Characterizing Multiple Memory Systems in Moderate-Severe Traumatic Brain Injury

By

Natalie V. Covington

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Hearing and Speech Sciences

August 31, 2019

Nashville Tennessee

Approved:

Melissa Duff, Ph.D.

Stephen Wilson, Ph.D.

Michael de Riesthal, Ph.D.

Neal Cohen, Ph.D.

To Nana: brilliant & strong,
compassionate & faithful –
everything I aspire to be

ACKNOWLEDGEMENTS

Apparently it takes a village to raise a grad student, and I am incredibly thankful for mine. First, thank you to each member of my committee, for providing such valuable support and feedback to this project specifically, and for the myriad ways you have supported my development as a scholar over the past few years. I would especially like to thank Dr. Melissa Duff for her brilliant mentorship and commitment to developing her students into independent researchers. Thank you for teaching me to think big thoughts, hunt down interesting ideas, and respond to setbacks with resilience - grateful everyday for your support and friendship. Thanks also to Dr. Sarah Brown-Schmidt for her considerable contributions to my growth as a scientist.

Thanks to the members of the Communication and Memory Lab, past and present, for supporting this work and for each of your contributions to my development as a scientist. Thanks in particular to Caitie Hilverman, Dave Warren, Emily Morrow, Sharice Clough, and Nirav Patel. Thanks to my wonderful officemates, Pumpki and Carlos, for seeing me at my most-stressed and liking me anyway. Thanks to my Vanderbilt cohort, for adopting a transfer student. Thanks to Sarah and Deb in particular for keeping me sane these past few months, and for being there for me in the ups and downs of the past three years. Thanks to the Iowa Writing Crew (Hope, Kelsey, Jess, Larissa, Naomi) for staying in touch, encouraging me daily, and being such an inspiring group of female scientists. Thanks to my undergraduate mentors, Drs. Tiffany Hogan, Jordan Green, and Julie Wolter, for encouraging my early research training and convincing me that this path was for me.

Thanks to all our friends for supporting David and I so well. Thanks in particular to the Dietzmans, Evans, Langers, and Tutsons in Iowa City, the Sunday Night Dinner crew in

Nashville, Liz and Matt for your support and solidarity from afar, and our church communities at OAH and Midtown. Thanks to my amazing in-laws, for loving me like a daughter and for dropping everything when we needed you. I am immensely grateful and forever indebted to my wonderful parents, for their daily love and support and consistent example of how to live life well. Thanks for supporting my work while continuing to remind me that work itself is not the chief end of man. Thanks to my brothers for being your weird, lovable selves, and specifically to Josh for consistently calling me a scientist before I really felt like one. Finally, all of my love and gratitude to David, for your unfailing love and support each step of the way. It is such a joy to walk through life with you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Heterogeneity in patient characteristics and outcomes for individuals with traumatic brain injury (TBI) is a frequently cited problem in rehabilitative fields, impeding both research progress and clinical practice. Functional outcomes and recovery trajectories following TBI are highly variable across individuals, even when two patients present with the same injury severity and lesion location (Dahdah et al., 2016; Hart et al., 2014). This heterogeneity poses challenges for speech-language pathologists' clinical decision making while implementing cognitive rehabilitation post-TBI, and makes determining accurate prognoses for individual patients nearly impossible. Previous research demonstrates that not all individuals with TBI respond equally well to the same treatments (Cicerone et al., 2011). Despite frequent acknowledgement of the problem of heterogeneity, cognitive rehabilitation continues to take a one-size-fits-all approach to treating cognitive impairments following brain injury. Time spent on treatments that are ineffective for a given individual waste valuable healthcare resources and ineffectively use recovery time that might instead be dedicated to treatments tailored to the individual.

How might clinicians best match individual patients to particular treatments or predict which patients will go on to have good outcomes? I propose that one avenue for improving response to treatment in cognitive rehabilitation is by considering the learning and memory profiles of individual patients in treatment decision-making and prognostication. This proposal is motivated by recognition that all of cognitive rehabilitation (regardless of the particular impairment being targeted) rests on patients' ability to learn or relearn some set of knowledge or

skills. In particular, the current study sets a foundation for the development of learning and memory phenotypes, from a multiple memory systems perspective. This perspective takes into account the possibility that patients with TBI may have differential degrees of impairments within and across memory systems that may impact their responsiveness to different behavioral treatment methods. Ultimately, future studies in this area will allow clinicians to make decisions that promote alignment between the memory systems that are intact and available to the patient (the patient's memory profile) and the memory system(s) that support the acquisition and use of particular information, strategies, and skill (i.e. their chosen treatment strategy). From this perspective, memory and learning are more than potential targets of rehabilitation. Rather, they are the foundation upon which all other behavioral interventions rest, supporting the learning, relearning, and use of knowledge, skills, and strategies across cognitive domains.

The current study sets the foundation for this work by examining learning and memory abilities in patients with moderate-severe TBI across multiple memory systems (declarative, procedural memory). This study provides preliminary data that serves to inform several challenges in the development of useful memory and learning profiles in patients with TBI. First, clinical assessment of some memory systems is routine (i.e. working memory, declarative memory), with multiple clinically available methods for informal and formal assessment (e.g., Auditory Verbal Learning Test, Weschler Memory Scale, Rivermead Behavioral Memory Test). Conversely, procedural memory is not routinely assessed in the clinic; indeed, there are no well-validated or standardized measures that would allow for characterization of this memory system in a clinical setting. The current study tests the utility of several experimental measures of procedural memory in individuals with TBI. Furthermore, there are few studies that have examined procedural memory ability in TBI using these experimental tasks. Second, even for

declarative memory, for which many clinical assessments exist, currently available clinical measures do not fully capture the functional impairments in memory that patients report. As a result, mild impairments in declarative memory may go undetected and undermine treatment success. Third, assumptions about which memory systems are impaired and spared following TBI pervade the literature, but to date, no study has examine memory ability across memory systems in the same group of patients. Given the well-documented heterogeneity in the TBI population (i.e., cognitive profiles are highly variable within and across individuals with TBI), this lack of data represents a critical gap in the literature.

The long-term and overarching goal of this work is to develop a comprehensive memory and learning profile or phenotype for individuals with TBI. This study takes the first step in meeting this long-term goal by attempting to identify a set of tasks that are sensitive to memory deficits in individuals with TBI across distinct forms of memory and that individuals with TBI can perform and tolerate. In the rest of this introduction, I review the literature on traumatic brain injury and multiple memory systems concluding with the specific research questions for this study.

Traumatic Brain Injury

Traumatic brain injury (TBI) is defined as sudden physical damage to the brain caused by mechanical force (Frankowski, 1986; McDonald, Togher, & Code, 1999; Morton & Wehman, 1995). TBI can result from either penetrating or blunt forces. Penetrating head injuries are caused by penetration of the skull and dura by a foreign object (e.g. bullet, knife). Penetrating injuries are relatively uncommon and are more likely to be fatal (Togher, McDonald, & Code, 1999). Closed head injuries are caused by a blunt blow to the skull, without penetration of the

skull or dura. Closed head injuries result in multi-focal damage to the brain (in contrast to the more focal injuries caused by penetrating brain injury) and cause widespread deficits (Togher et al., 1999). The current study focuses on participants with closed head injury, as they are more representative of the population of TBI patients receiving rehabilitative services.

Demographic characteristics of TBI.

