2020-07-30

Heterogenity in Brain Injury: An Investigation of the Efficacy of Qualitative Comparative Analysis in Diffusion Tensor Imaging

Cooper Benton Hodges
Brigham Young University

Heterogeneity in Brain Injury: An Investigation of the Efficacy of Qualitative Comparative Analysis in Diffusion Tensor Imaging

Cooper Benton Hodges

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

C. Brock Kirwan, Chair
Rebecca A. Lundwall
Shawn D. Gale
Derin J. Cobia
Elisabeth A. Wilde

Department of Psychology

Brigham Young University

ABSTRACT

Heterogeneity in Brain Injury: An Investigation of the Efficacy of Qualitative Comparative Analysis in Diffusion Tensor Imaging

Cooper Benton Hodges
Department of Psychology, Brigham Young University
Doctor of Philosophy

Traumatic brain injury (TBI) and its associated neural and cognitive sequelae are of increasing interest in military populations. Blast-related TBI is becoming more commonplace in military Service Members and Veterans since Operation Iraqi Freedom/Operation Enduring Freedom/Operation New Dawn and their following conflicts. It is currently unclear whether blast-related injuries cause unique neural and cognitive deficits. The present investigation, in Study 1, aims to investigate the differences in blast-related and non-blast related TBI using traditional statistical techniques. In Study 2, this study will demonstrate the use of Qualitative Comparative Analysis (QCA) in diffusion tensor imaging data. QCA is a relatively new technique that examines configurations of variables that lead to a predefined outcome. QCA has the ability to uncover configurations of variables not yet considered in empirical literature, which may contribute new perspectives on the many different variables often associated with brain injury. Study 1 demonstrated no significant differences between uninjured and injured subjects in white matter integrity, and no differences between blast-related and non-blast related mechanisms. Study 2 demonstrated limited support for the use of QCA in diffusion tensor imaging. Evidence for the use of this method in other neuroimaging modalities is reviewed.

Keywords: traumatic brain injury, military, blast-related TBI, Qualitative Comparative Analysis

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Brock Kirwan, for his support during my graduate school career. Dr. Kirwan is an incredibly talented scientist who unquestioningly supports and inspires those who work with him. I was privileged to be a member of his laboratory, where I was encouraged to pursue my passions and taught to be an independent researcher. I would also like to acknowledge Dr. Elisabeth Wilde, who served as a guiding hand on this entire project. Dr. Wilde is one of the most dedicated and brilliant clinicians I have ever had the privilege to work with. Her support and guidance on this project have been an incredible influence. This work would not have been possible without the unwavering support of Dr. Wilde and her lab at the University of Utah.

I am indebted to the substantial contributions of my committee members: Dr. Rebecca Lundwall, Dr. Shawn Gale, and Dr. Derin Cobia. The time they committed and the dedication they showed for both feedback on my project and personal support cannot be understated.

I am incredibly grateful to my family: my mother, Christi, my father, Dale, my brother, Colton, and my grandmother, Kathleen. They have showed daily support on a difficult journey through my education. I simply would not have come this far without their love, advice, and support. It was a comfort to know that my family was always there to support me, no matter the circumstances.

My friends: Jakin Thomas, Dr. Kyle Whiteside, Dr. Zachary Reichert, Rhys Olson, Greg Simpson, Tiffany Thordarson, Dr. Christiana Atkins, Josephine Dimanche, Dr. Paula Johnson, Dr. Tyler Graff, Dr. Hannah Lindsey and Dr. Tara Austin. All of you have been supportive of me for years, and in some cases, well over a decade. All of you played an integral role in my journey through graduate school and have shown support for me in both my successes and my failures. I would not be the person I am today without your friendship, love, and support. To Christian, Ariana, Janet and Rick Ledesma: your family provided me a place to be home when I needed it most. I will never forget the influence you had over the person I've become. The impact you all have made on my life has been immeasurable, and I am deeply grateful for each one of you. I hope one day to be able to repay your kindness and support.

I would also like to acknowledge the Andrews University Department of Behavioral Science. Dr. Harvey Burnett, Dr. Karl Bailey, and Dr. Duane McBride- I am deeply thankful to all of you for the influence you had over my development as a scientist. When I first entered the department, I was unsure of what path my life would take. I'm certain that I wouldn't be where I am today without the guidance and support of all of you. You acted as mentors above and beyond what others would do– I am deeply grateful for the time and effort you took to make sure I had every bit of preparation needed to succeed in my career. Dr. Bailey, you were an inspiration to me and encouraged me to follow my dreams– I hope that I can pay even a fraction of that forward to students I mentor in the future.

To the /r/science moderation team: I never could have imagined how a science communication forum could have given me so many connections to so many wonderful people. Josh, Kelly, James, Max, Brad, Claudia, Piper and all the rest of you: I consider you some of my closest friends, even though we all live so far apart. Thank you for being there over the years I've known all of you. I am so grateful to have you all in my life.

Finally, to my partner, Daphne Batista: Daphne, when I started graduate school, I never anticipated finding a partner who would be as supportive and loving as you. You are an incredible clinician and the most driven, dedicated, and kind woman I have ever had the privilege of knowing. Your support during my graduate education has been nothing short of invaluable. You've offered your love and support unconditionally, from the beginning of my education until the end. I have been incredibly privileged to grow and learn with a partner who is there every step of the way. I am so thankful for you, and everything you've done to support me- I hope I can repay a little bit of that in the next chapter of our lives.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. HETEROGENEITY IN BRAIN INJURY: AN INVESTIGATION OF THE EFFICACY OF QUALITATIVE COMPARATIVE ANALYSIS IN DIFFUSION TENSOR IMAGING

Mental health among active-duty military Service Members and Veterans is growing in importance in military communities over the past several decades. In recent years, military combat theaters have had significant changes in wartime tactics and duration of deployment, resulting in an increase in both physical injury severity and more frequent wartime casualties (Kelly et al., 2008). These changes have, in part, caused a significant change in the downstream effects of wartime combat on the physical and psychological health of Service Members, and are largely responsible for the increase in concern for these groups. Active-duty military and veteran members show high rates of suicidality (AFHC et al., 2012; Bush et al., 2013; Blosnich et al., 2016), post-traumatic stress disorder (Richardson et al., 2010, PTSD), and mental health disorder difficulties associated with traumatic brain injury (Ginzburg et al., 2010; McKee and Robinson, 2014, TBI). Additionally, Veterans often report poorer physical health and overall functioning than individuals in the general population. One study (n=456,502) of Operation Iraqi Freedom/Operation Enduring Freedom Veterans founds that 11% of Veterans in the Veterans' Affairs (VA) healthcare system from 2001-2010 were diagnosed with a substance use disorder, defined as either alcohol abuse, drug abuse, or both (Seal et al., 2011). Of these people, 55-75% received a comorbid PTSD or depression diagnosis. Mental health concerns are not the only sequelae, however; studies have found a large portion of active-duty or veteran military subjects with a history of TBI report cognitive complaints (Helmick et al., 2015). Some studies have demonstrated associations between cognitive complaints and performance on objective measures of cognition (such as executive functioning), though evidence for this association is inconsistent (Schiehser et al., 2011) and often attributed to psychological distress (French et al., 2014). From this evidence, it appears as though significant heterogeneity exists in military populations.

Heterogeneity in brain injuries stands as the foremost problem with the development and implementation of treatment for TBI. TBI severity is generally classified as mild, moderate, or severe (Saatman et al., 2008). Classification often involves using a rating of voluntary and involuntary physical responses, such as the Glasgow Coma Scale scores (Teasdale and Jennett, 1974, GCS). However, even two brain injuries with similar GCS scores can be vastly different. This is due to the unique biomechanics of impact and individual differences in brain structure, which results in variation among brain injuries (Bigler et al., 2013). In addition to differences in brain structure, other complicating factors exist among those with a history of brain injury. One interesting phenomenon is how psychiatric factors relate to white matter (WM) integrity in TBI. Past studies have found that psychiatric disorder comorbidity, such as depression-TBI and PTSD-TBI comorbidity, have consistently been associated with compounding detrimental effects on WM integrity that result in more severe alterations in diffusivity than TBI groups alone (Alhilali et al., 2015; Davenport et al., 2016). These persistent effects of comorbidity across entire samples of TBI patients suggests some common factors are involved in the maintenance of WM integrity. This also highlights the potential for interaction effects between variables influencing neural structure that remain unknown.

Of all head injuries resulting in medical attention, approximately 75% of these injuries are diagnosed as mild (Dikmen et al., 2017). Most mild TBIs do not result in visible, macroscopic abnormalities (as detectable by MRI or other imaging modalities) which sometimes creates challenges for diagnosis (Koerte et al., 2016). Although most symptoms of mild TBI resolve quickly, within 90 days, 10-15% of patients with mild TBI experience chronic symptoms, often referred to as Post Concussive Syndrome (Eme, 2017). In military populations, postconcussive symptoms persist longer than normal, distinctly different from civilians with similar injuries (Cooper et al., 2017). It was proposed that this may be linked to psychological distress and trauma pre-deployment (Reid et al., 2018). This phenomenon can be challenging to study in military populations, given the unique injury mechanisms that are often involved in these head injuries. Mechanism of injury in military TBI patients can be largely related to the combat theater of service; those serving in Afghanistan and Iraq often report blast-related TBI (Goodrich et al., 2013). Blast-related injuries (both physical and brain specifically) are subcategorized into primary, secondary, and tertiary injuries. Primary injuries result from the shockwave of a blast

and changes in atmospheric pressure that injure the brain, secondary injuries are caused by objects that are put into motion as a result of the blast, and tertiary injuries are caused when the body strikes an object in the environment after it is put into motion from the blast (Goodrich et al., 2013). Furthermore, blasts can be controlled (low yield blasts intentionally caused in safe environments for the purposes of training or other military activity) or uncontrolled (accidental or intentionally caused by enemy military weaponry) (Caplan et al., 2015). While it is acknowledged that blast-induced TBI has a unique pathophysiology (Ling et al., 2009), few studies have examined structural differences between blast-related and non-blast related injuries. Some have found differences in measures of cognition (Luethcke et al., 2011) and symptoms of mental health problems (Kennedy et al., 2010) between patients with blast and non-blast related TBI, but others have found no evidence that these differences exist (Belanger et al., 2009). Further studies of blast-related TBI report a diffuse, global disruption of WM integrity (Davenport et al., 2012). It remains unclear whether this pattern would persist in military mild TBI, particularly between blast exposed and blast unexposed patients. Systematic reviews of the literature revealed significant variation in research findings. Greer and colleagues (2018) reported no consistently different patterns of clinical or functional outcomes between blast-related and non-blast related TBI among Service Members and Veterans (Greer et al., 2018). A large sample investigation of mild traumatic brain injury and outcomes post injury might allow for a more robust analysis of this heterogeneous population.

In the general population, moderate to severe TBI often impacts overall neural function and structure, and results in macroscopic lesions (Bigler, 2001), cortical volume decline (Bigler, 2013a), ventricular enlargement (Shenton et al., 2012), and atrophy in both white matter tracts and cortical regions (Anderson and Bigler, 1995; Gale et al., 1995; Johnson et al., 2013; Masel and DeWitt, 2010). In military populations, these effects are also seen; military TBI is usually distinct because there are multiple injury mechanisms (primary, secondary, tertiary, and quaternary) that can interact both at time of injury and post-injury. However, some of these physiological effects can be difficult to detect on a macroscopic scale through the use of traditional magnetic resonance imaging and similar modalities. Diffusion Tensor Imaging (DTI), a relatively new neuroimaging method, allows researchers to make inferences about the health and integrity of the brain's white matter tracts post-injury. This magnetic resonance imaging

(MRI) based acquisition sequence detects the diffusion of water molecules along white matter bundles in the brain (Alexander et al., 2007). Similar to functional magnetic resonance imaging, DTI requires researchers to make inferences about the state of the brain; analyses use measures of fractional anisotropy (FA) and diffusivity to indicate the relative microstructural integrity of the white matter tracts. Mean diffusivity (MD) measures average diffusion along axons (Basser and Pierpaoli, 2011). Fractional anisotropy is a measure that describes the degree of diffusion of the water molecules in the brain, and whether that diffusion is constrained to one direction (presumably along a white matter tract) or if there is no ”preferred” direction of diffusion (Basser, 1995). Although not a direct measure of neural integrity, DTI currently remains one of the best available options for in-vivo imaging of white matter structures.

Researchers have also used DTI to investigate whether there are specific white matter tracts that are frequently damaged as a result of brain injury. Indeed, there appear to be commonalities in injuries, where frontal association pathways, including the anterior corona radiata, uncinate fasciculus, superior longitudinal fasciculus, and the anterior corpus callosum are often damaged in subjects with a history of TBI (Niogi and Mukherjee, 2010). This would be consistent with past findings, as these association fibers are involved in executive function, attention, processing speed, memory, and learning, which all appear to be altered post-TBI (Catroppa et al., 2006; Little et al., 2010; Lajiness-O'Neill et al., 2010; Silver et al., 2009). Past research has demonstrated that diffuse axonal injury often occurs in frontal fibers, which is a result of deceleration and rotational forces causing shearing to occur in these tracts (Zappalà et al., 2012). Although frontal regions are readily susceptible to injury, damage often occurs in other regions as well, often in temporal regions (Wilde et al., 2007). Some temporal lobe areas, such as the temporal stem, display damage more frequently than other areas. The temporal stem contains segments of four major fasciculi, including the arcuate fasciculus, inferior branch of the cingulum bundle, the inferior occipitotemporal fasciculus, and the inferior occiptofrontal fasciculus (Bigler et al., 2010). Previous research has suggested cognitive deficits are not only caused by injury to the medial temporal lobe, but also from afferent/efferent pathways (Bigler and Maxwell, 2012), such as the temporal stem. Others have found that the anterior temporal regions are more likely to be injured in focal TBI, where inertial forces cause localized contusions (Levine et al., 2008). There is some conflicting evidence regarding injuries in these regions, however. While some have shown the

internal capsule and temporal projections, such as the uncinate fasciculus, show low FA post-injury, indicating damage (Lipton et al., 2008; Niogi et al., 2008), others have demonstrated patterns of high FA in temporal regions, hypothesizing the existence of a compensatory response in WM post-TBI (Lipton et al., 2012). Hippocampal damage has been shown to correlate with memory impairment post-TBI (Kinnunen et al., 2011; Tate and Bigler, 2000), but it is unknown whether certain injury mechanisms result in pervasive damage to temporal regions. It is worth investigating whether temporal regions show frequent damage, as evidence of this would help explain the frequent reports of memory deficits in patients with TBI (Draper and Ponsford, 2008). Some have found correlations between WM abnormalities in temporal regions and memory functioning (Kinnunen, 2011), but the evidence for this is far from conclusive.

Impaired WM integrity in both the frontal and temporal tracts has been associated with decreases in cognitive performance, namely executive function and memory (Kraus et al., 2007; Lockhart et al., 2012). Numerous clinical measures have been used to attempt to relate more broad functions of memory, cognition, and visual/motor processing to WM integrity post TBI. Among the most used are the Delis-Kaplan Executive Function System (D-KEFS), the California Vebal Learning Test (CVLT), the NIH Toolbox, and the Test of Premorbid Functioning (TOPF). Briefly, the D-KEFS is a standardized assessment of executive functioning. The D-KEFS has been used in military TBI samples to assess executive functioning deficits in relation to white matter integrity. Decreased performance on the D-KEFS has been linked to decreased fractional anisotropy in the corpus callosum and the cingulum bundle (Sorg et al., 2014). The CVLT is a prominent assessment of verbal learning, with batteries assessing cued and free recall after both a long and short delay between word administration and recall testing. This assessment has been associated with decreased FA in the corpus callosum and the uncinate fasciculus (Arenth et al., 2014; Bigler, 2013b). The NIH Toolbox is a digitized assessment testing functioning in a number of domains. Relevant to this investigation is the Cognition battery, which measures executive function, episodic memory, language, processing speed, working memory, and attention (Weintraub et al., 2013). Although few studies have considered how scores on the NIH Toolbox might be related to specific white matter tracts, there is evidence about similar links between the domains of measurement of the NIH Toolbox and white matter. For instance, there is significant evidence that processing speed is highly correlated with FA on both a whole brain and region of interest scale

(in the genu of the corpus callosum, cingulum bundle, and superior longitudinal fasciculus) (Kochunov et al., 2016). Further work is needed, as it is unclear whether these types of relationships are unique to the domains, or whether a similar pattern would emerge on an average of the scale score on the NIH Toolbox. Finally, the Test of Premorbid Functioning is a revised version of the Wechsler Test of Adult Reading, based on the concept that word reading is highly associated with intelligence and remains preserved after traumatic brain injury and neurodegenerative disorders (Berg et al., 2016). As this measure relies on preserved functioning, there has been little work relating decreased scores to white matter disruption. Although all of the tests described here have been used (in some capacity) in military TBI samples, few have considered whether differences in injury mechanisms might be reflected by scores on these measures.

In addition to cognitive measures, impaired WM in TBI patients appears to be related to motor and visual performance as well. One study of WM integrity found TBI patients performed worse on measures of motor function and had lower WM integrity in motor areas than control groups (Caeyenberghs et al., 2011). Similarly, one systematic review of DTI findings of white matter abnormalities in subjects with a history of mild TBI found that military samples generally showed decreased WM integrity in the frontotemporal WM tracts, including the genu of the corpus callosum, cingulum bundle, and uncinate fasciculus (Asken et al., 2018). Previous research does suggest that WM integrity is related to performance on cognitive, visual, and motor tasks, but it remains unclear whether there would be significant differences between blast-related and non-blast related TBI patients.

While literature in structural change resulting from TBI has increased substantially in recent years, significant gaps remain. First, there are few studies with large sample sizes studying TBI in veteran and active duty military members; therefore, it is important to determine whether trends in previous research findings hold in large samples. As previous research has noted the problem of poor statistical power due to small sample sizes in neuroimaging, utilizing data with a large cohort could help allay fears that previous findings were spurious (Button et al., 2013). Spurious results are becoming of increasing concern, in light of recent research showing that completely fabricated data can often show compelling results (Poldrack, 2012). Second, there appear to be no presently published studies using equifinal statistical methods in the analysis of structural

neuroimaging data in military populations. Keeping in mind the vast array of factors that have been shown to be influential in neural structure and function in Veterans with TBI, an equifinal approach could help shed light on variables not yet considered. Lastly, it is still unclear whether there are pervasive differences between blast-related and non-blast related injuries. Given the relatively high incidence of blast-related injuries in recent military cohorts (Greer et al., 2016), it is important to understand whether there are unique cognitive and neurological sequelae that should be considered in the development of treatment and future research. On a large scale, with greater statistical power than past studies, a goal of the present study is to find evidence that the unique injury mechanisms (primary, secondary, and tertiary) of blast-related TBI have differential impacts on neural structure.

Equifinality is the principle that many different methods/factors can lead to one specific end state (Grofman and Schneider, 2009). In practice, this can be an important (albeit infrequently studied) property of human subjects research. In neuroscience literature, many variables have been examined individually to determine their effect on white matter integrity and overall neural structure and function. However, there is no analysis published in the neuroimaging literature to date that allows for equifinality.

A newer analysis method, called Qualitative Comparative Analysis (QCA), may allow for this examination of equifinality in neuroimaging research. While many published DTI studies use univariate analyses to test how study variables are related to DTI outcome (diffusivity) measures, the data from the present study will be analyzed using Qualitative Comparative Analysis (Ragin, 1994). QCA allows researchers to answer the question, ”What sets/configurations of cases in the dataset are consistent with a target outcome?” Here, configurations of cases refers to a combination of variables in a dataset. This is quite different to the question one might ask with a traditional correlational approach, which typically asks, ”Is there a linear relationship between study variables and a target outcome?” One of the first salient aspects about this method is that it is quite different from the typical linear/correlational statistical method, which tends to look for variables accounting for the most variance in a certain outcome. QCA was constructed using mathematical set theory (based on Boolean algebra; the mathematics of set membership, involving intersections, unions, and disjunctions of sets), which allows for the identification sets of cases whose members consistently align with a particular outcome. In the case of this present

study, an outcome might be low FA or high RD. Once sets of cases with the target outcome are identified, they are analyzed for consistency (i.e., how often there are contradictory cases with the same set configuration but different outcomes) and coverage (i.e., how often there are cases with the target outcome but a different set configuration).

Keeping with the previous example, a QCA might find that subjects with an outcome of low global FA consistently suffered from blast-related injury mechanisms. This has been supported by previous work, as prior studies have found preliminary evidence that blast-related injury mechanisms have a very diffuse effect on WM integrity, affecting a wide number of tracts (Magnuson et al., 2012). From this, one can interpret that injury mechanisms likely play some part in the degree of FA in these tracts (through an indirect mechanism). The way QCA accomplishes this type of analysis is by allowing researchers to set thresholds. These thresholds are usually set based on empirical data and are then used to determine membership. Once these thresholds are determined, QCA finds configurations of variables that large portions of participants share who also have the pre-defined outcome. These configurations are often referred to as ”extracted pathways” (indicating that those who share this configuration of variables eventually end with the outcome). Given the diverse number of variables often cited in social science studies, multi-value QCA has a unique ability to look at these factors simultaneously. This method also has use in neuroimaging. Many studies have looked at a diverse number of factors and the influence these factors have on measures of white matter integrity. However, to date, there have been no equifinal statistical analyses performed to examine these factors in unison. A unique aspect to mvQCA is that this analysis does not require researcher specification of statistical models. Unlike traditional statistical methods, which require researchers to specify and construct statistical models, mvQCA only requires that researchers assign thresholds to variables. Here, the thresholds are set to assign up to 3 levels to a given variable. For instance, a continuous variable consisting of values between 0 and 1 (such as FA) would be turned into a variable with 3 categories, representing the lower, median, and upper quartiles of continuous values. The purpose of study 2 is to examine the efficacy of multi value QCA (mvQCA), and whether it would provide any additional use to neuroscientists in future data analyses. This method is often acknowledged in the literature as one of the few statistical techniques that allows

for equifinality (Zschoch, 2011). In the present study, QCA will allow for very specific analyses and determination of factors that may play a role in influencing neural structural integrity.
There are a number of federated neuroimaging consortia that maintain data repositories for various research purposes. Among these is the Chronic Effects of Neurotrauma Consortium (CENC). One of the primary strengths of the CENC dataset is the multiple longitudinal datapoints it has for each subject included in the study. A primary aim of CENC is to identify and study the longitudinal effects of chronic neurotrauma, particularly in active-duty military and military Veterans. For the purpose of this study, baseline assessments from CENC were used, given the breadth and depth of cognitive, health, injury, and visuomotor measures collected.
As it stands, there are a number of gaps in the scientific literature that need to be addressed. First, it is unclear whether there are persistent differences in white matter integrity, psychological symptoms, and cognitive functioning between blast-related and non-blast related mild TBI patients. Second, to date, no analysis of neuroimaging data has used equifinal statistical methods. Equifinal methods might be able to uncover configurations of variables influencing WM integrity not previously considered. Qualitative Comparative Analysis is a relatively new equifinal statistical method that would allow for the simultaneous examination of many factors and how these factors may relate to neural integrity. Addressing this gap may prove the efficacy of this new statistical tool and provide a launching point for other studies in neuroimaging to examine new configurations of variables that may play a role in certain specified outcomes.

1.1 Current Investigation

The present study has two aims: (1) to examine the large-scale effects of military-relevant TBI on white matter structure and (2) to investigate the efficacy of a new statistical analysis technique, QCA, to uncover previously unknown configurations of variables that may be influencing WM integrity. Although QCA has been used previously in the social sciences (Rihoux et al., 2013), it has never been used to analyze neuroimaging data. QCA, if proven to be efficacious in neuroimaging analyses, could be useful to researchers working with large datasets. It is anticipated this extensive analysis will elucidate the white matter tracts that are most often damaged following blast related injury, and whether certain tracts are more susceptible to specific

injury mechanisms. This will occur in two studies: the first study is dedicated to a traditional DTI analysis, and the second is dedicated to investigating the efficacy of QCA in neuroimaging.

1.2 Aim 1

The following hypotheses are offered for Aim 1:

1.2.1 Hypothesis 1.1

There will be evidence of microstructural damage in the cingulum, uncinate fasciculus, and the genu of the corpus callosum in military subjects with a history mild TBI relative to non-TBI, reflected by lower fractional anisotropy (FA) and alterations in mean diffusivity.

1.2.2 Hypothesis 1.2

There will be a differential effect of injury mechanism where blast-related injuries will show higher diffusivity and lower FA than non-blast related/impact related traumatic brain injuries in whole-brain WM comparisons.

1.3 Aim 2

The following hypotheses are offered for Aim 2:

1.3.1 Hypothesis 2.1

For military subjects with an outcome of low FA in frontal and temporal tracts, there will be high consistency and coverage for extracted configurations including blast-related injury mechanisms.

1.3.2 Hypothesis 2.2

For military subjects with an outcome of high FA in frontal and temporal tracts, there will be low consistency and coverage for extracted configurations with blast related injury mechanisms.

