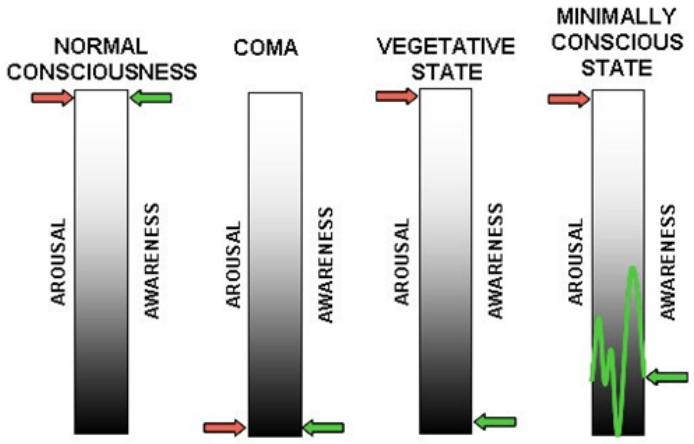



Figure 2. Disorders of consciousness represented as variations within the dimensions of wakefulness and awareness. (Adapted from Laureys, Owen, & Schiff, 2004)

The Glasgow Coma Scale (GCS) is the most widely used scale in coma assessment and attempts to provide a reliable, objective protocol for assessing the patient's level of consciousness (Teasdale & Jennett, 1974). The scale is comprised of three components: eye opening, verbal response and motor response (Figure 3).

Glasgow Coma Scale (GCS)

| 1. Eye Opening Response, E | E |
|---|---|
| Spontaneous | 4 points |
| Opens on verbal command, speech, or shouting | 3 points |
| Opens to pain not applied to face | 2 points |
| None | 1 point |
| **2. Verbal Response** | **V** |
| Oriented | 5 points |
| Confused conversation, but able to answer questions | 4 points |
| Inappropriate responses, words indecipherable | 3 points |
| Incomprehensible speech | 2 points |
| None | 1 point |
| **3. Motor Response** | **M** |
| Obeys commands for movement | 6 points |
| Purposeful movement to painful stimulus | 5 points |
| Withdraws from pain | 4 points |
| Abnormal (spastic) flexion, decorticate posture | 3 points |
| Extensor (rigid) response, decerebrate posture | 2 points |
| None | 1 point |

Figure 3. Examination and scoring criteria for assessment using the Glasgow Coma Scale. (From Teasdale & Jennett, 1974)

A score is given in each subsection of the test along with a global score composed of the sum of all subsections. A score of less than nine is considered severe, a score between nine and 12 is considered moderate and a score above 13 is considered minor. Any score of less than eight qualifies for a diagnosis of coma (Teasdale & Jennett, 1974).

Despite the widespread use of the GCS, it has several limitations. The use of sedating and paralyzing drugs often renders assessment difficult and the inter-rater reliability of the scale has been questioned (Buechler, Blostein, Koestner, Hurt, Schaars, & McKernan, 1998) (Crossman, Bankes, Bhan, & Crockard, 1998). The prognostic utility of the scale has also been criticized (Green, 2011). Nonetheless, the GCS is currently the gold standard of coma assessment.

Coma patients who recover typically begin to display signs of improvement within 2 to 4 weeks. While some patients go on a near complete recovery of consciousness, others remain in a state of disrupted awareness and receive a diagnosis of either VS or MCS depending on their level of recovery.

**Vegetative State**

Patients who emerge from coma typically transition into VS, a clinical diagnosis first introduced in 1972 (Jennett & Plum, 1972). Though some patients pass through VS on the path to more substantial recovery, others suffer profound impairments of consciousness for a prolonged or permanent period resulting in a diagnosis of Persistent Vegetative State, commonly referred to simply as VS (Figure 1). The important characteristics that distinguish VS from coma are eye opening and the presence of circadian sleep-wake cycles. Diagnosis of VS is made after repeated clinical examinations that yield no evidence of sustained, reproducible, purposeful or

voluntary behavioral response to stimuli presented across multiple sensory modalities (Jennett & Plum, 1972; Royal College of Physicians Working Group, 1996). Additionally, VS patients show no evidence of language comprehension or communication. Given the presence of sleep-wake cycles and the absence of volitional behavior indicating consciousness, VS is often understood as wakefulness without awareness (Figure 2).

Despite the clinical distinction made between purposeful and non-purposeful behavior, it can be difficult in practice to determine whether a given movement is purposeful (Andrews, Murphy, Munday, & Littlewood, 1996). Consequently, several objective rating scales have been introduced which intend to provide a standardized, reliable assessment protocol. In particular, the JFK Coma Recovery Scale-Revised (JFK CRS-R) is designed to assess level of function in VS and to distinguish VS from MCS (Kalmar & Giacino, 2006). The JFK CRS-R includes 23 operationally defined behaviors arranged into six subscales that assess auditory, visual, motor, oromotor, communication and arousal functions (Figure 4). Each subscale is arranged hierarchically, with lower-level items assessing reflexive activity and higher-level items assessing purposeful, cognitive-mediated behaviors. The JFK CRS-R has demonstrated good inter-rater reliability and prognostic utility and has been heralded as a promising new assessment tool (Schnakers, et al., 2008). However, as discussed below, recent research has shown that the JFK CRS-R may lack the sensitivity to detect awareness in some DOC patients, particularly those who suffer severe motor impairments (Owen A. M., Coleman, Boly, Davis, Laureys, & Pickard, 2006).

**JFK COMA RECOVERY SCALE - REVISED** ©2004
Record Form

| Patient: Date: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **AUDITORY FUNCTION SCALE** | | | | | | | | |
| 4 - Consistent Movement to Command * | | | | | | | | |
| 3 - Reproducible Movement to Command * | | | | | | | | |
| 2 - Localization to Sound | | | | | | | | |
| 1 - Auditory Startle | | | | | | | | |
| 0 - None | | | | | | | | |
| **VISUAL FUNCTION SCALE** | | | | | | | | |
| 5 - Object Recognition * | | | | | | | | |
| 4 - Object Localization: Reaching * | | | | | | | | |
| 3 - Visual Pursuit * | | | | | | | | |
| 2 - Fixation * | | | | | | | | |
| 1 - Visual Startle | | | | | | | | |
| 0 - None | | | | | | | | |
| **MOTOR FUNCTION SCALE** | | | | | | | | |
| 6 - Functional Object Use † | | | | | | | | |
| 5 - Automatic Motor Response * | | | | | | | | |
| 4 - Object Manipulation * | | | | | | | | |
| 3 - Localization to Noxious Stimulation * | | | | | | | | |
| 2 - Flexion Withdrawal | | | | | | | | |
| 1 - Abnormal Posturing | | | | | | | | |
| 0 - None/Flaccid | | | | | | | | |
| **OROMOTOR/VERBAL FUNCTION SCALE** | | | | | | | | |
| 3 - Intelligible Verbalization * | | | | | | | | |
| 2 - Vocalization/Oral Movement | | | | | | | | |
| 1 - Oral Reflexive Movement | | | | | | | | |
| 0 - None | | | | | | | | |
| **COMMUNICATION SCALE** | | | | | | | | |
| 2 - Functional: Accurate † | | | | | | | | |
| 1 - Non-Functional: Intentional * | | | | | | | | |
| 0 - None | | | | | | | | |
| **AROUSAL SCALE** | | | | | | | | |
| 3 - Attention | | | | | | | | |
| 2 - Eye Opening w/o Stimulation | | | | | | | | |
| 1 - Eye Opening with Stimulation | | | | | | | | |
| 0 - Unarousable | | | | | | | | |
| **TOTAL SCORE** | | | | | | | | |

Denotes emergence from MCS †
Denotes MCS *

Figure 4. Scoring protocol used during administration of the JFK Coma Recovery Scale. (From Kalmar & Giacino, 2006)

**Minimally Conscious State (MCS)**

The minimally conscious state was recently introduced as a new diagnostic category in DOC (Giacino, et al., 2002). Following more substantial recovery, patients often transition from VS into MCS, though others remain in MCS indefinitely. In general, MCS represents a partial recovery of consciousness. Like VS, patients display eye opening and intact circadian rhythms with the addition of partial recovery of volitional behavior. In order to receive a diagnosis of MCS using the JFK CRS-R, at least one of the following behaviors must be present:

- Consistent movement to command

- Object recognition
- Reaching toward object
- Visual pursuit
- Fixation
- Object manipulation
- Localization to noxious stimulation
- Intelligible verbalization
- Non-functional, intelligible communication

Due to the presence of occasional volitional behavior, MCS is understood as wakefulness with a low level of fluctuating awareness (Figure 2). Diagnosis of MCS is often not straightforward, however. Given that MCS patients often show inconsistent and ambiguous responses, the line between MCS and VS is often not clear cut in practice (Schnakers, et al., 2009).

Overall, prognosis for recovery is better for MCS than VS. Some patients who have remained in MCS for years have progressed to live meaningful lives (Bernat, 2006). As a milestone of recovery, demonstration of either functional object use or functional communication entails the emergence from MCS altogether (Figure 5).

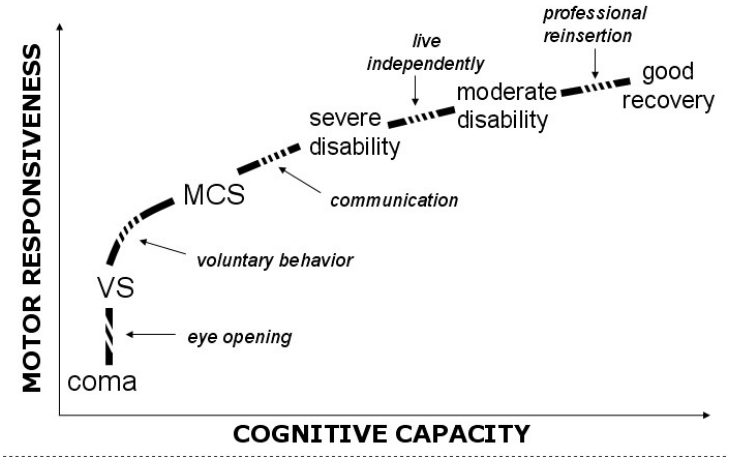



Figure 5. Milestones in recovery from coma to the re-emergence of awareness. (From Laureys, Perrin, Schnakers, Boly, & Majerus, 2005)

**Misdiagnosis in Disorders of Consciousness**

Given the complexity of diagnosing disorders of consciousness and the ambiguity that can accompany patient responses, the accuracy of diagnosis in DOC patients has been questioned. An early study of 49 patients diagnosed as VS concluded that 18 (37%) were misdiagnosed (Childs, Mercer, & Childs, 1993), with trauma patients misdiagnosed significantly more often. Misdiagnosis often resulted from confusion of terminology, lack of extended observation of patients, and a lack of skill and training in the diagnosis of severely brain injured patients.

A further study examined 40 patients referred to a rehabilitation unit with a diagnosis of VS (Andrews, Murphy, Munday, & Littlewood, 1996). The authors used standard command following protocols as well as a novel response mechanism where patients could use small movements to activate a buzzer. 43% of patients were considered to have been misdiagnosed as VS rather than MCS. Of the patients who were misdiagnosed, seven had been erroneously considered as VS for longer than one year and three were considered VS for between four and seven years. A subsequent study utilized electromyography to examine subthreshold motor

activity in DOC patients. Of the 10 VS patients examined, one patient reliably produced electromyographic activity in response to verbal command that was below the threshold required to elicit overt motor activity. Together, these studies suggest that a substantial proportion of DOC patients may be systematically misdiagnosed due to impairments in motor function (Bekinschtein, Coleman, Niklison III, Pickard, & Manes, 2009).

Proper diagnosis of DOC carries both medical and legal importance. As MCS has a much better prognosis than VS, accurate prediction of patient outcome can be significantly impacted by misdiagnosis (Ashwal & Cranford, 1995; Giacino, 2004). Importantly, these data may be due, in part, to what has been called the fallacy of the self-fulfilling prophesy; because the prognosis is poorer in VS, these patients potentially receive less medical care and are denied aggressive life-sustaining treatments, thus affirming the trend of poorer prognosis (Becker, et al., 2001). There are also important legal implications for diagnoses given by clinicians, especially with regard to end of life decision making. Life support may be withdrawn from a patient if brain function is severely compromised and there is no expected value in the administration of continued care. Given that the value of continued care is largely determined by prognostic predictions, which in turn are informed by the patient's diagnosis, proper clinical diagnosis and prognosis in DOC patients can mean the difference between life and death in some cases (Cribb, 2012).

**Measuring Neural Activity to Aid in the Diagnosis of DOC**

Aside from behavioral deficits, clinicians and scientists have investigated the patterns of neural activity and cerebral metabolism that characterize DOC. These efforts can be divided into

three categories: resting state neural activity, neural activity in response to external stimulation, and active paradigms that measure neural activity in response to a volitional cognitive task.

**Resting State Neural Activity in DOC**

Resting state neural activity includes patterns of brain activity that occur spontaneously in the absence of any specific behavior or cognitive task. Studies of resting state activity can therefore be administered to a wide variety of patients at any level of cognitive function. In general, they are effective in characterizing patterns of global brain activity, though less effective in assessing particular cognitive or neural deficits.

Electroencephalography (EEG) is the most prominent modality for assessing brain function in clinical environments (Demertzi, et al., 2009). The presence or absence of electrical activity in the brain, as measured by EEG, is often used to distinguish DOC from brain death with a sensitivity and specificity of approximately 90% (Buchner & Schuchardt, 1990). Resting state EEG patterns differ between VS and healthy individuals, with VS patients displaying a pattern predominantly characterized by low frequency theta (4-7.5hz) and delta (1-3.5hz) waves (Demertzi, et al., 2009). EEG has also been shown to have prognostic value in DOC patients, with burst-suppression patterns of EEG typically signaling a poor outcome (Posner, Saper, Schiff, & Plum, 2007).

Resting state functional neuroimaging has also been used to assess brain function in DOC patients. Positron emission tomography (PET) studies have shown a complete reduction of metabolism throughout the brain in patients suffering brain death and a reduction of up to 50% in DOC patients (Laureys, 2005). DOC patients also display a specific impairment of activity in polymodal association areas including the precuneus, Broca's area, and prefrontal,

parietotemporal and posterior parietal regions, areas associated with higher cognitive functions including attention, memory and language (Laureys, et al., 1999).

In addition to activity in specific brain regions, functional connectivity between distributed areas of the brain is predictive of the level of function in DOC. In a functional magnetic resonance imaging (fMRI) study of 14 DOC patients, a higher degree of functional connectivity between areas of the Default Mode Network (DMN) was correlated with behavioral diagnosis of patients (Vanhaudenhuyse, et al., 2010). Another fMRI study examining coma patients found that functional connectivity in the DMN predicted emergence from coma, while absence of default network connectivity indicated a poor prognosis for recovery (Norton, Hutchinson, Young, Lee, Sharpe, & Mirsattari, 2013). Thalamocortical connectivity is likewise disrupted in DOC. VS patients show significant reductions in connectivity between thalamic nuclei and the prefrontal cortex relative to healthy controls, with restorations in connectivity occurring with the recovery of awareness (Laureys, Faymonville, Luxen, Lamy, Franck, & Marquet, 2000).

Following from earlier functional connectivity research, Rosanova et al. (2012) developed a novel perturbation approach for examining resting state cortical connectivity using combined transcranial magnetic stimulation (TMS) and EEG. In a group of 12 DOC patients, various areas of the cortex were stimulated using TMS while neural activity was recorded using EEG. In VS patients, stimulation led to short term, localized neural activity. In contrast, when stimulation was applied to MCS patients, complex long-range activity was recorded throughout the cortex, demonstrating a wider network of functionally connected regions. Moreover, increasingly complex patterns of connectivity began to emerge as patients recovered consciousness and were displayed prior to the reemergence of communication and changes in

resting state EEG patterns. This technique also has the distinct advantage of demonstrating a causal relationship between activities in separate cortical areas compared to fMRI functional connectivity, which relies exclusively on correlation.

**Neural Responses to Passive External Stimulation**

In contrast to resting state studies, passive stimuli paradigms measure neural activity in response to sensory stimulation. Passive stimuli paradigms have several important features. In contrast to the active paradigms discussed below, they do not require active participation and are therefore applicable to a wider range of patients. Furthermore, by administering stimuli that generally evoke a response in a particular set of brain regions, they allow researchers and clinicians to make inferences about the integrity of specific neural systems. They therefore provide a more nuanced understanding of the unique neural disruptions in DOC populations as well as individual differences between patients.

Passive stimuli paradigms have several shortcomings, however. In general, the fact that an area becomes active in response to a class of stimuli does not necessarily mean that that area is processing those stimuli accurately (Coleman, et al., 2007). In addition, passive stimuli paradigms say little about whether a patient is conscious of the stimuli being presented, as many studies have shown that unconsciously perceived stimuli may nonetheless increase neural activity in stimulus specific regions (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006). Despite these shortcomings, passive stimuli paradigms have provided important insights into DOC and can aid in patient prognosis.

In ERP paradigms, stimuli are presented while neurophysiological activity is recorded using EEG. Segments of EEG activity accompanying the presentation of stimuli are known as

event related potentials (ERP). Typically, stimuli are varied along two or more experimental dimensions. A difference in the magnitude, frequency, or timing of the ERP response between conditions can be used to infer differential neural processing in each condition and sensitivity to the experimental dimension manipulated (Luck, 2005).

One of the most widely studied ERPs in DOC is the Mismatch Negativity (MMN). In a typical MMN task, a standard auditory stimulus is repeated with high frequency, while a different stimulus, known as a deviant, occurs infrequently. Relative to standard stimuli, deviant stimuli elicit a negative electrical potential on the scalp with a post-stimulus latency of 150-250ms (Näätänen, Pakarinen, Rinne, & Takegata, 2004). The presence of an MMN thus confirms the integrity of neural structures involved in auditory change detection.