The "classic" TBI profile is of a young adult male engaging in risk-taking behavior (Ylvisaker & Feeney, 1998). In general, this depiction is accurate – incidence of TBI is higher in males compared to females, young adults have increased risk of TBI, and roughly half of hospitalized patients with TBI are intoxicated on admission (Corrigan, Rust, & Lamb-Hart, 1995; Faul, Xu, Wald, & Coronado, 2010). Still, this portrayal does not illustrate the full range of demographics of TBI survivors. TBI incidence across the lifespan is a trimodal distribution, with young children between the ages of 0 and 4, older adolescents age 15 to 19, and adults older than 65 most likely to sustain a TBI (Faul et al., 2010). Within these age ranges, very young children are the most likely to visit the emergency room, while hospitalization and death occurs most frequently in adults over the age of 75 (Faul et al., 2010). Across all age groups, incidence of TBI is higher for males than for females, with males approximately 1.4 times more likely to sustain a TBI than females (Faul et al., 2010). However, males are even more likely than females (1.8 times) to sustain more severe injuries involving hospitalization (Faul et al., 2010). Historically research in TBI has reflected this uneven sex distribution, with many studies including a significantly larger proportion of male participants, or including no female participants at all. This has led to a call for increased attention to the study of TBI in female participants (El-Menyar et al., 2014; Farace & Alves, 2000), particularly since recent evidence

suggests sex differences in the effects of TBI on various cognitive domains (e.g. social cognition; Rigon et al. 2016). The current study aims to recruit an even sex distribution of individuals with TBI. To avoid the potentially confounding influences of cognitive aging, the current study includes patients between the ages of 18 and 55.

Mechanism of TBI.

The pathophysiology of TBI is often unique to the circumstances surrounding the injury. For example, consequences differ depending on whether the skull was moving versus immobile during injury and whether the linear versus rotational forces were present. Following the primary injury immediately following the traumatic event, each TBI also results in unique secondary events (Ylvisaker & Feeney, 1998). While each brain injury is unique, common pathophysiological characteristics are described below.

Brain damage following TBI can be classified into two phases: primary injury at the time of impact and secondary injury that unfolds over time in response to the initial traumatic event. Primary injuries include contusions, hemorrhages, and diffuse axonal injury. Secondary injuries include damage caused by elevated intracranial pressure, edema, hypoxia, and ischemia.

Primary injuries are caused by the immediate mechanical disruption of neural pathways (Mcintosh, 1993). Focal contusions can occur at the site of the blow (coup) and often in an area opposite the blow (contrecoup). Coup and contrecoup contusions can result in specific and localizable behavioral changes following TBI (Hannay, Howieson, Loring, Fischer, & Lezak, 2004). Similar localized damage can also occur in the absence in cases of rapid deceleration, with contusions occurring in areas where the cortex rests on the rough surface of the base of the skull (Courville, 1942; Hannay et al., 2004). In cases involving significant momentum on impact

(such as MVAs), clearly defined focal deficits are harder to distinguish, resulting in the description of these injuries as "multi-focal" (Hannay et al., 2004; Ponsford, Olver, & Curran, 1995).

The combination of translational forces, rotational acceleration, and inertial forces can put strain on nerve fibers, leading to diffuse axonal injury. Diffuse axonal injury (DAI) results from the shearing, twisting, tearing, and breakage of axons and is often considered the most common and important pathologic feature of TBI (Smith, Meaney, & Shull, 2003). Diffuse damage caused by DAI tends to be more pronounced in anterior compared to posterior regions and in deeper brain structures compared to surface areas (Hannay et al., 2004). The severity of damage due to DAI is influenced by the velocity, duration, rate of deceleration, and direction of head movement (Hannay et al., 2004). DAI severity has been shown to correspond with level of impairment and prognosis following TBI (Adams, Graham, Gennarelli, & Maxwell, 1991).

Secondary injury following TBI can develop hours and even weeks following the initial traumatic event (Cooper, 1985). Delayed damage is thought to result from changes in the brain's neurochemical systems brought caused by primary injuries. These changes in neurochemistry may result in brain swelling, changes in brain blood flow, and neurotoxic effects leading to the death of neuronal or glial cells (Hannay et al., 2004).

Swelling or edema following TBI can focal or generalized and can result in compression of the brain against the confines of the skull. Compression of lower brainstem structures devoted to vital functions due to swelling and increased intracranial pressure is the most frequent cause of death following TBI (Adams, Graham, & Gennarelli, 1985). Increased intracranial pressure can also lead to reduced blood flow to the cortex, leading to cell death due to hypoxia or ischemia. In severe TBI, some degree of hypoxic-ischemic injury is present in over eighty percent of cases

(Chang et al., 2009; Graham, Adams, & Doyle, 1978). The hippocampus, a structure central to the formation of new declarative memory, is particularly sensitive to hypoxic-ischemic injury (Bigler, 1990). Even in mild TBI, the structure and connectivity of the hippocampus is vulnerable to injury (Leh et al., 2017). While frank impairments in declarative memory are easy to detect, it is possible the more mild deficits may go undetected. In the experimental memory literature, new tasks (e.g., Spatial Reconstruction Task) have shown increased sensitivity in detecting hippocampal pathology and memory deficits over traditional neuropsychological measures (Clark et al., 2017; Watson, Voss, Warren, Tranel, & Cohen, 2013), which holds promise for detecting subtle memory deficits in TBI.

Public health impact of TBI.

TBI represents a considerable burden on individuals, families, and health care systems worldwide. In the United States, the Centers for Disease Control and Prevention estimate that more than 2.5 million traumatic brain injuries per year are serious enough to warrant an emergency department visit or hospitalization (Taylor, Bell, Breiding, & Xu, 2017). Worldwide, an estimated 10 million people sustain a TBI annually (Hyder, Wunderlich, Puvanachandra, Gururaj, & Kobusingye, 2007). These seemingly high numbers likely underestimate the problem by excluding those with milder injuries who do not immediately seek medical attention. TBI results in a constellation of physical, cognitive, and psychosocial deficits that have a wide-reaching negative impact on individuals and their families, and result in billions of dollars of medical management and lost income.

After sustaining a TBI, many individuals present with lasting changes in cognition, behavior, and communication. While some individuals with TBI have visible physical

impairments, many do not. Therefore, to lay observers, deficits in cognition may not be immediately evident. The tendency for deficits related to TBI to be missed during initial visual appraisal contributes to limited public awareness of the pervasiveness of the problem, leading to TBI's designation as a "silent epidemic" (Faul et al., 2010).

Improved medical management of acute TBI has resulted in a decrease in TBI-related mortality over time (Arabi et al., 2010; Gerber et al., 2013; Taylor et al., 2017). Of the 2.5 million yearly hospitalizations due to TBI, only 2% of cases result in death (Taylor et al., 2017). While TBI-related deaths have decreased, annual rates of emergency department visits and hospitalizations due to TBI are on the rise (Coronado et al., 2011; Faul et al., 2010; Langlois, Rutland-Brown, & Wald, 2006). As a result, TBI is a major cause of long-term disability, with an estimated 3.2-5 million Americans living with TBI-related long-term disability (Coronado et al., 2011; Zaloshnja, Miller, Langlois, & Selassie, 2008). Characterizing and rehabilitating disability is particularly important in TBI, given that a large proportion of patients are young adults (versus other neurological disabilities more common in older adults, such as stroke) who will survive for decades following their injury (Pagulayan, Temkin, Machamer, & Dikmen, 2006; Ponsford et al., 2014).

Societal costs of TBI.

In addition to its negative impact on individuals and their families, TBI also imposes high costs to society. While the *aggregated* societal costs of TBI are less costly compared to other disabling conditions requiring rehabilitative care (such as more frequently occurring conditions such as back pain or arthritis), the *per-capita* costs of TBI are among the highest of any disabling condition (Ma, Chan, & Carruthers, 2014). This is due in part to the potential severity of

disability following TBI, as well as the length of time disabled (due to higher prevalence during younger years) (Ma et al., 2014). Additionally, TBI results in productivity loss that is fourteen times that associated with spinal cord injury (Finkelstein, Corso, & Miller, 2006).

Losses in productivity, in combination with the high incidence of TBI, result in high costs to society. The cost of TBI in the United States is estimated to total $60 billion annually, as measured by the sum of medical costs and lost productivity (Finkelstein et al., 2006). Besides these direct monetary costs, TBI imposes additional burdens on society that are harder to quantify. A meta-analysis of epidemiologic studies estimates the prevalence of TBI in the overall offender population to be 48-72% (Shiroma, Ferguson, & Pickelsimer, 2010). In contrast, psychiatric disorders are reported in a much smaller 6-19% of incarcerated individuals (Ditton, 1999). Strikingly, a majority of head-injured inmates report that their head injury preceded their first criminal act, suggesting that head injury may increase the likelihood of criminal activity (Sarapata, Herrmann, Johnson, & Aycock, 1998). Thus, TBI results in significant costs to individuals and society that may be alleviated in part by improvements in rehabilitative care.