CHAPTER 2. STUDY 1

2.1 Methods

2.1.1 Participants

Data were acquired from the CENC data consortium. The final dataset received and used in the analysis contained neuroimaging and clinical data collected across eight acquisition sites for 946 subjects with a history of military service (126 female). Subjects in this dataset mainly identify as white (n=683), followed by black/African American (n=180), other (n=47), Asian (n=8), American Indian/Alaska native (n=9) and Pacific Islander (n=7). The remaining subjects either did not disclose their ethnicity or indicated they were unsure. This dataset also included age, sex, time since injury, number of TBIs, educational level, injury mechanisms, and comorbid diagnoses (see Tables 2.5 and 2.6). The CENC sample consists of subjects with mild TBI (acquired previously, during, or post-military service) and subjects who report no previous history of TBIs. On average, subjects reported 10 years since the worst TBI event (Walker et al., 2016). TBI was diagnosed through an in-depth structured interview that screened for potential injury events and, when events were identified, assessed for injury severity using the Virginia Commonwealth University Retrospective Concussion Diagnosis Interview (Walker et al., 2015).
Blast exposure information was collected from all participants. These variables were:

- Number of blast exposures not resulting in TBI
- Number of exposures to controlled blasts
- Number of exposures to uncontrolled blasts
- Total number of exposures to blast

A blast TBI variable (number of exposures to blasts resulting in TBI) was created by subtracting the number of blast exposures not resulting in TBI from the total number of exposures to blast.

2.1.2 Measures

CENC also contains a number of neuropsychological measures. For the purposes of the present study, which sought to both investigate blast injury mechanisms and relate WM integrity to a number of neuropsychological testing measures, a number of measures were used: The California Verbal Learning Test (CVLT-II) is a measure of verbal learning, retention, and retrieval, which involves the memorization of word list, as well as an interference trial (Delis et al., 2000). Both free and cued recall of the initial list is required after both a short and a long delay. The CVLT-II is the second form of the test. Statistical analyses use raw scores, as statistical models already account for age and premorbid functioning. Additionally, covariates were added to models that were not taken into account by the original sociodemographically adjusted scaled scores (namely psychological disorder screening measures).

The Delis-Kaplan Executive Function System (Delis et al., 2004, D-KEFS) is a measure of cognitive and executive function. The D-KEFS includes nine individually administered tests, which measure higher level cognitive functioning (specifically, executive function). The D-KEFS provides normative (based on census data), qualitative and performance data for each subject. For this study, the D-KEFS verbal fluency subscale is used, which consists of raw scores for both total category and total letter fluency. These subscales were chosen as they measure memory performance, and are most likely to be associated with the tracts chosen for this analysis.

The Test of Premorbid Functioning (TOPF) (Wechsler, 2009) is a measure designed to estimate premorbid cognitive function. The TOPF is an adaptation of the Wechsler Test of Adult Reading (Wechsler, 2001), consisting of 70 words that are difficult to translate from writing to speech. The test is designed to measure intellectual performance that is thought to be preserved after injury. TOPF results include a premorbid IQ estimate (adjusted for age and demographic information) which was used in the models in Study 1.

More clinical measures were obtained and included in data processing for use in Study 2. The Alcohol Use Disorders Identification Test (AUDIT-C) is an abbreviated version of the full scale, consisting of 3 items measuring hazardous alcohol use (Bush et al., 1998). The AUDIT-C has Likert-scale responses, allowing for discrete total scores between 0 and 12.

The Behavioral Risk Factor Surveillance System (BRFSS) is a telephone-based assessment, directed by the Centers for Disease Control, in conjuction with US state governments, to measure

a variety of health and risk factors in the American general population (Remington et al., 1988). Started in 1984, the BRFSS measures many risk factors, including seatbelt use, alcohol and cigarette use, drinking and driving, as well as preventable health problems, including high cholesterol and influenza vaccination. The analysis in Study 2 used subscales from the BRFSS measuring Angina/Heart Disease, self rating of general health, and physical activity.

The Brief Visuospatial Memory Test (BVMT) is a visual memory test, consisting of 3 administration trials followed by a delayed recall trial, in which participants are shown a stimulus with various geometric designs and asked to reproduce them from memory after viewing (Benedict et al., 1996). The revised BVMT measures both overall recall accuracy and delayed recall accuracy specifically, both of which are used in the Study 2 analysis.

The Drug Abuse Screening Test (DAST-10) is a 10 item scale measuring problematic drug use behaviors, a condensed version of the original 28 item scale (Skinner, 1982). The DAST-10 measures drug use over the past year, with binary ”yes/no” responses. Items regard concerns about use from family and friends, drug use experiences such as ”blackouts”, and personal concerns over use. Generally, more than 3-5 ”yes” responses to the DAST-10 is considered worthy of follow-up assessment by a qualified professional. Here, the DAST-10 total score (across all 10 items) was used in Study 2 analysis.

The EQ-5D-5L is a modified version of the original EuroQol measurement of health related quality of life (Group, 1990). This modified version consists of 5 domains (Mobility, Self Care, Usual Activities, Pain/Discomfort, and Anxiety/Depression) across 5 levels of severity (no problems, slight problems, moderate problems, severe problems, or extreme problems with each domain). The EQ-5D-5L produces a health state profile with 5 levels: Level 1 (No Problem), Level 2, (Slight Problems), Level 3 (Moderate Problems), Level 4 (Severe Problems), and Level 5 (Extreme Problems). Additionally, one item asks participants to rate their current health on a scale of 0-100 (0= the worst health you could imagine, 100= the best health you could imagine). Profile scores and health ratings were used in Study 2 analyses.

The Glasgow Outcome Scale-Extended (GOSE) is a scale of functional outcome assessment for patients with head injuries (Jennett and Bond, 1975). Initial assessments are made on objective physiological functioning, ranging from 1 (death) to 8 (upper good recovery). Functional outcome assessments are made with a structured interview, which measures consciousness,

Independence, working ability, social and leisure activities, family and friendships, and possibility of return to normal life. The total GOSE outcome score was used in Study 2 analyses.

The General Self-Efficacy Scale (GSE) is a measure of self-efficacy, or optimistic self-belief (Schwarzer et al., 1995). The GSE is a 10 item Likert-scale (ranging from 1=Not at all true to 4=Exactly true) with items such as, ”I can solve most problems if I invest the necessary effort”. The total score (a discrete value between 10 and 40) across all 10 items was used in Study 2 analyses.

The Hearing Handicap Inventory for Adults (HHIAS) is a 3 level Likert-scale across 25 items that measures hearing difficulties (Newman et al., 1991). Responses include ”Yes” (4 points), ”No” (0 points), and ”Sometimes” (2 points). Items measure social and emotional consequences of hearing problems. Scores range between 0 and 100 for the total scale, between 0 and 48 for the social subscale and 0 and 52 for the emotional subscale. The total score across all 25 items was used in Study 2 analyses.

The Headache Impact Test (HIT-6) is an abbreviated 6 item version of the full 54 item scale that measures severity and frequency of headaches/migraines (Kosinski et al., 2003). The HIT-6 is a 5 level Likert scale, with responses ranging from ”never” (6 points) to ”always” (13 points). The scale suggests clinical interventions for scores reaching or exceeding 50. Items include questions such as, ”How often do headaches limit your ability to do usual daily activities including household work, work, school, or social activities?”. The total score across all 6 items was used in Study 2 analyses.

The Mild Brain Injury Atypical Symptoms (MBIAS) is a scale intended to measure symptom validity in patients with mild TBI and used in conjunction with postconcussion and PTSD symptom checklists (Cooper et al., 2011). The MBIAS is meant to measure symptom exaggeration/over-reporting, originally developed for use in Operation Enduring Freedom/Operation Iraqi Freedom Service Members. The MBIAS consists of 5 items with a 5 level Likert scale. Items include symptoms unassociated with mild TBI, such as complete loss of feeling/sensation in arms and complete deafness for periods of time. Results from the original study suggest a score of 8 (out of 25) is sufficient for detecting symptom exaggeration. The total score on the MBIAS was used in Study 2 analyses.

The Mini International Neuropsychiatric Interview (MINI) is a structured diagnostic interview meant to be used in the diagnosis of DSM-IV and ICD-10 psychiatric disorders (Sheehan et al., 1998). The MINI, which takes between 15 and 25 minutes to administer, probes for the presence of symptoms that could be indicative of psychiatric disorders. Common disorders diagnosed through use of the MINI are Major Affective Disorder, Generalized Anxiety Disorder, and social phobia. The final diagnosis reached after administration of the MINI was used in Study 2 analyses.

The Neurobehavioral Symptom Inventory (NSI) was developed as a tool to measure symptoms occurring as a result of traumatic brain injury (Cicerone, 1995). The NSI contains 22 items across a 5 level Likert scale ranging from 0 (None) to 4 (Very Severe). Likert scale levels specifically measure symptom severity as well as general functioning (whether symptoms significantly impact normal life). Subscales of the NSI include affective, cognitive, somatosensory, and vestibular symptoms. The NSI includes items such as, ”Poor frustration tolerance, feeling easily overwhelmed by things”. Total scores for each subscale were used in the Study 2 analyses.

The PTSD Check List (PCL-5) is a 20 item self-report measure that assesses DSM-V symptoms of PTSD (Blevins et al., 2015). The PCL-5 uses a 4 level Likert Scale, ranging from 0 (”Not at All”) to 4 (”Extremely”). Scores are summed, with a possible range between 0 and 80. Scores between 31 and 33 indicate probable PTSD; clinical assessment is recommended for those meeting or exceeding this threshold. DSM-5 symptom clusters can be calculated by item groupings. The total score for the PCL-5 was used in Study 2 analyses.

The Patient Health Questionnaire-9 (PHQ-9) is a subset of the full PHQ scale pertaining to Major Depressive Disorder (Kroenke and Spitzer, 2002). The PHQ-9 is a 9 item Likert scale ranging from 0 (”Not at All”) to 3 (”Nearly Every Day”), with possible scores ranging between 0 and 27. Each question measures one of the DSM-V criteria of Major Depressive Disorder. Authors suggest a score of 15 or above on the PHQ-9 is indicative of nee for psychiatric assessment. The total score for the PHQ-9 was used in Study 2 analyses.

The Pittsburgh Sleep Quality Index (PSQI) is a 24 item 3-level Likert scale, consisting of questions regarding sleep in the past month (Bysse et al., 1989). 19 questions are self-reported, while 5 are reported by the bedpartner or roommate. Questions regard estimates of sleep duration and the frequency/severity of sleep-related problems. The total range on the PSQI ranges between

0 and 21, with 0 indicating no problem and 21 indicating severe sleep difficulties. The total global score was used in Study 2 analyses.

The Snore Tired Observed Pressure BMI Age Neck Gender (STOPBANG) is a 8 item measure of obstructive sleep apnea in adults (Chung et al., 2008). All 8 items are answered with Yes/No responses. Scores between 0 and 2 indicate low risk of sleep apnea. Scores of 3 or 4 indicate intermediate risk of sleep apnea, while scores between 5 and 8 indicate high risk of sleep apnea. Each item measures either a symptom or risk factor of obstructive sleep apnea (such as BMI, age, neck circumference, and gender). The total score on the STOPBANG, as well as the BMI question, were used in Study 2 analyses.

The TBI Quality of Life scale (TBIQOL) is 20 domain self report assessment of outcomes after a traumatic brain injury, with a question bank of 631 total items (Tulsky et al., 2016). Each item has a 5 level Likert scale ranging from 1 (”Never”) to 5 (”Cannot Do”). Administration of items can be the full question bank or a short form/fixed length version of the bank. The Study 2 analyses included total scores for the full question banks from the Anger, Emotional and Behavioral Dyscontrol, Anxiety, Cognition, Executive Function, Fatigue, General Concentration, Pain, and Ability to Participate domains.

The Trail Making Test (TMT) is a pen-and-paper motor skills test involving sequentially connecting a series of circles under a time constraint (Reitan, 1958). The TMT has two trials (Trial A and Trial B). Trial A involves sequentially connecting 25 circles numbered 1-25. Trial B involves sequentially connecting a combination of numbers and letters (e.g. 1 to A to 2 to B). Scores on the TMT are the time (in seconds) it takes to complete each trial. Scores for Trials A and B were included in Study 2 analyses.

The Unified Parkinsons' Disease Rating Scale (UPDRS) is a 42 item clinician administered assessment of Parkinsonian symptoms in a number of domains, including Mentation, Behavior and Mood, ADL, Motor Function, Hoehn and Yar Scale, and Schwab and England's ADL scale (Martínez-Martín et al., 1994). Possible scores range between 0 and 199, with 0 indicating no disability and 199 indicating the most severe, total disability. Each item is a categorical rating between 0 and 4 depending on severity of symptoms. The UPDRS total score was used in Study 2 analyses.

Finally, the Wechsler Adult Intelligence Scale (WAIS-IV) is a clinician administered test of cognitive ability (Drozdick et al., 2012). The WAIS-IV is the fourth edition of the WAIS, and consists of a total of 10 subtests, which measure verbal comprehension, perceptual reasoning, working memory, fluid reasoning, and processing speed. Scores on the WAIS-IV are scaled by age. Scores on the coding, digit span, and visual puzzle subtests were used in the Study 2 analyses.

2.1.3 Preprocessing and ROI Analysis Pipeline: ENIGMA

Data drawn from the CENC consortium include volumetric data derived from T1-weighted imaging as well as diffusion metrics derived from diffusion weighted imaging. Data were collected across eight sites on 3T MRI scanners of various makes and models, including:

1. Site 1: Virginia Commonwealth University. Richmond, VA. Scanner: Philips Ingenia. Parameters included 64 diffusion directions on diffusion weighted scans and a resolution of 2.73 x 2.73 x 2.7mm.

2. Site 2: Michael E. DeBakey Veterans Affairs Medical Center. Houston, TX. Scanner: Siemens TIM Trio. Parameters included 64 diffusion directions on diffusion weighted scans and a resolution of 2.73 x 2.73 x 2.7mm.

3. Site 3: James E. Haley Veterans Affairs Medical Center. Tampa, FL. Scanner: General Electric Medical Systems Signa HDxt. Parameters included 64 diffusion directions on diffusion weighted scans and a resolution of 1.3 x 1.3 x 2.7mm.

4. Site 4: Brooke Army Medical Center. Fort Sam Houston, TX. Scanner: Siemens Verio. Parameters included 64 diffusion directions on diffusion weighted scans and a resolution of 2.73 x 2.73 x 2.7mm.

5. Site 5: Fort Belvoir Community Hospital. Fort Belvoir, VA. Scanner: General Electric Medical Systems Discovery MR750. Parameters included 64 diffusion directions on diffusion weighted scans and a resolution of 1.4 x 1.4 x 2.7mm.

6. Site 6: Veterans Affairs Medical Center. Portland, OR. Scanner: Philips Achieva dStream. Parameters included 32 diffusion directions on diffusion weighted scans and a resolution of 2.73 x 2.73 x 2.73mm.

7. Site 7: Center for Magnetic Resonance Research, University of Minnesota. Minneapolis, MN. Scanner: Siemens Prisma Fit. Parameters included 79 diffusion directions on diffusion weighted scans and a resolution of 1.5 x 1.5 x 1.5mm.

8. Site 8: Jamaica Plain Veterans Affairs Medical Center. Boston, MA. Scanner: Siemens Prisma Fit. Parameters included 79 diffusion directions on diffusion weighted scans and a resolution of 1.5 x 1.5 x 1.5mm.

Acquisiton parameters followed recommended guidelines from CENC, which has been outlined previously (Walker et al., 2016). All data types (structural, functional, and diffusion-weighted images) were preprocessed according to the ENIGMA pipeline (Jahanshad et al., 2013). The first step in the diffusion imaging step of the pipeline was the conversion of DICOM image files to a diffusion-weighted imaging (DWI) set. This was accomplished using the *dcm2niix* package, which outputs images in the Neuroimaging Informatics Technology Initiative format (Li et al., 2016, NIFTI). This package automatically outputs gradient information in .bval and .bvec files. Data were then corrected for eddy current distortions and movement using affine registration. Eddy current distortions are caused by the rapid changing in magnetic fields characteristic of pulse sequences and can cause artifacts to appear on MR images. Registration was completed using FSL (FMRIB's Software Library), provided by the Oxford Centre for Functional MRI of the Brain (Smith et al., 2004).

FSL performs eddy and movement correction automatically through the eddy_correct command. Brain masks were constructed using FSL's *bet2* command, which performs skull-stripping on T1- and T2-weighted images. Correction for echoplanar imaging (EPI)-induced susceptibility artifacts was also performed. DWI images were adjusted by warping b0 to a high-resolution structural T1- or T2-weighted image to adjust for EPI distortions. A b0 image is essentially a T2-weighted image, as it is collected during a scanning sequence with all diffusion gradients off to serve as a baseline for further analyses. Finally, tensors were calculated using FSL's *dtifit* command. Diffusion tensors represent a three dimensional Gaussian model of the direction and

magnitude of the diffusion of water molecules in a given voxel. Each tensor consists of three eigenvectors and eigenvalues; together, these define an ellipsoid that represents the degree of anisotropy at a given position (O'Donnell and Westin, 2011). *dtifit* uses least-square fitting to determine tensors. Diffusivity measures were extracted using FSL's *tbss* command. This performs tract-based spatial statistics on DWI data using a skeletonized template. Skeletonization is an automated segmentation technique, developed to avoid the bias introduced by traditional hand segmentation. Studies suggest that skeletonized data contains less non-WM data and higher accuracy than hand-segmented data (Zhang and Arfanakis, 2014). Skeletonization involves warping subject-space images to a template and segmented by using pre-defined labels on the template skeleton. The template used in this analysis was developed specifically for ENIGMA (http://enigma.ini.usc.edu/protocols/dti-protocols/#analysis). Participant (original) space DWI images were then mapped onto the ENIGMA-DTI template, to allow for ROI extraction across all participants in the sample. ROI extraction was completed using an ENIGMA-specific model script (http://enigma.ini.usc.edu/wpcontent/uploads/DTI_Protocols/ENIGMA_ROI _protocol_USC.pdf). ROI extraction involves mapping original subject-space scans onto the template (through image registration), and using the template as a guide for the definition and extraction of WM tracts. The ROI extraction script script extracts FA values from all 20 ROIs defined using the Johns Hopkins University White-Matter Atlas. The JHU Atlas identified these 20 WM structures probabilistically through averaging the results of 28 deterministic tractography analyses (Mori et al., 2005). An average FA was also calculated across the entire skeleton. Other diffusivity scalars were extracted using a separate skeletonization and analysis (http://enigma.ini.usc.edu/protocols/dti-protocols/). The ENIGMA script then combines participant information and diffusivity measures into a .csv file.

2.1.4 ENIGMA Pipeline Replication

DTI analyses were replicated on a subset of participants in order to verify the accuracy of the data analysis. Fifty subjects from CENC Site 1 were chosen for this replication. These subjects were chosen as their images were high quality and had few artifacts. Where possible, scripts used were acquired directly from the ENIGMA consortium. All scripts used to complete this replication, statistical analyses in both Studies 1 and 2, and detailed documentation of each step can be found

in the Open Science Foundation (OSF) repository (https://osf.io/s89du/). Initial subject images were acquired from CENC as pre-constructed NIFTI files. Before preprocessing and after skull stripping, images were quality controlled and viewed for obvious errors or MRI artifacts. Subject images were then pre-processed and skull-stripped using FreeSurfer's (v6) *'recon-all'* pipeline. The full *'recon-all'* command was used, which performs 31 different functions, outlined on the NIMH website (*https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all*). The *mri_convert* function was used in FreeSurfer to convert processed images from the native .mgz format to .nii format. Next, ENIGMA's pipeline suggests using *eddy_correct*, a function in FSL software used to correct eddy current induced distortions and subject movement. However, *eddy_correct* is an outdated function and is no longer supported on recent versions of FSL. In order to conduct corrections, *eddy_cuda8.0* was used. *Eddy_cuda8.0* is an alternative of *eddy_openmp*, which uses Compute Unified Device Architecture (CUDA), an accelerated computing platform on NVIDIA graphics processor units, to parallelize analyses. *Eddy_openmp* was incompatible with the software on the remote supercomputing cluster used to complete this analysis. ENIGMA's next suggested step is rotation of .bval and .bvec files, as the original *eddy_correct* function did not have this capability. However, as *eddy_cuda8.0* does this automatically, this step was skipped. New B0 masks were generated for each subject using eddy current-corrected files through FSL's bet function. These new masks were then registered to T1 brain masks using ANTs affine registration. Any distortions between T1 and B0 images were corrected through a non-linear registration in ANTs. EPI distortions were also corrected, using ANTs to apply transformations to the entire eddy corrected image. DTI tensors were created and fitted to subject masks using FSL's dtifit function. At this point, ENIGMA's original suggested pipeline was followed. Registration of T1, diffusivity metrics, and eigenvalues/vectors followed the original ENIGMA pipeline, as did the skeletonization and averaging with the JHU atlas.

Extracted FA and MD values from the replicated pipeline were compared to original FA and MD values acquired directly from CENC using Pearson's and Intraclass Correlations (ICCs). Pearson's correlation coefficients indicated small and insignificant effects between replicated and original values on MD in the genu of the corpus callosum ($r = -.143, p = .33$), cingulum bundle ($r = .347, p = .015$), and the uncinate fasciculus ($r = .282, p = .05$). For FA data, only insignificant coefficients were found between FA values of replicated and original data in the

genu of the corpus callosum ($r = -.069, p = .647$), cingulum bundle ($r = .214, p = .14$), and the uncinate fasciculus ($r = -.100, p = .49$). ICC correlations indicated low similarity between replicated data and original data for MD and FA across all three tracts (Average raters absolute ICC=$-4.0 * 10^5$, $p = 1.00$).

There are a number of reasons the replicated data and original data were not more similar (as indicated by Pearson's and ICC coefficients). One reason relates to the differences in analysis pipelines between the current replication and the original used. Although efforts were made to ensure the replication pipeline was as similar as possible to the original ENIGMA pipeline, some deviations were made due to issues that have been outlined here. These deviations may account for the poor relations between the replication and original data. It is documented that when comparing different DTI analysis pipelines, such as a traditional tract-based spatial statistic and a symmetric image normalization approach, these different pipelines can result in statistical differences in the diffusivity metrics extracted from white matter tracts (Bergamino et al., 2017). Likewise, evidence from related fields like functional and structural neuroimaging supports the proposition that deviations in analysis pipelines can result in significant differences in MR images (Muncy et al., 2017; Carp, 2012). Differences have even been documented when running the same analysis pipeline between multiple computing systems (Glatard et al., 2015). It is likely that any one of these circumstances played a role in the lack of statistical relation between replicated and original data.

Another possibility for this result relates to the harmonization process that the original CENC data underwent before analysis. An issue that has affected multi-site neuroimaging consortia is that of variability introduced into data being collected at multiple sites. CENC data used in this analysis contains data collected across eight sites. ComBat harmonization is a machine learning algorithm that transforms data to reduce the site-to-site variability (Fortin et al., 2017). ComBat has efficacy in the reduction of unwanted variability introduced by individual sites, while maintaining biological differences and enhancing replicability in subsequent analyses (Fortin et al., 2018). In the process of reducing this unwanted variability, data undergoes multiple transformations. This may account for the lack of relation demonstrated between replicated and original datasets. While data in the replication appears to be similar to the original dataset, there

also appears to be no significant relations between replicated and original data. It may be necessary to both replicate and harmonize data to obtain more accurate results.

2.1.5 Statistical Analysis

For each identified tract (the uncinate fasciculus, the genu of the corpus callosum, and the cingulum bundle), fractional anisotropy (FA) and mean diffusivity (MD) were used in statistical analyses and were extracted using the analysis pipeline described above. Analysis models were constructed to answer specific aspects of the hypotheses of Study 1.

Due to the heterogeneity in the dataset, there were only rarely subjects who only had a blast-related TBI or only a non-blast related TBI. As a result, investigations of participant group and number of blast exposures were used to answer Hypotheses 1.1 and 1.2. In this dataset, study group consisted of five levels. Subjects were assigned to a group based on the injury histories they reported. Groups included mild TBI acquired during only combat, mild TBIs acquired both during combat and outside of combat, mild TBI acquired before combat only, mild TBI acquired after combat only, and unexposed. Here, unexposed refers to subjects who did not meet the requirements for diagnosis of a mild TBI after investigation of health history. Number of blast exposures resulting in TBI and not resulting in TBI were calculated using the information described in Section 2.1.1. Using participant group information allowed for the investigation of whether TBIs acquired where exposure to blasts was frequent caused different alterations in WM integrity than TBIs acquired outside of exposure to blasts (pre and post combat, specifically). It also allowed for the investigation of differences between unexposed (uninjured) and exposed (injured) subjects. To address Hypothesis 1.1, multiple linear regressions were performed on FA and MD with study group levels and blast variables (number of blast exposures resulting in TBI and number of blast exposures not resulting in TBI) as model variables, psychological disorder screening measures as covariates, and diffusivity outcomes across all 3 tracts as a dependent variable. Two additional multiple regressions were run using the same variables, but with global average FA and global average MD as outcomes.