In several studies, a preserved MMN response has been shown to predict successful recovery from coma (Kane, Curry, Butler, & Cummin, 1993; Fischer, Morlet, Bouchet, Luaute, Jordan, & Salord, 1999). Likewise, VS and MCS patients with an intact MMN are significantly more likely to show clinical improvement six months after testing (Kotchoubey, et al., 2005). In a longitudinal study that tested MMN responses every two weeks over a three and a half month period, MMN amplitude increased with the level of function in DOC patients and a sudden increase in MMN amplitude coincided with the return of functional communication (Winjin, van Voxtel, Eilander, & de Gelder, 2007).

Aside from low-level change detection, ERPs have been used to assess a number of higher cognitive functions in patients. In a large-scale study, ERPs were used to assess preserved language function in 98 DOC patients (50 VS, 48 MCS) (Kotchoubey, et al., 2005). Three separate paradigms were used: a semantic oddball which included streams of categorically related words with unrelated words interspersed (cat, dog, horse, hat), a word-pairs task where

pairs of words were either semantically related or unrelated (table-chair vs. table-peach), and a sentences task where seven word sentences were played where the last word was either highly expected or highly unexpected. The presence of an appropriate ERP in each task was significantly related to the level of function of patients as measured by the Disability Rating Scale (Rappaport, Hall, Hopkins, Belleza, & Cope, 1982). Moreover, it was also noted that some patients who scored low on behavioral assessments still exhibited ERPs to semantic violations, indicating that language processing systems may be preserved in the absence of overt linguistic behaviour.

In addition to EEG, PET and fMRI imaging studies have also used passive stimulation paradigms to assess a variety of functions in DOC patients. In a group of seven VS patients, five MCS patients and two patients who recently emerged from MCS, an fMRI-based hierarchical language task was developed to assess residual speech comprehension (Coleman, et al., 2007). Low-level auditory processing was measured by contrasting fMRI responses to auditory stimuli (consisting of both intelligible speech and unintelligible noise) and a silent baseline. Higher level, speech-specific processing was assessed by contrasting intelligible speech to amplitude-modulated noise. At the highest level of the hierarchy, neural responses to sentence meaning were assessed by comparing responses to sentences that contained ambiguous words to matched sentences without ambiguity. In the low-level auditory task, three VS, two MCS and all of the recently emerged patients showed significant activation in the temporal lobe. These patients also showed significant activation in more extensive regions of the temporal lobe to higher level, speech specific stimuli. At the highest-level contrast, ambiguous sentences elicited significant activation in one VS patient, one MCS patient, and one recently emerged patient, spanning the temporal lobe as well as the left inferior frontal gyrus, an area previously associated with the

resolution of linguistic ambiguity (Rodd, Davis, & Johnsrude, 2005). These results again demonstrated that patients who do not respond overtly to linguistic stimuli might nonetheless retain speech networks that support language comprehension. However, as recognized by the authors, the results do not necessarily suggest successful language processing *per se*, as it is possible that these patients unsuccessfully attempted to resolve ambiguous sentences, thereby still engaging the relevant brain networks. Nonetheless, differential activation to ambiguous sentences still implies that many lower level processing functions were successful, including segmenting an auditory stream and activating word meaning within language networks, despite the lack of linguistic behaviour in these patients.

In addition to language functions, neuroimaging paradigms have been used to examine whether DOC patients experience pain. Reducing pain is an important part of patient care and behavioural scales such as the Nociception Coma Scale have been developed to detect pain in DOC patients (Schnakers, et al., 2010). However, it is still unclear to what extent DOC patients process nociceptive stimuli beyond the pain reflexes typically exhibited. In a PET study, 15 VS patients received high-intensity electrical stimulation to the median nerve of the wrist. In all VS patients, noxious stimuli evoked activations in the midbrain, contralateral thalamus, and primary somatosensory cortex. Unlike healthy controls, however, secondary somatosensory, insular, posterior parietal and anterior cingulate cortices did not show increased activation. Moreover, in patients, the activated primary somatosensory cortex was functionally disconnected from secondary somatosensory, posterior parietal, premotor, superior temporal, and prefrontal cortices. Of the cortical areas involved in pain perception, only the anterior cingulate is consistently associated with the subjective experience of pain (Derbyshire, Jones, Clark, Townsend, & Firestone, 1997), which was notably absent in these patients. Likewise, it has been

shown in neurological patients that somatosensory stimuli below conscious threshold activate somatosensory areas but do not give rise to downstream activations (Libet, Alberts, Wright Jr, & Feinstein, 1967), and direct stimulation of somatosensory cortex does not elicit the feeling of pain (Penfield & Jasper, 1954). The results therefore suggest that the majority of the VS patients tested did not consciously experience pain, though it should be noted that these patients scored particularly low on behavioral assessments and the results may not extend to other DOC patients. Nonetheless, the study took important first steps in understanding the experience of pain in DOC as well as using the neural correlates of conscious experience in healthy participants to infer the conscious state of patients.

Emotional processing has also been investigated in DOC patients using neuroimaging paradigms. In an early fMRI case study, audio clips were played to an MCS patient of his mother reading a story followed by the same clips read by a stranger (Bekinschtein, et al., 2004). When neural activity was contrasted between the two conditions, it was found that listening to clips read by the patient's mother elicited greater activation in the insula as well as the amygdala, two regions robustly associated with emotional processing (Figure 6) (Phan, Wager, Taylor, & Liberzon, 2002).

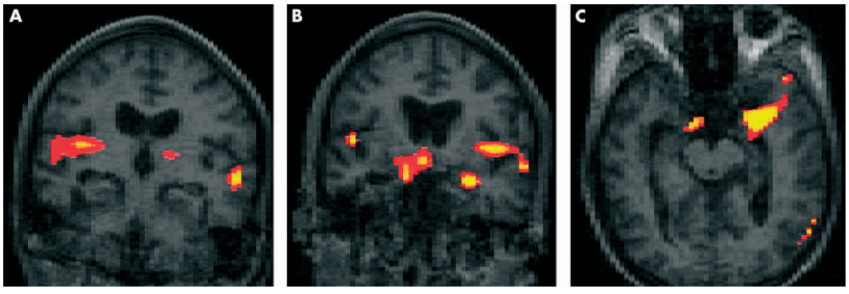



Figure 6. Patterns of neural activity elicited by stimuli with greater emotional significance. (From Bekinschtein, et al., 2004)

A subsequent study further examined emotional processing in DOC using visual stimuli and a larger sample of patients (Zhu, et al., 2009). Nine MCS patients and 10 controls were shown personal family photos as well as never-before-seen images that varied in emotional valence. Of the nine patients, six showed cortical activation patterns to personal family portraits that were highly similar to controls, including occipital, parietal, orbitofrontal and prefrontal cortices as well as the fusiform gyrus. These patients also generally showed more activation to highly emotional novel images than to novel images with less emotional valence. These results suggest that, despite their lack of outward emotional expression, DOC patients may nonetheless retain aspects of emotional processing.

**Active Paradigms**

In contrast to the paradigms outlined above, in active imaging paradigms the patient is given an instruction to perform a cognitive process that requires volitional intention and control (Bruno, et al., 2010). An appropriate pattern of neural activity observed after the instruction provides evidence that the patient was carrying out that cognitive process and therefore exhibiting volitional behavior. Active paradigms are largely motivated by the recognition that some patients who retain consciousness may have nervous system damage that prevents them from signifying their awareness using a motor response (Figure 7). Unlike passive and resting state paradigms, active paradigms have the distinct property of potentially challenging a patient's clinical diagnosis. Consistent, accurate performance within an active paradigm can be viewed as a form of command following, one of the basic criteria that denotes emergence from DOC. Active tasks also have the potential to be used to implement brain-computer interfaces (BCIs) in non-communicative patients, allowing the direct translation of systematic neural activity into the

control of an external communication device (Kotchoubey B. , 2007). Active tasks therefore have tremendous potential for assessing high-functioning, non-communicative patients.

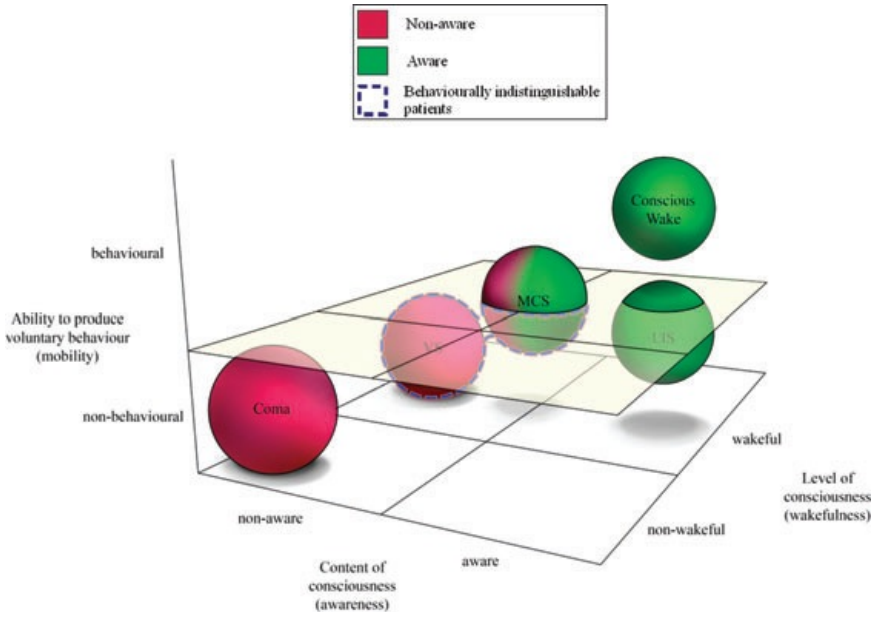



Figure 7. DOC reconsidered with the additional dimension of motor response capability. (From Monti, Coleman, & Owen, 2009)

In a pioneering study, the first active fMRI paradigm was implemented with a VS patient who had suffered traumatic brain injury (Owen, Coleman, Boly, Davis, Laureys, & Pickard, 2006). During the fMRI scan, she was instructed to either imagine playing tennis or visiting all of the rooms in her house. During periods of tennis imagery, significant activation was observed in the supplemental motor area, an important area for motor imagery (Jeennerod, 1994). Conversely, during periods of spatial imagery, significant activation was found in the parahippocampal gyrus, posterior parietal cortex and lateral premotor cortex, areas previously associated with spatial imagery (Maguire, Burgess, Donnett, Frackowiak, Frith, & O'Keefe, 1998) (Figure 8). This data provided strong evidence that the patient was able to understand spoken commands and volitionally modulate her brain activity, thereby challenging her clinical diagnosis. In a subsequent discussion of the study, commentators proposed that the last word of

the instructions in each condition ("tennis" or "home") may have unconsciously triggered the brain activity found in the study (Nachev & Husain, 2007; Greenberg, 2007). However, as argued by the original authors, activity in the two conditions persisted for a full 30 seconds, much longer than activity found in unconscious priming experiments. Furthermore, activity extended well beyond the word recognition areas that typically activated to unconscious stimuli (Owen A. M., et al., 2007).

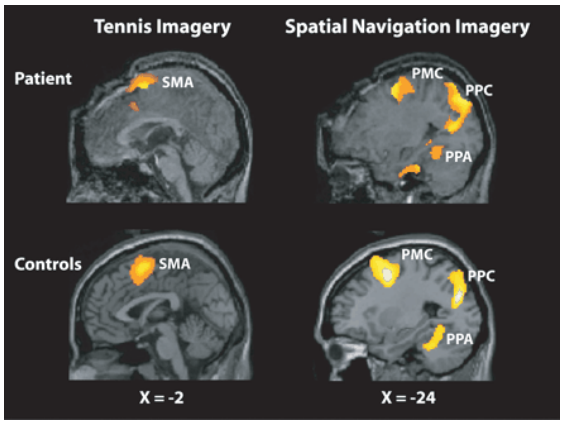



Figure 8: fMRI activation in response to imagery instructions in a behaviourally unresponsive VS patient and healthy controls. (From Owen, Coleman, Boly, Davis, Laureys, & Pickard, 2006)

In a follow-up large scale study, 23 VS and 31 MCS patients were assessed with a similar imagery paradigm (Monti, et al., 2010). Patients were again instructed to imagine playing tennis or navigating their house while undergoing fMRI. Among the 54 patients, 4 VS and 1 MCS could willfully modulate their brain activity in the imagery conditions, demonstrating that a significant minority of patients diagnosed as VS may retain cognitive function and awareness undetected during standard DOC assessment. The authors also stressed that the fMRI protocol likely underestimates the number of patients that retain some level of awareness, as false negatives may have resulted from the lack of adequate statistical power as well as deficits in language comprehension, working memory and other cognitive faculties that prevented some

patients from performing the task. Aside from replicating the initial study in a larger population, the authors implemented a novel communication paradigm with one of the patients who had demonstrated a consistent fMRI response. The patient was given six yes-or-no autobiographical questions (i.e., "is your father's name Alexander?") and instructed to respond by thinking of one type of imagery for an affirmative answer and the other for a negative answer. In five of the six questions, fMRI activation closely matched one of the imagery conditions from the previous scan and the corresponding yes-or-no answer was taken as the patient's response (Figure 9). The response was correct for all five questions. The study provided the first demonstration that functional communication could be established in a non-communicative DOC patient using neural activity, suggesting that a similar approach could be used to address important clinical questions, such as asking if the patient was in pain or allowing some degree of autonomy in determining their clinical treatment.

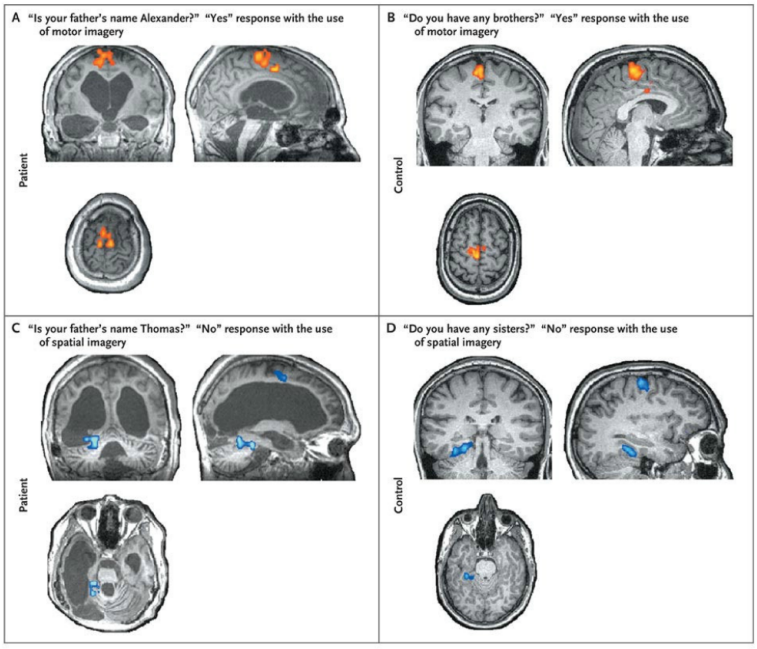



Figure 9. Functional communication established through covert motor imagery in a VS patient. (From Monti, et al., 2010)

Following the success of fMRI motor imagery paradigms, several studies have attempted to implement a similar approach using EEG. EEG has several important advantages over fMRI in clinical settings including lower cost, higher availability, and the possibility of testing patients with medical implants that may rule out fMRI scanning. In addition, EEG portability allows testing to be performed at the bedside, avoiding the physical stress involved in transporting patients and raising the possibility of BCIs for communication in the long term.

In a cohort study of 16 VS patients, an EEG motor imagery task was developed where patients imagined either squeezing their right hand or wiggling their toes (Cruse, et al., 2011). Due to the complexity of EEG spectral changes during motor imagery, a support vector machine-learning algorithm was used to examine whether it was possible to classify neural activity in the two conditions. The EEG patterns of three of the patients could be classified significantly above chance with classification accuracies ranging from 61-78%. The EEG spectral patterns that provided the highest classification accuracy were highly similar to those of healthy controls, providing evidence that similar neural generators were contributing to the signal in VS patients (Figure 10). None of the patients who showed significant changes in EEG spectral power had shown consistent responses to commands during standard behavioural assessments. The study provided the first evidence that command following could be assessed in non-communicative VS patients using EEG.

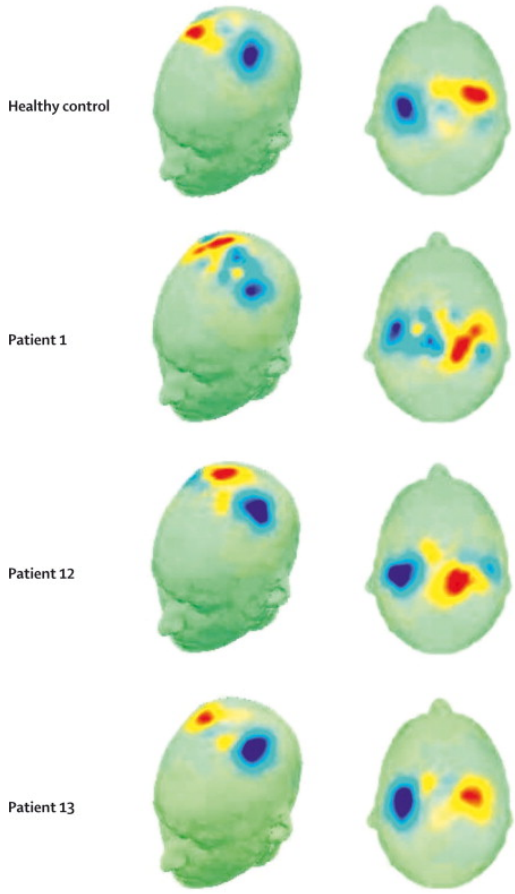



Figure 10. Topographic maps of EEG activity during motor imagery in VS patients and a healthy control. (From Cruse, et al., 2011)

The aforementioned studies established BCIs by instructing the participant to generate an internal mental state while attempting to detect neural activity associated with that state. Another predominant approach in BCI research has been the use of volitional changes in attention (Wolpaw, et al., 2000). Within these approaches, several stimuli are presented, either concurrently or sequentially. By actively paying attention to a single stimulus or stimulus category, a differential brain pattern is generated in response to those stimuli, which in turn can be used to signal a response.