Resolution of cognitive impairments following TBI.

Critically, the cognitive consequences of TBI are some of the last to resolve, if they resolve at all. While patients with TBI make significant physical gains in the first months following injury, cognitive and communication difficulties are more resistant to change and often show little to no improvement (Pagulayan et al., 2006). Impairments in cognition and communication persist in a large percentage of patients with TBI, with approximately 60% of patients reporting memory impairments even 10 years post-injury (Ponsford et al., 2014). Memory problems were the most highly reported impairment in a longitudinal study of self-

reported outcome following TBI, and these problems did not lessen over time (at 2-, 5-, and 10-years post injury) (Ponsford et al., 2014).

Cognitive problems often persist long into the chronic phase of recovery. In a study in which patients were followed for 10 years following injury, Ponsford and colleagues note that, "by and large, problems that were present at two years post-injury were still present at 10 years post-injury" (Ponsford et al., 2014). More than 50% of patients reported continuing difficulty concentrating and word-finding, as well as slowed thinking and cognitive fatigue. Additionally, reports of difficulties with planning, social interaction, speech intelligibility, and comprehending discourse increased over time post-injury, perhaps reflecting increased awareness of these deficits over time (Ponsford et al., 2014). Patients report increasing social isolation with the degree of difficulty in personal relationships increasing over time. In particular, while individuals sustaining a mild TBI largely cease to report a need for additional support by 10 years post injury, participants with moderate-severe TBI report a continued need for additional support at 10 years post-injury (Ponsford et al., 2014). Thus, outcomes research suggests that memory impairments are common following TBI, these impairments continue to negatively impact quality of life into the chronic phase of recovery, and individuals with moderate-severe TBI, in particular, demonstrate lasting cognitive difficulties that may be amenable to personalized interventions.

Efficacy of Cognitive Rehabilitation

Given the substantial risks cognitive impairments pose to maintaining gainful employment and regaining a high quality of life, maximizing gains in cognitive rehabilitation is essential. Evidence suggests that cognitive rehabilitation is generally effective, with moderate

effect sizes (Cicerone et al., 2011). On the whole, participation in behavioral interventions appears to improve functional outcomes for patients with TBI, compared to those who do not receive such treatment (Goranson, Graves, Allison, & La Freniere, 2003). However, it has been suggested that heterogeneity in the patient population of individuals with TBI may be hampering effective intervention at the individual level (Cicerone et al., 2011; Dahdah et al., 2016; Goranson et al., 2003; Hart et al., 2014; Lu, Gary, Neimeier, Ward, & Lapane, 2012). In addition, systematic reviews and meta-analyses frequently find that heterogeneity in the TBI sample is not adequately described (Elliott & Parente, 2014).

Rehabilitation services have been shown to be ultimately cost-saving but are still very resource intensive (Andelic et al., 2014). In addition, research suggests that patients receive significantly more hours of physical rehabilitation compared to cognitive rehabilitation (Ponsford et al., 2014), despite the disproportionate negative impact of cognitive impairment on employment and life satisfaction outcomes. Therefore, it is critical that rehabilitation maximizes gains for each patient. I propose that characterizing the memory and language profiles of individuals with moderate-severe TBI may be one fruitful means of addressing the heterogeneity in the population, allowing for maximal response to cognitive interventions.

Notably, evidence from the Cicerone and colleagues reviews suggest that cognitive rehabilitation continues to be effective in the post-acute period, including several years post-injury (Cicerone et al., 2011). This supports the utility of evaluating memory and learning profiles in the post-acute state of recovery, the current study's aim.

Multiple Memory Systems and Traumatic Brain Injury

Consideration of memory and learning abilities in rehabilitation decision-making in TBI is not new. Indeed, memory and learning deficits are a frequent consequence of TBI (Vakil, 2005) and are the most commonly recognized and treated deficit (Murray, Ramage, & Hopper, 2001; Wilson, 1998). However, untested assumptions about memory and learning have permeated the design and development of interventions for individuals with TBI. For example, a common assertion in clinical texts is that *declarative* (episodic) memory is highly vulnerable to impairment following TBI whereas *non-declarative* (procedural) memory is intact, or relatively preserved (Levin, 1990; Wilson, 2009). This assertion, as it relates to declarative memory, likely stems from how easily moderate-to-severe deficits can be detected (e.g., failure to learn a therapist's name or recall events from earlier in the day). Frank impairments are also easy to detect with current clinical measures (e.g., Auditory Verbal Learning Test free recall). The focus on declarative memory following TBI is likely the product of its obvious impact on daily life, and our ability to measure it and clearly demonstrate impairments. Declarative memory impairments also reflect frequently occurring pathophysiological consequences of TBI: hypoxia and seizure activity are common secondary effects of TBI, which disproportionally affect the volume of the hippocampus (Chang et al., 2009). In fact, structure of the hippocampus is vulnerable even in mild TBI (Leh et al., 2017).

In contrast, while non-declarative (procedural) memory has received significantly less study and is not routinely assessed in the clinic, it is often presumed to be preserved compared to declarative memory (Vakil, Biederman, Liran, Groswasser, & Aberbuch, 1994; Watt, Shores, & Kinoshita, 1999). That non-declarative memory is spared in TBI and can be leveraged in rehabilitation serves as the foundation of a number of therapy approaches (e.g., errorless

learning; Glisky, 1993). These approaches seek to compensate for declarative memory deficits but have also been recommended as techniques for treating other abilities (e.g., facial affect recognition, Radice-Neumann et al., 2007).

The problem with the assertion that declarative memory is impaired in TBI while non-declarative memory is spared is that it does not hold true for the population as a whole. Consistent with other work (Kraus, Little, Wojtowicz, & Sweeney, 2010; Vakil, Kraus, Bor, & Groswasser, 2002), preliminary examination of three different procedural memory tasks from our lab suggests that this form of memory is not uniformly intact in TBI: 61% of participants with moderate-severe TBI were impaired on at least one procedural memory task (Crooks et al., 2015). These deficits are compatible with physiological effects of TBI. The white matter tracts within and around structures supporting procedural memory (e.g., basal ganglia) are susceptible to shearing forces and diffuse axonal injury. Indeed, motor speech disorders such as dysarthria are common in TBI (30% of individuals with TBI have a form of dysarthria) (Duffy, 2013), and many are the product of injury to the same neural systems (BG) that support procedural memory (Reber et al., 1996).

The commonly held belief that procedural memory is preserved in TBI has likely created a blind spot for researchers and clinicians when considering why a patient does not respond to treatment. For example, one explanation for poor response to treatment is that there is a *mismatch* between the memory and learning system the intervention was designed to engage and the memory and learning systems available to the patient. If an intervention to treat facial affect recognition is developed to leverage intact procedural memory but a given patient has deficits in that form of memory, the intervention may fail; not because the intervention for facial affect

recognition is not effective but because the patient did not have the underlying capacity to learn from that technique.

To date, there has been no systematic study of procedural memory in TBI. While there are a number of experimental procedural memory tasks in the literature, a battery of tasks have never been administered to a single sample of individuals with TBI. There is little data regarding the utility of a given task to detecting procedural memory ability or impairment in TBI or the relationships among tasks. In fact, the use and reliability of experimental procedural memory tasks in clinical populations more broadly is understudied. Thus, determining which procedural memory tasks are sensitive to measuring procedural memory in a TBI population, as part of the long-term goal of developing a comprehensive memory and learning profile or phenotype, represents the most significant challenge and undertaking of the current work.