Further investigation of Hypothesis 1.2 was completed by using both generalized linear models and multilevel generalized linear models. The only difference between the generalized linear model and the multilevel generalized linear model was that the multilevel model used participant

group as a random effect. Due to the heterogeneity issue previously mentioned, there were rarely subjects with a history of only blast-related TBI and only non-blast related TBI. Hypothesis 1.2 specifically asks about differences based on blast exposure, and whether blast exposure itself would result in any significant differences in white matter integrity. In order to indirectly answer this question, number of blast exposures were used in generalized linear models. Model variables included number of blasts resulting in TBI and number of blasts not resulting in TBI. Although this does not specifically answer Hypothesis 1.2, it does allow for the investigation of whether diagnosed blast injuries result in significantly different alterations in WM integrity than exposure to blasts alone. Assuming there is a unique bio-mechanical mechanism underlying blast-related TBI, one would expect that diagnosed injuries would show decreased WM integrity than blast exposures without diagnosed injuries. When constructing models to answer this question, each generalized linear model additionally included error terms (including translation, rotation, and acquisition site) as covariates. A blast exposure measure was constructed by combining variables for exposure to uncontrolled and controlled blasts that possibly resulted in TBI. These models were constructed for each diffusion measure for each of the pre-specified bilateral tracts (uncinate fasciculus, cingulum, and the genu of the corpus callosum), including the blast exposure variables, the CVLT-II and the D-KEFS summary scores in the model, and using volume-by-volume translation, rotation, premorbid functioning (as measured by the Test of Premorbid Functioning), psychological disorder screening measures, and age as covariates. The CVLT-II measures included were the scores on the long and short delay cued recall subscales. The D-KEFS measure included was the score on the letter correct subscale. The CVLT-II cued recall subscale was chosen as evidence shows cued recall is likely more effective in ensuring similar learning rates between TBIs and controls (Vakil and Oded, 2003). The letter fluency correct trial from the D-KEFS was chosen because this trial relies on phonetic word retrieval, rather than semantic knowledge retrieval, reducing possible collinearity between this test and the CVLT-II (Strong et al., 2010). Additionally, these covariates were selected as they are commonly associated with white matter findings from DTI. Volume-by-volume translation was extracted from eddy current-corrected .ecclog and .nii files using a script from the FMIRB, available on the OSF repository. This script distinguishes between eddy-current induced and subject-induced volume translation and rotation. All statistical analyses were performed using R's (v3.6.0) *lme4*

package and the *glm* function. The *lme4* package is a set of statistical techniques to conduct generalized linear models (GLMs). Although many variables in the dataset were not normally distributed, the *glm* function compensates for this by transforming the distribution of the data to a normal curve (specified by the *family* option). Model fit was assessed using Akaike information criterion (AIC) estimators, generated by *lme4* at the end of each analysis. AIC estimates measure out of sample prediction error through use of a maximum likelihood estimation. Both a generalized linear model and a multilevel generalized linear model were run using the previously described model variables and covariates. The multilevel generalized linear model used TBI study group as a nesting variable.

2.1.6 Multiple Imputation

Inspection of the dataset revealed 1,200 missing observations in the CENC dataset (a missing data rate of .006%). In order to conduct final analyses, R requires complete observations for each subject. Multiple imputation was used to estimate values for these missing observations. Multiple imputation is regarded as an effective technique to deal with missingness in psychological science, as it is comparable to similar methods (like maximum likelihood estimations) and allows for complexities in data (Enders, 2017). A multiple imputation would allow for the regressions and generalized linear models to be conducted. A number of preparatory steps were completed in order to perform multiple imputation. To ensure data was appropriate for multiple imputation, missingness was dummy coded (0= missing observation, 1=complete observation) for intended model variables (Age, premorbid IQ, scores on the D-KEFS and the CVLT, blast exposure variables, ethnicity, and education level) using the *complete.cases* function in R. Once dummy coded missingness variables were generated, t-tests were performed between the dummy coded missing variable and the normal variable using the *t.test* function. Theoretically, any dummy coded missing variables should be statistically different from their normal counterparts, assuming data is either missing at random or missing completely at random. Table 2.1 shows the results of t tests between these two variable types. P values were adjusted for multiple comparisons using a Bonferroni correction through the *p.adjust* function in R. Even after adjusting for multiple comparisons, all t tests produced significant differences between dummy coded missing variables and normal (variables with both complete and missing observations) variables. This suggests that

Table 2.1: T Tests between Dummy Coded Missing Variables and Normal Variables

Variable	t Value	df	Adjusted p Value	95% CI
Age	176.02	945	$p = 2.64 * 10^{-15}$	.960 - .981
TOPF Premorbid IQ	332.87	945	$p = 2.64 * 10^{-15}$	.985 - .997
D-KEFS Category Correct	667.86	945	$p = 2.64 * 10^{-15}$	.995 - 1.00
D-KEFS Letter Correct	667.86	945	$p = 2.64 * 10^{-15}$	.995-1.00
CVLT-II Long Delay Cued Recall	384.77	945	$p = 2.64 * 10^{-15}$	.988-.998
CVLT-II Long Delay Free Recall	384.77	945	$p = 2.64 * 10^{-15}$	.988-.998
CVLT-II Short Delay Cued Recall	384.77	945	$p = 2.64 * 10^{-15}$	.988-.998
CVLT-II Short Delay Free Recall	384.77	945	$p = 2.64 * 10^{-15}$	.988-.998
Num Exposures to Controlled Blasts	209.17	945	$p = 2.64 * 10^{-15}$	.969-.988
Num Exposures to Uncontrolled Blasts	209.17	945	$p = 2.64 * 10^{-15}$	.969-.988
Num Exposures to Blasts not resulting in TBI	209.17	945	$p = 2.64 * 10^{-15}$	.969-.988
Total Num of Exposures to Blasts	209.17	945	$p = 2.64 * 10^{-15}$	.969-.988
Ethnicity	-13.277	945	$p = 2.64 * 10^{-15}$	-.69 - -.51
Education	–	–	–	–

*Note that education was not missing any observations, which is why no results for this variable were reported here. Additionally, the results for the subscales on the D-KEFS, CVLT, and Blast exposure variables are the same because subjects who were missing data on these measures did not have partial data on other subscales.

data were at least missing at random, as there is no association between missingness and the normal variables. Data missing at random is a good fit for multiple imputation (Donders et al., 2006).

Scaled/standard scores were dropped from the dataset in favor of raw scores, as the final statistical model would account for the same factors as the standard scores (e.g., age, education, and gender). Some scales in the dataset, such as the PHQ-9 and PCL-5, included cluster scores for subscales in addition to total raw scores. Cluster scores were dropped from these scales in favor of the total raw scale scores, as multiple raw scale scores would allow for more specific analyses than one cluster. Last, variables that were missing the majority of responses (>50%) were removed. Among the scales dropped for missing data were the EQ-5D-5L, the Hearing Handicap Inventory for Adults, the Pittsburgh Sleep Quality Index, and subscales from the Behavioral Risk Factor Surveillance System. There were a number of variables that only applied to a small subset of the entire sample, and thus contained responses for a small portion of all subjects. These variables were dropped for majority missingness. After removing variables based on these

criteria, the final dataset consisted of 207 variables across 946 observations. A full list of variables with the number of missing observations can be found in Appendix B.

Multiple imputation was completed using the *mice* (Multivariate Imputation by Chained Equations) package in R (van Buuren and Groothuis-Oudshoorn, 2011). This package has the ability to conduct imputation for each variable in a dataset, using Fully Conditional Specification (Van Buuren, 2007, FCS). FCS specifies an imputation model for each incomplete variable and generates imputed values iteratively. To complete imputation, mice requires a researcher-specified imputation model (through the required -meth flag). For the purposes of this investigation, the CART method was chosen, which uses classification and regression trees to model missingness. Regression trees are a machine learning algorithm in which existing data within a dataset is summarized (by taking the mean of response values for complete observations) and used to predict, through use of regression, the value of missing observations while also minimizing the residual sum of squares. As the existing variables in the CENC dataset do not have the same structure (e.g., different numbers of levels per variable), other imputation methods, such as predictive mean matching, were deemed inappropriate. The described CART imputation was conducted on a remote supercomputing cluster. However, in order to parallelize computing, a variation of *mice* called *parlmice* was used, which is a joint package using the imputation functions from *mice* and the parallelization functions from the *parallel* package. *Parlmice* allows R to parallelize analyses across multiple nodes and CPU cores. This package also requires specifications for number of iterations, number of cores, number of imputations per core, and clustering method. This imputation was run across 5 iterations to improve accuracy. It is generally recommended that the number of iterations should reflect the amount of missing data (1 iteration per 1% missing data) (White et al., 2011). The final analysis used 10 CPU cores and ran 155 imputations per core. In order to optimize the analysis for the Linux-based compute cluster, the FORK parallelization method was used. Inspection of missing data revealed no individual subject was missing more than 17% of observations in the dataset.

2.2 Results

2.2.1 Descriptive Information

The sample with neuroimaging data included 946 subjects (820 Male, 126 Female). Ages of subjects were highly variable (mean=40.47, SD=10.04). Education among subjects varied, with most holding a high school diploma (n=122), Associates Degree (n=96), Bachelors' Degree (n=247), or Master's Degree (n=138). Portions of the subjects had attended some college but held no degree (n=192). Most identified as White (n=683) or African American (n=180). Most subjects in the final sample reported a history of traumatic brain injury (n=769), although a small subset of the sample did not report exposure consistent with a traumatic brain injury (n=180). Subjects exhibited a wide range of injury characteristics and injury mechanisms. Over half of the sample reported acquiring a brain injury during military service deployment (n=577). Portions of these injuries were blast exposure-related (n=350) TBI or TBIs resulting from general exposure (n=432). Additionally, most reported exposure to blast, either controlled or uncontrolled, during their military service (mean=3.35 exposures, SD=2.28). Of those reporting a history of TBI, most reported post-traumatic amnesia (n=717) and loss of consciousness after injury (n=647). Subjects reporting a history of TBI also reported a wide ranging number of TBI events (min=0, max=12, mean=2.33, SD=2.1). Subjects averaged a premorbid IQ of 102 (SD=7.88). A full list of descriptive information can be found in Tables 2.5-2.7

Three one-way ANOVAs were performed using age, premorbid IQ, and years of education as outcome variables and study group as a model variable to test for differences between study groups (described above). The ANOVAs for age ($F(4,941) = 5.13, p = .0004$), education ($F(4,919) = 4.104, p = .003$), and premorbid IQ ($F(4,941) = 9.45, p = 1.72 * 10^{-7}$) indicated significant differences between study groups. Posthoc Tukey's HSD tests were conducted on the output from each ANOVA. The results from each Tukey's HSD can be found below.

Table 2.2: Tukey Group Comparisons: Education

Comparison	Difference	Lower Interval	Upper Interval	Adjusted p value
Combat mTBI and Non-combat mTBI *X* Combat mTBI only	.597	.038	1.16	$p = .03^{*}$
Post-combat mTBI *X* Combat mTBI only	.709	-.043	1.46	$p = .075$
Pre-combat mTBI only *X* Combat mTBI only	.938	.276	1.60	$p = .001^{***}$
Unexposed *X* Combat mTBI only	.643	.003	1.28	$p = .05^{*}$
Post-combat mTBI only *X* Combat mTBI and Non-combat mTBI	.111	-.55	.778	$p = .990$
Pre-combat mTBI only *X* Combat mTBI and Non-combat mTBI	.341	-.222	.903	$p = .464$
Unexposed *X* Combat mTBI and Non-combat mTBI	.045	-.491	.582	$p = .999$
Pre-combat mTBI only *X* Post-combat mTBI only	.229	-.526	.984	$p = .922$
Unexposed *X* Post-combat mTBI only	-.066	-.801	.669	$p = .999$
Unexposed *X* Pre-combat mTBI only	-.295	-.938	.348	$p = .719$

Table 2.3: Tukey Group Comparisons: Premorbid IQ

Comparison	Difference	Lower Interval	Upper Interval	Adjusted p value
Combat mTBI and Non-combat mTBI *X* Combat mTBI only	4.04	2.01	6.07	$p = 7 * 10^{-7***}$
Post-combat mTBI *X* Combat mTBI only	3.53	0.793	6.27	$p = .004^{**}$
Pre-combat mTBI only *X* Combat mTBI only	4.77	2.36	7.19	$p = 8 * 10^{-7***}$
Unexposed *X* Combat mTBI only	2.80	.472	5.13	$p = .009^{**}$
Post-combat mTBI only *X* Combat mTBI and Non-combat mTBI	-.516	-2.93	1.91	$p = .978$
Pre-combat mTBI only *X* Combat mTBI and Non-combat mTBI	.730	-1.33	2.79	$p = .869$
Unexposed *X* Combat mTBI and Non-combat mTBI	-1.24	-3.19	.709	$p = .410$
Pre-combat mTBI only *X* Post-combat mTBI only	1.25	-1.51	4.00	$p = .731$
Unexposed *X* Post-combat mTBI only	-.728	-3.41	1.95	$p = .946$
Unexposed *X* Pre-combat mTBI only	-1.97	-4.33	.380	$p = .148$

Table 2.4: Tukey Group Comparisons: Age

Comparison	Difference	Lower Interval	Upper Interval	Adjusted p value
Combat mTBI and Non-combat mTBI *X* Combat mTBI only	2.93	.378	5.49	$p = .02^{*}$
Post-combat mTBI *X* Combat mTBI only	3.22	-.228	6.66	$p = .080$
Pre-combat mTBI only *X* Combat mTBI only	4.96	1.92	8.00	$p = .00009^{***}$
Unexposed *X* Combat mTBI only	2.83	-.106	5.76	$p = .065^{*}$
Post-combat mTBI only *X* Combat mTBI and Non-combat mTBI	.283	-2.77	3.34	$p = .999$
Pre-combat mTBI only *X* Combat mTBI and Non-combat mTBI	2.03	-.562	4.62	$p = .204$
Unexposed *X* Combat mTBI and Non-combat mTBI	-.107	-2.57	2.35	$p = .999$
Pre-combat mTBI only *X* Post-combat mTBI only	1.74	-1.73	5.22	$p = .645$
Unexposed *X* Post-combat mTBI only	-.390	-3.76	2.98	$p = .998$
Unexposed *X* Pre-combat mTBI only	-2.13	-5.09	.828	$p = .282$

As seen in Tables 2.1-2.3, there are a number of significant differences between study groups, particularly between combat acquired TBI and pre/post combat acquired TBI. These differences should be kept in mind when viewing any significant results from Study 1 and Study 2. It indeed appears there are statistically significant differences between at least 3 groups on education, premorbid IQ, and age. All of these factors have the potential to influence results.

Table 2.5: Descriptive Information of the CENC Cohort

Variable	n	Mean	Standard Deviation
Gender: Male	820	–	–
Gender: Female	126	–	–
Ethnicity: White	683	–	–
Ethnicity: Black/African-American	180	–	–
Ethnicity: American Indian or Native Alaskan	8	–	–
Ethnicity: Asian	9	–	–
Ethnicity: Pacific Islander	9	–	–
Ethnicity: Other	47	–	–
Ethnicity: Don't Know/Not Sure	5	–	–
Ethnicity: Refused	5	–	–
Age	–	40.47	10.04
TOPF Demographic Adjusted IQ	–	102	7.88
History of TBI	769	–	–
No History of TBI	177	–	–
Blast Exposure TBI	350	–	–
General Exposure TBI	432	–	–
Number of TBI with LOC	647	–	–
Number of TBI with PTA	717	–	–
Education: High School	122	–	–
Education: Some College, No Degree	192	–	–
Education: Associates Degree	96	–	–
Education: Bachelors' Degree	247	–	–
Education: Master's Degree	138	–	–
Number of TBI Events	–	2.33	2.10
Number Meeting PCL-5 PTSD Criteria	171	–	–
Number Meeting PHQ-9 Major Depression Criteria	184	–	–
Number Meeting PHQ-9 Minor Depression Criteria	103	–	–
MINI: Has Not Experienced Traumatic Event	177	–	–
MINI: PTSD Positive	241	–	–
MINI: PTSD Provisionally Positive	12	–	–
MINI: Criteria for PTSD Not Met	511	–	–
DAST: No Problems Reported	812	–	–
DAST: Low Level Problems	105	–	–
DAST: Moderate Level Problems	17	–	–
DAST: Substantial Problems	5	–	–
DAST: Severe Problems	4	–	–

Table 2.6: Descriptive Information of the CENC Cohort: Continued

Variable	n
AUDIT: Never Use Alcohol	186
AUDIT: Monthly Alcohol Use	262
AUDIT: Alcohol Use 2-4 Times per Month	255
AUDIT: Alcohol Use 2-3 Times per Week	147
AUDIT: Alcohol Use 4 or More Times per Week	91
BRFSS: General Health is Excellent	62
BRFSS: General Health is Very Good	218
BRFSS: General Health is Good	379
BRFSS: General Health is Fair	254
BRFSS: General Health is Poor	33
GOSE: Lower Severe Disability	62
GOSE: Upper Severe Disability	65
GOSE: Lower Moderate Disability	148
GOSE: Upper Moderate Disability	211
GOSE: Lower Good Recovery	156
GOSE: Upper Good Recovery	301

Table 2.7: Descriptive Information of the CENC Cohort: Neuropsych Scores

Test	Combat Only	Combat+NonCombat	Pre-combat Only	Post-combat Only	Unexposed
CVLT-II Long Delay Cued Recall	10.48	10.99	11.36	11.33	10.89
CVLT-II Short Delay Cued Recall	10.64	11.01	11.41	11.35	11.03
CVLT-II Long Delay Free Recall	9.87	10.20	10.44	10.57	10.31
CVLT-II Short Delay Free Recall	9.77	9.98	10.15	10.38	10.28
D-KEFS Letter Correct	38.47	39.61	41.85	40.09	37.75
D-KEFS Category Correct	40.37	42.36	41.53	41.67	40.92
TOPF Demographically Adjusted IQ	98.78	102.84	102.33	103.57	101.60

2.2.2 Multiple Imputation

Before conducting the final analysis, data imputation was completed through the *mice* package in R. The imputation was completed over five iterations. The final iteration was used in both Study 1 and Study 2 analyses. The final dataset contained 1,262 imputed datapoints out of 195,822 total observations (946 observations * 207 variables), indicating a missing data rate of 0.006%. Using the final imputed dataset, generalized linear models were built using each diffusivity metric in each pre-determined tract as an outcome variable. Tract variables were the average of the left and right hemispheric tracts, as there was no expectation about hemispheric differences. All analyses were completed using the *glm* function of the *lme4* package (Bates et al., 2015). Each model of fractional anisotropy used a binomial distribution, as the outcome variable is discrete (bounded between 0 and 1). The *outliersKD* function was used to identify outliers +/- 1.5 interquartile ranges outside of the distribution of the movement variables. Forty-one outliers were identified and dropped from the final model. These observations were dropped instead of Winsorized because evidence suggests that subjects with a significant amount of motion during MRI scanning renders any data extracted from these scans unusable. The final analysis was conducted on 946 subjects with complete, unique observations.

2.2.3 Analysis 1: Linear Regressions

In order to answer whether participant group had any significant statistical effect on ROI-specific diffusivity outcomes, linear regressions were conducted using study group and two blast exposure variables (number of blast exposures resulting in TBI and number of blast exposures not resulting in TBI) as model variables and psychological disorder screening variables as covariates with FA and MD in the uncinate fasciculus, the genu of the corpus callosum, and the cingulum bundle as outcomes. Code used to generate these results can be found in Appendix A.

Linear regressions infrequently showed significance for some group and blast variables. As seen in Tables 2.8-2.15, age, total score on the DAST, number of blast exposures resulting in TBI and combat mild TBI were significant in the statistical models. Although this appears to partially support *a priori* hypotheses, these results should be viewed with caution. These models in most cases accounted for roughly 4% of the variance in diffusivity outcomes. Given this incredibly

Table 2.8: Linear Regression: FA in the Uncinate Fasciculus

Variable	Coefficient	SE	T-Value	Significance
Intercept	.538	.004	138.54	$p = 2 * 10^{-16}$
Combat Mild TBI Only	.004	.005	.847	$p = .397$
Pre-Combat Mild TBI Only	.003	.005	.618	$p = .536$
Post-Combat Mild TBI Only	.003	.006	.544	$p = .586$
Combat TBI + Non-combat TBI	-.003	-.006	-.584	$p = .586$
Unexposed	.007	.005	1.453	$p = .147$
Num Blast Exposures Resulting in TBI	-.003	.002	-1.40	$p = .163$
Num Blast Exposures Not Resulting in TBI	.0003	.0007	.455	$p = .163$
Age	-0.0007	0.001	-4.633	$p = 4.13 * 10^{-6}$
PCL-5 Total	$3.14 * 10^{-5}$	$1.44 * 10^{-4}$	.219	$p = .827$
PHQ-9 Total	$1.79 * 10^{-6}$	$4.45 * 10^{-4}$	.004	$p = .997$
AUDIT-C Total	$-4.94 * 10^{-3}$	$3.41 * 10^{-3}$	-1.45	$p = .147$
DAST-10 Total	$-2.77 * 10^{-3}$	$1.67 * 10^{-3}$	-1.66	$p = .098$
Adjusted R-squared = .026				

Table 2.9: Linear Regression: FA in the Genu of the Corpus Callosum

Variable	Coefficient	SE	T-Value	Significance
Intercept	.655	.003	229.79	$p = 2 * 10^{-16}$
Combat Mild TBI Only	.008	.004	.959	$p = .016^{**}$
Pre-Combat Mild TBI Only	-.001	.004	-.303	$p = .762$
Post-Combat Mild TBI Only	.004	.004	.959	$p = .338$
Combat TBI + Non-combat TBI	-.008	.004	-2.41	$p = .016$
Unexposed	-.0005	.003	-.139	$p = .890$
Num Blast Exposures Resulting in TBI	-.001	.002	-.776	$p = .438$
Num Blast Exposures Not Resulting in TBI	-.0002	.0006	-.492	$p = .623$
Age	-0.002	-.0001	-13.215	$p = 2 * 10^{-16}$
PCL-5 Total	$-6.91 * 10^{-5}$	$9.85 * 10^{-5}$	-.702	$p = .483$
PHQ-9 Total	$-7.94 * 10^{-5}$	$3.06 * 10^{-4}$	-.260	$p = .795$
AUDIT-C Total	$-1.89 * 10^{-3}$	$2.35 * 10^{-3}$	-.802	$p = .423$
DAST-10 Total	$5.99 * 10^{-4}$	$1.12 * 10^{-3}$	.535	$p = .593$
Adjusted R-squared = .160				

Table 2.10: Linear Regression: FA in the Cingulum Bundle

Variable	Coefficient	SE	T-Value	Significance
Intercept	.570	.003	194.86	$p = 2 * 10^{-16}$
Combat Mild TBI Only	.005	.003	1.52	$p = .129$
Pre-Combat Mild TBI Only	.005	.004	1.25	$p = .212$
Post-Combat Mild TBI Only	.004	.004	.849	$p = .396$
Combat TBI + Non-combat TBI	-.003	.004	-.849	$p = .396$
Unexposed	-.001	.003	-.523	$p = .601$
Num Blast Exposures Resulting in TBI	-.0008	.002	-.538	$p = .591$
Num Blast Exposures Not Resulting in TBI	.0006	.0006	1.10	$p = .270$
Age	-0.0007	0.0001	-6.430	$p = 2.08 * 10^{-10}$
PCL-5 Total	$2.20 * 10^{-5}$	$1.04 * 10^{-4}$	.212	$p = .832$
PHQ-9 Total	$-4.52 * 10^{-4}$	$3.22 * 10^{-4}$	-1.40	$p = .161$
AUDIT-C Total	$8.95 * 10^{-4}$	$2.48 * 10^{-3}$	-361	$p = .718$
DAST-10 Total	$1.89 * 10^{-3}$	$1.18 * 10^{-3}$	1.60	$p = .110$
Adjusted R-squared = .044				

Table 2.11: Linear Regression: MD in the Uncinate Fasciculus

Variable	Coefficient	SE	T-Value	Significance
Intercept	.0007	$6.41 * 10^{-6}$	122.85	$p = 2 * 10^{-16}$
Combat Mild TBI Only	$4.05 * 10^{-7}$	$7.58 * 10^{-6}$	.053	$p = .366$
Pre-Combat Mild TBI Only	$7.88 * 10^{-6}$	$8.33 * 10^{-6}$	-.946	$p = .344$
Post-Combat Mild TBI Only	$8.64 * 10^{-6}$	$9.56 * 10^{-6}$	.905	$p = .366$
Combat TBI + Non-combat TBI	$-8.64 * 10^{-6}$	$9.56 * 10^{-6}$	-.905	$p = .366$
Unexposed	$-2.92 * 10^{-6}$	$7.94 * 10^{-6}$	-.368	$p = .712$
Num Blast Exposures Resulting in TBI	$6.92 * 10^{-6}$	$3.47 * 10^{-6}$	1.992	$p = .046^{*}$
Num Blast Exposures Not Resulting in TBI	$-1.39 * 10^{-6}$	$1.26 * 10^{-6}$	-1.11	$p = .267$
Age	$1.25 * 10^{-6}$	$2.17 * 10^{-7}$	4.609	$p = 4.63 * 10^{-6}$
PCL-5 Total	$2.64 * 10^{-7}$	$2.30 * 10^{-7}$	1.15	$p = .250$
PHQ-9 Total	$4.06 * 10^{-7}$	$7.13 * 10^{-7}$	.570	$p = .569$
AUDIT-C Total	$3.68 * 10^{-6}$	$5.49 * 10^{-6}$	.670	$p = .503$
DAST-10 Total	$-3.27 * 10^{-6}$	$2.61 * 10^{-6}$	-1.26	$p = .210$
Adjusted R-squared = .029				