One particular pattern of attention-related neural activity used extensively in DOC research is the P300b (Figure 11. Typical time course of P300b ERP, elicited in response to a

target stimulus embedded within a stream of distractors (From Polich, 2007).). The P300b is an EEG component defined as a positive change in scalp voltage occurring approximately 300ms after the presentation of a target stimulus in the context of several distractor stimuli (Picton, 1992). A target stimulus must be both infrequent and actively attended in order to elicit a P300b. There is still widespread debate about the cognitive processes reflected by the P300b. An early model proposed that it reflects a context updating operation, when the brain revises an active hypothesis about the current context, explaining why infrequent stimuli are necessary for its generation. More recently, it has been proposed that the P300b reflects a template matching process, where an external stimulus is matched with an internal representation. Converging neuroscience research has localized its neural generators to temporal and parietal cortices, with potential contributions from the anterior cingulate, relying upon dopamine and norepinephrine neurotransmitter systems. Despite over four decades of human and animal research, the nature of the P300b is still a heavily debated area of research (Polich, 2007).

Due to the selective, attention-dependent nature of the P300, it has been widely exploited in BCI research. Stimuli presented to the participant can be selected to represent various choices (i.e., “yes” or “no”). The participant decides which stimulus to attend to, thereby determining which stimulus will evoke a P300b. The EEG is then analyzed to determine the stimulus that evoked P300b ERPs, signaling the particular choice made by the participant and allowing the implementation of a BCI (Mak & Wolpaw, 2009). Likewise, the attended stimulus can be selected by the experimenter, allowing for the implementation of covert command following paradigms.

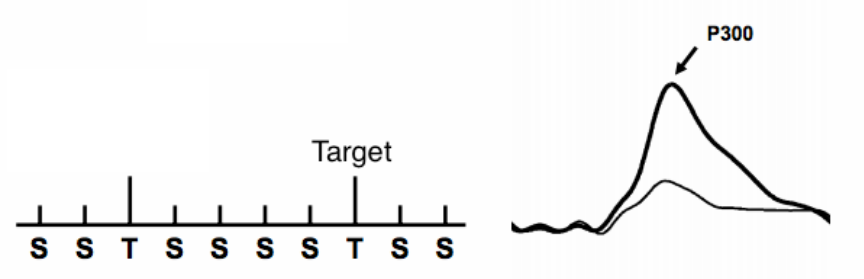



Figure 11. Typical time course of P300b ERP, elicited in response to a target stimulus embedded within a stream of distractors (From Polich, 2007).

A P300b-based auditory paradigm was successfully used to detect command following in a locked-in patient with no behavioural signs of awareness (Schnakers, et al., 2009). Auditory sequences of the patient's own name and unfamiliar names were presented in three separate conditions. In the passive condition, the patient was instructed to simply listen to the names. In the first active condition the patient was instructed to count her own name in an auditory sequence, while in the second she was asked to count a specified unfamiliar name. The authors found a significantly larger P300b response to the patient's name relative to the other name in the passive condition. Moreover, there was a larger P300b in the active condition than the passive condition, indicating that she had willfully modulated her attention in response to task instructions, a form of command following. The same experiment was run at earlier times relative to brain injury with no significant result, while the positive result came two weeks before the first behavioral signs of conscious recovery.

Lulé et al. (2012) further tested a voluntary, auditory P300b paradigm on healthy controls as well as locked-in, MCS and VS patients. The auditory stimuli "yes", "no", "pass", and "go" were presented to patients who were instructed to count either "yes" or "no." In the majority of healthy controls, as well as a locked-in and MCS patient, a significantly larger P300b was

observed to attended stimuli, demonstrating volitional behaviour in typically unresponsive patients.

Aside from EEG paradigms, active approaches relying on attention-based neural responses have been piloted using fMRI. In one study, an fMRI-based test of executive function was developed that assessed the ability to maintain task relevant stimuli over time and in the face of distractor stimuli without requiring an overt response (Monti, Coleman, & Owen, 2009). Using a within-participant, block design, 20 healthy controls as well as one MCS patient completed alternating tasks where they either rested or counted stimuli. In the counting task, a new target was given at the beginning of each block and participants were instructed to subvocally count each time it occurred. In the passive condition, participants were instructed to simply listen to stimuli. FMRI activations were first contrasted between counting and passive conditions at the group level in healthy participants. Significantly greater activation was found in frontal, temporal, and parietal cortex during counting. When analyzed at the single participant level, the contrast was significant for all 20 healthy participants. The same contrast in the MCS patient likewise revealed significantly greater activation during the counting task, with a pattern of neural activation similar to healthy controls. These results suggested that the patient was able to understand and maintain task instructions and experimental stimuli for a prolonged period of time, demonstrating the capacity for working memory, a key component of executive function.

In a recent study, an fMRI imagery paradigm was used to directly assess reasoning capacity in a VS patient, another key component of executive function (Hampshire, et al., 2013). The experimenters utilized a verbal reasoning test that requires participants to infer the relative position of two items as described by sentences that are varied in complexity (Baddeley, 1968). The VS patient was presented with a problem at the beginning of each trial where the correct

answer was either “face” or “house.” After solving the problem, the patient was instructed to generate a mental image of either a face or house depending on their answer. To assess whether the patient had correctly solved the reasoning problems, trials where the correct answer was face were contrasted with trials where the correct answer was house. In trials where house was the correct answer, there was significantly greater activation in areas specific to spatial processing, including parahippocampal place area, left superior occipital gyrus and left lateral premotor cortex, indicating that the patient had imagined a house when it corresponded to the correct answer and was able to solve a significant number of reasoning problems (Figure 11). There was also a significant main effect of level of difficulty on the amount of activation, demonstrating that the patient had greater difficulty with more complex sentences. The study also found significant activation in several areas of the frontal lobes during the period when the patient was actively solving problems, with the level of activation in these regions again modulated by the difficulty of the sentence (Duncan & Owen, 2000). This study provided the first evidence of successful reasoning in a patient diagnosed as VS.

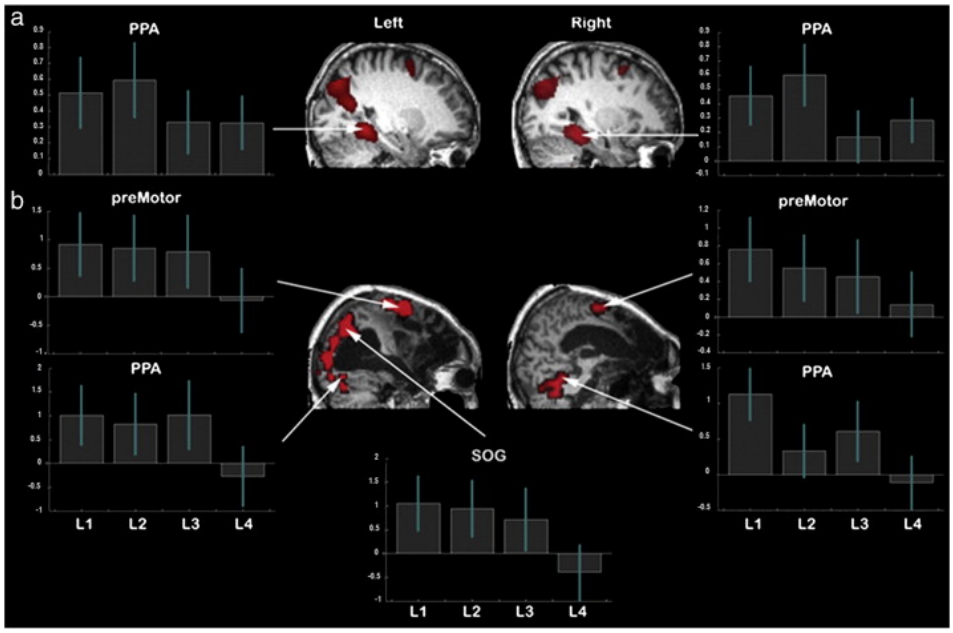


Figure 12. fMRI activations to trials where the correct answer to a reasoning problem was "house." (From Hampshire, et al., 2013)

**Current Study**

Recent research has shown that both neuroimaging and EEG paradigms are complementary to behavioral tests of DOC patients and may allow for better assessment of patients with deficits in motor function. Accordingly, several authors have argued that neuroimaging paradigms should be routinely used to aid in clinical diagnosis (Coleman, et al., 2009), making use of hierarchically structured batteries that sequentially assess lower level sensory functions, language comprehension, command following, and executive control.

In particular, assessment of executive function in DOC patients has been largely understudied despite carrying particular clinical and ethical importance. As BCI systems become more widely utilized in DOC, questions will naturally arise as to what degree of autonomy communicative patients should be given in determining their medical care and managing their lives. In general, the ethical provision of medical autonomy assumes the capacity for understanding complex information and appreciating foreseeable consequences to decisions, both of which are closely tied to executive functioning (Etchells, Sharpe, Elliot, & Singer, 1996; Marson & Harrell, 1999). Although important steps have been made in developing assessment paradigms for executive function (Hampshire, Highfield, Parkin, & Owen, 2012), further study is required in order to develop assessment protocols that are validated and can be administered easily on a wide variety of patients.

The current study sought to develop and validate a battery of executive function tasks in healthy participants that can be used to assess DOC patients without requiring an overt motor response. Due to the motor limitations of DOC patients, the paradigm was designed to be

capable of assessing executive function using the P300b response validated in previous studies (Guger, Edlinger, Harkam, Niedermayer, & Pfurtscheller, 2003). Additionally, because DOC patients characteristically have difficulty maintaining eye fixation, the task was administered using auditory stimuli. Three tasks were chosen for the battery: a basic command following paradigm, a modified memory span task, and a verbal reasoning task. These tasks were chosen based on recent studies that highlight memory and reasoning as fundamental but dissociable components of executive functioning (Hampshire, Highfield, Parkin, & Owen, 2012).

The main hypothesis of the experiment was that the battery would be able to detect command following in the first task as well as correct performance in the executive function tasks at the group and single participant level using the P300b response. It was also hypothesized that the battery would be capable of determining individual differences in performance using the P300b. Lastly, exploratory analyses were conducted examining the sensitivity of the paradigm to detect correct performance as a function of testing time and participant accuracy.

## 2. Methods

### Participants

All experiments were approved by the Psychology Research Ethics Board at Western University. 16 participants (eight females, age: 21.1 ± 2.2 years) were recruited from Western University in London, Canada. Written, informed consent was given. All participants were right-handed, native speakers of English with no history of neurological disorders. Participants were paid $15 per hour. Data from two participants was excluded from analyses due to excessive movement artifacts.

## Experimental Paradigm

### Stimuli

The word stimuli used for the P300b paradigm consisted of auditory recordings of "boss", "cake", "dot", "fan", "map", "pen", "leaf", and "seal," subsequently referred to as word stimuli. All word stimuli were monosyllabic, concrete nouns with different onset consonants, matched for frequency and imagability using the MRC Psycholinguistics Database (Wilson, 1988). All auditory stimuli were recorded from the same speaker with a sampling rate of 44100hz and normalized for peak amplitude. Word stimuli were 400ms in duration. All auditory stimuli were delivered through EEG compatible headphones at a clear but comfortable volume.

### Task1

Task 1 attempted to implement a basic command following paradigm using the P300b ERP. During Task 1, subsequently referred to as auditory attention (AT), participants were given a target word at the beginning of each trial (Figure 13: Organization of experiment and tasks.). The target word was given using the auditory phrase "the word you will be counting is x, begin counting," where x was replaced by the target. After a two second pause, a sequence of word stimuli was played (referred to subsequently as the "stream"), including the target word as well as all seven non-target words. The participant's task was to internally count the number of occurrences of the target word while ignoring non-target words. Subjects were instructed to maintain fixation on a cross centred on the screen during the presentation of the stream.

The stream consisted of an equal number of occurrences of all eight word stimuli, played 7-10 times each, so that the total stream length was either 56, 64, 72 or 80 words. For each trial, the stream was generated by repeatedly appending randomized sets of all eight words. For

example, words 1-8 in the stream consisted of all word stimuli in random order, words 9-16 included all word stimuli in random order, and so forth. Streams were also generated with the condition that no word could immediately follow itself. Word stimuli within the stream were presented with an interstimulus interval that randomly varied between 50 to150ms.

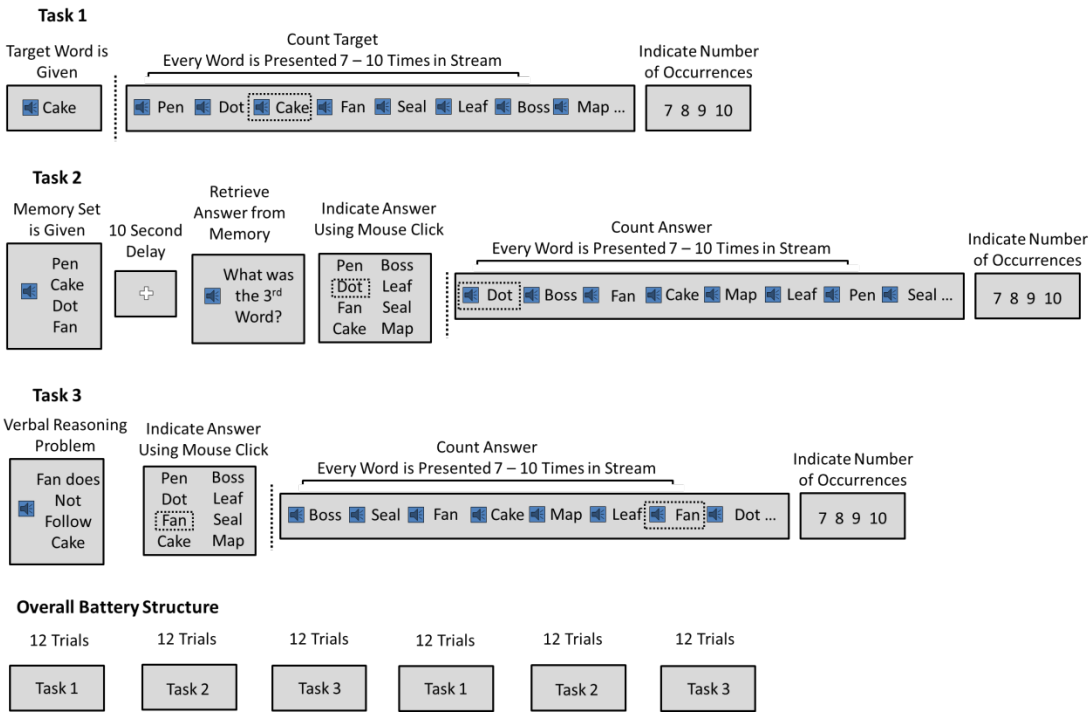



Figure 13: Organization of experiment and tasks.

At the end of each stream, an auditory instruction played, asking, “how many times did the word occur?” Four buttons appeared on the screen, with the choices 7, 8, 9, and 10. The participant clicked the button corresponding to their response. The purpose of the behavioral response was to test whether the participant was performing the counting task. A 10 second rest delay was given before the start of the next trial.

Each block of the auditory attention task consisted of 12 trials. There were two blocks in the experiment, resulting in 24 trials total. Each word stimulus was a target three times.

**Task2**

Task 2 used the P300b to assess working memory. During task 2, subsequently referred to as working memory (WM), participants were given a memory set at the beginning of each trial (Figure 13). The auditory phrase "remember the following words in order" was played, followed by a memory set of four, six or eight word stimuli. A delay period of 10 seconds followed, during which the participant attempted to maintain the memory set. Word stimuli were counterbalanced so that each word was in a memory set an equal number of times. Each word stimulus was a correct answer in three trials.

Following the delay, the auditory phrase "what was the xth word" was played, where x could be any ordinal position from one to the length of the memory set (i.e., $1^{st}$, $2^{nd}$, $3^{rd}$, etc.). Within each set memory length, there was an equal probability that the word to be remembered would appear in any ordinal position. The participant attempted to recall the word in that position from the set given. Eight buttons appeared on screen, one for each of the word stimuli. Participants indicated their answer by clicking the appropriate button with a mouse. The purpose of the behavioral response was to provide a benchmark to test the validity of analyzing participants' performance using the P300b ERP.

After indicating their answer, a two second delay followed, followed by the auditory phrase "count that word." A stream of word stimuli were played, arranged exactly as described in task 1, with 7-10 repetitions of each word stimulus. Participants counted the word stimulus corresponding to their answer and indicated its number of occurrences after the completion of the stream.

Each WM block contained 12 trials with two blocks in the experiment for a total of 24 WM trials. Each memory set size was used in eight trials.