Taken together, the literature suggests that all forms of long-term memory can be disrupted in TBI. Given the hallmark heterogeneity in TBI, it also stands to reason that the severity of deficit across distinct memory systems will also vary both within and across individuals. Yet, deficits in any one form of memory and learning, and of any severity, threaten rehabilitation of knowledge and skills across broad domains of functioning.

There have been no attempts to measure the behavioral integrity of distinct memory systems within an individual to create an overall memory and learning profile or phenotype. Yet, the development of memory phenotypes could represent a significant breakthrough in addressing challenges associated with improving responsiveness to treatment in TBI. If we could establish a relation between individual phenotypes and long-term outcome, then the phenotype may serve as a biomarker for treatment success or predicting independence and reintegration.

In sum, all memory systems are vulnerable in TBI. However, empirical study and clinical assessment of some memory systems (declarative) have been privileged over others (procedural). We also have very little information about distinct patterns of spared and impaired memory abilities within individuals. The heterogeneity that is hallmark in cognitive profiles in TBI likely extends to variable disruption across memory systems. I propose that an individual's memory and learning profile may be a critical factor in determining the appropriateness and success of a given intervention, is the key to understanding mechanisms of change in rehabilitation, and is linked to a range of real world abilities critical for societal reintegration. The current project represents the first steps in demonstrating these relations in the same sample of individuals with TBI.

The Current Study

I propose that a helpful framework for classifying TBI into categories useful for behavioral health professionals (e.g. speech-language pathology, occupational therapy, physical therapy) is through consideration of patterns of impaired and intact learning and memory systems from a multiple memory systems framework. Existing treatment techniques across allied health fields differ in which memory system they are likely to engage. For example, treatments involving didactic teaching of explicit strategies (e.g. clear speech strategies for dysarthria) likely tap into the declarative memory system. In contrast, errorless learning-based treatments (e.g. for memory or word-retrieval) likely tap into procedural memory processes.

A multiple memory systems framework for classifying impairments in TBI has several benefits. At the root of this framework is the understanding that any memory dysfunction likely impacts all behavioral treatment efforts. That is, the success of any behavioral intervention that

requires (re)learning new information, a new strategy, or new skill may be impacted by the presence of memory impairment. If this is the case, treatment decisions should be made with these considerations in mind. Memory itself is the most frequently persisting complaint in individuals with TBI (Paniak et al., 2002). But beyond the utility of characterizing memory profiles for addressing memory complaints themselves, a characterization of a patient's memory and learning strengths and weakness bears directly on all other aspects of their post-acute rehabilitation; most of which involves engaging in behavioral therapies. These behavioral therapies (whether in physical, occupational, or speech-language therapy) each require that patients learn and retain new or previously-learned information. Thus, reliable and valid characterization of memory profiles would provide wide-reaching benefits to patients and allied healthcare providers in choosing appropriate interventions. Development of well-validated measures of multiple memory systems profiles has the added advantage of potential to be administered in other clinical populations, as the importance of preserved memory and learning capabilities for successful behavior-based rehabilitation is unlikely to be limited solely to the TBI population.

Importantly, it should be noted that the current project represents a first step in accounting for heterogeneity in TBI patients. As described above, memory and learning abilities have been selected for this first attempt at investigating individual differences in the population of patients with TBI because I believe that this framework may have significant implications for treatment success. However, I do not propose that memory and learning profiles are the only significant predictors of treatment success. Future work should expand into other cognitive, psychosocial, and motivational domains that likely play a unique role in treatment effectiveness.

The research questions to be addressed by the current study include the following:

1a. Are individuals with moderate-severe TBI impaired on the Spatial Reconstruction Task relative to healthy comparison participants?

I predict that individuals with moderate-severe TBI will be impaired relative to comparison participants.

1b. Does the Spatial Reconstruction Task have increased sensitivity in detecting and characterizing declarative memory ability over traditional neuropsychological assessment in individuals with moderate-severe TBI?

I predict that more individuals with TBI will be impaired on the SR task than on the declarative memory assessment in the NIH Toolbox Cognitive Battery. If this prediction is correct, it would suggest that the SR task is able to detect a range of declarative memory deficits in TBI, including subtle impairments, and should be included in a larger memory-systems battery for the development of memory phenotypes in TBI.

2a. Are individuals with TBI impaired on a battery of experimental procedural memory tasks relative to healthy comparison participants?

I predict that individuals with moderate-severe TBI will be impaired relative to comparison participants.

2b. Is there a relationship in performance across procedural memory tasks?

Significant correlations between individuals' performance on each procedural memory task would support the validity of these experimental measures in the measurement of procedural learning and memory. I predict that motor skill-based procedural memory tasks will be significantly correlated with one another, and that cognitive-perceptual skill learning will be significantly, but more weakly, correlated with motor-skill procedural memory tasks.

2c. Is there a relationship between the traditional, photoelectric implementation of the rotor pursuit task and a freely-available, computerized version of the task?

The rotor pursuit task is perhaps the most widely used measure of procedural memory across disorders. Traditionally, this task has been implemented using a specialized piece of equipment that records time-on-target using a photoelectric cell. The machine is bulky and expensive, impeding translation to clinical practice. A mouse-tracking version of the task has been available since 2012, but has not been directly tested against the traditional photoelectric version of the task. I predict that individual scores on the mouse-tracking version of the rotor pursuit task will be significantly correlated with scores on the photoelectric version.

CHAPTER II

METHODS

Participants

Participants were 25 individuals with moderate-severe TBI and 25 healthy comparison participants. Patients with TBI were recruited from the Vanderbilt Brain Injury Registry, from fliers posted in the greater Nashville area, and from outpatient clinics at Vanderbilt. Inclusionary criteria for TBI patients included: 1) moderate-severe TBI as determined by the Mayo Classification System (Malec et al., 2007); 2) chronic epoch of recovery (6 months-10 years post-injury); 3) 18-55 years of age. The 6 month criterion was set to ensure that participants are not experiencing post-traumatic amnesia, that edema and other acute processes have resolved, and that participants have passed the steepest trajectory of spontaneous recovery. Participants older than 55 were excluded to avoid the potential confounding influence of age-related cognitive changes. Children were excluded because of the interaction of cognitive development with recovery from TBI. All participants were screened for normal or corrected-to-normal vision. Patients with TBI were excluded if they had any other history of medical or neurological disease affecting the brain, or developmental or premorbid learning disability.

Available clinical data from medical records and patient interviews were used to determine whether participants sustained a moderate-severe TBI, as defined by the Mayo classification system (Malec et al., 2007). The Mayo classification system utilizes commonly reported clinical measures and indicators of TBI severity. The Mayo classification system allows for injury severity to be assigned even in cases of missing clinical data, by defining injury

severity as follows. Patients are classified as having sustained a moderate-severe TBI if one or more of the following conditions is met: 1) Loss of consciousness of 30 minutes or more; 2) Post-traumatic amnesia of 24 hours or greater; 3) Glasgow Coma Scale full score of less than 13 in the first 24 hours following injury; 4) The presence of any of the following: intracerebral hematoma, subdural hematoma, epidural hematoma, cerebral contusion, hemorrhagic contusion, penetrating injury (dura penetrated), subarachnoid hemorrhage, brain stem injury. If none of the previous criteria apply, the patient was not classified as moderate-severe and was not eligible for inclusion in the study.

Healthy comparison participants were matched to patients with TBI on sex, age, and years of education. Potential comparison participants were excluded if they reported any history of brain injury or concussion, any other history of medical or neurological disease affecting the brain, developmental disability, chronic medical condition known to impact memory or learning (e.g. diabetes, heart disease), or psychiatric disease.

Patients with TBI (16 female, 9 male) had an average age of 34.64 (SD = 9.29) and an average of 15.04 years of education (SD = 2.17). Healthy comparison participants (16 female, 9 male) had an average age of 35.92 (SD = 9.97) and an average of 15.76 years of education (SD = 1.94). Patients were in the chronic epoch of recovery, and were an average of 3.48 years post-injury (range: 1-10 years).