Table 2.12: Linear Regression: MD in the Genu of the Corpus Callosum

Variable	Coefficient	SE	T-Value	Significance
Intercept	.0007	$7.48 * 10^{-6}$	95.76	$p = 2 * 10^{-16}$
Combat Mild TBI Only	$-3.68 * 10^{-6}$	$8.83 * 10^{-6}$	-.417	$p = .676$
Pre-Combat Mild TBI Only	$7.54 * 10^{-6}$	$9.71 * 10^{-6}$	.776	$p = .437$
Post-Combat Mild TBI Only	$1.51 * 10^{-5}$	$1.11 * 10^{-5}$	1.40	$p = .161$
Combat TBI + Non-combat TBI	$3.69 * 10^{-6}$	$8.38 * 10^{-6}$	.417	$p = .676$
Unexposed	$8.23 * 10^{-6}$	$9.25 * 10^{-6}$	.889	$p = .374$
Num Blast Exposures Resulting in TBI	$6.81 * 10^{-6}$	$4.05 * 10^{-6}$	1.68	$p = .093$
Num Blast Exposures Not Resulting in TBI	$-2.56 * 10^{-6}$	$1.47 * 10^{-6}$	-1.74	$p = .081$
Age	$2.86 * 10^{-6}$	$3.08 * 10^{-7}$	9.29	$p = 2 * 10^{-16}$
PCL-5 Total	$4.14 * 10^{-7}$	$2.58 * 10^{-7}$	1.60	$p = .109$
PHQ-9 Total	$4.62 * 10^{-8}$	$8.01 * 10^{-7}$	.059	$p = .954$
AUDIT-C Total	$-4.93 * 10^{-6}$	$6.17 * 10^{-6}$	-.800	$p = .424$
DAST-10 Total	$-5.54 * 10^{-6}$	$2.93 * 10^{-6}$	-1.89	$p = .059$
Adjusted R-squared = .001				

Table 2.13: Linear Regression: MD in the Cingulum Bundle

Variable	Coefficient	SE	T-Value	Significance
Intercept	.0007	$7.12 * 10^{-6}$	95.76	$p = 2 * 10^{-16}$
Combat Mild TBI Only	$-1.09 * 10^{-5}$	$9.09 * 10^{-6}$	-1.21	$p = .227$
Pre-Combat Mild TBI Only	$1.13 * 10^{-5}$	$8.79 * 10^{-6}$	-1.29	$p = .200$
Post-Combat Mild TBI Only	$1.59 * 10^{-5}$	$8.16 * 10^{-5}$	1.96	$p = .051$
Combat TBI + Non-combat TBI	$-1.60 * 10^{-6}$	$8.16 * 10^{-6}$	-1.29	$p = .200$
Unexposed	$-9.96 * 10^{-6}$	$8.53 * 10^{-6}$	-1.17	$p = .243$
Num Blast Exposures Resulting in TBI	$8.28 * 10^{-6}$	$2.97 * 10^{-6}$	2.79	$p = .005^{**}$
Num Blast Exposures Not Resulting in TBI	$-2.56 * 10^{-6}$	$1.47 * 10^{-6}$	-1.74	$p = .081$
Age	$1.08 * 10^{-6}$	$2.32 * 10^{-7}$	4.66	$p = 3.70 * 10^{-6}$
PCL-5 Total	$3.98 * 10^{-7}$	$1.95 * 10^{-7}$	2.04	$p = .042^{*}$
PHQ-9 Total	$1.34 * 10^{-7}$	$6.06 * 10^{-7}$	.221	$p = .825$
AUDIT-C Total	$1.37 * 10^{-6}$	$4.66 * 10^{-6}$	.294	$p = .769$
DAST-10 Total	$-5.08 * 10^{-6}$	$2.21 * 10^{-6}$	-2.29	$p = .022^{*}$
Adjusted R-squared = .040				

Table 2.14: Linear Regression: Global Average FA

Variable	Coefficient	SE	T-Value	Significance
Intercept	.436	.001	315.59	$p = 2 * 10^{-16}$
Combat Mild TBI Only	.002	.002	.963	$p = .336$
Pre-Combat Mild TBI Only	-.002	.002	-.848	$p = .397$
Post-Combat Mild TBI Only	.0007	.002	.342	$p = .732$
Combat TBI + Non-combat TBI	-.002	.002	-.963	$p = .336$
Unexposed	-.001	.001	-.723	$p = .470$
Num Blast Exposures Resulting in TBI	-.001	.0007	-1.53	$p = .125$
Num Blast Exposures Not Resulting in TBI	.0001	.0002	4.62	$p = .644$
Age	$-6.48 * 10^{-4}$	$5.43 * 10^{-5}$	-11.94	$p = 2 * 10^{-16}$
PCL-5 Total	$-5.85 * 10^{-5}$	$4.65 * 10^{-5}$	-1.26	$p = .209$
PHQ-9 Total	$-4.32 * 10^{-5}$	$1.44 * 10^{-4}$	-.300	$p = .765$
AUDIT-C Total	$-1.46 * 10^{-3}$	$1.11 * 10^{-3}$	-1.31	$p = .189$
DAST-10 Total	$-5.84 * 10^{-5}$	$5.28 * 10^{-4}$	-0.111	$p = .912$

Adjusted R-squared = .134

Table 2.15: Linear Regression: Global Average MD

Variable	Coefficient	SE	T-Value	Significance
Intercept	.0007	$4.79 * 10^{-6}$	146.09	$p = 2 * 10^{-16}$
Combat Mild TBI Only	$4.59 * 10^{-6}$	$5.66 * 10^{-6}$	.810	$p = .418$
Pre-Combat Mild TBI Only	$1.56 * 10^{-6}$	$6.22 * 10^{-6}$	.251	$p = .801$
Post-Combat Mild TBI Only	$1.30 * 10^{-5}$	$7.13 * 10^{-6}$	1.83	$p = .070$
Combat TBI + Non-combat TBI	$-1.30 * 10^{-5}$	$7.14 * 10^{-6}$	-1.83	$p = .068$
Unexposed	$2.10 * 10^{-6}$	$5.93 * 10^{-6}$	.355	$p = .722$
Num Blast Exposures Resulting in TBI	$7.02 * 10^{-6}$	$2.60 * 10^{-6}$	2.70	$p = .007^{**}$
Num Blast Exposures Not Resulting in TBI	$-9.45 * 10^{-7}$	$9.40 * 10^{-7}$	-1.00	$p = .315$
Age	$1.007 * 10^{-6}$	$2.03 * 10^{-7}$	4.96	$p = 8.3 * 10^{-7}$
PCL-5 Total	$3.64 * 10^{-7}$	$1.70 * 10^{-7}$	2.14	$p = .032^{*}$
PHQ-9 Total	$1.89 * 10^{-7}$	$5.29 * 10^{-7}$	.358	$p = .720$
AUDIT-C Total	$-1.39 * 10^{-6}$	$4.07 * 10^{-6}$	-.343	$p = .732$
DAST-10 Total	$-4.17 * 10^{-6}$	$1.93 * 10^{-6}$	-2.16	$p = .031^{*}$

Adjusted R-squared = .046

small portion of variance accounted for, significant coefficients should be viewed skeptically. It is well documented that significance statistics are inflated in large samples, and are virtually meaningless (Smith and Nichols, 2018). Given the large sample this analysis was conducted on, and the small amount of variance explained, it is increasingly likely that these significance values are the result of inflation. Any inferences made about effects reported here should be viewed within this context. The strongest model was one considering FA in the genu of the corpus callosum, which accounted for 16% of the variance in FA. Even this model should be viewed skeptically, as significant variables still accounted for a very small linear association between independent and dependent variables.

2.2.4 Analysis 2: Generalized Linear Models

Both a multilevel mixed effects generalized linear model and generalized linear model were built using the *glmer* function of the *lme4* package in R (Bates et al., 2015). A mixed effects method was chosen as it allows for the specification of a nesting variable. A nested model is a model that allows for a lower level factor to only appear within a higher level factor– in this case, subjects within TBI groups. This effectively allows for the investigation of unique group effects. This model used study group as a nesting variable, but did not produce different results from the previously described generalized linear model. All results reported here are from the generalized linear model, not the multilevel mixed effects model. Code used to generate all statistical models can be found in Appendix A.

Tables 2.16-2.22 show results indicating poor predictive ability of all the covariates in the model, with the exception of subject-induced volume translation. Although these models demonstrate minimal deviance (the measure of how well the specified model fits when compared to a saturated model), the AIC statistic generated for each suggests that these models are far from being a parsimonious solution. This, in turn, suggests a smaller model with fewer covariates would be more optimized for the current data structure. Consider the guidelines proposed previous studies (Burnham and Anderson, 2004). When comparing the AIC from multiple models, one does not compare the absolute AIC values produced, but the difference between the two. The equation $\Delta_i = AIC_i - AIC_{min}$ is suggested for comparing two AIC values, where AIC_i is the ith model and AIC_{min} is the minimum value obtained from any model. A Δ_i greater than 10 is considered to have

no support. Using this equation with the $AIC_{min} = 23.25$, one can see that the differences between each model run and the minimum AIC value averages much greater than 10, often hundreds or thousands of points higher. This greatly undermines the support for these models being an appropriate fit for the underlying data structure. This is an unexpected result, given that there were strong theoretical reasons to include all the model variables and covariates present in the specified generalized linear models. Collinearity does not seem to be a pressing issue, as the correlations between items reported in Table 2.22 appear to be fairly small. None of the neuropsychological testing measures or injury mechanisms resulted in model coefficients significant enough to make any meaningful inferences about the diffusivity outcomes. Based on the current results, Hypotheses 1.1 and 1.2 were not supported, as no evidence was found indicating a difference between unexposed (non-TBI) and exposed (TBI) subjects in the post-hoc tests conducted on the previously conducted one-way ANOVA (see Section 2.2.1). Further, Hypothesis 1.2 was unsupported, as the variables representing blast TBI and non-blast TBI (participant group) in the generalized linear models had no significant predictive effect on global diffusivity outcome.
It is concerning that the only coefficient with any predictive power appears to be subject induced volume translation during the MRI sequences. Although the neuroimaging data used in the final model was corrected for both eddy current-induced and subject-induced volume translation and rotation, it appears this data is still significantly affected by the presence of these variables. It should be noted that these results appear after significant data cleaning, including imputation of missing datapoints, correction of neuroimaging data for subject induced and eddy current induced volume translation and rotation, and removal of outlying datapoints.

Table 2.16: Fractional Anisotropy in the Uncinate Fasciculus

Variable	Coefficient	Standard Error	z-Value	p-Value
Intercept	.518	.995	.521	$p = .602$
CVLT-2 Long Delay Cued Recall	$-4.74 * 10^{-3}$	.054	-.087	$p = .931$
CVLT-2 Short Delay Cued Recall	$3.84 * 10^{-3}$	.057	.067	$p = .946$
D-KEFS Letter Correct	$2.35 * 10^{-4}$	$6.47 * 10^{-3}$	.036	$p = .971$
TOPF Demographic Adjusted IQ	$-1.76 * 10^{-3}$	$9.28 * 10^{-3}$	-.190	$p = .850$
Age	$-3.01 * 10^{-3}$	$7.58 * 10^{-3}$	-.397	$p = .691$
Num Blast Exposures Resulting in TBI	$-1.41 * 10^{-2}$	$7.94 * 10^{-2}$	-.177	$p = .859$
Num Blast Exposures Not Resulting in TBI	$7.27 * 10^{-4}$	$3.33 * 10^{-2}$	.022	$p = .983$
Average Subject Translation	-1.84	21.4	-.086	$p = .931$
Average Subject Rotation	$-2.86 * 10^{-2}$	.209	-.137	$p = .891$
Acquisition Site	$-1.60 * 10^{-3}$	$3.19 * 10^{-2}$	-.050	$p = .960$
PCL-5 Total	$-4.15 * 10^{-6}$	$6.11 * 10^{-3}$	-.001	$p = .999$
PHQ-9 Total	$-2.55 * 10^{-4}$	$1.88 * 10^{-2}$	-.014	$p = .989$
AUDIT-C Total	$-1.67 * 10^{-2}$	.146	-.115	$p = .909$
DAST-10 Total	$-1.08 * 10^{-2}$	$7.15 * 10^{-2}$	-.151	$p = .880$
AIC=1158.8				

*Note that values displayed in this table are the result of a generalized linear model, using a binomial distribution, with average fractional anisotropy in the cingulum bundle as an outcome variable.

Table 2.17: Fractional Anisotropy in the Cingulum Bundle

Variable	Coefficient	Standard Error	z-Value	p-Value
Intercept	.620	.965	.643	$p = .520$
CVLT-2 Long Delay Cued Recall	$-3.66 * 10^{-3}$	.052	-.070	$p = .945$
CVLT-2 Short Delay Cued Recall	$1.37 * 10^{-4}$	.055	.002	$p = .998$
D-KEFS Letter Correct	$8.91 * 10^{-4}$	$6.28 * 10^{-3}$	.142	$p = .887$
TOPF Demographic Adjusted IQ	$-1.69 * 10^{-3}$	$9.02 * 10^{-3}$	-.188	$p = .851$
Age	$-2.71 * 10^{-3}$	$7.34 * 10^{-3}$	-.370	$p = .712$
Num Blast Exposures Resulting in TBI	$-1.83 * 10^{-3}$	$7.78 * 10^{-2}$	-.024	$p = .981$
Num Blast Exposures Not Resulting in TBI	$2.28 * 10^{-3}$	$3.26 * 10^{-2}$	.070	$p = .944$
Average Subject Translation	-2.72	20.4	-.133	$p = .894$
Average Subject Rotation	$-1.27 * 10^{-2}$	.203	-.062	$p = .950$
Acquisition Site	$-1.86 * 10^{-3}$	$3.14 * 10^{-2}$	-.059	$p = .953$
PCL-5 Total	$-8.30 * 10^{-5}$	$5.93 * 10^{-3}$	-.014	$p = .989$
PHQ-9 Total	$-1.56 * 10^{-3}$	$1.83 * 10^{-2}$	-.086	$p = .932$
AUDIT-C Total	$-1.08 * 10^{-3}$	.142	-.008	$p = .994$
DAST-10 Total	$8.23 * 10^{-3}$	$6.80 * 10^{-2}$	.121	$p = .904$
AIC=1092.5				

*Note that values displayed in this table are the result of a generalized linear model, using a binomial distribution, with average fractional anisotropy in the uncinate fasciculus as an outcome variable.

Table 2.18: Fractional Anisotropy in the Genu of the Corpus Callosum

Variable	Coefficient	Standard Error	z-Value	p-Value
Intercept	1.19	1.00	1.18	$p = .237$
CVLT-2 Long Delay Cued	-.008	.054	-.142	$p = .887$
CVLT-2 Short Delay Cued Recall	.005	.060	.087	$p = .930$
D-KEFS Letter Correct	.0002	.007	.040	$p = .970$
TOPF Demographic Adjusted IQ	-.002	.009	-.207	$p = .836$
Age	-.006	.007	-.826	$p = .409$
Num Blast Exposures Resulting in TBI	-.0007	.-81	-.009	$p = .993$
Num Blast Exposures Not Resulting in TBI	-.003	.034	-.075	$p = .940$
Average Subject Translation	-.400	21.22	-.020	$p = .985$
Average Subject Rotation	-.054	.211	-.257	$p = .797$
Acquisition Site	-.003	.033	-.084	$p = .933$
PCL-5 Total	-.0005	.006	-.078	$p = .938$
PHQ-9 Total	-.0002	.019	-.008	$p = .994$
AUDIT-C Total	-.004	.148	-.025	$p = .980$
DAST-10 Total	.003	.071	.043	$p = .965$
AIC=827.36				

*Note that values displayed in this table are the result of a generalized linear model, using a binomial distribution, with average fractional anisotropy in the genu of the corpus callosum as an outcome variable.

Table 2.19: Mean Diffusivity in the Uncinate Fasciculus

Variable	Coefficient	Standard Error	z-Value	p-Value
Intercept	-7.12	16.98	-.419	$p = .675$
CVLT-2 Long Delay Cued Recall	$2.52 * 10^{-3}$	.929	.003	$p = .998$
CVLT-2 Short Delay Cued Recall	$-1.09 * 10^{-3}$	.975	-.001	$p = .999$
D-KEFS Letter Correct	$-3.73 * 10^{-5}$	.110	.000	$p = 1.00$
TOPF Demographic Adjusted IQ	$-1.50 * 10^{-5}$	.158	.000	$p = 1.00$
Age	$1.18 * 10^{-3}$	.129	.009	$p = .993$
Num Blast Exposures Resulting in TBI	$8.11 * 10^{-3}$	1.36	.006	$p = .995$
Num Blast Exposures Not Resulting in TBI	$1.77 * 10^{-4}$	.574	.000	$p = 1.00$
Average Subject Translation	2.95	359.3	.001	$p = .999$
Average Subject Rotation	$-1.81 * 10^{-2}$	3.59	-.005	$p = .996$
Acquisition Site	$-2.73 * 10^{-4}$	.567	.000	$p = 1.00$
PCL-5 Total	$-2.20 * 10^{-5}$	.104	.000	$p = 1.00$
PHQ-9 Total	$-1.08 * 10^{-3}$	.322	.003	$p = .997$
AUDIT-C Total	$5.66 * 10^{-3}$	2.51	.002	$p = .998$
DAST-10 Total	$-3.64 * 10^{-6}$	1.21	.000	$p = 1.00$
AIC=31.49				

*Note that values displayed in this table are the result of a generalized linear model, using a binomial distribution, with mean diffusivity in the uncinate fasciculus as an outcome variable.

Table 2.20: Mean Diffusivity in the Cingulum Bundle

Variable	Coefficient	Standard Error	z-Value	p-Value
Intercept	-7.32	18.01	-.406	$p = .685$
CVLT-2 Long Delay Cued Recall	$1.39 * 10^{-3}$	.984	.001	$p = .999$
CVLT-2 Short Delay Cued Recall	$4.47 * 10^{-4}$	1.04	.000	$p = 1.00$
D-KEFS Letter Correct	$1.35 * 10^{-4}$	.117	.001	$p = .999$
TOPF Demographic Adjusted IQ	$6.43 * 10^{-4}$	.168	.004	$p = .997$
Age	$8.77 * 10^{-4}$	$1.37 * 10^{-3}$	.006	$p = .995$
Num Blast Exposures Resulting in TBI	$6.28 * 10^{-3}$	1.44	.004	$p = .997$
Num Blast Exposures Not Resulting in TBI	$2.85 * 10^{-3}$	.609	.005	$p = .996$
Average Subject Translation	.971	377.1	.003	$p = .998$
Average Subject Rotation	$6.88 * 10^{-3}$	3.78	.002	$p = .999$
Acquisition Site	$-3.14 * 10^{-2}$	.604	-.055	$p = .956$
PCL-5 Total	$6.64 * 10^{-5}$	.110	.001	$p = 1.00$
PHQ-9 Total	$9.42 * 10^{-4}$	.341	.003	$p = .998$
AUDIT-C Total	$2.87 * 10^{-3}$	2.66	.001	$p = .999$
DAST-10 Total	$-2.51 * 10^{-3}$	1.30	-.002	$p = .998$
AIC=31.32				

*Note that values displayed in this table are the result of a generalized linear model, using a binomial distribution, with mean diffusivity in the cingulum bundle as an outcome variable.

Table 2.21: Mean Diffusivity in the Genu of the Corpus Callosum

Variable	Coefficient	Standard Error	z-Value	p-Value
Intercept	-7.38	17.78	-.415	$p = .678$
CVLT-2 Long Delay Cued Recall	$1.77 * 10^{-3}$	.969	.002	$p = .999$
CVLT-2 Short Delay Cued Recall	$2.08 * 10^{-3}$	1.02	-.002	$p = .998$
D-KEFS Letter Correct	$1.56 * 10^{-4}$	.116	.001	$p = .999$
TOPF Demographic Adjusted IQ	$1.65 * 10^{-3}$	.166	.010	$p = .992$
Age	$2.68 * 10^{-3}$	.135	.020	$p = .984$
Num Blast Exposures Resulting in TBI	$2.94 * 10^{-3}$	1.43	.002	$p = .998$
Num Blast Exposures Not Resulting in TBI	$4.06 * 10^{-5}$	.603	.000	$p = 1.00$
Average Subject Translation	1.75	370.00	.005	$p = .996$
Average Subject Rotation	$8.41 * 10^{-3}$	3.74	-.002	$p = .998$
Acquisition Site	$-4.60 * 10^{-2}$	.603	-.076	$p = .939$
PCL-5 Total	$-1.05 * 10^{-4}$	.109	-.001	$p = .999$
PHQ-9 Total	$1.04 * 10^{-3}$	.338	.003	$p = .998$
AUDIT-C Total	$-3.74 * 10^{-3}$	2.64	-.001	$p = .999$
DAST-10 Total	$-1.89 * 10^{-3}$	1.30	-.001	$p = .999$
AIC=31.35				

*Note that values displayed in this table are the result of a generalized linear model, using a binomial distribution, with mean diffusivity in the genu of the corpus callosum as an outcome variable.

Table 2.22: Correlation Matrix between Model Variables

Variable	CVLT-2 Long Delay Cued Recall	CVLT-2 Short Delay Cued Recall	D-KEFS Letter Correct	TOPF IQ
CVLT-2 Long Delay Cued Recall	1.00	.917	.244	.131
CVLT-2 Short Delay Cued Recall	.917	1.00	.242	.134
D-KEFS Letter Correct	.244	.242	1.00	.152
TOPF Demographic Adjusted IQ	.131	.134	.152	1.00

CHAPTER 3. STUDY 2

3.1 Methods

The participants and measures used in Study 2 are the same as those used in Study 1. No MRI preprocessing steps are required for Study 2, as all structural and diffusion data used in Study 2 will be taken from that which was processed in Study 1. More variables were included in this analysis, which included many cognitive, psychological, injury and health measures. A full list of variables included in Study 2 can be found in Appendix B. Note that subscales were used in lieu of total scores where possible. Additionally, not all subscales were used, as some were missing large portions of data and were dropped from the model.

3.1.1 Data Dimensionality Reduction

One of the drawbacks to QCA is that the method is most effective on small numbers of variables (that contain multiple levels). This is because as variable count increases, exponentially more subjects are required to fill each cell. To circumvent this issue, it was decided that data dimensionality reduction would be used. Dimensionality reduction is the process of selecting or extracting important features from data such that only a small number of features are used relative to the full set of features from the dataset. Data dimensionality reduction can be accomplished through a number of methods, the most common being principal component analysis (PCA). PCA is a reductive method that summarizes a data table in an attempt to extract the most important information and compress it into pieces called principal components (Abdi and Williams, 2010). Principal components are linear combinations of the original variables in the dataset. Principal components are additive– the first component explains the largest variance, with more components adding to the explanatory power of the model. Principal components are orthogonalized, or differentiated, from each other, rendering them functionally independent from

Table 3.1: Principal Component Analysis

Metric	PC 1	PC 2	PC 3	PC 4
Standard Deviation	3.6711	2.4837	1.84449	1.65097
Proportion of Variance Explained	0.2174	0.0995	0.05487	0.04396
Cumulative Proportion of Variance Explained	0.2174	0.3169	0.3174	0.41570

other components. Components are extracted from a covariance matrix, which plots the common variance between variables in the dataset being analyzed. Magnitude and direction of variability are extracted from eigenvectors and eigenvalues. The PCA algorithm uses these metrics to identify important characteristics to keep based on direction and magnitude of variance. Although PCA is one of the most common reductive techniques, it does not prove useful in QCA. Although PCA reduces the dimensionality of a dataset, it also suppresses the individual items that were transformed into the component. To illustrate this, PCA was performed on the full imputed dataset. Table 3.1 displays the results from this analysis. The PCA results indicate that one component explains 21% of the variance in the dataset, with cumulative proportion of variance reaching 41% by the fourth principal component. The full PCA indicated that 100% of the variance in the dataset could be explained by 65 components. After the fourth component, explanatory power increases slowed to 1-2% per component. In the current investigation, QCA is being used to investigate how specific domains (including clinical and health variables) might be related to neuroimaging outcomes. Using PCA would render this investigation unable to answer the questions at hand. Once variables are transformed into components, there is no way to determine how much a specific domain is influencing the variance in a specific component. Therefore, EFA was chosen, as it allows for the construction of factors that would be theoretically domain related, and allow one to make inferences about the relationships between specific study domains and neuroimaging outcomes.

Statistical Analysis

Using the previously extracted diffusion measures of FA and MD, a multi-value QCA was performed. The mvQCA was conducted using R's (v 3.5.1) *QCApro* package (Thiem, 2018). The scripts used to conduct this analysis can be found on the same OSF repository as the previous study (https://osf.io/s89du/). In order to specify a mvQCA, an outcome measure must be

identified. In this case, each diffusivity measure (FA and MD) was used for each previously identified frontotemporal white matter tract as an outcome measure.

Before the QCA, an exploratory factor analysis (EFA) was conducted using the *fa* function in R. The imputed dataset described in Experiment 1 was used in the factor analysis to avoid problems with missing observations. The function was specified to use maximum likelihood factoring and an oblimin rotation. Maximum Likelihood (ML) factoring estimates model parameters especially well for non-normal data, which is the primary reason it was chosen for this analysis (Myung, 2003). Factoring methods, including ML, attempt to determine the underlying probability distribution of the researcher-specified factor analysis model. ML finds the distribution that makes the observed data most likely, and thus searches for parameters that maximize the likelihood of the specified model being true (Myung, 2003). While most rotational methods involve orthogonalizing factors in order to better differentiate components that may be highly correlated with each other, oblimin rotation is an oblique method that allows factors to be correlated with one another, but searches for the simplest factor structure (Rennie, 1997). In the EFA model constructed, many factors would be expected to be correlated with one another, so it was decided that Oblimin rotation would be the most efficacious method to use.