**Task3**

The purpose of task 3 was to assess reasoning ability using the P300b. During task 3, referred to as auditory reasoning (AR), participants were given an auditory reasoning problem at the beginning of each trial. The auditory phrase "in the sentence" was played, followed by a problem sentence (i.e., "The cake follows the pen"), followed by the phrase "which of the two items should be first?" (Figure 13). The task of the participant was to determine the word in 1$^{st}$ position as described by the sentence (Figure 14). This verbal reasoning task has been used in the past to measure general reasoning abilities, as it requires the participant to maintain the sentence in mind and manipulate it according to logical rules in order to arrive at a solution (Hampshire, et al., 2013). 24 unique sentences were generated, one for each AR trial. Sentences were manipulated according to the verb used (precede VS follow), negation (positive VS negative) and whether the sentence was active or passive ("follows" VS "is followed by"). In total, eight sentence types were generated and each was played three times over the course of the experiment (Figure 14). Word stimuli were counterbalanced so that each word formed the correct answer in three sentences and an incorrect answer in three sentences.

Following the auditory phrases, eight buttons appeared on screen, one for each of the word stimuli. Participants indicated their answer by clicking the corresponding button using the computer mouse. The purpose of the behavioral response was again to provide a benchmark to test the validity of analyzing participants' performance using the P300b. Participants were given an unlimited amount of time to solve each problem. A two second delay followed the response, followed by the auditory phrase "count that word." A stream of word stimuli was played,

arranged exactly as described in task 1, with 7-10 occurrences of each word stimulus. Participants counted the word stimulus corresponding to their answer and indicated its number of occurrences after the completion of the stream with the computer mouse.

Each AR block contained 12 trials, and two blocks were run in the experiment, resulting in a total of 24 AR trials.

| Example | Precede/Follow | Active/Passive | Negative/Positive | Correct Answer |
|---|---|---|---|---|
| Cake precedes dot | Precede | Active | Positive | Cake |
| Cake does not precede dot | Precede | Active | Negative | Dot |
| Cake is preceded by dot | Precede | Passive | Positive | Dot |
| Cake it not preceded by dot | Precede | Passive | Negative | Cake |
| Cake follows dot | Follow | Active | Positive | Dot |
| Cake does not follow dot | Follow | Active | Negative | Cake |
| Cake is followed by dot | Follow | Passive | Positive | Cake |
| Cake it not followed by dot | Follow | Passive | Negative | Dot |

Figure 14. Sentence types used in AR task

**Overall Experimental Design**

The experiment contained six blocks total with two blocks of each task (Figure 13). Each block contained 12 trials, for a total of 72 trials. The arrangement of the experiment was task 1, task 2, task 3, task 1, task 2, task 3, for all participants. The arrangement was not counterbalanced across participants as it is common practice to give neuropsychological tests with a fixed sequence (Tombaugh & McIntyre, 1992; Fray, Robbins, & Sahakian, 1996). Participants were given as much time as needed to rest between blocks. The total experiment time was typically one hour and 15 minutes, plus the time required for setup of the EEG recording system.

**Behavioral Analysis**

To test for systematic fluctuations of attention over the course of the task, the accuracy of counting target stimuli (as given by the experimenter in task 1, as indicated by participant in task 2 and 3) was compared across blocks. Within each block, the number of trials where the participant indicated the correct number of occurrences of the target was divided by the total number of trials in order to calculate accuracy. A repeated measures ANOVA was conducted with block number as the independent variable and accuracy as the dependent variable.

To test for systematic fluctuations of attention between tasks, the accuracy of counting target stimuli (as given by the experimenter in task 1, as indicated by participant in task 2 and 3) was compared across task type. Within both blocks of each task, the number of trials where the participant indicated the correct number of occurrences of the target was divided by the total number of trials to calculate accuracy. A repeated measures ANOVA was conducted with task type as the independent variable and accuracy as the dependent variable.

To assess the effect of set size on recall accuracy in the WM task, a one-way repeated measures ANOVA was conducted with three levels: four item sets, six item sets and eight item sets. For each participant, the number of correctly recalled sets was divided by the total number of questions within each set size to calculate accuracy for that set size.

To assess the effect of sentence type on performance in the AR task, a three-way repeated measures ANOVA was conducted. Verb type (precedes vs. follows), negation (negative vs. positive sentences) and form (active vs. passive) were used as factors, each with two levels. Within each sentence type, the number of correctly solved problems was divided by the total number of problems of that type to calculate accuracy rate.

## EEG Analysis

### EEG recording.

EEG recording was performed using a G.Tec amplifier and G.Tec gel-based active electrode system (G.Tec Medical Engineering, GMBH). Electrodes were placed using the 10-10 convention and recorded from locations FC3, C3, CP3, FCZ, CZ, CPZ, FC4, C4, CP4, T7, T8, PZ, POZ, OZ, P7, P8. Data was analogue filtered with a passband of 0.1-100 Hz and a notch filter at 60 Hz to reduce interference. Sampling was performed at 256 Hz, with impedances kept below 5k Ω. Scalp voltages were referenced to the right earlobe.

### EEG Preprocessing

All EEG processing was performed using Matlab with EEGLAB and FieldTrip toolboxes (Delorme & Makeig, 2004; Oostenveld, Fries, Maris, & Schoffelen, 2011). EEG was digitally filtered from 0.5-10hz using the EEGLAB finite impulse response filter, with these parameters selected based on previous p300 BCI research (Guo, Gao, & Hong, 2010). Eye and muscle artifacts were rejected using independent component analysis (ICA), a form of blind source separation that decomposes neural activity recorded at EEG electrodes into independent sources that are mutually independent (Delorme & Makeig, 2004). Components that were likely the result of movement, blink and saccade artifacts were rejected using a previously validated method utilizing kurtosis, extreme value thresholding, data improbability and linear trending (Delorme, Sejnowski, & Makeig, 2010). Remaining independent components were back-projected to electrodes.

ERPs were generated by dividing trials into epochs from -200ms to 1000ms relative to word stimuli onsets. ERPs were baseline corrected by subtracting the average pre-stimulus magnitude from the epoch.

**Cluster Mass Permutation Test**

For the following ERP analyses, a cluster mass permutation test (CMPT) was used, adapted from Cruse, Chennu, Fernádez-Espejo, Payne, Young, & Owen, 2012. The general motivation behind this procedure is the recognition that EEG recordings generate a large number of time samples within each ERP, with this number multiplied by the number of electrodes. It would require a substantial number of comparisons to compare each time-electrode sample between the two conditions, each increasing the probability of type 1 errors. Due to the need to correct for multiple comparisons, the sensitivity of the test is severely diminished. Instead, the cluster mass approach provides a test statistic that is based on clustering adjacent spatial-temporal samples. This approach was first developed for fMRI (Bullmore, Suckling, Overmeyer, Rabe-Hesketh, Taylor, & Brammer, 1999) and has since been adapted for analysis of MEG and EEG data (Maris & Oostenveld, 2007). In order to generate the test statistic, the following procedure is used:

1. For every time sample in a predefined window at every electrode, compare the EEG signal between the two conditions.

   For each participant, there will be n trials from condition one and m trials from condition two. Each trial is a matrix of time x electrode EEG voltage samples. Therefore, for each time-electrode point, there will be n samples from condition one and m samples from condition two. Perform an independent samples t-test for each time-electrode point in the

matrix between the two conditions. A temporal analysis window of 300ms to 800ms post stimulus onset was used following based on previous P300b research (Guo, Gao, & Hong, 2010).

2. Select all time-electrode points whose p-value is lower than a predefined threshold. In this study, $p<0.05$ was used, following conventional practice (Maris & Oostenveld, 2007).
3. Cluster significant points that are both spatially and temporally adjacent. Points must be temporally adjacent by immediately following one another and spatially adjacent by virtue of being recorded from neighboring electrodes
4. For each cluster, sum all t-values of significant time-electrode points.
5. Select the cluster with largest summed t-values. This sum forms the test statistic.

From this analysis, a single value is generated, referred to subsequently as the cluster mass value (CMV). In order to perform statistical analyses on the differences between conditions, a non-parametric permutation approach is taken. For comparisons between conditions at the single participant level, the following procedure is used:

1. Collect all trials of the two experimental conditions in a single set. Each trial includes the time-varying voltage recorded at all electrodes.
2. Randomly draw as many trials from this combined dataset as there are trials in condition one. Place those trials into subset one. Place the remaining trials in subset two. This results in a random partition.
3. Calculate the CMV on this random partition.
4. Repeat steps 2 and 3 10000 times. This large number of permutations allows a more precise characterization of the probability distribution and dilutes the effects of statistical anomalies.
5. Place the test statistic that was actually observed into the histogram created in step 4.

6. Calculate the proportion of random partitions that resulted in a larger CMV than the observed one to derive a p-value.

A similar approach is used for CMPT group level analysis. In order to generate the test statistic, the following procedure is used:

1. For each participant, an average ERP for each condition is calculated. Each participant is given a single matrix of time x electrode values for each condition. Time windows were restricted to 300-800ms post stimulus onset similar to single-participant level analysis.
2. Each condition consists of a set of time x electrode matrices, one for each participant. Conduct a paired-samples t-test at each time-electrode point to determine points that differ significantly between the two conditions.
3. Cluster significant points ($p<0.05$) that are both spatially and temporally adjacent. Points must be temporally adjacent by preceding or following one another and spatially adjacent by virtue of being recorded from neighboring electrodes
4. For each cluster, sum all t-values.
5. Take the largest of the cluster-level statistics.

From this analysis, a CMV is generated. In order to perform statistical analyses on the differences between conditions, a non-parametric permutation approach is used, albeit differing slightly from the previous method:

1. Within individual participants, permute the average ERPs in each condition. For example, within participant one, the value of the average ERP from condition one is reassigned to condition two, and vice versa. The participants for which this exchange takes place are selected randomly
2. Calculate the CMV on this permuted data set.

3. Repeat steps 1 and 2 10000 times. This number of permutations allows a more precise characterization of the probability distribution and dilutes the effects of statistical anomalies.
4. Place the test statistic that was actually observed into the histogram created in step 3.
5. Calculate the proportion of random partitions that resulted in a larger CMV than the observed one to derive a p-value.

In both group level and single participant level CMPTs, a p-value less than a predefined alpha level ($p<0.05$ in the current study) resulted in a rejection of the null hypothesis. As mentioned above, this test is useful because a single statistical test is conducted for each comparison between conditions, controlling for the multiple comparisons problem. In contrast to parametric approaches, this approach does not make assumptions about the distribution of the test statistic. Importantly, CMPT allows spatiotemporal localization of significant changes in electrophysiological activity in a data-driven manner.

**Group-Level Analysis**

For analysis of the AT task, ERPs were averaged within participants, as described above, for all attended word stimuli and separately averaged for all unattended word stimuli, generating 28 ERPs total including two for each participant. CMPT group level analysis was conducted in order to test for significant differences between conditions.

For analysis of the WM task, ERPs were averaged, as described above, for all correct word stimuli and separately averaged for all incorrect word stimuli within each participant. For example, in a particular WM trial, if “dot” was the correct answer, the ERP responses to “dot” would be added to the correct condition regardless of whether it was also the stimuli that the participant attended to, while ERP responses to all other word stimuli were added to the incorrect

condition. The logic of this approach is that, should a participant solve problems significantly above chance, they will attend to the correct word stimulus and P300b ERPs will accumulate in the correct condition while non-P300b ERPs (of mean magnitude zero) will accumulate in the incorrect condition, leading to a significant difference between the two conditions. However, should a participant solve problems at chance, P300b ERPs will be assigned with equal probability to both correct and incorrect bins, leading to a null result.

In the AR task, ERPs to correct word stimuli were compared to ERPs to the word stimuli that formed the other possible answer within the sentence. For example, if the sentence used was "cake precedes dot," ERPs to the "cake" word stimuli were added to the correct stimuli condition while ERP responses to "dot" were added to the incorrect condition. This approach was used because a participant could potentially cheat and listen for both words in each trial. If all stimuli other than the correct word stimuli were added to the incorrect bin, even though one of the incorrect stimuli ERPs contained a P300b it would be diluted by the other incorrect stimuli and be significantly lower in magnitude, giving rise to a positive result when the participant did not solve the problem correctly. Adding only the ERP to the incorrect stimuli in the sentence to the incorrect condition avoids this issue.

**Single-participant level EEG analysis**

For analysis of the AT task at the single participant level, ERPs to attended word stimuli were compared to ERPs to the unattended word stimuli using the single participant CMPT outlined above. For analysis of the WM task, ERPs to correct word stimuli were compared to ERPs to incorrect word stimuli using single participant CMPT analysis. For analysis of the AR

task, ERPs to correct word stimuli were compared to ERPs to incorrect word stimuli within the sentences using single participant CMPT analysis.

**Prediction of individual differences from P300b responses**

The purpose of this analysis was to explore whether the magnitude of P300b responses predicted performance as indicated by behavioral measures. As discussed below, participants performed largely at a ceiling level with no significant differences between conditions in the AR task. Consequently, differences between participants were predicted from ERPs only in the WM task.

The theory motivating this analysis is that, should a participant perform at 100%, their ERP magnitude to correct word stimuli would be equal to their ERP magnitude to attended stimuli, as all correct stimuli would also be attended stimuli. Conversely, if the participant performed at 50%, only half of the correct word stimuli would also be attended. Given that half of the ERPs in the correct condition will be P300b ERPs while half will be non-P300b ERPs of average magnitude zero, the correct ERP will be half as large as the attended ERP magnitude. If a performance coefficient is calculated between zero and one, the resulting correct ERP magnitude should be the attended ERP magnitude multiplied by this coefficient. For example, if the attended P300b magnitude is 4μv for a participant who performs at 75%, the correct P300b magnitude should be 3μv.

In order to test this predictive model, ERP magnitudes were first calculated within each participant at each WM difficulty level (four, six, or eight item memory sets), resulting in three ERPs for each participant. Time x electrode values that were in the spatiotemporal regions

selected by the CMV were averaged to calculate a mean magnitude within each of the three conditions for each participant. This magnitude was then divided by the average ERP magnitude to attended stimuli from the AT task to calculate normalized ERP magnitude (NM). A NM was calculated for each condition in each participant, resulting in 42 NM total. Each NM magnitude had a paired performance score calculated from behavioural data.

The unique and combined relationship between set size, NM and performance was analyzed using a generalized linear model with memory set size as a factor and NM as a covariate. The purpose of this test was to first examine whether NM could be used predict individual differences in performance overall. The analysis also modeled the prediction of performance from NM within difficulty levels (by factoring out the effects of difficulty level on performance). Secondly, this test was able to analyze whether interactions existed between set size and NM such that NM was more predictive of performance depending on the level of difficulty. Following this analysis, a correlation test was performed between NM and performance within each difficulty level to examine the nature of the linear relationship at each level.

Due to the large differences in variance of participant performance in memory sets with eight items (discussed below), an additional single participant level analysis CMPT was conducted with only four and six item memory set trials included.

**Relationship between P300b Significance, Time and Performance**

As discussed above, the normalized magnitude of the P300b response to correct stimuli should vary linearly with performance. Therefore, the p-value for differences between ERP responses to correct vs. incorrect word stimuli should also vary with performance such that better

performance decreases the p-value and increases statistical confidence that the participant is able to perform the task. Similarly, as task time increases, the number of stimuli in each condition likewise increases, also decreasing the p-value and adding to statistical confidence. Furthermore, these two variables are related. Better performance decreases the amount of time required to attain a significant p-value, while worse performance increases the time required to achieve the same p-value. The nature of this relationship is crucial to the purposes of this paradigm, as these parameters determine the sensitivity of the test to detect accurate performance, or lack thereof, as well as the length of time required for the battery to reach a significant level of confidence.

The relationship between task time, performance and p-value in the overall task was modeled using a Monte Carlo simulation. The Monte Carlo method is broadly used to simulate complex stochastic processes, and is useful in this case to explore hypothetical experimental outcomes and their effects on statistical results using the existing data. The following Monte Carlo procedure was used:

1. Within each participant, ERP responses to stimuli that were actually attended were collected by selecting ERP responses to word stimuli that corresponded to the participants' behavioral response in each trial.
2. A random selection of n (where n increases with time on task) attended ERPs were selected from the attended ERP set, while 7*n ERP responses were randomly selected from the unattended ERP set.
3. A single-participant level CMPT was performed between these two sets to attain a p-value.
4. For each value of n, steps two and three were repeated 100 times to decrease the effect of particular selections on the resultant p-value.
5. These 100 p-values were averaged to attain mean p-value at that n for the given participant.

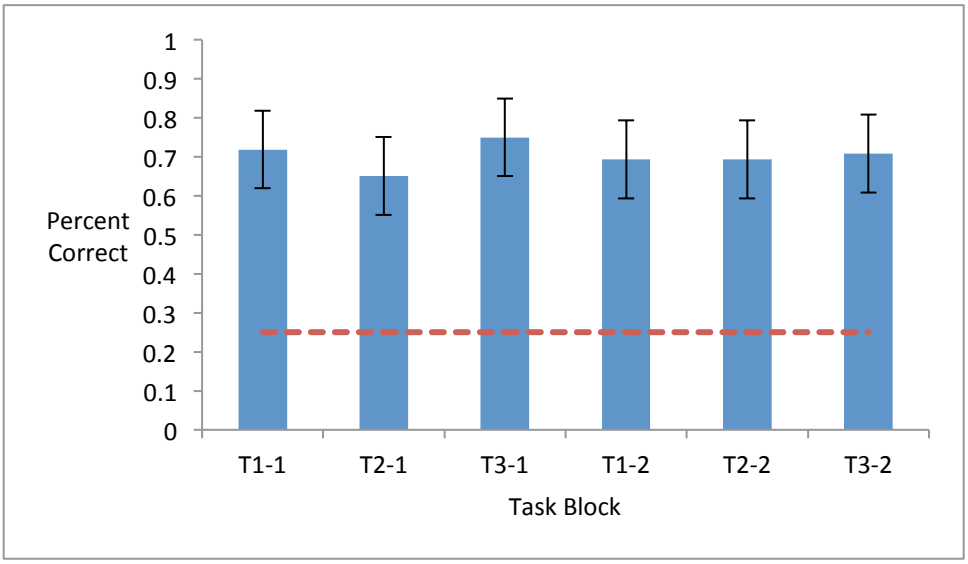
6. The value of n was increased in multiples of 25 to simulate increasing time, with steps 2-5 repeated at each value of n.