Clinical Profile

Available clinical data are presented in Table 1.

Participant	Time Post-Injury	Etiology	Clinical Imaging Findings	GCS	LOC	Retrograde amnesia	Post-traumatic amnesia	Occupation
5003	2 years	Pedestrian hit by car	SDH	12	N/A	1 week	3 weeks	Student/ Nutritionist
5011	5 years	Fall	SAH; Frontotemporal contusion; Epidural hematoma	N/A	> 30 mins	10 minutes	2 months	Enrolled in vocational rehab
5013	1 year	Pedestrian hit by car	SAH	15	None	Minutes	Minutes	Unemployed
5016	2 years	MVA	SAH; SDH	13	> 30 mins	Minutes	4 days	N/A
5018	10 years	MVA	Multiple areas of SAH	3	None	8 months	3 weeks	Nurse
5019	3 years	Pedestrian hit by car	SAH; SDH	6	> 30 mins	1 month	2 months	Unemployed
5020	6 years	Helmeted motorcycle accident	N/A	N/A	N/A	Minutes	1 month	Addiction counselor
5021	3 years	MVA	Epidural hematoma; SAH	3	> 30 mins	Hours	27 days	Unemployed
5027	1 year	Fall	Extensive SAH right hemisphere	9	N/A	Hours	1 month	Golf caddy
5028	2 years	MVA	SAH	6	> 30 mins	Hours	5 days	Student
5029	1 year	Unhelmeted bike hit by car	SDH; SAH	14	< 30 mins	Minutes	Hours	Nurse
5031	1 year	Struck by tree limb	SDH; SAH; Intraparenchymal	13	N/A	N/A	N/A	Antique dealer

			hemorrhage					
5034	3 years	MVA	N/A	3	N/A	Hours	2 months	Part-time work from home
5037	4 years	MVA	Diffuse intracranial swelling	3	N/A	1 week	21 days	Student
5038	2 years	Fall	SDH; Multifocal hemorrhages in frontal and temporal lobes; Hemorrhagic contusions	N/A	> 30 mins	None	3 days	Buyer for distribution company
5039	5 years	MVA	Intraventricular hemorrhage; Subcortical shear, Scattered SAH	3	N/A	1 week	28 days	N/A
5040	6 years	MVA	SDH; Scattered SAH	3	N/A	5 months	3 weeks	Assembly production
5041	5 years	MVA	Negative	10	None	None	1 week	Student
5044	6 years	Unhelmeted longboarding accident	SDH	N/A	None	Seconds	3 days	Heavy machinery operator
5046	4 years	Bike accident	SAH	14	> 30 mins	Seconds	24 hours	Asset manager
5047	2 years	Assault	SDH	15	> 30 mins	Half an hour	12 hours	IT
5053	1 year	Snowmobile accident	Intraparenchymal hemorrhage; SDH	5	> 30 mins	24 hours	2 weeks	Head camp counselor
5054	3 years	MVA	SDH	N/A	N/A	24 hours	Hours	Sells clothes online
5055	6 years	MVA	Hemorrhagic shear injury; Scattered SAH;	4	> 30 mins	Minutes	Several days	Contractor

			SDH					
5056	3 years	Unhelmeted skateboarding accident	Hemorrhagic shear injury	11	> 30 mins	24 hours	2 weeks	Rehab tech

Table 1. Clinical Characteristics of Patients with TBI

Standardized Cognitive Assessment

Wide Range Achievement Test (WRAT).

The Wide Range Achievement Test is a standardized measure of basic academic skills. The Word Reading subtest of the WRAT was implemented as a measure of participant's exposure to low-frequency words, word decoding ability, and word recognition ability. Participants were asked to read a list of 55 words out loud and were scored on whether each word was accurately pronounced. Words increase in difficulty across the task (e.g. from "see" to "synecdoche").

NIH Toolbox: Cognition Battery.

The NIH Toolbox Cognition Battery was implemented to characterize participants' general cognitive profiles (Weintraub et al., 2013). The Cognition Battery is comprised of seven different tasks. Constructs tested include: declarative memory, executive functioning, attention, working memory, vocabulary, and processing speed. Each task in the Cognition Battery is described below:

Picture Sequence Memory Test: The Picture Sequence Memory Test measures declarative memory. Participants view sequences of pictures and then reproduce these sequences on the screen.

Flanker Inhibitory Control and Attention Test: The Flanker task measures attention and executive functioning. Participants are asked to attend to a particular stimulus while inhibiting responses to adjacent stimuli.

List Sorting Working Memory Test: The List Sorting task measures working memory. Stimuli are presented orally or visually. Participants are asked to recall these stimuli and order them.

Picture Vocabulary Test: The Picture Vocabulary test measures receptive vocabulary. The task is in Computer Adaptive Test (CAT) format. Participants select the picture that most closely matches the meaning of a provided word.

Dimension Change Card Sort Test: The Dimension Change Card Sort task measures executive function and attention. Participants sort cards that vary in two dimensions (color and shape). The screen indicates by which dimension participants should sort the cards. The sorting dimension changes throughout the task.

Pattern Comparison Processing Speed Test: The Pattern Comparison Processing Speed task measures processing speed. Participants are asked to determine whether two side-by-side pictures are identical or not. Participants attempt to respond to ask many picture pairs as possible over 85 seconds. The two pictures are not complex in order to best measure processing speed.

Declarative Memory

Spatial Reconstruction.

The Spatial Reconstruction task is an experimental task of declarative (relational) memory. This task was chosen for this study because of its ease of administration (e.g., could be used as a quick bedside assessment) and its promise of having sensitivity to detecting a range of declarative memory ability and impairment. A tablet-based version of the Spatial Reconstruction task, similar to that described in (Clark et al., 2017) was administered. The task consists of three practice trials, followed by 25 trials in which participants were asked to remember the locations of novel objects on the screen (see Figure 1).

Figure 1. Example Spatial Reconstruction Display

Each test trial consisted of a study phase followed by a placement phase. The number of items to be studied in each trial (set size) varied from two to ten in multiples of two. For the study phase of each trial, participants were allotted three seconds of study time per item. Thus, depending on the set size, the study phase varied in duration from 6 to 30 seconds. Following the study phase, the studied novel objects disappeared from the screen for four seconds, and then reappeared at

the top of the screen. Participants then attempted to replace the novel objects as close as possible to their original, studied locations by clicking and dragging them using a stylus.

The methods for determining participants' placement accuracy are based on point set registration methods. The process by which error metrics are determined is described in Horecka et al., 2018. Briefly, the accuracy of participants' placements are assessed based both on their distance from each object's actual studied location (with placements near enough to an object's studied location deemed an "exact hit") and based on their distance from *any* previously studied location (that is, irrespective of object identity). I will refer to these locations as "valid locations." In this framework, if object A had been studied in location A, and the participant places object A in location A, they have made an "exact hit". If they instead place object A in location B (the previously studied location of object B), they have placed object A in a "valid location" but have not made an "exact hit". The Euclidean distance between an object's placement by the participant and a studied location is its "misplacement," measured in pixels. These metrics are assessed in two ways: on participants' "pre-processed" data (participants' actual placements are compared directly to studied locations) and on data that has undergone "global transformation" on a by-participant basis (participants' placement undergo a set of global transformations, if such transformations improve the degree to which participant placements correspond to studied locations). These global error corrections (in which global translation, scaling, and rotation errors are subtracted from the placed item locations) address noise at the individual participant level that may not reflect their declarative memory ability. A schematic illustrating how global transformations are performed is presented in Figure 2.