Qualitative Comparative Analysis

QCAPro requires a number of preprocessing steps to be done before conducting a mvQCA. First, variables to be included in the analysis must be placed in a concatenated list and assigned to a placeholder variable. For this analysis, extracted factors were assigned from the previous EFA to one variable, and the diffusivity outcomes to other variables. Next, the researcher must specify consistency and case number thresholds. Consistency thresholds specify the proportion contradictory cases that the researcher will consider acceptable. Contradictory cases are cases that have neither the specified outcome (diffusivity metrics) or any configuration of study variables (the aforementioned EFA factors). For this analysis, a threshold of 0.75 was set. This means that the QCA algorithm will disregard any configurations that do not represent at least 75% of the subjects in the dataset having the specified outcome (diffusivity metrics). Case number thresholding informs the algorithm of the minimum number of cases that must be observed with a condition for that condition to be added to a truth table. A truth table is a binary table, based on

Boolean algebra, which determines whether an expression is true or false. In the present analysis, a truth table would be constructed for extracted configurations meeting both the previously set consistency threshold and case number threshold. For each configuration in the truth table, a TRUE (1) or FALSE (0) value is assigned. This process allows the QCA algorithm to pare down the number of configurations to only those that represent the best portion of subjects. For instance, a case with 5 exogenous factors and 2 levels each would have 32 possible study conditions. However, if 23 of these conditions are not observed with at least 4 cases, they are dropped from the analysis and only 9 of the 32 conditions would be represented in the truth table. This is a measure of diversity, or the number of possible conditions included in the truth table and subsequent final analysis. The previous example would have a sample diversity of 28% (9 conditions present / 32 total conditions = 28% diversity). Once the truth table is constructed, the table is input into the *eQMC* (Thiem et al., 2016, minimization with enhanced Quine-McCluskey algorithm) function in *QCAPro*. *eQMC* uses an alternate, optimized version of the QMC algorithm to identify prime implicants represented in the truth table. This process is referred to as minimization, as it reduces the number of conditions to the minimum possible. A prime implicant is a condition that cannot be left out of any configuration on the previously constructed truth table. For a summary of this process, see (Dusa and Thiem, 2018). Data visualizations were generated using the corners package (Bailey, 2019), which creates plots to visualize the coverage and consistency of the parsimonious solution produced by eQMC package. A four corners plot shows the relative distribution of cases into consistent (configuration leads to outcome), unexplained (not Configuration, but outcome), contradictory (Configuration, but not the outcome), and irrelevant (not Configuration and not outcome). The factors extracted from the EFA were then calibrated for use in QCA using the *calibrate* function from *QCApro* (Thiem, 2018). This function can calibrate data into crisp (0 or 1 dichotomous categories), fuzzy (values bounded on a discrete spectrum between 0 and 1), or thresholded values. For this study, this function was used to threshold the model variables (in this case, the extracted factors from the EFA). This function requires researcher specification to set thresholds for partial inclusion and full inclusion. Partial inclusion was set as the median, and full inclusion was set as the upper quartile. Quartiles were determined using the *quantile* function in R. Any values not meeting the partial inclusion threshold are automatically excluded. The thresholds used for each can be found in the code in Appendix A.

Once data were calibrated, the *QCApro* package was used to construct a truth table, using the *truthTable* function. Truth tables were constructed for each diffusivity outcome (FA and MD) in each of the predetermined white matter tracts (the uncinate fasciculus, the genu of the corpus callosum, and the cingulum bundle). Diffusivity metrics also were calibrated to three categories, using the *calibrate* function. These thresholds were also determined using the lower, median, and upper quartiles. Once truth tables were constructed, the number of possible patterns was calculated by multiplying all the levels of all the variables included in the analysis. This resulted in 1,048,576 possible patterns in the dataset. Patterns present in the dataset were determined by counting the number of patterns in the truth table, using the *nrow* function in R. The minimum number of cases needed to extract a configuration to an outcome was set at 5, and the consistency threshold (the proportion of contradictory cases) was set at 0.75. This was to ensure that no more than 15% of the cases reported were contradictory, an acceptable threshold set by past work (Schneider and Wagemann, 2010).

3.2 Results

Before conducting the final EFA, a scree plot was generated to determine the number of factors to extract for the final model (see Figure 3.1). A scree plot is a visualization of eigenvalues plotted along an axis to visualize the eigenvalues of component factors as the total number of factors increases. A commonly used heuristic is to determine the location of the ”elbow” of the scree plot, or where the eigenvalues for component factors drops off sharply. The number of factors located before the elbow of the scree plot is thought to be an appropriate number of factors to use in exploratory factor analysis (D’agostino Sr and Russell, 2005).

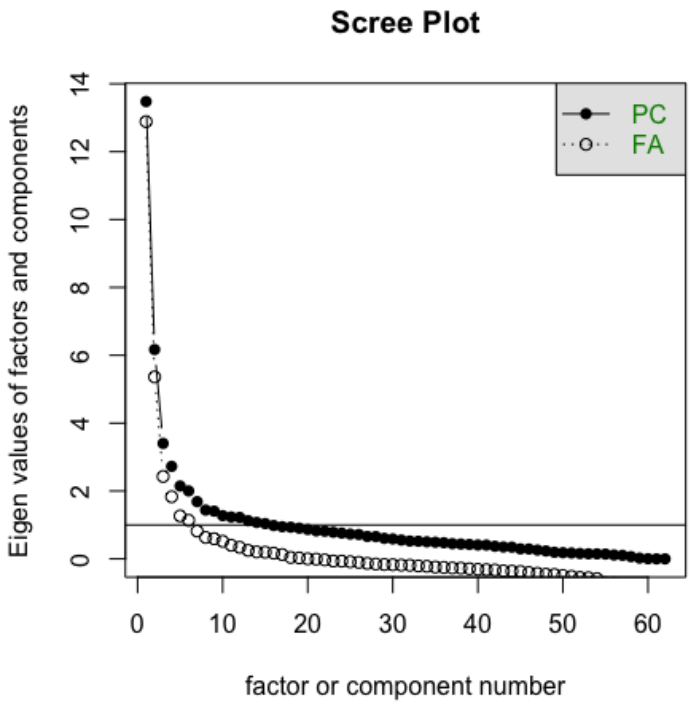
After determining the number of factors that would be appropriate to extract, a Kaiser-Meyer-Olkin (KMO) test of factor adequacy was performed. A KMO test determines the amount of common variance between items in a dataset that are intended for use in an EFA. The KMO test indicated sufficient sampling adequacy (overall MSA=0.88). A Bartlett’s test of sphericity was also conducted, which tests datasets to determine whether the underlying correlation matrix is an identity matrix, or a matrix in which there is little correlation between items. Results from this test indicated suitable correlations between items

($\chi^2 = 687609, p < .0001$). The final factor analysis consisted of all the study variables outlined in Study 1, with diffusivity variables and character variables removed from the analysis.

3.2.1 Exploratory Factor Analysis

An exploratory factor analysis was conducted using the *fa* function in the *psych* library in R (Revelle, 2019). A full spreadsheet containing all factor loadings can be found on this project's OSF repository.

Figure 3.1: Scree Plot for Factor Extraction



This figure depicts a scree plot of eigenvalues for component factors. This is used as a heuristic for determining the number of factors to extract for an exploratory factor analysis. Based on this scree plot, it was determined that 4 factors would be an appropriate number for subsequent analysis. In the figure legend, FA represents the plot of component factors for a factor analysis, and PC represents the plot of component factors for a principal components analysis.

The extracted factor solution consisted of 4 factors, containing 893 observations across 62 items. In order to filter out low item loadings, a minimum threshold was set at .30. This threshold allows

for some items to load onto multiple factors, but also eliminates weak factor loadings. Indeed, in the remaining factor structure, a number of items loaded onto multiple factors. In order to simplify the factor solution as much as possible, when multiple factor loadings occurred, the higher loading was chosen and the lower loading was disregarded. This factor solution appears to be a good fit for the underlying data structure, as indicated by fit statistics (RMSEA=.093, Tucker Lewis Index=.065). The final model explained 38% of the variance in the dataset. Factors were then constructed using the results of the EFA. I accomplished this by averaging across all the items that loaded onto a factor. For an example, a factor with three items would be created using the following code:

```
data.frame$Factor = (Item1 + Item2 + Item3)/3
```

The item loadings seemed to ordered by domain. For accessibility, these factors were assigned the names General Functioning, the CVLT-II, the NIH Toolbox, and Cognition. The code used to generate factors can be found in Appendix A. There were a few items that did not meet the threshold of 0.30 and did not load onto any factor. Among these items were subscales from the Behavioral Risk Factor Surveillance System, the Mini International Neuropsychiatric Interview, the Drug Abuse Screening Test, the Unified Parkinsons Disease Rating Scale, the digit span and visual puzzle subscales from the WAIS-IV, and the STOPBANG questionnaire for obstructive sleep apnea. These items were excluded from the final QCA analysis. Additonally, some factors contained items from multiple scales. In these cases, all items were transformed to z-scores from raw scores using the *scale* function in R. Reference values for this transformation were automatically derived from the sample by the function.

3.2.2 Qualitative Comparative Analysis

The results from each QCA on each diffusivity outcome can be found in Tables 3.1-3.6. Diffusivity outcomes were binarized (i.e., set to crisp configuration), with 0 representing subjects falling below the median of the outcome and 1 representing subjects falling above the median of the outcome.

There was good diversity in the number of configurations extracted by QCA. The number of configurations extracted varied between three and ten, depending on the tract set as the outcome

Table 3.2: QCA Analysis: Uncinate Fasciculus FA

Configuration	Inclusion	Coverage: Unique
NIH Toolbox[Medium] + General Function[High] + Cognition[Medium]	0.500	0.011
CVLT-II[Low] + NIH Toolbox[Medium] + General Function[Medium] + Cognition[High]	0.800	0.008
CVLT-II[Low] + NIH Toolbox[High] + General Function[Medium] + Cognition[Medium]	0.800	0.008
CVLT-II[Low] + NIH Toolbox[High] + General Function[Medium] + Cognition[Medium]	0.750	0.013
CVLT-II[Low] + NIH Toolbox[High] + General Function[High] + Cognition[Medium]	0.750	0.013
CVLT-II[Medium] + NIH Toolbox[High] + General Function[High] + Cognition[High]	0.750	0.006
CVLT-II[High] + NIH Toolbox[Medium] + General Function[High] + Cognition[Low]	0.750	0.013

*Note that the results of the table should be read as following: Each row represents a unique configuration extracted by QCA. FA=Fractional Anisotropy, MD= Mean Diffusivity. "High" refers to the upper quartile being met for inclusion for the subjects included in this pattern, while "Medium" refers to each subject in the pattern meeting the Median criteria. Outcome = "Low" refers to subjects outcomes meeting the lower quartile for the variable represented. Inclusion is the proportion of the configuration included in other extracted configurations. Coverage: Unique is a proportion representing the number of subjects who appear only on this configuration and this outcome, without ever being represented by any other configurations.

measure. In each analysis, inclusion and unique coverage values were extracted. Presently, unique coverage is of more theoretical interest, as this represents subjects who appear on only one configuration. A high unique coverage of one configuration would indicate this configuration is particularly important for subjects sharing the same diffusivity outcome. Unfortunately, unique coverage across all configurations in each analysis rarely represented above 5% of the sample. This makes it difficult to claim with any certainty that any of these configurations might be representative of military samples as a whole. However, the sheer number of differences in configurations between analyses should be noted. Depending on tract and diffusivity metric (FA

Table 3.3: QCA Analysis: Cingulum Bundle FA

Configuration	Inclusion	Coverage: Unique
NIH Toolbox[Medium] + General Function[High] + Cognition[Medium]	0.600	0.013
CVLT-II[Low] + NIH Toolbox[High] + General Function[Medium] + Cognition[Medium]	1.000	0.010
CVLT-II[Medium] + NIH Toolbox[Medium] + General Function[Low] + Cognition[Low]	0.750	0.019
CVLT-II[Medium] + NIH Toolbox[Medium] + General Function[Low] + Cognition[Medium]	0.800	0.017
CVLT-II[Medium] + NIH Toolbox[High] + General Function[Medium] + Cognition[Low]	0.750	0.006
CVLT-II[Medium] + NIH Toolbox[High] + General Function[High] + Cognition[High]	0.750	0.006
CVLT-II[High] + NIH Toolbox[Low] + General Function[High] + Cognition[Low]	0.857	0.013
CVLT-II[High] + NIH Toolbox[Medium] + General Function[Medium] + Cognition[Medium]	0.750	0.013
CVLT-II[High] + NIH Toolbox[High] + General Function[Low] + Cognition[High]	0.773	0.035
CVLT-II[High] + NIH Toolbox[High] + General Function[High] + Cognition[Medium]	0.750	0.035

Table 3.4: QCA Analysis: Genu of the Corpus Callosum FA

Configuration	Inclusion	Coverage: Unique
CVLT-II[High] + NIH Toolbox[Medium] + Cognition[Medium]	0.724	0.044
CVLT-II[Low] + NIH Toolbox[Medium] + General Function[Medium] + Cognition[High]	0.800	0.008
CVLT-II[High] + NIH Toolbox[Low] + General Function[Medium] + Cognition[High]	0.750	0.006

or MD), the number and configuration of configurations extracted by QCA varied widely. This speaks to the inherent heterogeneity of the sample– although there is no configuration covering a majority of subjects, it is interesting that there was so much variability between analyses. It is also notable that these configurations covered both outcomes (of falling below the median or above). Theoretically, one would expect to find differences between subjects with higher diffusivity and lower diffusivity. Indeed, this was a difference that was originally hypothesized. However, it appears subjects share the same configurations, regardless of outcome.

When considering the differences between FA and MD, it appears there are no significant fluctuations in configurations extracted. While traditional linear analyses often find differences in FA and MD by tract, there is currently no evidence that this is the case when using QCA. Configurations were not exactly the same between the two, but the inclusion and unique coverage metrics were very similar for each outcome and indicate no differences that should be noted. One explanation for this result is that the structural changes that altered diffusivity metrics represent occur on a microscopic level, making it unlikely that either metric would represent a unique behavioral alteration. It should also be noted that the inclusion metric (a measure of redundant configurations appearing in more than one analysis) for each configuration was high, ranging from 0.60 to 1.00– this is likely due to the reduction in items resulting from the exploratory factor analysis. A more effective method to use in the future might be including items with strong theoretical links to diffusivity metrics in a QCA analysis, with remaining items being factored together using exploratory or confirmatory factor analysis.

Table 3.5: QCA Analysis: Uncinate Fasciculus MD

Configuration	Inclusion	Coverage: Unique
CVLT-II[High] + General Function[Medium] + Cognition[High]	0.769	0.021
NIH Toolbox[Low] + General Function[Medium] + Cognition[Medium]	0.818	0.038
CVLT-II[Low] + NIH Toolbox[Medium] + General Function[Low] + Cognition[Low]	0.786	0.023
CVLT-II[Medium] + NIH Toolbox[High] + General Function[High] + Cognition[Medium]	0.750	0.006
CVLT-II[High] + NIH Toolbox[Medium] + General Function[High] + Cognition[High]	0.800	0.008

Table 3.6: QCA Analysis: Cingulum Bundle MD

Configuration	Inclusion	Coverage: Unique
CVLT-II[Medium] + NIH Toolbox[Low] + General Function[Medium] + Cognition[Medium]	0.833	0.011
CVLT-II[Medium] + NIH Toolbox[Medium] + General Function[Medium] + Cognition[High]	0.750	0.006
CVLT-II[Medium] + NIH Toolbox[High] + General Function[High] + Cognition[Medium]	0.750	0.006
CVLT-II[High] + NIH Toolbox[Low] + General Function[Medium] + Cognition[High]	0.750	0.006
CVLT-II[High] + NIH Toolbox[Low] + General Function[High] + Cognition[High]	0.900	0.019
CVLT-II[High] + NIH Toolbox[Low] + General Function[High] + Cognition[Medium]	0.750	0.006

Table 3.7: QCA Analysis: Genu of the Corpus Callosum MD

Configuration	Inclusion	Coverage: Unique
CVLT-II[Low] + NIH Toolbox[Low] + General Function[High] + Cognition[Low]	0.788	0.087
CVLT-II[Medium] + NIH Toolbox[Low] + General Function[High] + Cognition[Medium]	0.800	0.008
CVLT-II[Medium] + NIH Toolbox[Medium] + General Function[Medium] + Cognition[Medium]	0.875	0.015
CVLT-II[Medium] + NIH Toolbox[Medium] + General Function[Medium] + Cognition[High]	0.750	0.006
CVLT-II[Medium] + NIH Toolbox[Medium] + General Function[High] + Cognition[Low]	0.889	0.017
CVLT-II[High] + NIH Toolbox[Low] + General Function[Low] + Cognition[High]	0.875	0.015
CVLT-II[High] + NIH Toolbox[Low] + General Function[High] + Cognition[High]	0.800	0.017
CVLT-II[High] + NIH Toolbox[Low] + General Function[Low] + Cognition[Medium]	0.750	0.006

CHAPTER 4. DISCUSSION

4.1 General Discussion

The current investigation sought to accomplish two tasks. First, this study sought to determine whether blast-related injury mechanisms in TBI would have any unique, detectable effects on white matter structure in a sample of military Veterans. Second, this study used Qualitative Comparative Analysis in a neuroimaging and clinical dataset to determine whether QCA would offer any unique insights into the factors that play a role in TBI and functioning post-injury. In Study 1, it was hypothesized that there would be evidence of microstructural damage in the cingulum, genu of the corpus callosum, and the uncinate fasciculus in subjects with a history of TBI when compared to subjects without history of TBI. Additionally, it was expected that there would be a differential effect of injury mechanism when including blast-related injuries in analyses. Neither of these hypotheses were supported. In Study 2, it was hypothesized that evidence would emerge, through use of QCA, supporting the implication of blast related injury mechanisms in configurations of variables that had an outcome of tract diffusivity. These hypotheses were also unsupported.

4.2 Study 1

It was surprising to find none of the variables in the generalized linear models had any predictive power on diffusivity metrics. There is strong reason to expect each one of these variables is linked to each tract specified here. Performance on cognitive measures like the CVLT-II and D-KEFS has shown to decrease post-injury (Kennedy et al., 2009). Likewise, premorbid functioning has been linked as a strong predictor of injury characteristics and outcomes (Novack et al., 2001; Green et al., 2008). The neuropsychological measures included in these models were chosen due to their high internal reliability and predictive validity (DeJong and Donders, 2009; Shunk et al.,

2006). However, some evidence suggests that scores on the CVLT can differ significantly between patients depending on severity of brain injury (Jacobs and Donders, 2007). This is unlikely to be the case in Study 1, as all subjects with TBI were diagnosed with mild injuries. It is possible that differing injury mechanisms may have played a role in these results, however. Overall, results from Study 1 were unexpected, given past work finding blast exposure is linked to global disruptions in white matter structure (Davenport et al., 2012).

When put into context with previous literature, the results of Study 1 align with previous null results. A recent systematic review found increased FA in military patients with mild TBI in the acute phase of injury (Asken et al., 2018). One study failed to find differences in diffusivity between military veterans with and without a history of mild TBI, instead finding ”potholes” (pockets of exceptionally low FA) in WM tracts of veterans with mild TBI (Jorge et al., 2012). One reason for these differences might be mild TBI-psychological disorder comorbidity. For instance, one study showed differentially altered FA in subjects with comorbid mild TBI and PTSD diagnosis when compared to mild TBI subjects alone (Lepage et al., 2018). Another study found significantly lower FA in a PTSD group when compared to mild TBI alone and mild TBI + PTSD (Bolzenius et al., 2018). Although these results were not replicated in Study 1 (as screening measures for depression, PTSD, substance use disorder, and alcohol use disorder were included in models as covariates), it is possible covariates were not sensitive enough to account for comorbidities. At best, the screening measures used in the models are only intended to screen for possible comorbidities. Individuals can score high on a measure like the PCL-5 without being diagnosed with PTSD. Thus, without an in-depth, clinician guided assessment of psychological comorbidities, it is possible that the screening measures included may have not been completely sufficient in controlling for comorbid disorders.

Another possible explanation for the null findings of Study 1 relate to the statistical analyses. The most surprising finding was the lack of significant coefficients for blast injury variables. The initial assumption underlying the inclusion of blast injury information is that blast injury is additive– that is, as the number of blast exposures increased, one would expect to find more severe WM disruptions. This assumption was tested using generalized linear models. It is possible, however, that this assumption was flawed. One might logically expect to find an additive effect of blast exposure, but there may be a different underlying mechanism. For instance, there

may not be direct additive effects of blast injury, but in biological mechanisms that might not have strong effects on white matter integrity. If this is the case, it would help to explain the lack of significant findings in Study 1. Additionally, subjects in the sample who reported history of injury were all classified as mild TBI. Although a portion of patients with mild TBI exhibit long lasting deficits post injury (often referred to as postconcussive symptoms), the vast majority of those with mild TBI make a full recovery from deficits (Losoi et al., 2016). It is possible that the selected measures of cognitive functioning did not have adequate sensitivity to detect changes in this cohort. It is also possible that time since injury played a role in these findings. The majority of subjects in this sample were scanned and tested far post-injury (on average, 10 years), rather than soon after injury. Given the good recovery trajectories for most patients with mild TBI, it is possible that the longer time interval between injury and testing played a role in the null findings of Study 1. The long time interval between injury and testing also allows for confounding factors to occur. Aging or development of health problems could have occurred during this gap and may explain the findings reported here.

It should be noted that the highest coefficients found in statistical models were covariates for subject-induced volume-by-volume translation and rotation in the MRI scanner. Subject motion has long plagued the field of neuroimaging, as this can cause false positives in voxel-based fMRI and misalignment of diffusion weighted images in DTI sequences (Lemieux et al., 2007; Tijssen et al., 2009). Although the analysis pipeline of Study 1 included corrections for subject-induced and eddy-current induced volume translation and rotation, it appears that the data is still significantly affected by these factors. This is especially concerning, given the tendency for false positives and false negatives in neuroimaging data with large motion artifacts (Lemieux et al., 2007). More aggressive measures may need to be taken to deal with subject data including these artifacts, such as dropping those datapoints from statistical analyses altogether.

Neuroimaging data collected across sites and on different scanners has shown significant variability, especially on DTI diffusivity metrics (Fortin et al., 2017). Data consortia, including CENC, have made efforts to harmonize data and reduce the effects of acquisition site on neuroimaging data. One popular method is called ComBat harmonization, a batch-effect correction tool originally developed for genomics. ComBat harmonization has shown good efficacy in both preserving biological variability and reducing the influence of acquistion site

(Fortin et al., 2017). Neuroimaging data in the present study were harmonized using this tool. Box plots were generated for neuroimaging metrics across acquisition sites and can be found in Appendix B. Starting in Figure B.1, one can see that FA values seem to be quite similar across acquisiton sites. MD values appear to be more variabile, with sites 7 and 8 having a lower average MD than other sites. This is unlikely to be a significant factor, however, given that the variation occurs at the millionth decimal point. Although ComBat harmonization corrects for variability in neuroimaging data, it has not been used to harmonize neuropsychological measures. It may be possible that there is site variability within the neuropsychological measures used in the present analysis. If this is the case, it might explain the null findings of all six generalized linear models conducted in Study 1. More investigation is needed to determine whether neuropsychological data is susceptible to acquisition site variability, and if so, whether harmonization tools like ComBat can be used to eliminate this problem.

4.3 Study 2

In Study 2, the ability of QCA was examined to determine whether it might be useful in neuroimaging analyses. Due to limitations of the method, large numbers of variables are unable to be used unless the overall *n* of the study increases with variable number. EFA was used to work around this issue and insert factors into the analysis, which consisted of all study variables previously specified. The EFA indicated a good factor structure and model fit. Using the suggested factor structure, 4 factors were extracted for use in QCA. These factors were then input as exogenous factors and used in six models, one for each diffusivity metric on each tract. QCA was able to extract between three and 10 configurations per outcome, varying by tract and diffusvity metric.