The effects of performance were simulated by inserting an intervening step between steps 2 and 3. Attended and unattended ERPs were swapped between conditions depending on simulated performance. For example, if the simulated performance was 0.6, 40% of the attended ERPs were randomly swapped for an equal number of unattended ERPs between conditions. Performance levels of 0.3 to 1 were used in increments of 0.1.

## 3. Results

### Behavioral Results

Overall, participants' counting accuracy was 70% across all blocks (SD=0.03). Across all blocks, all participants scored above chance. A repeated measures ANOVA showed that block number did not have a significant effect on counting accuracy, $F(5,65) = 1.985$, $p = .092$ (Figure 15). Repeated measures ANOVA also showed that task type did not have a significant effect on counting accuracy, $F(2,26) = 1.985$, $p = 0.15$.




.

Figure 15: Counting accuracy as a function of block number and task type. Chance performance is represented by the dotted red line. Error bars represent standard error.

On the WM task, participants' averaged 81% correct across all three set sizes (SD = .16). Participants averaged 97% correct (SD = 0.05) for four item sets, 80% correct (SD = .16) for 6 item sets, and 64% correct (SD = .23) for eight item sets (Figure 16). All participants scored above chance at all difficulty levels.

Repeated measures ANOVA showed a significant effect of memory set size on the accuracy of recall, $F(2,26) = 22.701$, $p < 0.001$ (Figure 16). Mauchly's test indicated that the assumption of sphericity had been violated. $X^2(2) = 7.268$, $p = 0.026$. Data were Greenhouse-Geisser corrected ($\varepsilon = 0.68762$) to account for this violation, again showing memory set size to have a significant effect on accuracy of recall ($p < 0.001$).

Individual comparisons of accuracy between memory set sizes also revealed significant differences, with 4 item sets recalled significantly better than 6 items ($p < 0.001$) and 8 items ($p<0.001$), and 6 items sets remembered significantly better than 8 items ($p = .005$).

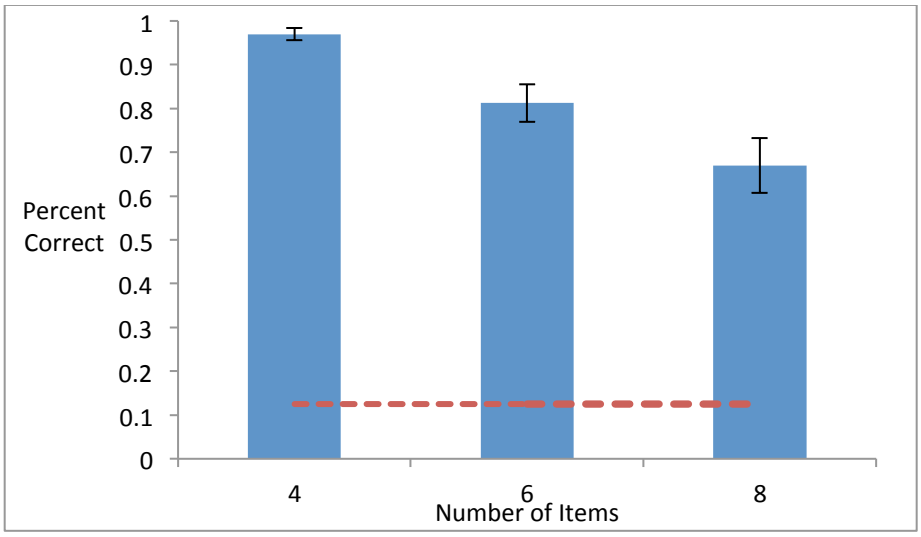



Figure 16. Working memory performance within each memory set size. Chance performance is represented by the dotted red line. Error bars represent standard error.

For the AR task, participants averaged 95% correct (SD = .03) across all sentence types. All participants scored above chance at all levels, other than participant 15 who answered 2/3 active, negative sentences using "follows" incorrectly. A 3-way, repeated measures ANOVA revealed no significant main effect of the verb used (precedes VS follows), $F(1,13) = -.11$ , $p > 0.05$, negation, $F(1,13) = .51$, $p > 0.05$, or passive VS active sentences, $F(1,13) = 0.21$, $p > 0.05$. There were no significant 2 or 3-way interactions (Figure 17). The absence of significant differences in this task was likely the result of ceiling performance.

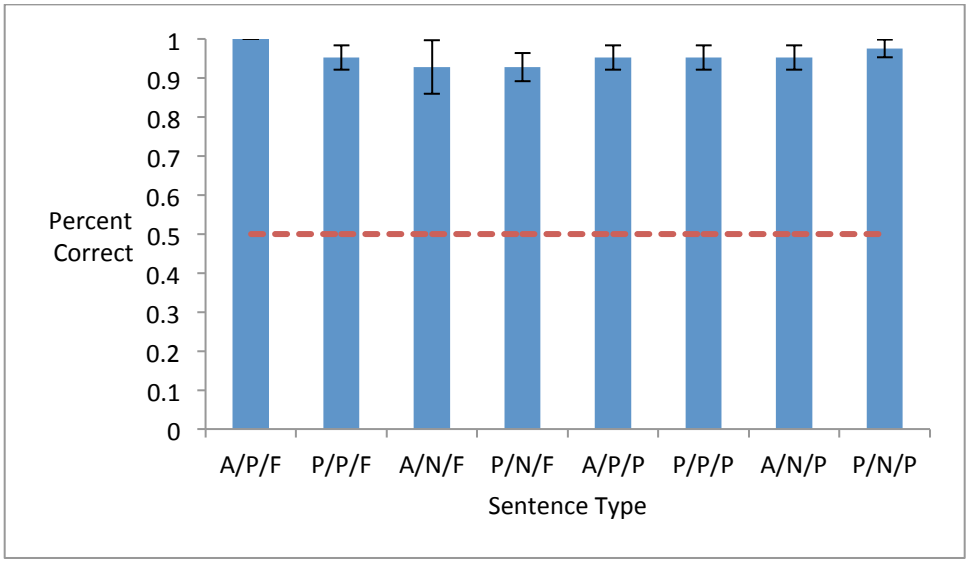



Figure 17. Performance in verbal reasoning task. Sentences are coded in figure as active/passive, positive/negative and follows/precedes. Chance performance is represented by the dotted red line. Error bars represent standard error.

## EEG Results

### Group level

At the group level, the CMPT revealed that ERPs were significantly larger to attended word stimuli than unattended word stimuli in the AT task ($p < 0.001$). CMPT also revealed significantly larger ERPs to correct word stimuli compared to incorrect word stimuli in both the

WM task (p = 0.002) and the AR task (p = 0.003). Group-averaged topomaps showed that for all tasks the P300b response was most prominent in posterior electrodes (Figure 18).

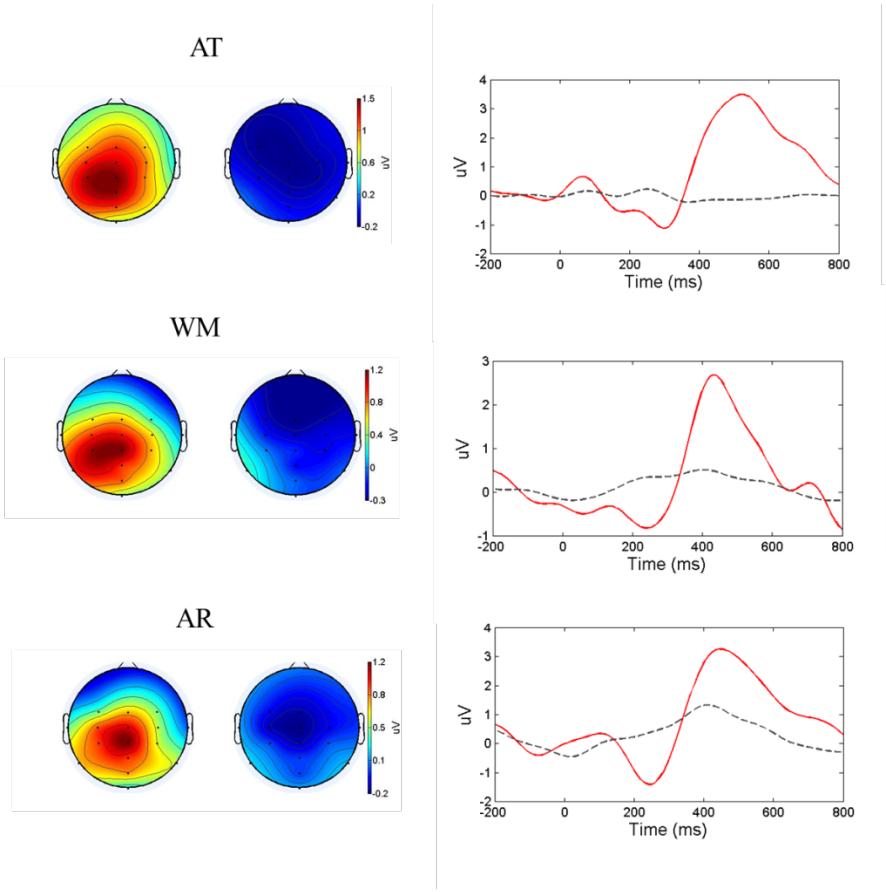



Figure 18. Group average scalp maps in all three tasks at 450ms post stimulus onset. Attended (AT) and correct (WM and AR) topomaps on left side, unattended (AT) and incorrect (WM and AR) topomaps on right. ERP time courses for attended and correct (red line) vs. unattended and incorrect (black line) over electrode CPz.

**Single-participant level**

CMPT at the single participant level for the AT task revealed a significant difference between ERPs to attended stimuli vs. unattended word stimuli for all participants (Figure 19).

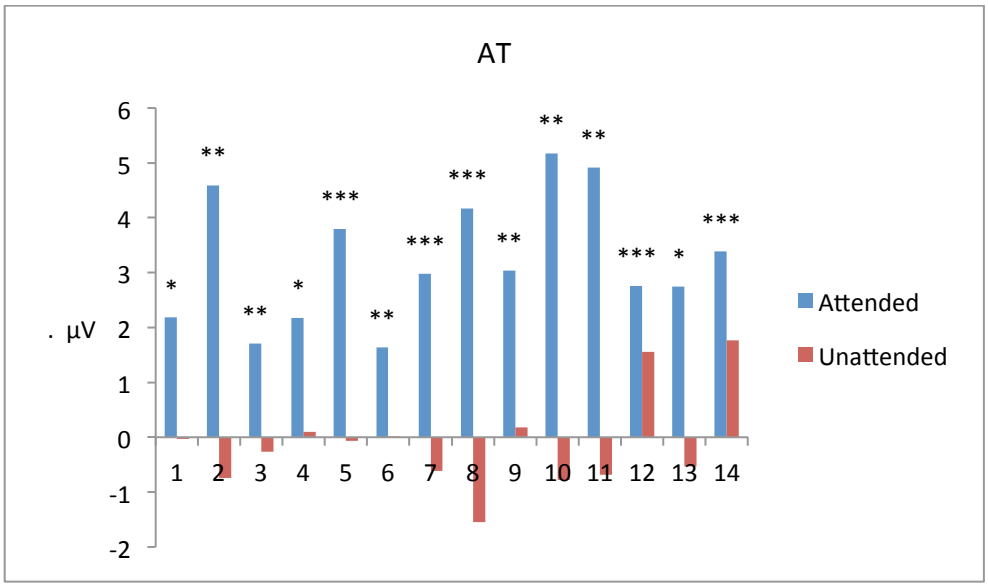



Figure 19. Mean voltage magnitude within largest CMPT cluster for each participant in AR task. *** p <0.001, **p<0.01, *p<0.05. µ

Results from CMPT at the single-participant level for the WM task revealed a significant difference between ERPs to correct VS incorrect word stimuli in 11/14 participants (Figure 20).

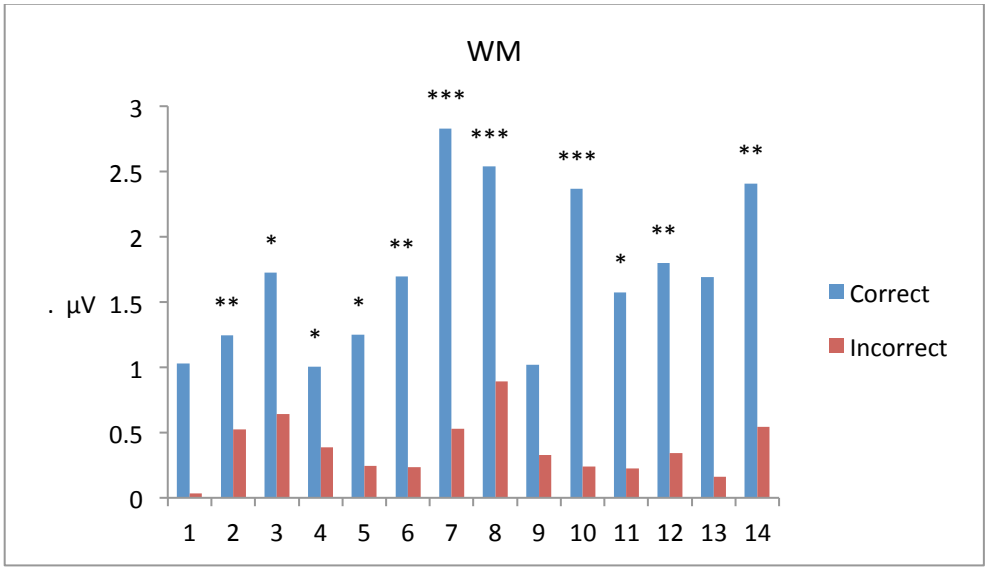



Figure 20. Mean voltage magnitude within largest CMPT cluster for each participant in WM task. *** p <0.001, **p<0.01, *p<0.05.

Results from CMPT at the single-participant level for the AR task revealed a significant difference between ERPs to correct VS incorrect word stimuli in 13/14 participants (Figure 21).

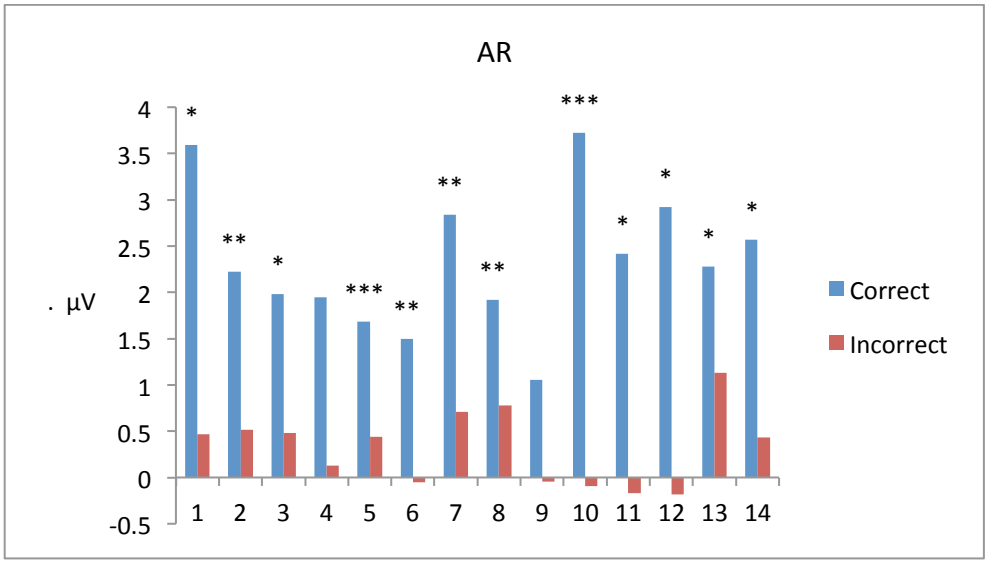



Figure 21. Mean voltage magnitude within largest CMPT cluster for each participant in AR task. *** p <0.001, **p<0.01, *p<0.05.

**Prediction of Individual Differences from P300b Components**

Results of the generalized linear model demonstrated that memory set size significantly predicted performance, $X^2(2, n = 42) = 15.123$, $p < 0.001$, with performance decreasing as memory set size increased. Importantly, normalized ERP magnitude (NM) predicted performance even when the general effect of set size was factored out, thus demonstrating that NM predicted participant differences within individual memory set sizes; $X^2(1,n = 42) = 6.742$, $p = 0.009$ (Figure 22). The generalized linear model also revealed a significant interaction between memory set size and normalized ERP magnitude, $X^2(2,n = 42) = 6.149$, $p = 0.049$, suggesting that the predictive power of ERP magnitude was modulated according to memory set size.