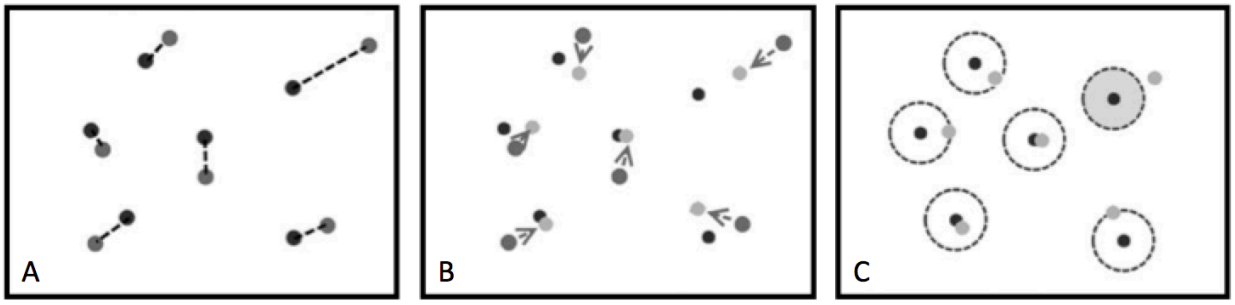

Figure 2. Global Transformations and Accuracy Windows. Panel A demonstrates mapping of participant places to valid locations on identity-stripped data. Dark dots indicate studied locations and lighter dots represent participant placements. Each participant placement is mapped to the closet valid location. Panel B illustrates global transformations. Here, participant placements have been slightly rotated and translated to reduce global error. Dark dots indicate studied locations, medium dots indicate participant placements, and the lightest dots indicate participant placements post-transform, after global transform error has been removed. Panel C illustrates determination of accuracy on post-transformed data. If a placement falls within an accuracy circle, it is counted as a placement to a valid location.

The accuracy of each of a participant's placements is determined based on whether their placement is within a calculated "accuracy window." To determine placement accuracy, an accuracy circle was computed based on aggregated data within a participant across all trials. The radius of the accuracy circle varied across individuals based on a particular individual's variability in local misplacement (that is, misplacement that is not shared across items or caused by item-identity errors). Thus, each individual's statistical distribution of misplacement, aggregated across all trials (following item-identity and global error corrections), was used to determine 95% confidence intervals that represented that participant's individual accuracy threshold. One critical decision in setting these accuracy windows is whether to determine accuracy windows on a by-participant basis for both healthy comparison participants and individuals with TBI, or whether to use matched healthy participant accuracy windows for the patients with TBI (as was done in Horecka et al., 2018 for amnesic patients). Allowing calculation of individual accuracy windows (irrespective of diagnostic group; NC vs. TBI) rests

on an assumption that participants' item placements will generally be in the studied locations. As noted by Horecka and colleagues, this assumption may not hold for patients compared to healthy adults. In comparing SR performance between healthy adults and patients with profound hippocampal amnesia, Horecka and colleagues opted to equate accuracy windows across groups by assigning each patient participant the accuracy window calculated on their matched comparison's data. Here, however, we opt to calculate accuracy windows within individual participants, irrespective of diagnostic group. We do so for the following reasons: 1) basing accuracy windows on a single matched comparison participant does not seem reasonable (it would be better to average across multiple comparisons with equivalent demographics, to get "normative" accuracy windows and then apply these norms to a new set of comparisons and patients); and 2) it is the more conservative approach: all subsequent analyses are based on these accuracy windows. If it is the case that TBI patients *are* more variable overall in their placements, leading to wider accuracy windows, then any group differences in other error metrics (e.g., number of exact hits, amount of misplacement) will index impairments above and beyond this generally more significant "noisiness" in the patient data.

To summarize, participants' placements will be compared to the original studied locations, and the following metrics will be assessed: 1) the amount of misplacement (distance in pixels) between participants' placements and studied locations, on data both pre- and post-transform; 2) the number of exact hits, on both pre-processed and globally transformed data; and 3) the number of placements to valid locations, irrespective of object identity. Metric 1 (pre-processed) and 2 (both pre- and post-transform) index participants' ability to bind object identity to its studied location, while Metric 2 (post-transform) and Metric 3 indexes participants' ability to recreate the gestalt relations between studied locations.

Procedural Memory

Photoelectric Rotor Pursuit.

The rotor pursuit task is a procedural learning task that requires participants to use a wand or stylus to track the circular movement of a turntable-like disk while keeping the tip of the stylus on a lighted target on the disk. The dependent measure is the amount of "time on target", which is automatically recorded by the rotor pursuit device. Evidence of procedural learning comes from increases in time-on-target across trials. This task was chosen because of its long history of use in the memory and amnesia literatures and in clinical populations.

To equate task difficulty for participants with differing baseline motor abilities, baseline testing was conducted to determine the speed with which the target would move during test trials. Participants completed 4 baseline trials at each of the following speeds: 15 rotations per minute (rpm), 30 rpm, 45 rpm, and 60 rpm. The speed for which the participant's time-on-target was closest to 5 seconds was their individual "baseline". All test trials were conducted at the individual participant's established baseline speed.

The rotor pursuit task consisted of 16 total trials, divided into two sets that were separated by intervening tasks (see Procedure below). Each set of 8 trials was divided into two 4-trial blocks, with a short (1-minute) break between blocks. On each trial, the target spun at a consistent speed set based on the participant's baseline testing for 20 seconds. After the first 20-second trial, the target paused for 8 seconds, and then started again automatically for a second 20-second trial. After the fourth trial, the machine was reset for the next 4-trial block. A photoelectric rotor pursuit machine (Lafayette Instruments) was utilized to capture participants' time-on-target on each trial. Participants held a stylus in their preferred hand and were instructed

that the goal of the task was to keep the stylus in contact with a white (lighted) target on the rotor pursuit machine, as it spun. The rotor pursuit machine recorded amount of time the stylus was in contact with the target for each 20 second trial.

Mouse-Tracking Rotor Pursuit.

A mouse-tracking computerized version of the rotor pursuit task was chosen for this study to determine its relationship to the traditional, photoelectric implementation of the rotor pursuit task described above. The mouse-tracking computerized version of the rotor pursuit task required participants to attempt to keep their mouse cursor directly over a red target as it moved around a larger gray circle. The dependent measure was the amount of "time-on-target" for each trial. In contrast to the photoelectric rotor pursuit task, no baseline testing was completed. All participants completed the task at the same speed: 0.13 rotations per second. Participants completed a total of 16 trials, divided into two conditions: Feedback and No Feedback. In the Feedback condition, the red target was illuminated when the participant's mouse cursor was directly over top of the target, and dimmed when they were not in contact with the target. In the No Feedback condition, the luminance of the red target did not change based on the participant's accuracy. The Feedback version of the task is part of an open-source psychological test battery called Psychology Experiment Building Language (PEBL) (Mueller, 2012; Mueller & Piper, 2014). The No Feedback version was created by modifying this code to more closely approximate the photoelectric rotor pursuit task (where participants do not receive online feedback during the task). Participants completed trials in 4-trial sets.

Mirror-Reversed Reading.

The Mirror-Reversed Reading (MRR) task was selected in order to assess procedural memory ability across all potential subdomains of the construct: while the rotor pursuit tasks examine motor skill learning, the MRR task assesses the ability of participants to acquire a new cognitive-perceptual skill. The MRR task was developed based on Cohen & Squire (1980). On each trial, three low-frequency words (e.g. "ambition insistence prevalent", see Figure 3) that have been mirror-reversed were displayed on a computer screen. Stimulus words are low in frequency, high in familiarity, three syllables in length, and comprised of 8-10 graphemes. The experiment is divided into 10 blocks, with each experimental block containing 10 "repeating" word triads and 10 "novel" word triads.

Figure 3. Example Mirror-Reversed Reading Trial

Participants were asked to read each mirror-reversed word triad aloud and press a button after reading the final word. The task consisted of two separate sessions, with approximately 30 minutes of intervening tasks between sessions. In each session, participants read five blocks of twenty words triads.

Stimuli creation.

Stimulus word selection.

Consistent with Cohen & Squire (1980), word stimuli for the MRR task were selected to be low frequency (to ensure that the task was difficult enough to require skill learning/avoid automatic recognition of stimulus words), three syllables in length, and comprised of eight to ten

letters. In addition, because participants would have varying levels of educational attainment and low-frequency words also tend to be low in familiarity, an additional constraint of relatively-high familiarity was imposed.