Study 2 originated as an investigation of whether equifinal statistical methods could be used successfully in a neuroimaging dataset. The results suggest that, indeed, equifinal analyses could be of some use to future studies conducting statistical analysis of neuroimaging datasets. However, there are a number of considerations that should be made for studies utilizing similar methods in the future. First, QCA might be more suited for use in normative populations. I hoped that the heterogeneity in military samples might be reduced somewhat through the use of QCA to identify commonalities between subjects sharing the same neuroimaging outcome. From the

results of Study 2, it appears that this sample may be too heterogeneous for QCA to be of any unique use. The configurations extracted only covered a fraction of the subjects in the full sample, and failed to find any commonalities shared between a significant portion of subjects. Using QCA to analyze a normative sample would avoid the issues posed by the diverse and heterogeneous military sample in Studies 1 and 2. Second, any study considering use of QCA should consider the size of the dataset and number of variables they wish to analyze. QCA was initially developed as a method of analyzing small samples– in most cases, it was used in case studies where the researcher had intimate case knowledge of the subject being analyzed (Schneider and Wagemann, 2010). Recent studies demonstrated that although case knowledge is helpful, it is not necessary for the successful implementation of QCA in analysis of large samples (Emmenegger et al., 2014; Cooper, 2005; Cooper and Glaesser, 2016). Cooper and Glaesser (2016) performed QCA on a dataset of 6,666 subjects, one of the largest studies to date to use this method. I sought to extend this work by using it in a large sample of neuroimaging data. To date, this is the only work I am aware of that has used QCA to analyze neuroimaging data. At the outset of this investigation, a primary goal was to use QCA to uncover configurations of study variables not yet carefully considered together in unison. For this reason, all of the CENC study variables were included in the final analysis. Unfortunately, this is also one of the primary issues in the results from Study 2. Although QCA successfully extracted configurations varying by outcome, these configurations were common (i.e., had higher inclusion values) across each analysis. It is likely that this occurred due to the EFA I was required to conduct. The EFA did allow me to circumvent the variable number issues presented by QCA, but effectively reduced any significant variation that might have been present due to items being combined into factor solutions. An orthogonal approach like PCA or orthogonal factor solutions in EFA indeed might help avoid this issue. However, using these sorts of methods pose the risk of losing interpretability. While PCA finds solutions that are orthogonal to one another, it also effectively makes it very difficult to draw meaningful conclusions about the influence of specific variables or factors. It might be more prudent to specifically design samples for QCA in the future, administering items to subjects that have strong theoretical reason to be linked to diffusivity outcome. These results might also be due to small variations between subjects on diffusivity metrics. These subjects are mild TBI patients, with injury occurring before, during, and after military service. Time between injury and testing allows

confounding factors to occur which may negate larger differences in integrity of WM structure between subjects. In some cases, diffusivity metrics differed only at the millionth decimal point. It is probable that QCA has difficulty differentiating between conditions on such a small scale. Even when the entire decimal value was input as a threshold (in some cases, as far down to the ten millionth decimal point), QCA extracted the same configurations as previously reported.

It is also noteworthy that QCA found the same common configurations across both outcomes (low and high diffusivity), and did not find any configurations unique to one outcome or the other. This goes against the primary hypotheses of Study 2, as it was expected that unique configurations would be extracted for subjects with either high or low diffusivity. One explanation for this is the sensitivity of QCA on DTI diffusivity metrics. While DTI has shown to be sensitive to minute microstructural changes, even in longitudinal samples, the differences being detected may be on a scale small enough that it causes problems for QCA. Although the method appeared to perform without issue, the differences between subjects speak to a sensitivity issue with QCA. Investigation with traditional statistical methods appears to be far more efficacious in the analysis of DTI data than QCA. It is likely this is due to QCA'S Boolean approach to analysis– the separation of subjects into distinct categories based on an outcome measure is not sensitive enough to detect the differences that other methods routinely find in analysis of similar datasets.

While the use of QCA proved to be limited in analysis of DTI data in this investigation, there are other applications in neuroimaging it may show some use for. For instance, a DTI analysis of a normative sample might be prudent, to determine whether this method is effective in DTI analyses as a whole, rather than in a specific subpopulation. Other fields, such as functional imaging, might also find QCA useful. Functional MRI studies often quantify the hemodynamic response in the brain through regression betas. These betas are much more likely work in a QCA model, given that they tend to be much larger than the values produced by diffusivity analyses. Likewise, studies using measures of total brain volume, cortical thickness, or region-based volumetric analysis would likely be more useful in a QCA model. These metrics are quantified on a much larger scale and would avoid the issues Study 2 faced with diffusivity metrics. Researchers using functional or structural imaging should consider QCA as a way to quantify their data and visualize relationships not yet analyzed.

4.4 Future Recommendations

There are certainly lessons to be taken away from the present study. First, it is likely that work remains to be done in the development of procedures to harmonize clinical, behavioral, and neuroimaging data collected at multiple sites, especially in large cohorts like CENC. Work is currently underway to limit the influence of multisite collection of neuroimaging data. Previous studies using machine learning have found that multisite neuroimaging data are easily predictable and distinguishable from data collected at other sites (Fortin et al., 2018), with datapoints from each site clustering nearly perfectly together across the distribution. Although efforts are being made to reduce the variability included by multisite data collection, no studies have currently introduced recommendations for the harmonization of clinical or behavioral data collected at multiple sites. There is almost certainly the possibility that clinical or behavioral data are being influenced by site collection, through differences in test administrators, digital data collection (such as software versions of digital testing materials) and scoring methods. Box plots of study measures across collection sites were generated after primary analyses and included in Appendix B. From these plots, it appears that there is significant variability in both mean scores and minimum/maximum scores across aquisition sites. It is likely that this variability between sites played some role in the findings of Studies 1 and 2. Future work should try to determine whether this variability exists and, if so, whether harmonization techniques can work to eliminate some of this site variability.

Second, future work should consider whether existing analysis pipelines can be improved to ensure subject motion is not having an undue effect on neuroimaging data. Although a robust pipeline was used to correct for eddy current-induced and subject-induced volume-by-volume translation and rotation, the largest predictor of diffusivity outcomes in the generalized linear models of Study 1 was average subject-induced translation. As recent work has shown (Button et al., 2013), neuroscience (and more specifically, neuroimaging) has faced considerable issues when attempting to replicate previous work. Multisite variations and uncorrected subject and eddy-current induced motion could possibly play a role in replication failures. Some techniques, such as completely immobilizing the tissue being imaged, can combat some of these issues, but may be unfeasible due to prohibitive cost or time requirements.

Third, future work should consider whether analysis techniques from similar fields might be suited to answer a question that traditional statistical methods cannot. QCA was used in this study in part because it allows for a different research question to be asked– rather than asking questions about differences between subjects, it allows one to ask questions about commonalities between subjects. Although this method proved limited in this particular dataset, there is an important conclusion to draw from this work. Study 2 has effectively shown that it is entirely possible to adapt methods traditionally used outside of neuroscience for use within the field. I believe that leveraging tools from other fields to answer questions within one's own field is essential in the progression of scientific knowledge and practices. Future work should seriously consider whether other fields might have analysis techniques to offer ongoing research. New analysis techniques might serve as a lens through which to view current work in different ways and encourage the development of new, novel research questions.

Finally, future work should consider the use of configurational approaches like QCA in neuroimaging modalities. Although the current project did not show QCA to be efficacious in diffusion imaging, it remains to be seen whether this would be the case in other modalities, like volumetric or functional-based approaches. This technique has unique potential to offer broad overviews of commonalities in large multisite datasets. When used in large datasets, QCA has the potential to highlight variables common across large numbers of subjects, which may indicate factors that should be more closely examined. QCA is a novel way of modeling relationships between study variables and outcome variables, and offers a unique perspective through which to view these relationships. In inherently heterogeneous populations (such as clinical or military samples), this technique would offer a way to minimize heterogeneity and view common relationships. It is also likely a better way to deal with individual differences than averaging across all subjects, as this causes the loss of individual differences.

4.5 Limitations

The findings of the present investigation should be viewed within the context of its limitations. The foremost limitation is that some allowances had to be made due to the structure of the archival dataset used in this analysis. For instance, although I hypothesized initially about differences between blast-related and non-blast related injuries, this was difficult to answer with

the information available in the CENC dataset. It was rare for subjects to have only a non-blast TBI, as there is considerable heterogeneity in military injury histories. This made it difficult to effectively filter subjects into either blast-related TBI or non-blast TBI. Furthermore, although I had information about the number of blast-related, combat acquired, and general exposure TBIs, there was no way to parse whether some of these TBIs appeared in multiple categories (for instance, blast exposure and non-combat TBIs). To investigate this question, I had to create a blast TBI variable by subtracting blast exposures not resulting in TBI from total blast exposures. Although this did not directly answer the questions posed initially, it is a roundabout way of determining whether blast injury has some differential effect on brain structure. Another limitation regarded missing data. Although I was missing less than 1% of total observations, every subject was missing at least one datapoint, requiring multiple imputation to estimate missing values. Likewise, I was required to drop a number of variables from the final analyses because they were missing over half of their observations. Finally, this study was limited by its access to more severe forms of TBI. Most subjects in the dataset who had a history of injury were diagnosed with mild TBI, whether acquired pre, post, or during combat tours. While mild injuries do often result in deficits, these may not have been severe or acute enough for the analyses used in this investigation to detect. In most cases, these injuries occurred years before participation. This may have played a role in the null findings of both Studies 1 and 2, as it allows for both recovery and confounding factors to occur. Future studies using similar techniques should endeavor to recruit both acute and chronic stages of injury.

REFERENCES

Abdi, H. and Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):433–459.

AFHC et al. (2012). Deaths by suicide while on active duty, active and reserve components, us armed forces, 1998-2011. *MSMR*, 19(6):7. PMID:22779434.

Alexander, A. L., Lee, J. E., Lazar, M., and Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics*, 4(3):316 – 329.

Alhilali, L. M., Delic, J. A., Gumus, S., and Fakhran, S. (2015). Evaluation of white matter injury patterns underlying neuropsychiatric symptoms after mild traumatic brain injury. *Radiology*, 277(3):793–800.

Anderson, C. V. and Bigler, E. D. (1995). Ventricular dilation, cortical atrophy, and neuropsychological outcome following traumatic brain injury. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 8(3):307–315.

Arenth, P. M., Russell, K. C., Scanlon, J. M., Kessler, L. J., and Ricker, J. H. (2014). Corpus callosum integrity and neuropsychological performance after traumatic brain injury: A diffusion tensor imaging study. *Journal of Head Trauma Rehabilitation*, 29(2):E1.

Asken, B. M., DeKosky, S. T., Clugston, J. R., Jaffee, M. S., and Bauer, R. M. (2018). Diffusion tensor imaging (dti) findings in adult civilian, military, and sport-related mild traumatic brain injury (mtbi): A systematic critical review. *Brain Imaging and Behavior*, 12(2):585–612.

Bailey, K. G. D. (2019). *Simulating and visualizing sets*. R Package Version 0.9.13.

Basser, P. J. (1995). Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine*, 8(7):333–344.

Basser, P. J. and Pierpaoli, C. (2011). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. *Journal of Magnetic Resonance*, 213(2):560–570.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Belanger, H. G., Kretzmer, T., Yoash-Gantz, R., Pickett, T., and Tupler, L. A. (2009). Cognitive sequelae of blast-related versus other mechanisms of brain trauma. *Journal of the International Neuropsychological Society*, 15(1):1–8.

Benedict, R. H., Schretlen, D., Groninger, L., Dobraski, M., and Shpritz, B. (1996). Revision of the brief visuospatial memory test: Studies of normal performance, reliability, and validity. *Psychological Assessment*, 8(2):145.

Berg, J.-L., Durant, J., Banks, S. J., and Miller, J. B. (2016). Estimates of premorbid ability in a neurodegenerative disease clinic population: Comparing the test of premorbid functioning and the wide range achievement test. *The Clinical Neuropsychologist*, 30(4):547–557.

Bergamino, M., Farmer, M., Yeh, H.-w., Paul, E., and Hamilton, J. P. (2017). Statistical differences in the white matter tracts in subjects with depression by using different skeletonized voxel-wise analysis approaches and dti fitting procedures. *Brain Research*, 1669:131–140.

Bigler, E. (2013a). Traumatic brain injury, neuroimaging, and neurodegeneration. *Frontiers in Human Neuroscience*, 7:395.

Bigler, E. D. (2001). The lesion (s) in traumatic brain injury: Implications for clinical neuropsychology. *Archives of Clinical Neuropsychology*, 16(2):95–131.

Bigler, E. D. (2013b). Neuroimaging biomarkers in mild traumatic brain injury (mtbi). *Neuropsychology Review*, 23(3):169–209.

Bigler, E. D., Abildskov, T. J., Petrie, J., Farrer, T. J., Dennis, M., Simic, N., Taylor, H. G., Rubin, K. H., Vannatta, K., Gerhardt, C. A., and Stancin, T. (2013). Heterogeneity of brain lesions in pediatric traumatic brain injury. *Neuropsychology*, 27(4):438.

Bigler, E. D. and Maxwell, W. L. (2012). Neuropathology of mild traumatic brain injury: Relationship to neuroimaging findings. *Brain Imaging and Behavior*, 6(2):108–136.

Bigler, E. D., McCauley, S. R., Wu, T. C., Yallampalli, R., Shah, S., MacLeod, M., Chu, Z., Hunter, J. V., Clifton, G. L., Levin, H. S., and Wilde, E. A. (2010). The temporal stem in traumatic brain injury: Preliminary findings. *Brain Imaging and Behavior*, 4(3-4):270–282.

Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., and Domino, J. L. (2015). The post-traumatic stress disorder checklist for dsm-5 (pcl-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*, 28(6):489–498.

Blosnich, J. R., Brenner, L. A., and Bossarte, R. M. (2016). Population mental health among us military veterans: Results of the veterans health module of the behavioral risk factor surveillance system, 2011–2012. *Annals of Epidemiology*, 26(8):592–596.

Bolzenius, J. D., Velez, C. S., Lewis, J. D., Bigler, E. D., Wade, B. S., Cooper, D. B., Kennedy, J. E., Reid, M. W., Ritter, J. L., York, G. E., et al. (2018). Diffusion imaging findings in us service members with mild traumatic brain injury and post-traumatic stress disorder. *The Journal of Head Trauma Rehabilitation*, 33(6):393–402.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: Understanding aic and bic in model selection. *Sociological Methods & Research*, 33(2):261–304.

Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., and Bradley, K. A. (1998). The audit alcohol consumption questions (audit-c): An effective brief screening test for problem drinking. *Archives of Internal Medicine*, 158(16):1789–1795.

Bush, N. E., Reger, M. A., Luxton, D. D., Skopp, N. A., Kinn, J., Smolenski, D., and Gahm, G. A. (2013). Suicides and suicide attempts in the us military, 2008–2010. *Suicide and Life-Threatening Behavior*, 43(3):262–273.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., and Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5):365–376.

Bysse, D., Reynolds III, C., Monk, T., et al. (1989). The pittsburgh sleep quality index (psqi): A new instrument for psychiatric research and practice. *Psychiatry Research*, 28(2):193–213.

Caeyenberghs, K., Leemans, A., Geurts, M., Linden, C. V., Smits-Engelsman, B. C., Sunaert, S., and Swinnen, S. P. (2011). Correlations between white matter integrity and motor function in traumatic brain injury patients. *Neurorehabilitation and Neural Repair*, 25(6):492–502.

Caplan, B., Bogner, J., Brenner, L., Carr, W., Polejaeva, E., Grome, A., Crandall, B., LaValle, C., Eonta, S. E., and Young, L. A. (2015). Relation of repeated low-level blast exposure with symptomology similar to concussion. *Journal of Head Trauma Rehabilitation*, 30(1):47–55.

Carp, J. (2012). On the plurality of (methodological) worlds: Estimating the analytic flexibility of fmri experiments. *Frontiers in Neuroscience*, 6:149.

Catroppa, C., Anderson, V. A., Morse, S. A., Haritou, F., and Rosenfeld, J. V. (2006). Children's attentional skills 5 years post-TBI. *Journal of Pediatric Psychology*, 32(3):354–369.

Chung, F., Yegneswaran, B., Liao, P., Chung, S. A., Vairavanathan, S., Islam, S., Khajehdehi, A., and Shapiro, C. M. (2008). Stop questionnaire: A tool to screen patients for obstructive sleep apnea. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 108(5):812–821.

Cicerone, K. (1995). The neurobehavioral symptom inventory. *Journal of Head Trauma Rehabilitation*, 10(3):1–17.

Cooper, B. (2005). Applying ragin's crisp and fuzzy set qca to large datasets: Social class and educational achievement in the national child development study. *Sociological Research Online*, 10(2):1–20.

Cooper, B. and Glaesser, J. (2016). Exploring the robustness of set theoretic findings from a large n fsqca: An illustration from the sociology of education. *International Journal of Social Research Methodology*, 19(4):445–459.

Cooper, D. B., Bowles, A. O., Kennedy, J. E., Curtiss, G., French, L. M., Tate, D. F., and Vanderploeg, R. D. (2017). Cognitive rehabilitation for military service members with mild traumatic brain injury: A randomized clinical trial. *Journal of Head Trauma Rehabilitation*, 32(3):E1–E15.

Cooper, D. B., Nelson, L., Armistead-Jehle, P., and Bowles, A. O. (2011). Utility of the mild brain injury atypical symptoms scale as a screening measure for symptom over-reporting in operation enduring freedom/operation iraqi freedom service members with post-concussive complaints. *Archives of Clinical Neuropsychology*, 26(8):718–727.

D'agostino Sr, R. B. and Russell, H. K. (2005). Scree test. *Encyclopedia of Biostatistics*, 7.

Davenport, N. D., Lamberty, G. J., Nelson, N. W., Lim, K. O., Armstrong, M. T., and Sponheim, S. R. (2016). Ptsd confounds detection of compromised cerebral white matter integrity in military veterans reporting a history of mild traumatic brain injury. *Brain Injury*, 30(12):1491–1500.

Davenport, N. D., Lim, K. O., Armstrong, M. T., and Sponheim, S. R. (2012). Diffuse and spatially variable white matter disruptions are associated with blast-related mild traumatic brain injury. *NeuroImage*, 59(3):2017 – 2024.

DeJong, J. and Donders, J. (2009). A confirmatory factor analysis of the california verbal learning test—second edition (cvlt-ii) in a traumatic brain injury sample. *Assessment*, 16(4):328–336.

Delis, D., Kramer, J. H., Kaplan, E., and Ober, B. (2000). The california verbal learning test (CVLT-II)– second edition.

Delis, D. C., Kramer, J. H., Kaplan, E., and Holdnack, J. (2004). Reliability and validity of the delis-kaplan executive function system: An update. *Journal of the International Neuropsychological Society*, 10(2):301–303.

Dikmen, S., Machamer, J., and Temkin, N. (2017). Mild traumatic brain injury: Longitudinal study of cognition, functional status, and post-traumatic symptoms. *Journal of Neurotrauma*, 34(8):1524–1530.

Donders, A. R. T., Van Der Heijden, G. J., Stijnen, T., and Moons, K. G. (2006). A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, 59(10):1087–1091.

Draper, K. and Ponsford, J. (2008). Cognitive functioning ten years following traumatic brain injury and rehabilitation. *Neuropsychology*, 22(5):618.

Drozdick, L. W., Wahlstrom, D., Zhu, J., and Weiss, L. G. (2012). The wechsler adult intelligence scale—fourth edition and the wechsler memory scale—fourth edition.

Dusa, A. and Thiem, A. (2018). Qca: A package for qualitative comparative analysis. *Dept. Sociology, Univ. of Bucharest, Bucharest, Romania*.

Eme, R. (2017). Neurobehavioral outcomes of mild traumatic brain injury: A mini review. *Brain Sciences*, 7(5):46.

Emmenegger, P., Schraff, D., and Walter, A. (2014). Qca, the truth table analysis and large-n survey data: The benefits of calibration and the importance of robustness tests. In *2nd International QCA Expert Workshop, November, Zurich, Switzerland*.

Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour Research and Therapy*, 98:4–18.

Fortin, J.-P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., Adams, P., Cooper, C., Fava, M., McGrath, P. J., et al. (2018). Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, 167:104–120.

Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., et al. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, 161:149–170.

French, L. M., Lange, R. T., and Brickell, T. A. (2014). Subjective cognitive complaints and neuropsychological test performance following military-related traumatic brain injury. *Journal of Rehabilitation Research & Development*, 51(6).

Gale, S. D., Johnson, S. C., Bigler, E. D., and Blatter, D. D. (1995). Nonspecific white matter degeneration following traumatic brain injury. *Journal of the International Neuropsychological Society*, 1(1):17–28.

Ginzburg, K., Ein-Dor, T., and Solomon, Z. (2010). Comorbidity of posttraumatic stress disorder, anxiety and depression: A 20-year longitudinal study of war veterans. *Journal of Affective Disorders*, 123(1-3):249–257.

Glatard, T., Lewis, L. B., Ferreira da Silva, R., Adalat, R., Beck, N., Lepage, C., Rioux, P., Rousseau, M.-E., Sherif, T., Deelman, E., et al. (2015). Reproducibility of neuroimaging analyses across operating systems. *Frontiers in Neuroinformatics*, 9:12.

Goodrich, G. L., Flyg, H. M., Kirby, J. E., Chang, C.-Y., and Martinsen, G. L. (2013). Mechanisms of tbi and visual consequences in military and veteran populations. *Optometry and Vision Science*, 90(2):105–112.

Green, R. E., Colella, B., Christensen, B., Johns, K., Frasca, D., Bayley, M., and Monette, G. (2008). Examining moderators of cognitive recovery trajectories after moderate to severe traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, 89(12):S16–S24.

Greer, N., Sayer, N., Koeller, E., Velasquez, T., and Wilt, T. J. (2018). Outcomes associated with blast versus nonblast-related traumatic brain injury in us military service members and veterans: A systematic review. *Journal of Head Trauma Rehabilitation*, 33(2):E16–E29.

Greer, N., Sayer, N., Kramer, M., Koeller, E., and Velasquez, T. (2016). Prevalence and epidemiology of combat blast injuries from the military cohort 2001-2014. PMID:28813129.

Grofman, B. and Schneider, C. Q. (2009). An introduction to crisp set qca, with a comparison to binary logistic regression.

Group, T. E. (1990). Euroqol-a new facility for the measurement of health-related quality of life. *Health Policy*, 16(3):199–208.

Helmick, K. M., Spells, C. A., Malik, S. Z., Davies, C. A., Marion, D. W., and Hinds, S. R. (2015). Traumatic brain injury in the us military: Epidemiology and key clinical and research programs. *Brain Imaging and Behavior*, 9(3):358–366.

Jacobs, M. L. and Donders, J. (2007). Criterion validity of the california verbal learning test (cvlt-ii) after traumatic brain injury. *Archives of Clinical Neuropsychology*, 22(2):143–149.

Jahanshad, N., Kochunov, P. V., Sprooten, E., Mandl, R. C., Nichols, T. E., Almasy, L., Blangero, J., Brouwer, R. M., Curran, J. E., de Zubicaray, G. I., et al. (2013). Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the enigma–dti working group. *NeuroImage*, 81:455–469.

Jennett, B. and Bond, M. (1975). Assessment of outcome after severe brain damage: A practical scale. *The Lancet*, 305(7905):480–484.

Johnson, V. E., Stewart, J. E., Begbie, F. D., Trojanowski, J. Q., Smith, D. H., and Stewart, W. (2013). Inflammation and white matter degeneration persist for years after a single traumatic brain injury. *Brain*, 136(1):28–42.

Jorge, R. E., Acion, L., White, T., Tordesillas-Gutierrez, D., Pierson, R., Crespo-Facorro, B., and Magnotta, V. A. (2012). White matter abnormalities in veterans with mild traumatic brain injury. *American Journal of Psychiatry*, 169(12):1284–1291.

Kelly, J. F., Ritenour, A. E., McLaughlin, D. F., Bagg, K. A., Apodaca, A. N., Mallak, C. T., Pearse, L., Lawnick, M. M., Champion, H. R., Wade, C. E., et al. (2008). Injury severity and causes of death from operation iraqi freedom and operation enduring freedom: 2003–2004 versus 2006. *Journal of Trauma and Acute Care Surgery*, 64(2):S21–S27.

Kennedy, J. E., Leal, F. O., Lewis, J. D., Cullen, M. A., and Amador, R. R. (2010). Posttraumatic stress symptoms in oif/oef service members with blast-related and non-blast-related mild tbi. *NeuroRehabilitation*, 26(3):223–231.

Kennedy, M. R., Wozniak, J. R., Muetzel, R. L., Mueller, B. A., Chiou, H.-H., Pantekoek, K., and Lim, K. O. (2009). White matter and neurocognitive changes in adults with chronic traumatic brain injury. *Journal of the International Neuropsychological Society*, 15(1):130–136.

Kinnunen, K. M. (2011). *Traumatic brain injury: Relationships between brain structural abnormalities and cognitive function*. PhD thesis, Goldsmiths, University of London.

Kinnunen, K. M., Greenwood, R., Powell, J. H., Leech, R., Hawkins, P. C., Bonnelle, V., Patel, M. C., Counsell, S. J., and Sharp, D. J. (2011). White matter damage and cognitive impairment after traumatic brain injury. *Brain*, 134(2):449–463.

Kochunov, P., Thompson, P. M., Winkler, A., Morrissey, M., Fu, M., Coyle, T. R., Du, X., Muellerklein, F., Savransky, A., Gaudiot, C., et al. (2016). The common genetic influence over processing speed and white matter microstructure: Evidence from the old order amish and human connectome projects. *NeuroImage*, 125:189–197.

Koerte, I. K., Hufschmidt, J., Muehlmann, M., Lin, A. P., and Shenton, M. E. (2016). Advanced neuroimaging of mild traumatic brain injury. In *Translational Research in Traumatic Brain Injury*. CRC Press/Taylor and Francis Group.

Kosinski, M., Bayliss, M., Bjorner, J., Ware, J., Garber, W., Batenhorst, A., Cady, R., Dahlöf, C., Dowson, A., and Tepper, S. (2003). A six-item short-form survey for measuring headache impact: The hit-6™. *Quality of Life Research*, 12(8):963–974.