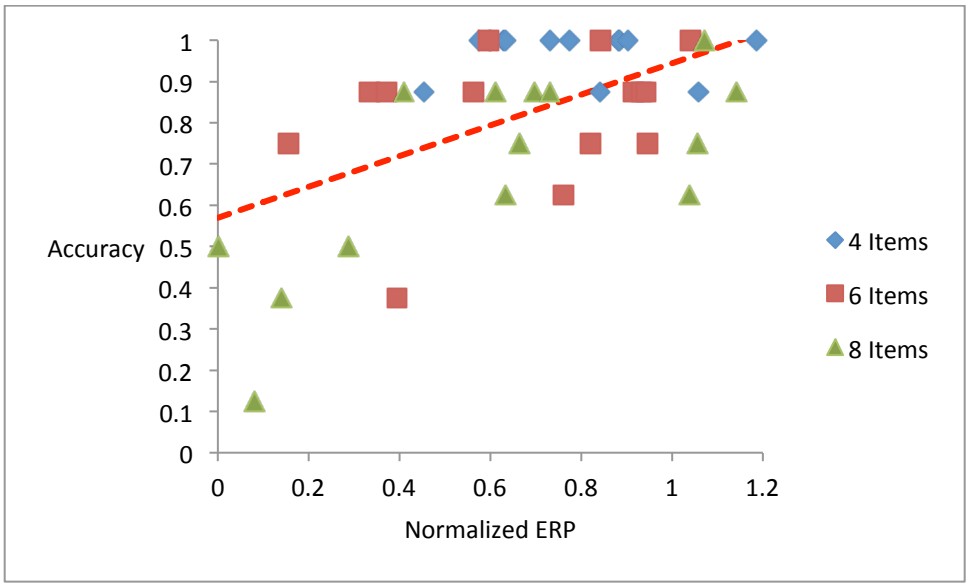



Figure 22. Relationship between accuracy within memory set sizes and normalized ERP magnitude. Each data point represents the overall accuracy of a single participant in a single WM set size condition. Red dotted line represents line of best fit.

In order to explain the linear relationship between NM and performance within each memory set size, three correlation tests were performed, one at each memory set size. NM was not significantly correlated with accuracy within four item sets (Figure 23), $r(12) = -0.44$, $p>0.05$ or six item sets (Figure 24), $r(12) = 0.321$, $p > 0.05$. NM and accuracy were significantly correlated within the eight item set size (Figure 25), $r(12) = .712$, $p = 0.002$.

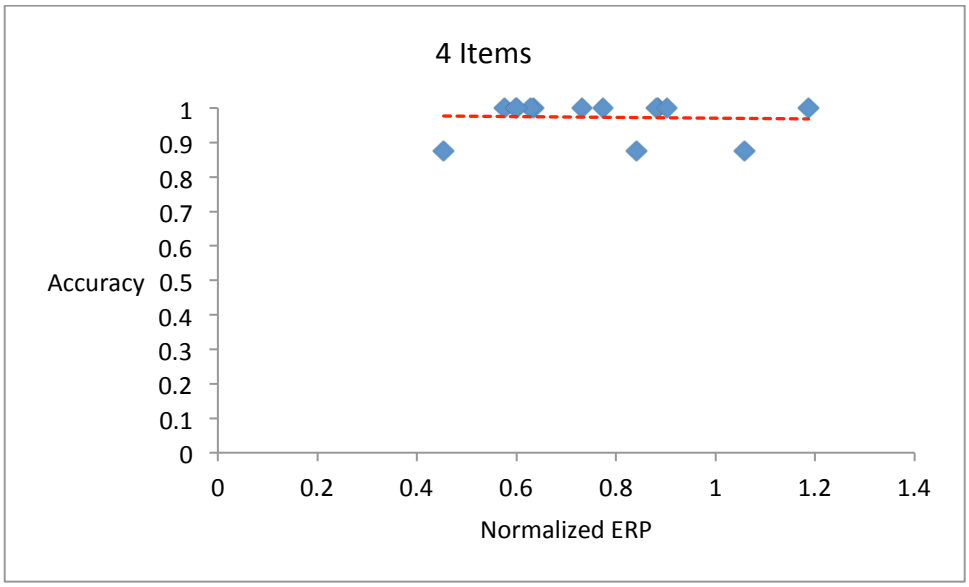



Figure 23. Relationship between accuracy within four item memory sets and normalized ERP magnitude. Each point represents data from one participant. Red line represents line of best fit.

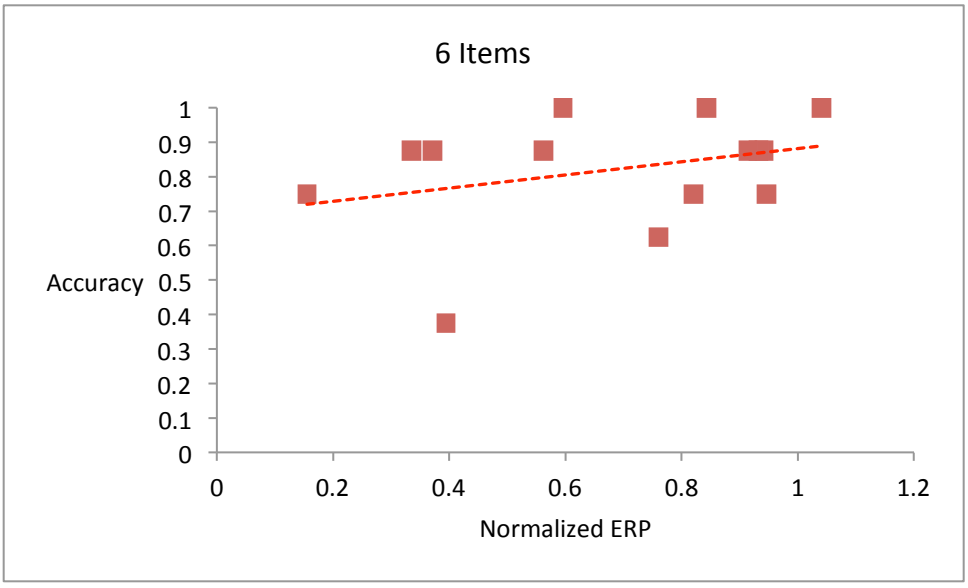



Figure 24. Relationship between accuracy within six item memory sets and normalized ERP magnitude. Each point represents data from one participant. Red line represents line of best fit.

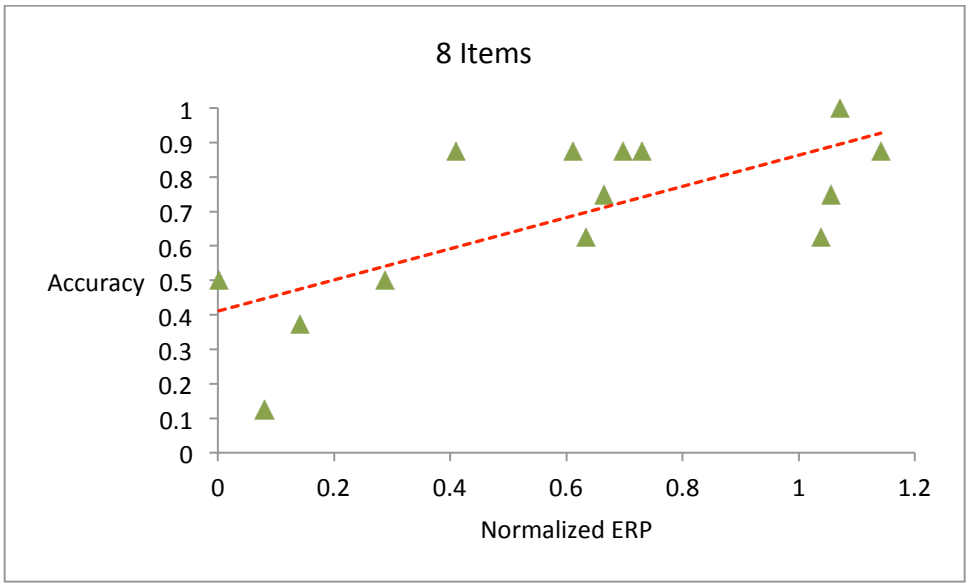



Figure 25. Relationship between accuracy within eight item memory sets and normalized ERP magnitude. Each point represents data from one participant. Red line represents line of best fit.

Due to increased variance of participant performance in the eight item memory sets, in the initial CMPT analysis the attended word stimuli often did not correspond to the correct word stimuli. The CMPT analysis was not significant for three participants, which may have been caused by the inclusion of incorrect trials from the eight item set. In order to explore this possibility, a second CMPT analysis was conducted with the inclusion of word stimuli from only four and six item memory sets. The restricted CMPT revealed a significant difference between ERPs to correct and incorrect word stimuli in all 14 participants (Figure 26).

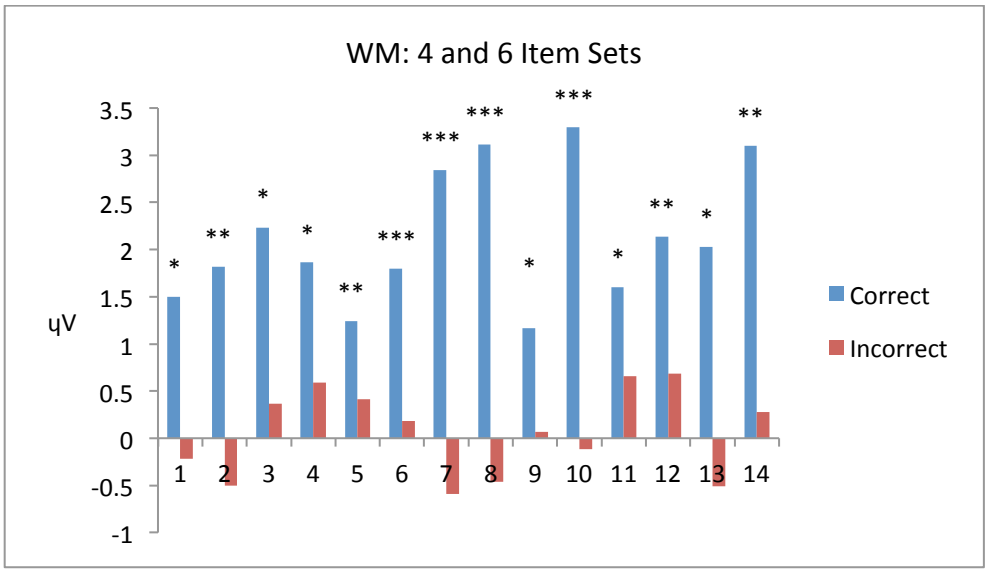
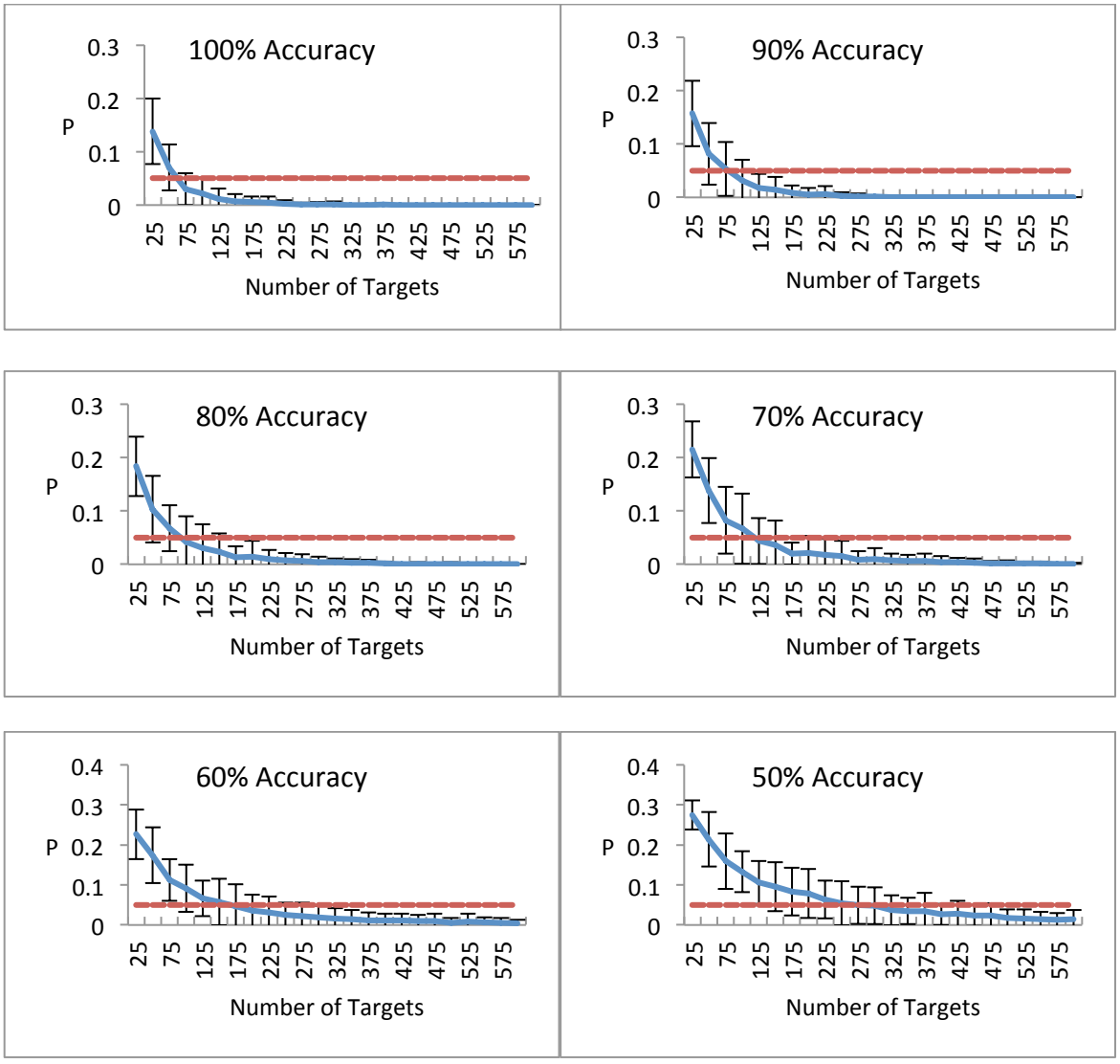



Figure 26. Mean voltage magnitude within largest CMPT cluster for each participant in WM task, restricted to four and six item sets. *** p <0.001, **p<0.01, *p<0.05.

**Significance as a function of number of Targets**

An exploratory analysis was conducted to investigate the combined effects of task duration and subject accuracy on the statistical sensitivity of the test (Figure 27). At 100% accuracy, mean p-value was less than 0.05 after 75 attended stimuli, with large variation between participants. As the number of stimuli increased, p-value decreased asymptotically to zero while variance likewise decreased. Given that each attended stimuli plus interstimulus interval is 500ms and it always accompanies seven unattended stimuli of the same duration, the average amount of time per attended stimuli is approximately four seconds. 75 attended stimuli therefore take approximately five minutes to deliver, not counting the time taken to pose questions within trials. As accuracy decreased, a larger number of stimuli were required to reach the same level of significance. At 90% accuracy, 100 attended stimuli were required to achieve the same p-value. At 60% accuracy, 175 attended stimuli were required, while at 30% accuracy a p-value of less than 0.05 was unattainable with 600 attended stimuli. In principle, because chance performance

was 12.5%, any accuracy above this threshold should be detectible with an arbitrarily large number of stimuli. However, factors such as participant fatigue and changes in electrode placement and impedance place an upper limit on the number of stimuli that can be delivered.



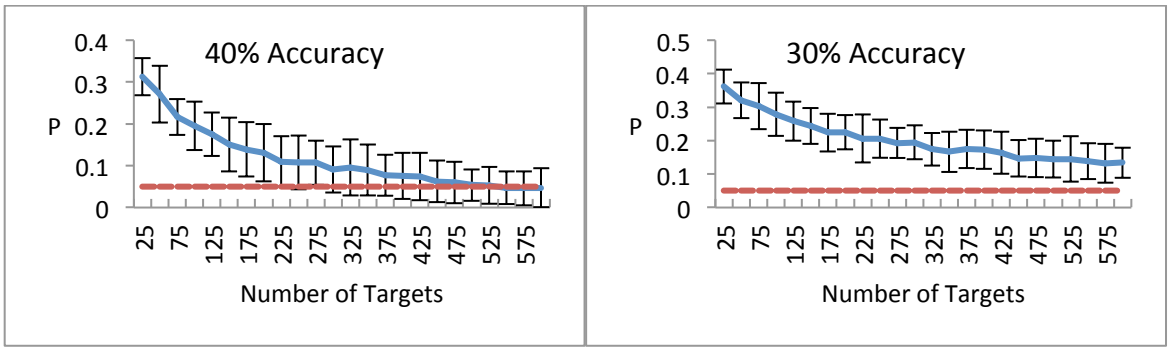



Figure 27. Mean CMPT p-value for all participants as a function of number of targets. Delivery of 25 targets takes approximately one minute and forty seconds. Error bars represent standard error. Red dotted line represents p=0.05.

## 4. Discussion

### General Battery Performance

Overall, the battery was largely successful and has the potential to be used for directly assessing executive function in DOC patients, making it the first battery of its kind. At the group level, both the WM and AR tasks generated a significant P300b to correct word stimuli, demonstrating that overall performance across participants was significantly above chance. At the single-participant level results were the most promising. A significant P300b was found in 11/14 participants in the WM task and 13/14 participants in the AR task. When 8 item memory sets were removed from analysis, P300b responses from all 14 participants were significant in the WM task. The adjusted battery was therefore able to detect correct performance in tasks requiring executive function in 27/28 cases using the P300b ERP. In the AT task, a significant P300b response was found in all participants analyzed, demonstrating that the P300b has the potential to be used as a command following paradigm for detecting residual awareness in DOC patients, with detection rates significantly higher than those found using other BCI approaches (Dias, Kamrunnahar, Mendes, Schiff, & Correia, 2007). This approach could be used as a

complementary means of detecting volition in addition to motor imagery paradigms developed previously. In standard behavioral DOC assessment batteries, a variety of motor output channels are used to examine command following to rule out the possibility that damage to specific channels obscures the patient's ability to follow instructions (Kalmar & Giacino, 2006). An analogous approach should be taken when using active paradigms that employ volitional modulations in neural activity. Damage to the motor system might prevent the patient from performing motor imagery. Likewise, damage to top down attention systems could prevent the patient from attending to target stimuli. By utilizing several command following paradigms to target a broad set of neural systems, a larger proportion of patients can be adequately assessed.