Thus, stimulus items were chosen in the following way. An initial candidate set of words was drawn from the MRC Psycholinguistic Database (Coltheart, 1981), with filters set for words that had: 1) 8-10 letters; 2) 3 syllables; 3) a maximum Kucera-Francis written frequency of 50 (Kucera & Francis, 1967); and 4) a minimum familiarity rating of 200. Familiarity ratings in the MRC database come from a merged sets of familiarity norms and range from 100 to 700 (M = 488, SD = 99) (Gilhooly & Logie, 1980; Pavio, Yuille, & Madigan, 1968; Toglia & Battig, 1978). This initial set of candidate stimulus words was then run through the SUBTLEXus database, a newer set of word frequency norms that better predicts word processing time than older normative sets, including Kucera and Francis. Using these word frequency norms, the candidate stimulus list was reduced to words with a $SUBTL_{WF}$ (word frequency per million words) of less than 10. Next, the candidate list was sorted by familiarity, and the 330 words with the highest familiarity were selected.

Task development.

During a pilot run of the MRR task, it was discovered that a small subset of stimulus items (n = 22) were inaccurately labeled in the MRC database as containing 3 syllables, when in fact they were comprised of 2 or 4. Thus, these 22 items were removed and replaced with the next 22 candidate stimulus words that were high in familiarity but low in frequency.

A random subset of 30 words from the original 330 were assigned to the *repeat* condition (to be repeated in each experimental block); the remaining 300 were assigned to the *novel* condition (to be encountered only once during the entire experiment). There were no significant

differences in word frequency (repeat M = 2.29, novel M = 2.66, t(35) = -0.79, p = 0.43), word familiarity (repeat M = 461.90, novel M = 462.95, t(35) = -0.10, p = 0.92), or number of letters (repeat M = 8.90, novel M = 8.87, t(35) = 0.18, p = 0.86) between the two conditions.

Next, stimulus words within conditions were assigned to word triads using the *minDiff* package in R, to minimize differences between triads (Papenberg, 2018). For each triad, the average word frequency and word familiarity was calculated. Word triads were then assigned to blocks, again using *minDiff* to minimize differences in average word frequency and familiarity across blocks. Across blocks, there were no significant differences in average word frequency ($F(1,108) = 0.02$, $p = 0.89$) or average familiarity ($F(1,108) = 0.07$, $p = 0.79$). Average word frequency and familiarity by block are reported in Table 2.

Condition	Block	Average Word Frequency (SD)	Average Familiarity Rating (SD)
Repeat	1-10	2.29 (0.57)	462 (16.7)
Novel	1	2.80 (1.08)	462 (18.5)
Novel	2	2.57 (1.02)	467 (25.5)
Novel	3	2.71 (0.93)	467 (20.2)
Novel	4	2.67 (1.29)	459 (23.5)
Novel	5	2.67 (1.33)	464 (30.4)
Novel	6	2.77 (0.58)	463 (28.3)
Novel	7	2.57 (0.94)	465 (21.9)
Novel	8	2.75 (1.51)	461 (35.6)
Novel	9	2.47 (0.89)	460 (34.1)
Novel	10	2.57 (1.04)	463 (28.8)

Table 2. Average Word Frequency and Familiarity by Block

Serial Interception Sequence Learning.

The Serial Interception Sequence Learning (SISL) task is an adaptation of the widely-used serial reaction time task (SRT), that examines procedural sequence learning over time. The SISL task was chosen for this study because of its similarity to the SRT task, and because it offers a number of advantages compared to SRT. The SISL task has been shown to be a more

"process-pure" procedural memory task (Sanchez, Gobel, & Reber, 2010). The task adaptively adjusts in speed, allowing for testing across varying motor abilities. Finally, the SISL task is more engaging than SRT, and pilot testing suggested it would be better tolerated by patients with TBI. A three-button version of the SISL task similar to (Gobel et al., 2013) was administered. Participants made responses during the task using the index, middle, and ring finger of their dominant hand on a mechanical gaming keypad. Participants were asked to push the corresponding button on their keypad when a blue circle (the cue) entered one of three target zones at the bottom of the computer screen (see Figure 4).

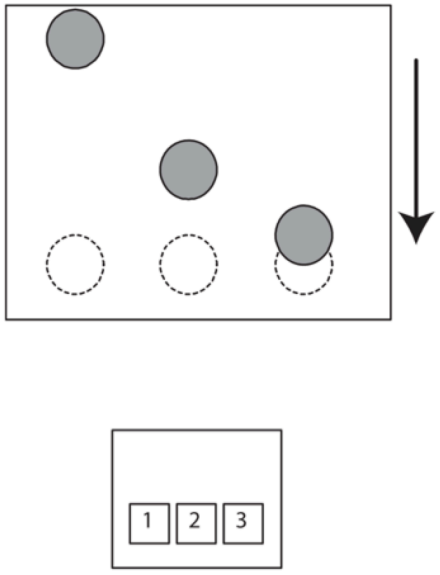

Figure 4. Example Serial Interception Sequence Learning Display

As demonstrated in Figure 4, during the task multiple cues are present and moving down the screen simultaneously. Participants responded to the cue closest to the bottom of the computer screen, timing their key press to the moment when the cue entered the target zone. Participants

received immediate feedback on their performance: 1) If the correct button was pressed as the cue entered the target, the cue turned green; and 2) If the wrong button was pressed, or it was pressed before or after the cue entered the target, the target turned red. Cues and target zones were 1.96 cm in height, and cues moved 11.8 cm from the top of the screen to the target zone.

The speed of the moving cues adaptively adjusted to participants' accuracy during the first four blocks of the task (training). Initial cue speed was set to 5.9 cm/sec (2 sec of travel time). Cue speed was then adaptively adjusted, holding performance between 75 and 92% accuracy in the following way. After each 12 trials, cue speed was increased by 5% if the participant had been 100% accurate. If the participant made three or more errors (< 75% accuracy), cue speed was reduced by 5%. Cue speed remained constant if the participant made between one and three errors.

Embedded within the task was a repeating sequence of 12 items. The repeating sequence was the same for all participants: 2-1-2-3-2-1-3-1-3-2-3-1. In addition, the relative timing of the cues was consistent across all participants. That is, half of the intercue intervals were short (initially 350 ms) and half were twice as long (700 ms). The timing sequence for patterned trials was: 0.35, 0.35, 0.70, 0.70, 0.35, 0.70, 0.70, 0.35, 0.35, 0.70, 0.70, 0.35. As the speed of the task changed adaptively, these intervals were adjusted relative to the overall speed, such that the ratio was maintained. During training trials, 20% of trials were sequences that did not share the cue order or timing structure of the repeating sequence. For each block of 60 trials, the target sequence was presented four times, and the novel 12-item sequence was randomly inserted between repetitions. After the novel sequence, the starting location of the target sequence was randomly selected, reducing the amount of explicit sequence knowledge acquired during practice.

The SISL task is comprised of 6 blocks of 360 trials each. The first four blocks comprised the "training" phase. The 1,440 training trails contained 96 repetitions of the target sequence and 24 novel 12-item non-repeating sequences. Participants were not informed of the underlying pattern embedded in the task and were given short breaks after each 360-trial block to minimize fatigue. Blocks five and six comprised the "implicit test" phase. Participants continued with the task as before and were not informed of any change in the task between blocks four and five. During the implicit test blocks, the relative proportion of target sequence repetitions and foil sequences changes, with 40 repetitions of the target sequence and 20 repetitions of each of a foil sequence. During the implicit test blocks, the speed set in the last training block was maintained.

Procedure

The WRAT and the NIH toolbox were obtained during separate sessions. Participants with TBI completed experimental memory tasks across two sessions. The spatial reconstruction task was implemented during the first session. During the second session, patients with TBI completed the remaining tasks in the following order: 1) WRAT Reading subtest; 2) photoelectric rotor pursuit trials 1-8; 3) mirror-reversed reading blocks 1-5; 4) serial interception sequence learning; 5) photoelectric rotor pursuit trials 9-16; 6) mirror-reversed reading blocks 6-10; 7) mouse-tracking rotor pursuit. Healthy comparison participants completed experimental memory tasks over either one or two sessions. Tasks were completed in the same order as in patients with TBI; if all tasks were administered in the same session, spatial reconstruction preceded the WRAT Reading subtest.