Kraus, M. F., Susmaras, T., Caughlin, B. P., Walker, C. J., Sweeney, J. A., and Little, D. M. (2007). White matter integrity and cognition in chronic traumatic brain injury: A diffusion tensor imaging study. *Brain*, 130(10):2508–2519.

Kroenke, K. and Spitzer, R. L. (2002). The phq-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, 32(9):509–515.

Lajiness-O'Neill, R., Erdodi, L., and Bigler, E. D. (2010). Memory and learning in pediatric traumatic brain injury: A review and examination of moderators of outcome. *Applied Neuropsychology*, 17(2):83–92.

Lemieux, L., Salek-Haddadi, A., Lund, T. E., Laufs, H., and Carmichael, D. (2007). Modelling large motion events in fmri studies of patients with epilepsy. *Magnetic Resonance Imaging*, 25(6):894–901.

Lepage, C., de Pierrefeu, A., Koerte, I. K., Coleman, M. J., Pasternak, O., Grant, G., Marx, C. E., Morey, R. A., Flashman, L. A., George, M. S., et al. (2018). White matter abnormalities in mild traumatic brain injury with and without post-traumatic stress disorder: A subject-specific diffusion tensor imaging study. *Brain Imaging and Behavior*, 12(3):870–881.

Levine, B., Kovacevic, N., Nica, E., Cheung, G., Gao, F., Schwartz, M., and Black, S. (2008). The toronto traumatic brain injury study: Injury severity and quantified mri. *Neurology*, 70(10):771–778.

Li, X., Morgan, P. S., Ashburner, J., Smith, J., and Rorden, C. (2016). The first step for neuroimaging data analysis: Dicom to nifti conversion. *Journal of Neuroscience Methods*, 264:47–56.

Ling, G., Bandak, F., Armonda, R., Grant, G., and Ecklund, J. (2009). Explosive blast neurotrauma. *Journal of Neurotrauma*, 26(6):815–825.

Lipton, M. L., Gellella, E., Lo, C., Gold, T., Ardekani, B. A., Shifteh, K., Bello, J. A., and Branch, C. A. (2008). Multifocal white matter ultrastructural abnormalities in mild traumatic brain injury with cognitive disability: A voxel-wise analysis of diffusion tensor imaging. *Journal of Neurotrauma*, 25(11):1335–1342.

Lipton, M. L., Kim, N., Park, Y. K., Hulkower, M. B., Gardin, T. M., Shifteh, K., Kim, M., Zimmerman, M. E., Lipton, R. B., and Branch, C. A. (2012). Robust detection of traumatic axonal injury in individual mild traumatic brain injury patients: Intersubject variation, change over time and bidirectional changes in anisotropy. *Brain Imaging and Behavior*, 6(2):329–342.

Little, D., Kraus, M., Joseph, J., Geary, E., Susmaras, T., Zhou, X., Pliskin, N., and Gorelick, P. (2010). Thalamic integrity underlies executive dysfunction in traumatic brain injury. *Neurology*, 74(7):558–564.

Lockhart, S., Mayda, A. B., Roach, A., Fletcher, E., Carmichael, O., Maillard, P., Schwarz, C., Yonelinas, A., Ranganath, C., and DeCarli, C. (2012). Episodic memory function is associated with multiple measures of white matter integrity in cognitive aging. *Frontiers in Human Neuroscience*, 6:56.

Losoi, H., Silverberg, N. D., Wäljas, M., Turunen, S., Rosti-Otajärvi, E., Helminen, M., Luoto, T. M., Julkunen, J., Öhman, J., and Iverson, G. L. (2016). Recovery from mild traumatic brain injury in previously healthy adults. *Journal of Neurotrauma*, 33(8):766–776.

Luethcke, C. A., Bryan, C. J., Morrow, C. E., and Isler, W. C. (2011). Comparison of concussive symptoms, cognitive performance, and psychological symptoms between acute blast-versus nonblast-induced mild traumatic brain injury. *Journal of the International Neuropsychological Society*, 17(1):36–45.

Magnuson, J., Leonessa, F., and Ling, G. S. (2012). Neuropathology of explosive blast traumatic brain injury. *Current Neurology and Neuroscience Reports*, 12(5):570–579.

Martínez-Martín, P., Gil-Nagel, A., Gracia, L. M., Gómez, J. B., Martinez-Sarries, J., Bermejo, F., and Group, C. M. (1994). Unified parkinson's disease rating scale characteristics and structure. *Movement Disorders*, 9(1):76–83.

Masel, B. E. and DeWitt, D. S. (2010). Traumatic brain injury: A disease process, not an event. *Journal of Neurotrauma*, 27(8):1529–1540. PMID: 20504161.

McKee, A. C. and Robinson, M. E. (2014). Military-related traumatic brain injury and neurodegeneration. *Alzheimer's & Dementia*, 10:S242–S253.

Mori, S., Wakana, S., Van Zijl, P. C., and Nagae-Poetscher, L. (2005). *MRI atlas of human white matter*. Elsevier.

Muncy, N. M., Hedges-Muncy, A. M., and Kirwan, C. B. (2017). Discrete pre-processing step effects in registration-based pipelines, a preliminary volumetric study on t1-weighted images. *PloS One*, 12(10):e0186071.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47(1):90–100.

Newman, C. W., Weinstein, B. E., Jacobson, G. P., and Hug, G. A. (1991). Test-retest reliability of the hearing handicap inventory for adults. *Ear and Hearing*, 12(5):355–357.

Niogi, S., Mukherjee, P., Ghajar, J., Johnson, C., Kolster, R., Sarkar, R., Lee, H., Meeker, M., Zimmerman, R., Manley, G., and McCandliss, B. (2008). Extent of microstructural white matter injury in postconcussive syndrome correlates with impaired cognitive reaction time: A 3t

diffusion tensor imaging study of mild traumatic brain injury. *American Journal of Neuroradiology*, 29(5):967–973.

Niogi, S. N. and Mukherjee, P. (2010). Diffusion tensor imaging of mild traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, 25(4):241—255.

Novack, T. A., Bush, B. A., Meythaler, J. M., and Canupp, K. (2001). Outcome after traumatic brain injury: Pathway analysis of contributions from premorbid, injury severity, and recovery variables. *Archives of Physical Medicine and Rehabilitation*, 82(3):300–305.

O'Donnell, L. J. and Westin, C.-F. (2011). An introduction to diffusion tensor image analysis. *Neurosurgery Clinics*, 22(2):185–196.

Poldrack, R. A. (2012). The future of fmri in cognitive neuroscience. *NeuroImage*, 62(2):1216 – 1220.

Ragin, C. C. (1994). *Introduction to Qualitative Comparative Analysis*, volume 92 of *The Comparative Political Economy of the Welfare State*. Cambridge University Press: New York, NY.

Reid, M. W., Cooper, D. B., Lu, L. H., Iverson, G. L., and Kennedy, J. E. (2018). Adversity and resilience are associated with outcome after mild traumatic brain injury in military service members. *Journal of Neurotrauma*, 35(10):1146–1155.

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8(3):271–276.

Remington, P. L., Smith, M. Y., Williamson, D. F., Anda, R. F., Gentry, E. M., and Hogelin, G. C. (1988). Design, characteristics, and usefulness of state-based behavioral risk factor surveillance: 1981-87. *Public Health Reports*, 103(4):366.

Rennie, K. M. (1997). Exploratory and confirmatory rotation strategies in exploratory factor analysis. *Annual Meeting of the Southwest Educational Research Association*, January 23-25, 1997.

Revelle, W. (2019). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University, Evanston, Illinois. R package version 1.9.12.

Richardson, L. K., Frueh, B. C., and Acierno, R. (2010). Prevalence estimates of combat-related post-traumatic stress disorder: Critical review. *Australian and New Zealand Journal of Psychiatry*, 44(1):4–19. PMID: 20073563.

Rihoux, B., Álamos Concha, P., Bol, D., Marx, A., and Rezsöhazy, I. (2013). From niche to mainstream method? a comprehensive mapping of qca applications in journal articles from 1984 to 2011. *Political Research Quarterly*, 66(1):175–184.

Saatman, K. E., Duhaime, A.-C., Bullock, R., Maas, A. I., Valadka, A., and Manley, G. T. (2008). Classification of traumatic brain injury for targeted therapies. *Journal of Neurotrauma*, 25(7):719–738.

Schiehser, D. M., Delis, D. C., Filoteo, J. V., Delano-Wood, L., Han, S. D., Jak, A. J., Drake, A. I., and Bondi, M. W. (2011). Are self-reported symptoms of executive dysfunction associated with objective executive function performance following mild to moderate traumatic brain injury? *Journal of Clinical and Experimental Neuropsychology*, 33(6):704–714. PMID: 21958432.

Schneider, C. Q. and Wagemann, C. (2010). Standards of good practice in qualitative comparative analysis (qca) and fuzzy-sets. *Comparative Sociology*, 9(3):397–418.

Schwarzer, R., Jerusalem, M., Weinman, J., Wright, S., and Johnston, M. (1995). *Measures in health psychology: A user's portfolio. Causal and control beliefs*. Windsor, UK: Nfer-Nelson.

Seal, K. H., Cohen, G., Waldrop, A., Cohen, B. E., Maguen, S., and Ren, L. (2011). Substance use disorders in iraq and afghanistan veterans in va healthcare, 2001–2010: Implications for screening, diagnosis and treatment. *Drug and Alcohol Dependence*, 116(1):93 – 101.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., and Dunbar, G. C. (1998). The mini-international neuropsychiatric interview (mini): The development and validation of a structured diagnostic psychiatric interview for dsm-iv and icd-10. *The Journal of Clinical Psychiatry*, 59(Supp20):22–33.

Shenton, M. E., Hamoda, H., Schneiderman, J., Bouix, S., Pasternak, O., Rathi, Y., Vu, M.-A., Purohit, M. P., Helmer, K., Koerte, I., et al. (2012). A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. *Brain Imaging and Behavior*, 6(2):137–192.

Shunk, A. W., Davis, A. S., and Dean, R. S. (2006). Test review: Dean c. delis, edith kaplan & joel h. kramer, delis kaplan executive function system (d-kefs), the psychological corporation, san antonio, tx, 2001. $415.00 (complete kit). *Applied Neuropsychology*, 13(4):275–27.

Silver, J. M., McAllister, T. W., and Arciniegas, D. B. (2009). Depression and cognitive complaints following mild traumatic brain injury. *American Journal of Psychiatry*, 166(6):653–661. PMID: 19487401.

Skinner, H. A. (1982). The drug abuse screening test. *Addictive Behaviors*, 7(4):363–371.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., et al. (2004). Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23:S208–S219.

Smith, S. M. and Nichols, T. E. (2018). Statistical challenges in "big data" human neuroimaging. *Neuron*, 97(2):263–268.

Sorg, M. S. F., Delano-Wood, L., Luc, M. N., Schiehser, D. M., Hanson, K. L., Nation, D. A., Lanni, M. E., Jak, A. J., Lu, K., Meloy, M., et al. (2014). White matter integrity in veterans with mild traumatic brain injury: Associations with executive function and loss of consciousness. *Journal of Head Trauma Rehabilitation*, 29(1):21.

Strong, C.-A. H., Tiesma, D., and Donders, J. (2010). Criterion validity of the delis-kaplan executive function system (d-kefs) fluency subtests after traumatic brain injury. *Journal of the International Neuropsychological Society*, 17(2):230–237.

Tate, D. F. and Bigler, E. D. (2000). Fornix and hippocampal atrophy in traumatic brain injury. *Learning & Memory*, 7(6):442–446.

Teasdale, G. and Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *The Lancet*, 304(7872):81 – 84. Originally published as Volume 2, Issue 7872.

Thiem, A. (2018). *Advanced functionality for performing and evaluating Qualitative Comparative Analysis*. R Package Version 1.1-2.

Thiem, A., Baumgartner, M., Duşa, A., and Spoehel, R. (2016). QCApro: Professional functionality for performing and evaluating qualitative comparative analysis. *R Package Version*, page 1.1.

Tijssen, R. H., Jansen, J. F., and Backes, W. H. (2009). Assessing and minimizing the effects of noise and motion in clinical dti at 3t. *Human Brain Mapping*, 30(8):2641–2655.

Tulsky, D. S., Kisala, P. A., Victorson, D., Carlozzi, N., Bushnik, T., Sherer, M., Choi, S. W., Heinemann, A. W., Chiaravalloti, N., Sander, A. M., et al. (2016). Tbi-qol: Development and calibration of item banks to measure patient reported outcomes following traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 31(1):40.

Vakil, E. and Oded, Y. (2003). Comparison between three memory tests: Cued recall, priming and saving closed-head injured patients and controls. *Journal of Clinical and Experimental Neuropsychology*, 25(2):274–282.

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16(3):219–242.

van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, 45(3):1–67.

Walker, W. C., Carne, W., Franke, L., Nolen, T., Dikmen, S., Cifu, D., Wilson, K., Belanger, H., and Williams, R. (2016). The chronic effects of neurotrauma consortium (cenc) multi-centre observational study: Description of study and characteristics of early participants. *Brain Injury*, 30(12):1469–1480.

Walker, W. C., Cifu, D. X., Hudak, A. M., Goldberg, G., Kunz, R. D., and Sima, A. P. (2015). Structured interview for mild traumatic brain injury after military blast: Inter-rater agreement and development of diagnostic algorithm. *Journal of Neurotrauma*, 32(7):464–473.

Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR.* Psychological Corporation.

Wechsler, D. (2009). Test of premorbid functioning. *San Antonio, TX: The Psychological Corporation*.

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., et al. (2013). Cognition assessment using the nih toolbox. *Neurology*, 80(11 Supplement 3):S54–S64.

White, I. R., Royston, P., and Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine*, 30(4):377–399.

Wilde, E. A., Bigler, E. D., Hunter, J. V., Fearing, M. A., Scheibel, R. S., Newsome, M. R., Johnson, J. L., Bachevalier, J., Li, X., and Levin, H. S. (2007). Hippocampus, amygdala, and basal ganglia morphometrics in children after moderate-to-severe traumatic brain injury. *Developmental Medicine & Child Neurology*, 49(4):294–299.

Zappalà, G., de Schotten, M. T., and Eslinger, P. J. (2012). Traumatic brain injury and the frontal lobes: What can we gain with diffusion tensor imaging? *Cortex*, 48(2):156 – 165. Frontal lobes.

Zhang, S. and Arfanakis, K. (2014). White matter segmentation based on a skeletonized atlas: Effects on diffusion tensor imaging studies of regions of interest. *Journal of Magnetic Resonance Imaging*, 40(5):1189–1198.

Zschoch, M. A. (2011). *Configurational comparative methods: Qualitative Comparative Analysis (QCA) and related techniques*. Canadian Journal of Political Science/Revue Canadienne de Science Politique. Sage Publications: Thousand Oaks, CA.

APPENDIX A. CODE USED TO GENERATE RESULTS FOR STUDIES 1 AND 2

A.1 Study 1 Analysis

A.1.1 Multiple Regression

Multiple regressions were generated with the following code:

- Model = lm(data=data.frame, WMTract $\sim$ STUDYGROUP + Blast_No_TBI + Blast_TBI)

A.1.2 Generalized Linear Model

Code used to generate each glm model can be found below. Note that family denotes the distribution to which the model is assigned, followed by a link function in parentheses. Link is a specification indicating the expected response to the predictors in the model. In these models, a logistic ("logit") link was chosen, as it is suited for models with dependent variables that are continuous but bounded between two values. This code was used on FA and MD values for each of the three specified white matter tracts.

- Model = `glm(Dependent Variable` $\sim$ `Covariate 1 + Covariate 2 + Covariate 3, family="binomial"(logit))`

- Model = `glm(Tract FA/MD` $\sim$ `CVLT Long Delay Cued Recall + CVLT Short Delay Cued Recall + D-KEFS Letter Correct + Age + TOPF Demographic Adjusted IQ + Average Translation + Average Rotation + Blast TBI + Blast No TBI +`
  `Acquisition Site, family="binomial"(logit))`

A.2 Study 2 Analysis

A.2.1 Exploratory Factor Analysis

Code used to generate factors for the QCA analysis can be found below:

```
FINAL_DATASET$GENERAL_FUNCTIONING =
(NSI_AFF_SUM + PHQ9_TOT +
PCL5_TOT + NSI_COG_SUM + TBIQOL_ANXIETY_TOT +
TBIQOL_FATIGUE_TOT + TBIQOL_PAIN_TOT +
TBIQOL_EXEC_FUNC_TOT + TBIQOL_GEN_CONC_TOT +
NSI_SOMA_SUM + TBIQOL_ABIL_PART_TOT +
TBIQOL_ANGER_TOT + TBIQOL_EMO_BEH_TOT + PSQI_TOT +
NSI_VEST_SUM + GSE_TOT + GOSE_OUTCOME +
HIT6_TOT + BRFSSGENHEALTH + HHIAS_TOT +
EQ5D5L_PROFILE + MBIAS_TOT)/22



FINAL_DATASET$CVLT=
(CVLTIILONGDELAYCUEDRECALLRAWSCOR +
CVLTIISHORTCUEDRECALLRAWSCORE +
CVLTIILONGDELAYFREERECALLRAWSCOR +
CVLTIISHORTDELAYFREERECALLRAWSCO)/4



FINAL_DATASET$NIH=
(FLUID_CORRECTTSCORE +
DCCS_CORRECTTSCORE + PATTERN_CORRECTTSCORE +
FLANKER_CORRECTTSCORE + LISTSORT_CORRECTTSCORE +
PICSEQ_CORRECTTSCORE + PICVOCAB_CORRECTTSCORE)/7
```

```
FINAL_DATASET$COG=
(TMTABSCOREB + TMTABSCOREA +
WAISIVDSFORWARDRAWSCORE + WAISIVDSBACKWARDRAWSCORE +
WAISIVCODINGTOTALRAWSCORE + DKEFSLETTERTTLCORRECTRAWSCORE +
BVMTRRECALLRAWSCORE + DKEFSCATEGORYTTLCORRECTRAWSCORE +
BVMTRDELAYEDRECALLRAWSCORE + WAISIVVPTOTALRAWSCORE)/10
```

A.2.2 Calibration: QCA

Code used to threshold factors for QCA can be found below:

- Output Calibrated Variable = `calibrate(Dataset, thresholds=c(partial inclusion, full inclusion))`
- CVLT Calibrated = `calibrate(CVLT, thresholds=c(10.75, 13.00))`
- NIH Calibrated = `calibrate(NIH, thresholds=c(49.29, 55.70))`
- General Function Calibrated = `calibrate(GENERALFUNC, thresholds=c(528.54, 987.59))`
- Cognition Calibrated = `calibrate(COG, thresholds=c(30.40, 32.50))`

APPENDIX B. SUPPLEMENTAL TABLES

Table B.1: Number of Missing Observations in DTI Variables

Variable	**n**	Num Missing
Subject ID	946	0
Average FA ACR	946	0
L ACR FA	946	0
R ACR FA	946	0
Average FA ALIC	946	0
L ALIC FA	946	0
R ALIC FA	946	0
Average Global FA	946	0
FA in Body of CC	946	0
FA in Whole CC	946	0
Average CGC FA	946	0
L CGC FA	946	0
R CGC FA	946	0
Average CGH FA	946	0
L CGH FA	946	0
R CGH FA	946	0
Average CR FA	946	0
L CR FA	946	0
R CR FA	946	0
Average FA CST	946	0
L CST FA	946	0
R CST FA	946	0

Variable	**n**	Num Missing
Average FA EC	946	0
L EC FA	946	0
R EC FA	946	0
Average FX FA	946	0
L FX:ST FA	946	0
R FX:ST FA	946	0
Average FA FXST	946	0
GCC FA	946	0
Average FA IC	946	0
L IC FA	946	0
R IC FA	946	0
Average IFO FA	946	0
L IFO FA	946	0
R IFO FA	946	0
Average PCR FA	946	0
L PCR FA	946	0
R PCR FA	946	0
Average PLIC FA	946	0
L PLIC FA	946	0
R PLIC FA	946	0
Average PTR FA	946	0
L PTR FA	946	0
R PTR FA	946	0
Average RLIC FA	946	0
L RLIC FA	946	0
R RLIC FA	946	0
SCC FA	946	0
Average SCR FA	946	0

Variable	**n**	Num Missing
L SCR FA	946	0
R SCR FA	946	0
Average SFO FA	946	0
L SFO FA	946	0
R SFO FA	946	0
Average SLF FA	946	0
L SLF FA	946	0
R SLF FA	946	0
Average SS FA	946	0
L SS FA	946	0
R SS FA	946	0
Average UNC FA	946	0
L UNC FA	946	0
R UNC FA	946	0
Average ACR MD	946	0
L ACR MD	946	0
R ACR MD	946	0
Average ALIC MD	946	0
L ALIC MD	946	0
R ALIC MD	946	0
Average Global MD	946	0
BCC MD	946	0
CC MD	946	0
Average CGC MD	946	0
L CGC MD	946	0
R CGC MD	946	0
Average CGH MD	946	0
L CGH MD	946	0

Variable	**n**	Num Missing
R CGH MD	946	0
Average CR MD	946	0
L CR MD	946	0
R CR MD	946	0
Average CST MD	946	0
L CST MD	946	0
R CST MD	946	0
Average EC MD	946	0
L EC MD	946	0
R EC MD	946	0
FX MD	946	0
L FX:ST MD	946	0
R FX:ST MD	946	0
Average FX:ST MD	946	0
GCC MD	946	0
Average IC MD	946	0
L IC MD	946	0
R IC MD	946	0
Average IFO MD	946	0
L IFO MD	946	0
R IFO MD	946	0
Average PCR MD	946	0
L PCR MD	946	0
R PCR MD	946	0
Average PLIC MD	946	0
L PLIC MD	946	0
R PLIC MD	946	0
Average PTR MD	946	0

Variable	**n**	Num Missing
L PTR MD	946	0
R PTR MD	946	0
Average RLIC MD	946	0
L RLIC MD	946	0
R RLIC MD	946	0
SCC MD	946	0
Average SCR MD	946	0
L SCR MD	946	0
R SCR MD	946	0
Average SFO MD	946	0
L SFO MD	946	0
R SFO MD	946	0
Average SLF MD	946	0
L SLF MD	946	0
R SLF MD	946	0
Average SS MD	946	0
L SS MD	946	0
R SS MD	946	0
Average UNC MD	946	0
L UNC MD	946	0
R UNC MD	946	0

*Note: ACR= Anterior Corona Radiata, ALIC=Anterior Limb of the Internal Capsule, CC=Corpus Callosum, CGC=Cingulum in the Cingulate Gyrus, CGH=Cingulum adjoining the Hippocampus, CR=Corona Radiata, CST=Corticospinal Tract, EC=External Capsule, FX=Fornix, FX:ST=Fornix Stria Terminalis, GCC=Genu of Corpus Callosum, IC=Internal Capsule, IFO=Inferior Fronto-Occipital Fasciculus, PCR=Posterior Corona Radiata, PLIC=Posterior Limb of the Internal Capsule, PTR=Posterior Thalamic Radiation, Retrolenticular Limb of the Internal Capsule, SCC=Splenium of Corpus Callosum,

SCR=Superior Corona Radiata, SFO=Superior Fronto-Occipital Fasciculus, SLF=Superior Longitudinal Fasciculus, SS=Saggital Stratum, UNC=Uncinate Fasciculus.