The second hypothesis of the study was that the battery would be able to predict individual differences in performance using the P300b response and results here were somewhat mixed. In the AR task, ceiling performance precluded meaningful variation in performance to predict. In the WM task, there was a significant relationship between ERP magnitude and accuracy. This result was largely driven by the variance in the 8 item memory sets, however, and ceiling performance in 4 item and 6 item memory sets again reduced individual differences. Within the 8 item memory set, there was a reasonably high correlation, suggesting that more difficult tasks that increase variability should result in higher predictive accuracy, though further research is required in order to better confirm this hypothesis.

**Insignificant participant level results in WM and AR tasks**

One of the shortcomings of the battery was that, despite showing a positive result for all participants in the AT task, results from three participants were not significant in the WM task and the result from one participant was not significant in the AR task. The battery did not have

the sensitivity to detect covert cognition across all participants in all tasks, and may be somewhat prone to false negatives, given that all participants can be assumed to have been performing the tasks as requested. This result is unsurprising however, as other BCI EEG paradigms have also failed to be effective for all participants studied, even among healthy individuals (Cruse, et al., 2012; Guger, Edlinger, Harkam, Niedermayer, & Pfurtscheller, 2003). Even taking these false negatives into account, this battery had an excellent rate of detection relative to similar paradigms, and when 8 item working memory sets were removed from analysis, the sensitivity was near perfect. Furthermore, the potential for type 2 errors is common to all active paradigms that attempt to use neural responses to signal cognitive abilities in patients. A null result does not definitively demonstrate a deficit in the faculty tested but could instead be due to lack of statistical power, lack of participant cooperation, fatigue, or deficits in other faculties such as language comprehension or sustained attention. Rather than being able to decisively show that patients lack specific abilities, the power of these approaches is that they can demonstrate cognitive faculties in patients who are assumed to lack higher levels of cognition altogether. In the vast majority of cases, the approach developed here successfully detected cognitive function in healthy participants.

**CMPT analysis used in this study**

One methodological departure from past research was the use of the cluster mass permutation test (CMPT) to examine statistical differences in neural activity between conditions. The virtue of this approach is that it makes fewer assumptions about the spatial and temporal nature of ERP responses. DOC patients often have significantly altered brain morphology that changes the spatial pattern of neural activity. Focusing analysis on a single electrode or set of

electrodes may result in insensitivity to atypical yet significant changes in electrophysiology. Likewise, examining activity at all electrodes has concomitant drawbacks. If statistics are not corrected for multiple comparisons, the likelihood of a false positive increases. Correcting for multiple comparisons, on the other hand, severely diminishes the sensitivity of the test. The CMPT test allows for differences in spatial patterns of activity by selecting the region of analysis in a data-driven manner without compromising statistic rigor. In addition, temporal patterns of EEG activity may be altered in patients. In many ERP studies, time windows are predetermined based on previous research. If this approach is taken with patients, systematic differences in ERP latencies may lead to type 2 errors. One approach that attempts to account for these differences uses ERP grand averages across experimental conditions to find a local maximum and centre the analysis window (Perrin, et al., 2006). Given that the grand averaged ERP magnitude is orthogonal to the magnitude of individual conditions, this approach allows the localization of the temporal window without increasing type 1 errors. However, because it relies on a single temporal sample, spurious increases in magnitude can lead to inappropriate analysis windows. The CMPT approach allows a suitable temporal window to be found using clusters of activity rather than single data points, incorporating the advantages of the grand average approach while avoiding its drawbacks.

**Assessment of Conscious Awareness in DOC**

An important question in DOC research and cognitive science revolves around the particular cognitive abilities that are sufficient to demonstrate consciousness. A wide variety of cognitive processes can potentially occur unconsciously. Demonstration of these processes in patients therefore does not confirm that conscious awareness remains. On the other hand, if a

cognitive process requires conscious awareness to occur, demonstration of this process in a patient is sufficient to establish awareness. In a review of a wide range of studies in cognitive psychology, two types of cognitive operations were identified that appear to robustly require conscious processing (Dehaene & Naccache, 2001). The first was durable information maintenance. In contrast to unconscious priming effects, which tend to decay exponentially and lose a majority of their influence within 100 milliseconds, information that is consciously attended can be held online for much longer periods (Greenwald, Drain, & Abram, 1996). Secondly, combining several cognitive operations to perform a novel or atypical task also appears to require conscious control.

The tasks used in the current study require both of these process types for successful completion. The WM task directly requires the maintenance of multiple items over a 10 second period and likely cannot be performed using unconscious priming mechanisms. In both tasks, using the solution to a problem to determine the locus of selective attention is a highly novel task that requires the combination of several cognitive processes. Due to the complexity of the tasks and their satisfaction of established criteria for conscious processing, demonstrating successful completion of the battery would provide a powerful argument for conscious awareness in DOC patients.

**Ethical Considerations**

The development of the current battery as well as others like it raises important ethical and legal considerations. As shown in several studies, a subset of patients are able to command follow by modulating neural activity, and efforts are underway to develop long term BCI systems for communication with these patients (Cruse, et al., 2012; Owen A. M., Coleman, Boly,

Davis, Laureys, & Pickard, 2006). In tandem with these advances, questions will naturally arise as to the degree of autonomy that should be granted to communicative patients in determining their medical care and managing their lives. The scope of these decisions is wide, ranging from details of routine medical care to estate planning and even the potential for ceasing life-sustaining medical interventions. Autonomy is not a binary phenomenon, but is granted to patients in proportion to their level of function, as assessed through a combination of expert opinion and standardized testing procedures (Etchells, Sharpe, Elliot, & Singer, 1996). Given the limited bandwidth of BCI systems and absence of many of the behavioral cues used in capacity assessments, standardized testing procedures carry increased importance for assessing DOC patients. Moreover, executive function has long been a crucial construct used in determining capacity, as complex ideas must be comprehended and integrated with information presented at multiple time points for decisions to be made competently (Etchells, Sharpe, Elliot, & Singer, 1996; Marson & Harrell, 1999). By directly assessing reasoning and working memory, the battery developed here represents a first attempt and proof of principle that these higher cognitive functions can potentially be detected in DOC patients using widely available EEG systems. Due to the limited sample size and homogeneity of participants tested, further work is required before this approach can be applied confidently. In addition, although necessary, executive function in itself is not sufficient for demonstrating capacity, but must exist in parallel with proper orientation in space and time and the ability to appreciate the consequences of decisions. Nonetheless, as part of a broader set of tests targeting a range of faculties, the battery developed here has the potential to provide a reliable tool for capacity assessment in behaviorally unresponsive patients, grounding ethical decision making in sound science.

**Extension and Optimization Of the Assessment Battery**

Although the main purpose of the battery was to examine executive function, the logic of this paradigm could be extended to many other neuropsychological tasks. The link between the task that generates the stimuli to be attended and the P300b counting task is arbitrary. In principle, any neuropsychological task where the answer is one of several options could be linked to a similar P300b counting task, allowing the assessment of the cognitive faculties necessary for its performance. Based on this method, a battery consisting of a wide variety of neuropsychological tasks could be developed, allowing the fine-tuned examination of cognitive deficits in DOC patients. Moreover, there is no reason that this approach should be limited to DOC. There is currently widespread debate about the extent of cognitive disruption in neurological disorders that primarily affect the motor system, in particular Amyotrophic Lateral Sclerosis (ALS) (Neary, Snowden, & Mann, 2000). It is in the final stages of ALS that cognition is most likely to be affected, when motor function is most severely compromised. Extending the battery developed here to other patient groups could allow fruitful insights into the nature and prevalence of cognitive dysfunction in ALS and other neurological conditions.

Another important avenue for further research involves optimizing experimental factors to maximize the information that can be gained from the battery in the shortest period of time. Many of these parameters were chosen from the outset based on past research (i.e., Guo, Gao, & Hong, 2010), such as interstimulus intervals and the particular word stimuli included. However, research directly manipulating and testing these parameters has the potential to significantly reduce the time required to deliver the battery. One parameter of particular importance is the number of word stimuli to use, each representing a potential multiple-choice answer. In general, as the number of choices increases, the statistical confidence that a participant is solving

problems skillfully rather than by chance increases. For example, if a participant solves a multiple-choice problem with two options, there is a 50% probability they arrived at the answer by chance. On the other hand, if there are eight potential options, there is only a 12.5% probability they solved the problem by chance. At the same time, if posing two options takes a quarter of the time of eight options, four questions could potentially be posed in the same amount of time, putting chance probability at $1/2^4 = 6.25\%$, assuming they were correct each time. However, in the context of P300b responses, research has shown that the amplitude of the ERP decreases as the number of non-target options decreases, complicating this relationship (Polich, Frequency, intensity, and duration as determinants of P300 from auditory stimuli, 1989). As with most parameters, it is difficult to determine the optimal specification a priori, and further research should be conducted in this direction to assist in reducing patient fatigue and maximizing the likelihood of detecting accurate performance.

**Detecting Significant Performance vs. Assessing Individual Differences**

In the current study, two related but distinct hypotheses were tested. The first was whether executive function could be detected using the P300b as a response mechanism. The presence of a significant difference between correct and incorrect ERPs confirms this hypothesis. In terms of accuracy in the task, near ceiling performance is ideal for attaining this result in a timely and robust manner, as demonstrated by the Monte Carlo simulation. In the AR task as well as the smaller WM set sizes, this difference was significant for almost all participants.

The second hypothesis was that individual differences in performance could be predicted from normalized ERP magnitude. In theory, the magnitude of the normalized ERP should vary linearly with performance, however, due to significant noise present in the ERP signal this

relationship is far from perfect. Consequently, unlike detection of above-chance performance, significant divergence in participant performance is necessary to assess individual differences.

In general, the ability to detect basic executive functions is much more important than characterizing normative performance in patients, at least at the outset. As DOC patients are presumed to have at most minimal levels of consciousness, demonstration of higher cognitive functions would profoundly challenge a patient's diagnosis. However, once the presence of these functions is established, providing a more fine-grained analysis of the patient's particular capacities is instrumental in determining the suitable amount and complexity of information to present, allowing an appropriate level of patient autonomy, and facilitating comparisons across patients.

Unfortunately, the particular conditions that maximize the likelihood of optimal detection vs. assessment are in tension in the current paradigm. Ceiling performance is required for detection, while variability in performance across conditions is necessary for assessment. In order to accommodate both of these motivations, a modified paradigm is recommended for future exploration. Rather than using a randomly presented, predefined number of trials at each difficulty level, working memory or reasoning problems should be presented in order of difficulty. Furthermore, the presentation of problems should be controlled dynamically in concert with real-time statistical analysis of ERPs. At the beginning of the neuropsychological battery, problems at the lowest level of difficulty should be presented first while ERPs to correct vs. incorrect word stimuli are compared online as data is collected. Once the statistical difference between conditions reaches a predefined threshold, the patient can proceed to a higher level of difficulty, with more challenging problems presented. Likewise, the statistical power for determining a lack of difference can be calculated in real-time and given a similar threshold for

determining that the patient cannot perform at that level. Similar to other dynamic neuropsychological tests, the last difficulty level at which a patient can perform satisfactorily can be taken as their capacity.

This approach has several advantages. Normalizing ERPs is not required for estimating performance, eliminating the need to establish baseline ERP responses during each testing session. Likewise, because difficulty is increased as soon as significance is reached, extraneous time need not be spent establishing performance at lower levels. Lastly, this procedure would better accommodate individual differences in ERP discernibility. As shown in Figure 27, participants varied significantly in the number of trials required for a significant result. Using a set number of trials therefore expends unnecessary time with some participants, while failing to detect a valid difference in others. Likewise, as demonstrated in the Monte Carlo simulation, a patient performing at 70-90%, though still acceptable, may require addition trials to reach significance. Given the variability in EEG activity and patient characteristics, a testing paradigm that adapts to the patient should be adopted, both in the present battery as well as other active neuroimaging assessment paradigms.

## 4. Conclusions

This study developed and evaluated a battery of neuropsychological tests that can be administered to behaviorally unresponsive patients using the P300b ERP component. In the majority of participants, the ability to perform tasks requiring executive function was detected without the need to rely on motor output. The magnitude of the P300b component was related to individual differences in performance, but only with sufficient variability between participants. Using Monte Carlo simulations, it was demonstrated that the battery could detect significant

performance with a mean time of five minutes, with the potential to be shortened with better optimization. As communication with DOC patients using BCIs becomes widespread, it will become increasingly necessary to assess residual cognitive function for both ethical and scientific purposes. As part of a larger battery of neuropsychological tests, the approach developed here has the potential to provide a DOC standardized assessment tool for clinicians and scientists.

## References


Andrews, K., Murphy, L., Munday, R., & Littlewood, C. (1996). Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit. *British Medical Journal , 313*, 13-16.

Ashwal, S., & Cranford, R. (1995). Medical Aspects of the Persistent Vegetative State: A Correction. *New England Journal of Medicine , 333*, 130.

Baddeley, A. D. (1968). A three-minute reasoning test based on grammatical transformation. *Psychometric Science , 10*, 341-342.

Becker, K. J., Baxter, A. B., Cohen, W. A., Bybee, H. M., Tirschwell, D. L., Newell, D. W., et al. (2001). Withdrawal of support in intracerebral hemorrhage may lead to self-fulfilling prophecies. *Neurology , 56* (6), 766-772.

Bekinschtein, T. A., Coleman, M. R., Niklison III, J., Pickard, J. D., & Manes, F. F. (2009). Can electromyography objectively detect voluntary movement in disorders of consciousness? *Journal of Neurology, Neurosurgery and Psychiatry , 79*, 826-828.

Bekinschtein, T., Niklison, J., Sigman, L., Manes, F. R., Leiguarda, R., Armony, J., et al. (2004). Emotional processing in the minimally conscious state. *Journal of Neurology, Neurosurgery and Psychiatry , 75*, 788.

Bernat, J. L. (2006). Chronic disorders of consciousness. *The Lancet , 367*, 1181-1192.

Bruno, M. A., Soddu, A., Demertzi, A., Laureys, S., Gosseries, O., Schnakers, C., et al. (2010). Disorders of consciousness: Moving from passive to resting state and active paradigms. *Cognitive Neuroscience , 1* (3), 193-203.

Buchner, H., & Schuchardt, V. (1990). Reliability of electroencephalagram in the diagnosis of brain death. *European Neurology , 30* (3), 138-141.

Buechler, C. M., Blostein, P. A., Koestner, A., Hurt, K., Schaars, M., & McKernan, J. (1998). Variation among trauma centers' calculation of Glasgow Coma Scale score: results of a national survey. *Journal of Trauma , 45* (3), 429-432.

Bullmore, E., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging , 32*, 32-42.

Childs, N. L., Mercer, W. N., & Childs, H. W. (1993). Accuracy of diagnosis of persistent vegetative state. *Neurology , 43* (8), 1465.

Coleman, M. R., Davis, M. H., Rodd, J. M., Robsin, T., Ali, A., Owen, A. M., et al. (2009). Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. *Brain , 132*, 2541-2552.

Coleman, M. R., Rodd, J. M., Davis, M. H., Johnsrude, I. S., Menon, D. K., Pickard, J. D., et al. (2007). Do vegetative patients retain aspects of language comprehension? Evidence from fMRI. *Brain , 130*, 2494-2507.

Cribb, R. (2012, December 10). *Supreme Court decision on Hassan Rasouli will clarify end-of-life medical decisions.* Retrieved from Toronto Star: http://www.thestar.com/news/canada/2012/12/10/supreme_court_decision_on_hassan_rasouli_will_clarify_endoflife_medical_decisions.html

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in The Neurosciences , 2*, 263-275.

Crossman, J., Bankes, M., Bhan, A., & Crockard, H. A. (1998). The Glasgow Coma Score: reliable evidence? *Injury , 29* (6), 435-437.

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. A., Fernández-Espejo, D., Pickard, J. D., et al. (2011). Bedside detection of awareness in the vegetative state: a cohort study. *The Lancet , 378* (9809), 2088–2094.

Cruse, D., Chennu, S., Fernádez-Espejo, D., Payne, W. L., Young, G. B., & Owen, A. M. (2012). Detecting awareness in the vegetative state: Electroencephalographic evidence for attempted movements to command. *Plos ONE , 7* (11).

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition , 79*, 1-37.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences , 10* (5), 204-211.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods , 134*, 9-21.

Delorme, A., Sejnowski, T., & Makeig, S. (2010). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage , 34* (4), 1443-1449.

Demertzi, A., Vanhaudenhuyse, A., Bruno, M.-A., Schnakers, C., Boly, M., Boveroux, P., et al. (2009). Is there anybody in there? Detecting awareness in disorders of consciousness. *Expert Rev. Neurother , 8* (11), 1719-1730.

Derbyshire, S. W., Jones, A. K., Clark, S., Townsend, D., & Firestone, L. L. (1997). Pain processing during three levels of noxious stimulation produces differential patterns of central activity. *Pain , 73* (3), 431-445.

Dias, N. S., Kamrunnahar, M., Mendes, P. M., Schiff, S. J., & Correia, J. H. (2007). Comparison of EEG pattern classification methods for brain-computer interfaces. *29th Annual International Conference of the IEEE* , 2540,2543.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences , 23* (10), 475-483.

Etchells, E., Sharpe, G., Elliot, C., & Singer, P. A. (1996). Bioethics for clinicians: 3. Capacity. *Canadian Medical Association Journal , 155*, 657-661.

Fischer, C., Morlet, D., Bouchet, P., Luaute, J., Jordan, C., & Salord, F. (1999). Mismatch negativity and late auditory evoked potentials in comatose patients. *Clinical Neurophysiology , 110*, 1601-1610.

Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuropsychiatric applications of CANTAB. *International Journal of Geriatric Psychiatry , 11* (4), 329-336.

Gentilucci, M., Chieffi, S., Deprati, E., Saetti, M. C., & Toni, I. (1996). Visual illusion and action. *Neuropsychologia , 34* (5), 369-376.

Giacino, J. T. (2004). The vegetative and minimally conscious states: consensus-based criteria for establishing diagnosis and prognosis. *NeuroRehabilitation , 19* (4), 293-298.

Giacino, J. T., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D. I., et al. (2002). The minimally conscious state: Definition and diagnostic criteria. *Neurology , 58* (3), 349-353.

Goldfine, A. M., Victor, J. D., Conte, M. M., Bardin, J. C., & Schiff, N. D. (2011). Determination of awareness in patients with severe brain injury using EEG power spectral analysis. *Clinical Neurophysiology , 122* (11), 2157–2168.

Green, S. M. (2011). Cheerio, Laddie! Bidding Farewell to the Glasgow Coma Scale. *Annals of Emergency Medicine , 58* (5), 427-430.

Greenberg, D. L. (2007). Comment on “detecting awareness in the vegetative state”. *Science , 315* (5816), 1221b.

Greenwald, A. G., Drain, S. C., & Abram, R. L. (1996). Three Cognitive Markers of Unconscious Semantic Activation. *Science , 273*, 1699-1701.

Guger, C., Edlinger, G., Harkam, W., Niedermayer, I., & Pfurtscheller, G. (2003). How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE Transactions on Neural Systems and Rehabilitation Sciences , 11* (2), 145-147.

Guo, J., Gao, S., & Hong, B. (2010). An auditory brain–computer interface using active mental response. *IEEE Transactions on Neural Systems and Rehabilitation Engineering , 18* (3), 230-235.

Hampshire, A., Highfield, R. R., Parkin, B. L., & Owen, A. M. (2012). Fractioning human intelligence. *Neuron , 76* (6), 1225-1237.

Hampshire, A., Parkin, B. L., Cusack, R., Fernández Espejoa, D., Allanson, J., Kamau, E., et al. (2013). Assessing residual reasoning ability in overtly non-communicative patients using fMRI. *NeuroImage: Clinical , 2*, 174-183.

Hombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society , 40* (9), 922-935.

Jack, A. I., & Shallice, T. (2001). Introspective physicalism as an approach to the science of consciousness. *Cognition , 79*, 161-196.

Jeennerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioural and Brain sciences , 17* (2), 187-202.

Jennett, B., & Plum, F. (1972). Persistent Vegetative state after brain damage: A syndrome in search of a name. *The Lancet , 1* (7753), 737-737.

Kalmar, K., & Giacino, J. T. (2006). The JFK coma recovery scale—revised. *Neuropsychological Rehabilitation: An International Journal , 15* (3-4), 454-460.

Kane, N. M., Curry, S. H., Butler, S. R., & Cummin, B. H. (1993). Electrophysiological indicator of awakening from coma. *Lancet , 341*, 688.

Kotchoubey, B. (2007). Brain–computer interfaces in the continuum of consciousness. *Current Opinion in Neurology , 20*, 643–649.

Kotchoubey, B., Lang, S., Mezger, G., Schmalohr, D., Schmeck, M., Semmler, A., et al. (2005). Information processing in severe disorders of consciousness: Vegetative state and minimally conscious state. *Clinical Neurophysiology , 116* (10), 2441-53.

Laureys, S. (2007). Eyes open, brain shut. *Scientific American , 296*, 86.

Laureys, S. (2005). Science and society: Death, unconsciousness and the brain. *Nature Rev. Neurosci , 6* (11), 899-909.

Laureys, S. (2005). The neural correlate of (un)awareness: Lessons from the vegetative state. *Trends in Cognitive Sciences , 12*, 556-559.

Laureys, S., Faymonville, M. E., Luxen, A., Lamy, M., Franck, G., & Marquet, P. (2000). Restoration of thalamocortical connectivity after recovery from persistent vegetative state. *The Lancet , 355* (9217), 1790-1791.

Laureys, S., Goldman, S., Phillips, C., Van Bogaert, P., Aerts, J., Luxen, A., et al. (1999). Impaired effective cortical connectivity in vegetative state: Preliminary investigation using PET. *NeuroImage , 9* (4), 377-382.

Laureys, S., Owen, A. M., & Schiff, N. D. (2004). Brain function in coma, vegetative state, and related disorders. *The Lancet Neurology , 3* (9), 537-546.

Laureys, S., Perrin, F., Schnakers, C., Boly, M., & Majerus, S. (2005). Residual cognitive function in comatose, vegetative and minimally conscious states. *Current Opinion in Neurology , 18*, 726-733.

Libet, B., Alberts, W. W., Wright Jr, E. W., & Feinstein, B. (1967). Responses of human somatosensory cortex to stimuli below threshold for conscious sensation. *Science , 158*, 1597-1600.

Luck, S. (2005). *An introduction to the event-related potential technique.* Massachusetts: MIT Press.

Lulé , D., Noirhomme, Q., Kleih, S. C., Camille, C., Halder, S., Demertzi, A., et al. (2012 (in press)). Probing command following in patients with disorders of consciousness using a. *Clinical Neurophysiology* .

Maguire, L. A., Burgess, N., Donnett, J. G., Frackowiak, R. S., Frith, C. D., & O'Keefe, J. O. (1998). Knowing where and getting there: A human navigation network. *Science , 280* (5365), 921-924.

Mak, J. N., & Wolpaw, J. R. (2009). Clinical applications of brain-computer interfaces: Current state and future prospects. *IEEE Review in Biomedical Engineering , 2*, 187-199.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods , 164*, 177–190.

Marson, D., & Harrell, L. (1999). Executive dysfunction and loss of capacity to consent to medical treatment in patients with Alzheimer's disease. *Seminars in Clinical Neuropsychiatry , 4* (1), 41-49.

Merikle, P. M., & Joordens, S. (1997). Parallels between perception without attention and perception without awareness. *Consciousness and Cognition , 6*, 219-236.

Metzinger, T. (2010). *The ego tunnel: The science of the mind and the myth of the self.* New York: Basic Books.

Monti, M. M., Coleman, M. R., & Owen, A. M. (2009). Executive functions in the absence of behavior: Functional imaging of the minimally conscious state. *Progress in Brain Research* , 249-260.

Monti, M. m., Coleman, M. R., & Owen, A. M. (2009). Neuroimaging and the vegetative state: Resolving the behavioral assessment dilemma? *Annals of the New York Academy of Sciences , 1157*, 81-89.

Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., et al. (2010). Wilfull modulation of brain activity in disorders of consciousness. *The New England Journal of Medicine , 362*, 579-589.

Moreno, R. D., Schiff, N. D., Giacino, J., Kalmar, K., & Hirsch, J. (2010). A network approach to assessing cognition in disorders of consciousness. *Neurology , 75* (21), 1971-1878.

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology , 115* (1), 140-144.

Nachev, P., & Husain, M. (2007). Comment on "detecting awareness in the vegetative state". *Science , 315* (5816), 1221a.

Neary, D., Snowden, J. S., & Mann, D. M. (2000). Cognitive change in motor neurone disease/amyotrophic lateral sclerosis (MND/ALS). *Journal of the Neurological Sciences , 180* (1-2), 15-20.

Norton, L., Hutchinson, M. R., Young, B. G., Lee, H. D., Sharpe, D. M., & Mirsattari, M. S. (2013). Disruptions of functional connectivity in the default mode network of comatose patients. *Neurology , 80* (9).

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience* .

Owen, A. M. (2008). Disorders of consciousness. *Annals of the New York Academy of Sciences , 1124* (1), 225-238.

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science , 313*, 1402.

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., Jolles, D., et al. (2007). Response to comments on "detecting awareness in the vegetative state". *Science , 5816*, 1221c.

Penfield, W., & Jasper, H. (1954). *Epilepsy and the Functional.* Boston: Little, Brown.

Perrin, F., Schnakers, C., Schabus, M., Degueldre, C., Goldman, S., Brédart, S., et al. (2006). Brain Response to One's Own Name in Vegetative State, Minimally Conscious State, and Locked-in Syndrome. *Journal of the Archives of Neurology , 63* (4), 562-569.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage , 16* (2), 331–348.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology , 9* (4), 456-479.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences , 10* (2), 59-63.

Polich, J. (1989). Frequency, intensity, and duration as determinants of P300 from auditory stimuli. *Journal of Clinical Neurophysiology , 6* (3).

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology , 118* (10), 2128-2148.

Posner, J. B., Saper, C. B., Schiff, N., & Plum, F. (2007). *The diagnosis of stupor and coma, 4th ed.* Oxford: Oxford University Press.

Posner, J., Schiff, N., & Plum, F. (2007). *Diagnosis of stupor and coma: 4th edition.* New York: Oxford University Press.

Pressey, A. W. (1967). A theory of the Mueller-Lyer Illusion. *Perceptual and Motor skills , 25*, 569-572.

Rappaport, M., Hall, K. M., Hopkins, H. K., Belleza, T., & Cope, D. N. (1982). Disability rating scale for severe head trauma: coma to community. *Archives of Physical Medicine and Rehabilitation , 63*, 118–123.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex , 15* (8), 1261-1269.

Rodriguez Moreno, D., Schiff, N. D., Giacino, J., Kalmar, K., & Hirsch, J. (2010). A network approach to assessing cognition in disorders of consciousness. *Neurology , 75* (21), 1871-1878.

Rosanova, M., Gosseries, O., Casarotto, S., Boly, M., Casali, A. G., Bruno, M.-A., et al. (2012). Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain , 135*, 1308-1320.

Royal College of Physicians Working Group. (1996). The permanent vegetative state. *Journal of the Royal College of Physicians of London , 30*, 113-121.

Schnakers, C., Chatelle, C., Vanhaudenhuyse , A., Majerus, S., Ledoux, D., Boly, M., et al. (2010). The Nociception Coma Scale: a new tool to assess nociception in disorders of consciousness. *Pain , 148* (2), 215-219.

Schnakers, C., Majerus, S., Giacino, J., Vanhaudenhuyse, M., Boly, M., Moonen, G., et al. (2008). A french validation of the coma recovery scale-revised (CRS-R). *Brain Injury , 22* (10), 786-792.

Schnakers, C., Perrin, F., Schabus, M., Hustinx, R., Majerus, S., Moonen, G., et al. (2009). Detecting consciousness in a total locked-in syndrome:. *NEUROCASE , 15* (4), 271-277.

Schnakers, C., Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., et al. (2009). Diagnostic accuracy of the vegetative and minimally conscious state: Clinical consensus versus standardized neurobehavioral assessment. *BMC Neurology , 9* (35).

Seth, A. K., & Baars, B. J. (2005). Neural Darwinism and consciousness. *Consciousness and Cognition , 14* (1), 140-168.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. *Lancet* .

The Multi-Society Task Force on PVS. (1994). Medical aspects of the persistent vegetative state. *New England Journal of Medicine , 330*, 1499-1508.

Tombaugh, T. N., & McIntyre, N. J. (1992). The Mini-Mental State Examination: A comprehensive review. *Journal of the American Geriatrics Society , 40* (9), 922-935.

Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L. J., Bruno, M.-A., Boveroux, P., Schnakers, C., et al. (2010). Default network connectivity reflects the level of consciousness in non-communicative brain damaged patients. *Brain , 133*, 161-171.

Weiskrantz, L. (1997). *Consciousness lost and found: A neuropsychological exploration.* New York: Oxford University Press.

Wilson, M. D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers , 20* (1), 6-11.

Winjin, V. M., van Voxtel, G. M., Eilander, H. J., & de Gelder, B. (2007). Mismatch negativity predicts recovery from the vegetative state. *Clinical Neurophysiology , 118*, 597-605.

Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Pecklam, P. H., Schalk, G., et al. (2000). Brain-computer interface technology: A review of the first international meeting. *IEEE Transations on Rehabilitation Engineering , 8* (2), 164-173.

Young, G. B., Ropper, A., & Bolton, C. (1997). *Coma and Impaired Consciousness: A Clinical Perspective.* New York: McGraw-Hill Professional.

Zhu, J., Wu, X., Liang, G., Mao, Y., Zhong, P., Tang, W., et al. (2009). Cortical activity after emotional visual stimulation in minimally conscious state patients. *Journal of Neurotrauma , 26* (5), 677-688.

## CURRICULUM VITAE

## AARON M. KIRSCHNER

### EDUCATION

Western University, London, Ontario, Canada
M.Sc. Candidate, Neuroscience 2011-present

University of British Columbia, Vancouver, Canada
B.A. Cognitive Systems. Specialization: Cognition and Brain 2007-2011

Capilano University, North Vancouver, British Columbia, Canada 2005-2007

### PUBLICATIONS AND CONTRIBUTIONS

Kirschner, A., Owen, A.M., Hampshire, A. (2012) A p300 based cognitive assessment battery for non-responsive patients. Manuscript in preparation.

Ranger, A., Yazdani, A., Grant,. A., Kirschner, A. (2012). An analysis of risk factors for minor brain injury in patients presenting with a facial fracture. Manuscript in preparation.

Kirschner, A., Owen, A.M., Hampshire, A. (2012) A p300 based cognitive assessment battery for severely motor impaired and overtly non-response patients. Poster presentation at 19th annual Cognitive Neuroscience Society Annual Meeting, 2012, New Orleans.

Kirschner, A., Kam, J.W.Y., Ward, L.M., Handy, T.C. (2012). Differential synchronization in default and task-specific networks of the human brain. *Frontiers of Neuroscience; 6:129.*

Ward, L.M., Kirschner, A., Emberson, L. & Kitajo, K. (2011). Endogenous neural noise and reaction time. In Y. Mama and D. Algom (Eds.) *Fechner Day 2011*. Tel-Aviv: International Society for Psychophysics

Synchronization in anti-correlated functional networks of the human brain. (2011) Slide presentation at Cognitive Neuroscience Society Annual Meeting, 2011, San Francisco.

Ward, L.M., MacLean, S.E. & Kirschner, A. (2010). Stochastic resonance modulates neural synchronization within and between cortical sources. *PLoS ONE*, 5(12), e14371 (1-12). doi:10.1371/journal.pone.001

Ward, L. M., MacLean, S., Kirschner, A. (2010) Neural Synchrony and Stochastic Resonance in the 40hz Transient Response. Poster presentation at 17th annual Cognitive Neuroscience Society Annual Meeting, 2010, San Diego.

## TEACHING EXPERIENCE

Cogs 401: Seminar in Cognitive Systems. — Spring 2011
[Teaching Assistant. Moderated discussions presented in seminar style capstone course. Evaluated assignments and guided research proposals. Maintained course website]

Cogs 300: Understanding and Designing Cognitive Systems. — Spring 2010
[Guest Lecture: “Force Dynamics in Human Language and Cognition”]

Cogs 300: Understanding and Designing Cognitive Systems — Spring 2010
[Conducted labs where students designed and built robots to test theories of cognition. Assisted students with course material. Evaluated lab reports and exams. Maintained course website]

Cogs 200: Introduction to Cognitive Systems — Fall 2009
[Program Appointed Tutor]

Phil 221: Formal Symbolic Logic — Fall 2008
[Tutor]

## ACADEMIC AND VOLUNTEER EXPERIENCE

Psychology Graduate Students Association — 2012 - present
[Executive Committee Member. Organizing social events, fundraising, liaising with faculty]

Department of Neurosurgery, University Hospital, Western University — 2012 - present
[Research Assistant. Supervisor: Dr. Adrianna Ranger. Assisting with literature review, statistics and editing publications]

Political Science Department, UBC. — 2010 - 2011
[Software Engineer. Supervisor: Dr. Christopher Kam, Political Science. Engineered software for extraction and analysis of online databases, text analysis and summarization of political documents. Consulted on statistical techniques]

Brain Research Centre, UBC. — 2008 - 2011
[Research Assistant. Supervisor: Dr. Lawrence M. Ward, Psychology. Designed, programmed and conducted experiments. Recorded and analyzed EEG data. Trained lab members and assisted in developing lab protocol. Assisted with writing scientific manuscripts]

Cogs 402: Student Directed Research Project — 2010 - 2011
[Supervisor: Dr. Lawrence M. Ward, Psychology. Conducted literature review. Designed and programmed experiment. Collected and analyzed EEG data. Wrote manuscript for publication]

Cognitive Systems Society, UBC 2010 - 2011
[Arts Undergraduate Society Representative. Represented Cognitive Systems Society at Arts Undergraduate Society meetings. Voted on policy within the Arts department. Liaised with representatives from Arts clubs to organize and oversee events. Conducted study sessions for students]

## AWARDS

| | |
|---|---|
| Ontario Graduate Scholarship | 2012-2013 |
| NSERC Alexander Graham Bell Canada Graduate Scholarship | 2011-2012 |
| The Russ Patrick Arts Undergraduate Student Research Award | Spring 2011 |
| York University Centre for Vision Science Summer School Selectee | Summer 2010 |
| NSERC Undergraduate Summer Research Award | Summer 2010 |
| Arts Undergraduate Research Award (AURA) | Spring 2010 |
| NSERC Undergraduate Summer Research Award | Summer 2009 |
| Trek Excellence Scholarship (UBC) | 2008 |
| Student Recognition Award: Psychology (Capilano University) | 2007 |
| Student Recognition Award: Sociology (Capilano University) | 2006 |