Table B.2: Number of Missing Observations in Clinical Variables

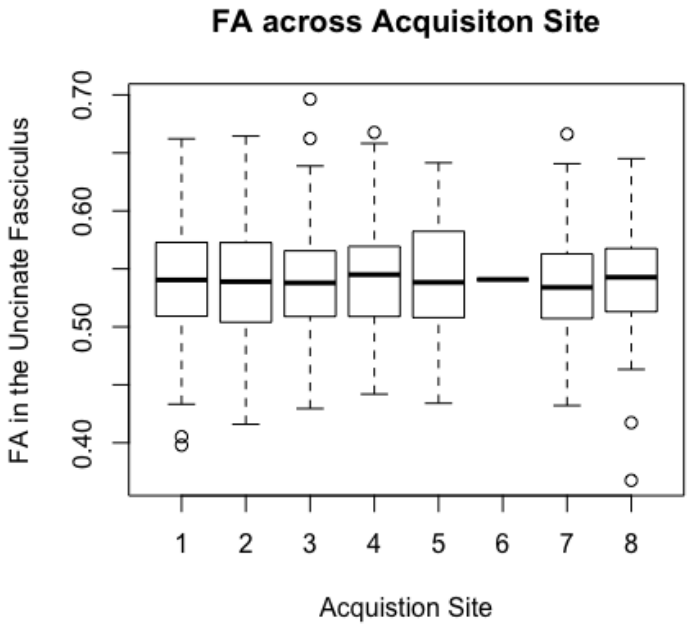
Variable	**n**	Num Missing
Age	918	28
AUDIT-C: Hazardous Alcohol Use	941	5
Num Exposures to Controlled Blasts	926	20
Num Exposures to Blasts Not Resulting in TBI	926	20
Num Exposures to Uncontrolled Blasts	926	20
Total Num of Blast Exposures	926	20
BRFSS: Diagnosed w/ Angina/Heart Disease	946	0
BRFSS: Rating of General Health	946	0
BRFSS: Times Taking Part in Physical Activity	946	0
BVMT: Delayed Recall Raw	943	3
BVMT: Total Recall Raw	943	3
CVLT: Long Delay, Cued Recall Raw	940	6
CVLT: Long Delay, Free Recall Raw	940	6
CVLT: Short Delay, Cued Recall Raw	940	6
CVLT: Short Delay, Free Recall Raw	940	6
DAST-10 Score	943	3
NIH Toolbox: DCCS Correct Raw Score	853	93
Best Description of Race/Ethnicity	946	0
D-KEFS Category Correct Raw	944	2
D-KEFS Letter Correct Raw	944	2
EQ-5D-5L Profile Score	944	2
EQ-5D-5L Rate Health Today	944	2
NIH Toolbox: Flanker Correct t Score	858	88
NIH Toolbox: Fluid Correct t Score	845	101
GOSE: Outcome Score	943	3
GSE Total Score	943	3
HHIAS Total	944	2

Variable	**n**	Num Missing
HIT-6 Total	944	2
NIH Toolbox: ListSort Correct t Score	856	90
MBIAS Total	944	2
MINI Diagnosis	945	1
NSI: Affective Sum	943	3
NSI: Cognitive Sum	944	2
NSI: Somatosensory Sum	943	3
NSI: Vestibular Sum	944	2
NIH Toolbox: Pattern Correct t Score	854	92
PCL-5 Total	941	5
PHQ-9 Total	939	7
NIH Toolbox: Picture Sequence Correct t Score	852	94
NIH Toolbox: Picture Vocab Correct t Score	859	87
PSQI Total	921	25
STOPBANG: BMI	940	6
STOPBANG: Total	921	25
TBIQOL: Ability to Participate	894	52
TBIQOL: Anger	902	44
TBIQOL: Anxiety	904	42
TBIQOL: Emotional and Behavioral Dyscontrol	896	50
TBIQOL: Executive Function	905	41
TBIQOL: Fatigue	899	47
TBIQOL: General Concentration	899	47
TBIQOL: Pain	897	49
TMT Trial A Score	942	4
TMT Trial B Score	942	4
TOPF Demographically Adjusted IQ	938	8
UPDRS Total	882	64

Variable	**n**	Num Missing
WAIS-IV Coding Score	942	4
WAIS-IV Backward Digit Span Score	943	3
WAIS-IV Forward Digit Span Score	943	3
WAIS-IV Visual Puzzles Score	939	7
Movement: Displacement Absolute	946	0
Movement: Displacement Relative	946	0
Movement: Rotation in X Direction	946	0
Movement: Rotation in Y Direction	946	0
Movement: Rotation in Z Direction	946	0
Movement: Translation in X Direction	946	0
Movement: Translation in Y Direction	946	0
Movement: Translation in Z Direction	946	0
Movement: Average Displacement	946	0
Movement: Average Rotation	946	0
Movement: Average Translation	946	0

*Note: AUDIT=Alcohol Use Disorders Identification Test, BRFSS= Behavioral Risk Factor Surveillance System, BVMT=Brief Visuospatial Memory Test, CVLT=California Verbal Learning Test, DAST-10=Drug Abuse Screening Test, DCCS=Dimensional Change Card Sort Test, D-KEFS=Delis Kaplan Executive Function System, GOSE=Glasgow Outcome Scale, HHIAS=Hearing Handicap Inventory for Adults, HIT-6=Headache Impact Test, MBIAS=Mild Brain Injury Atypical Symptoms, MINI= Mini International Neuropsychiatric Interview, NSI=Neurobehavioral Symptom Inventory, PCL-5=PTSD Check List, PHQ-9=Patient Health Questionnaire, PSQI=Pittsburgh Sleep Quality Index, TBIQOL=TBI Quality of Life, TMT=Trail Making Test, TOPF=Test of Premorbid Functioning, UPDRS=Unified Parkinsons' Disease Rating Scale, WAIS-IV=Wechsler Adult Intelligence Scale.

Figure B.1: FA in the Uncinate Fasciculus across Acquisition Sites



*Note: Site 6 appears without a box because only one subject from that site was included in the final analysis.

Figure B.2: FA in the Cingulum Bundle across Acquisition Sites

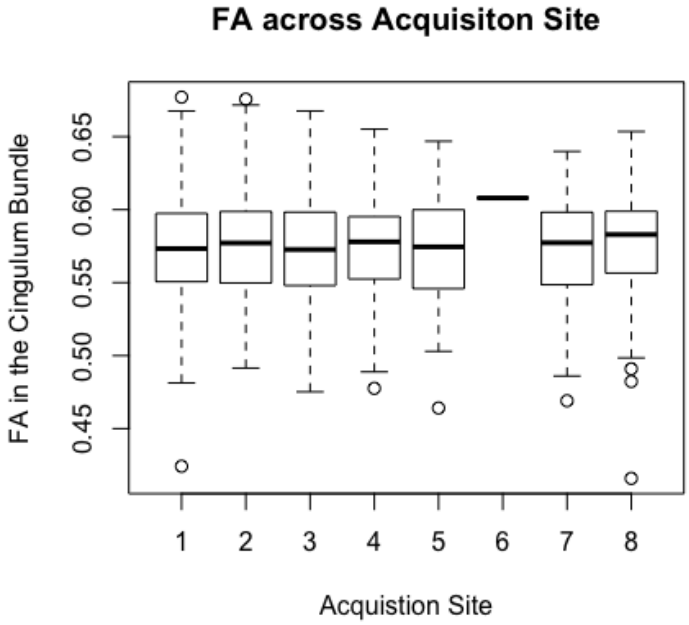


Figure B.3: FA in the Genu of the Corpus Callosum across Acquisition Sites

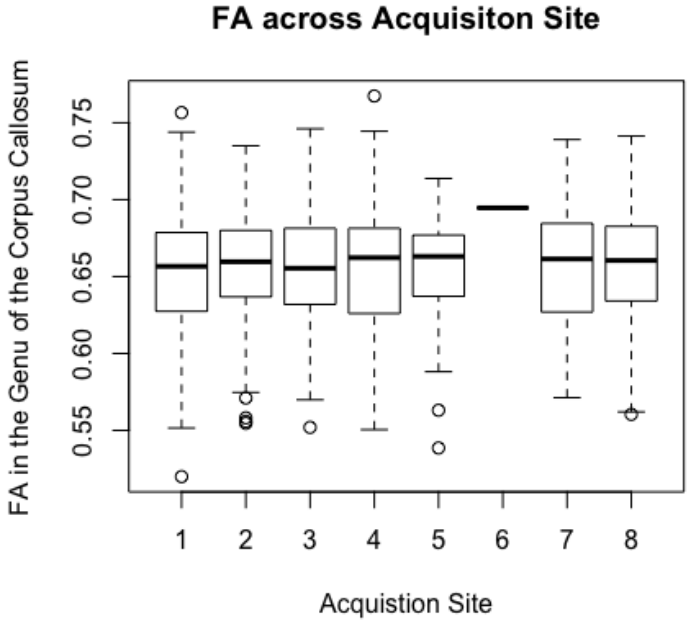


Figure B.4: MD in the Uncinate Fasciculus across Acquisition Sites

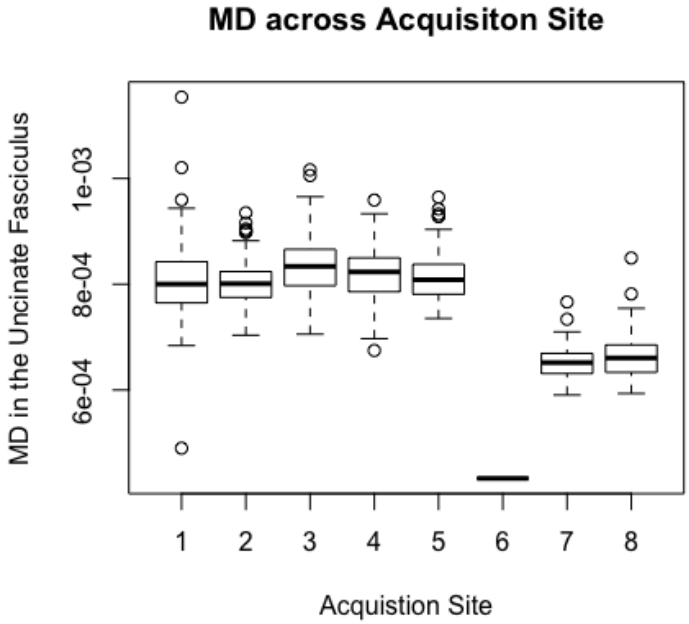


Figure B.5: MD in the Cingulum Bundle across Acquisition Sites

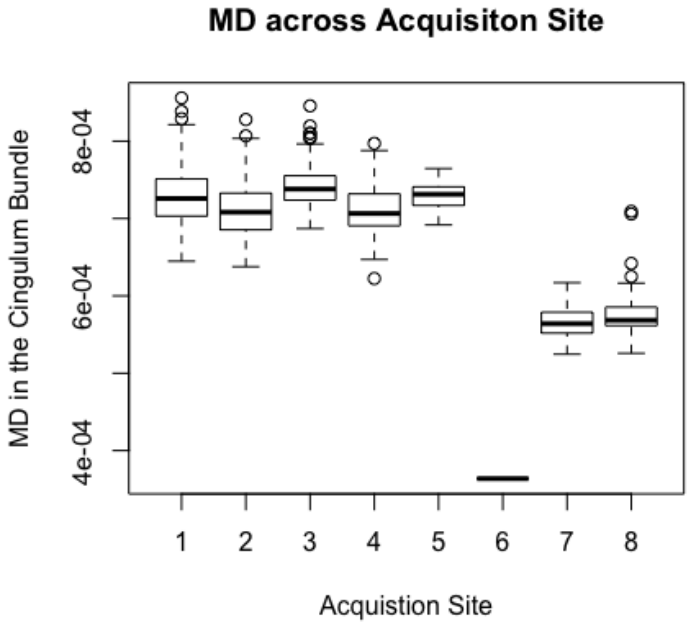


Figure B.6: MD in the Genu of the Corpus Callosum across Acquisition Sites

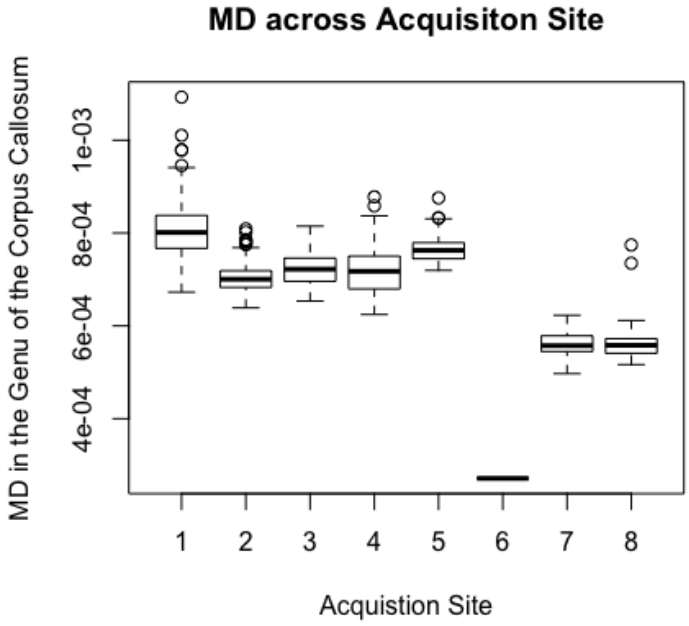
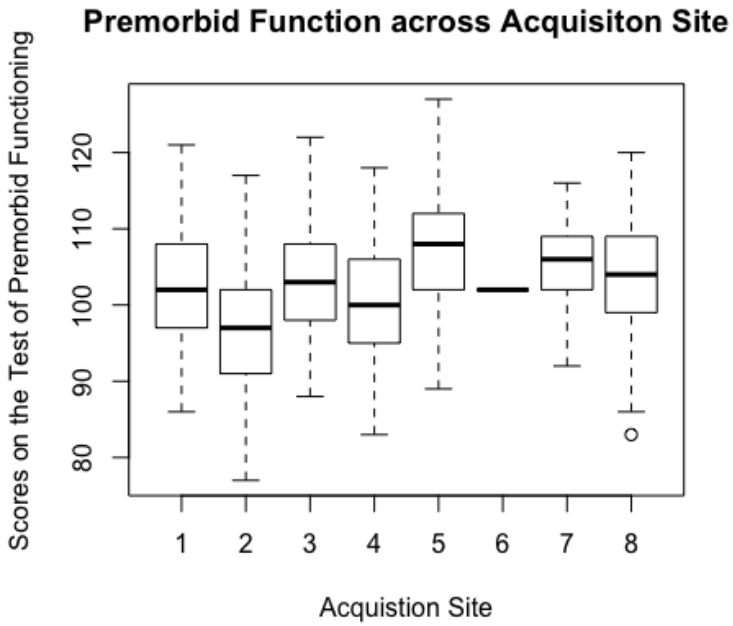
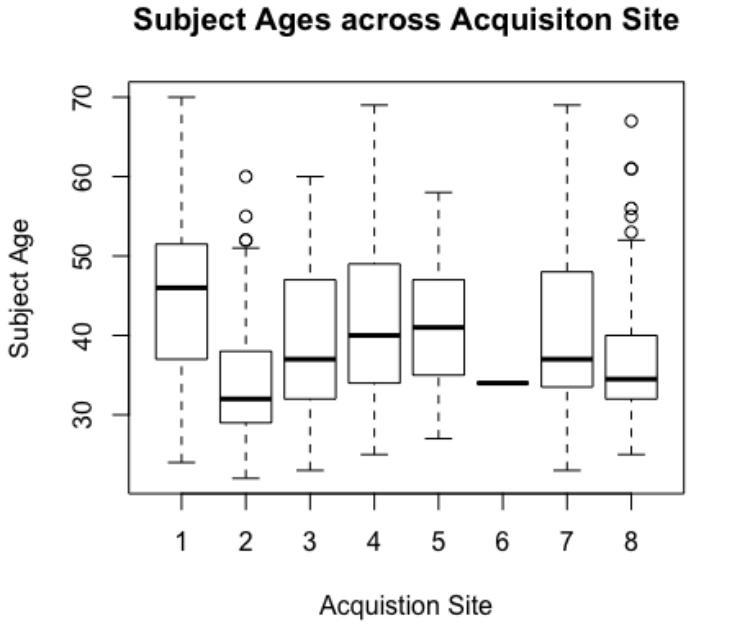


Figure B.7: Premorbid Functioning across Acquisition Sites



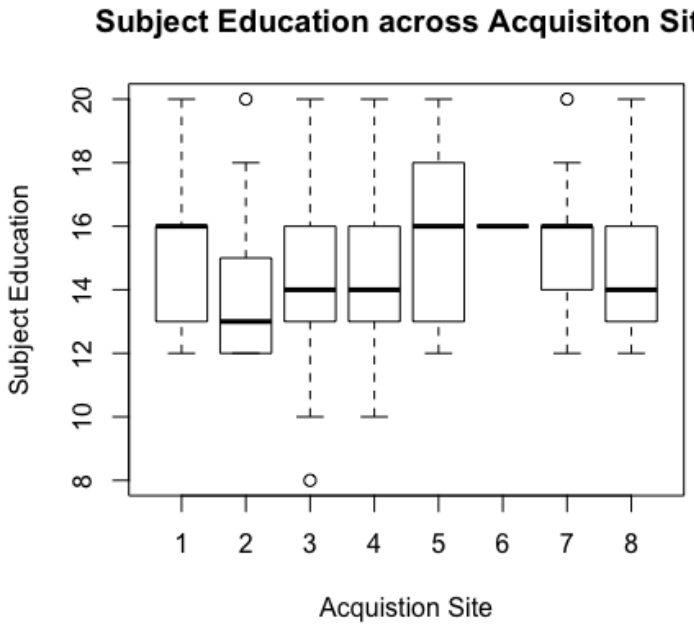
*Note: The Y-axis represents subject scores on the Test of Premorbid Functioning (TOPF).

Figure B.8: Age across Acquisition Sites



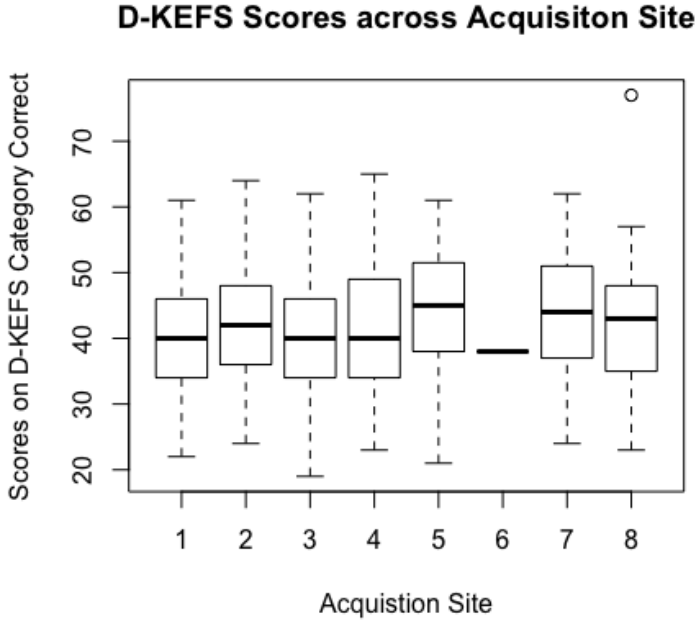
*Note: The Y-axis represents subject age, in years.

Figure B.9: Education across Acquisition Sites



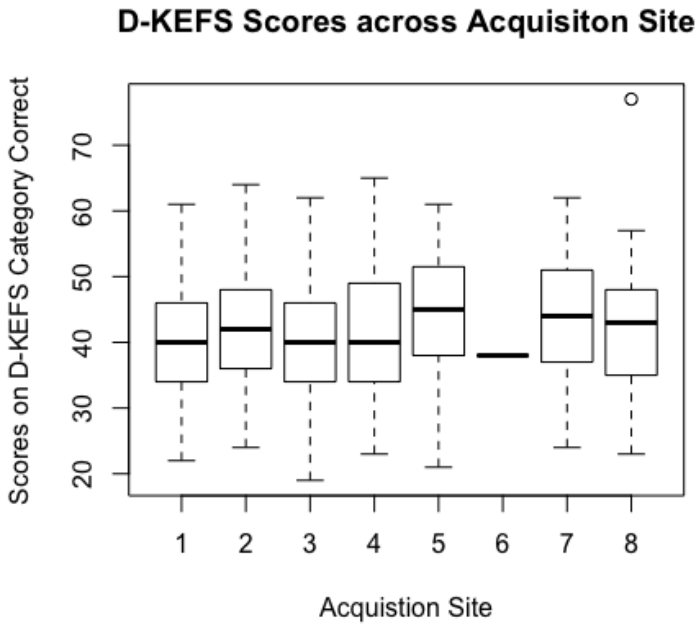
*Note: The Y-axis represents subject education, in years.

Figure B.10: D-KEFS Category Scores across Acquisition Sites



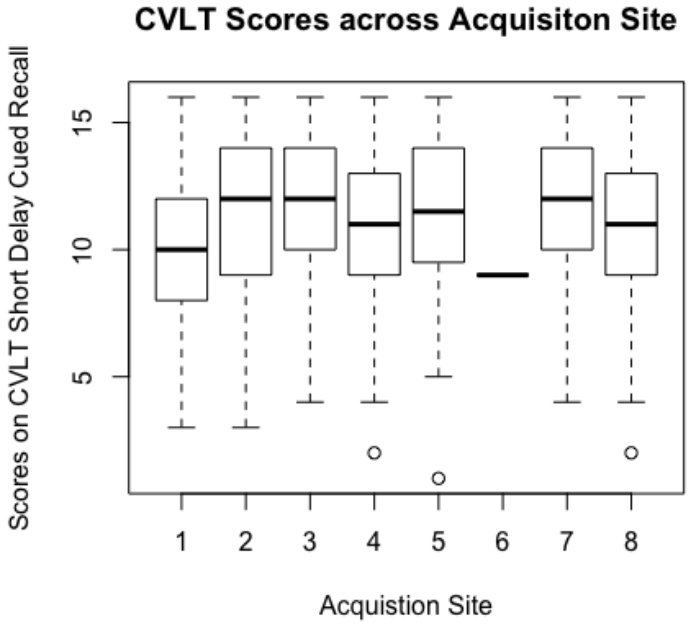
*Note: The Y-axis represents number of correct subject scores on the D-KEFS Category subscale.

Figure B.11: D-KEFS Letter Scores across Acquisition Sites



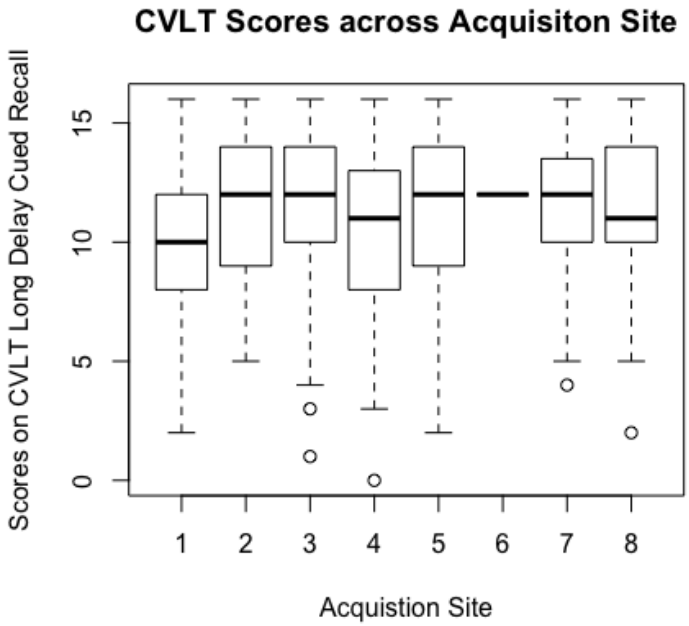
*Note: The Y-axis represents number of correct subject scores on the D-KEFS Letter subscale.

Figure B.12: CVLT Short Delay Cued Recall across Acquisition Sites



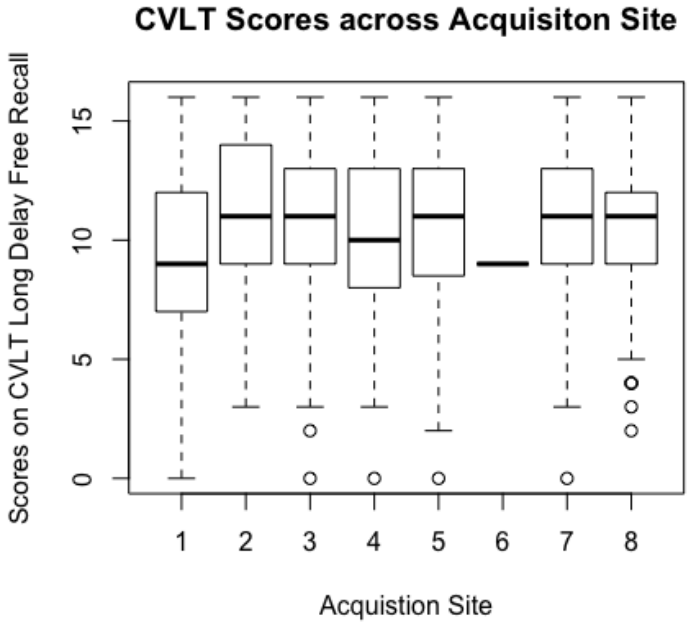
*Note: The Y-axis represents scores on the CVLT-II Short Delay Cued Recall subscale.

Figure B.13: CVLT Long Delay Cued Recall across Acquisition Sites



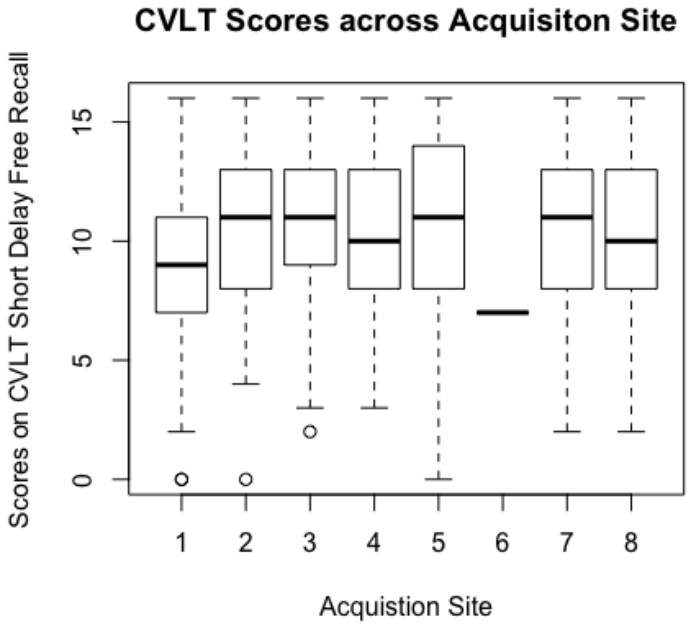
*Note: The Y-axis represents scores on the CVLT-II Long Delay Cued Recall subscale.

Figure B.14: CVLT Long Delay Free Recall across Acquisition Sites



*Note: The Y-axis represents scores on the CVLT-II Long Delay Free Recall subscale.

Figure B.15: CVLT Short Delay Free Recall across Acquisition Sites



*Note: The Y-axis represents scores on the CVLT-II Short Delay Free Recall subscale.