CHAPTER III

RESULTS

Overall Cognitive Profile

Data from the NIH Toolbox tasks were used to assess each group's overall cognitive profile. NIH Toolbox data was collected for 22 patients with TBI and 18 healthy comparison participants. All analyses were performed on standard scores for each subtest, and are presented in Table 3. Although patients' scores were lower than healthy comparison participants, few patients scored lower than a standard score of 85, one standard deviation below the normative mean, on any subtest.

Subtest	TBI Mean (SD)	NC mean (SD)	t statistic	p-value	# of TBI patients > 1 SD below normative mean
Executive Function	102 (9.23)	108.94 (7.80)	-2.58	0.01	1
Attention	93.18 (9.48)	100.17 (7.04)	-2.67	0.01	3
Working Memory	103.64 (12.77)	113.28 (7.86)	-2.93	0.006	1
Processing Speed	100.41 (19.38)	111.56 (19.19)	-1.87	0.07	5
Declarative Memory	108.68 (15.44)	116.33 (13.63)	-1.66	0.10	2
Vocabulary	108.55 (6.49)	112.28 (6.93)	-1.74	0.09	0

Table 3. NIH Toolbox Standard Scores

Declarative Memory

Spatial Reconstruction Task.

The pre-processed accuracy windows, computed on patients' across-trial data with no identity or transformation corrections, were significantly larger than those for healthy comparison participants ($t(40.85) = 3.06$, $p = 0.004$). When accuracy windows were computed with identity-stripped inputs, they were not significantly different between groups ($t(46.54) = 1.97$, $p = 0.055$). Patients and comparison participants did not significantly differ in any aspect (translation, scaling, rotation) in which their placements were transformed in the global transformation stage of processing (see Figure 5).

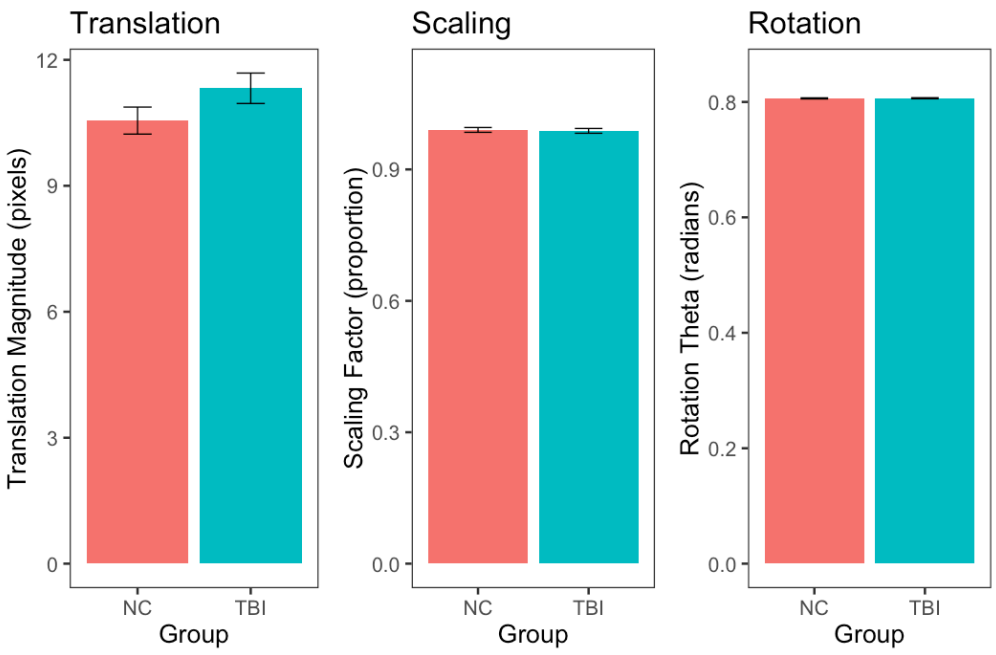

Figure 5. Global Transformation by Group.

To test Research Question 1a, patients with TBI were compared to healthy comparison participants on the following metrics: 1) the amount of misplacement (distance in pixels) between participants' placements and studied locations, on data both pre- and post-transform; 2) the number of exact hits, on both pre-processed and globally transformed data; and 3) the number of placements to valid locations, irrespective of object identity. Metric 1 (pre-processed) and 2 (both pre- and post-transform) index participants' ability to bind object identity to its studied location, while Metric 2 (post-transform) and Metric 3 indexes participants' ability to recreate the gestalt relations between studied locations.

As a group, patients with TBI demonstrated significantly more misplacement (sum of the Euclidean distance between each object placement and its studied location, TBI mean = 46.76 pixels, SD = 12.61) compared to healthy comparison participants (mean = 37.27, SD = 7.76; $t(39.88) = 3.21$, $p = 0.003$, Figure 6). When misplacement was calculated without regard for object identities, this group difference was maintained ($t(46.60) = 2.20$, $p = 0.03$), with patients with TBI placing objects an average of 26.08 pixels away from a valid location (SD = 5.12) and healthy participants placing objects an average of 23.13 pixels from a valid location (SD = 4.30, Figure 7). Finally, when misplacement was calculated without regard for object identities and after all global transforms (rotation, scaling, translation) had been performed, this group difference was maintained ($t(46.17) = 2.35$, $p = 0.02$), with the transformed object placements of patients with TBI an average of 23.63 pixels away from valid locations (SD = 4.76) and transformed object placements of healthy participants placing objects an average of 20.74 pixels from valid locations (SD = 3.89, Figure 8).

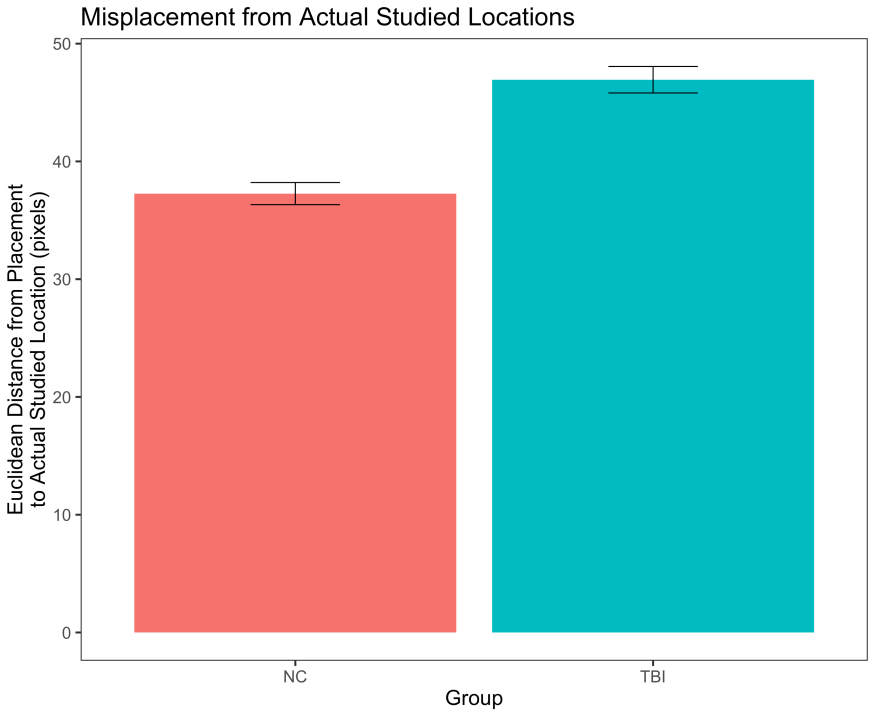

Figure 6. Average Misplacement Distance from Studied Locations

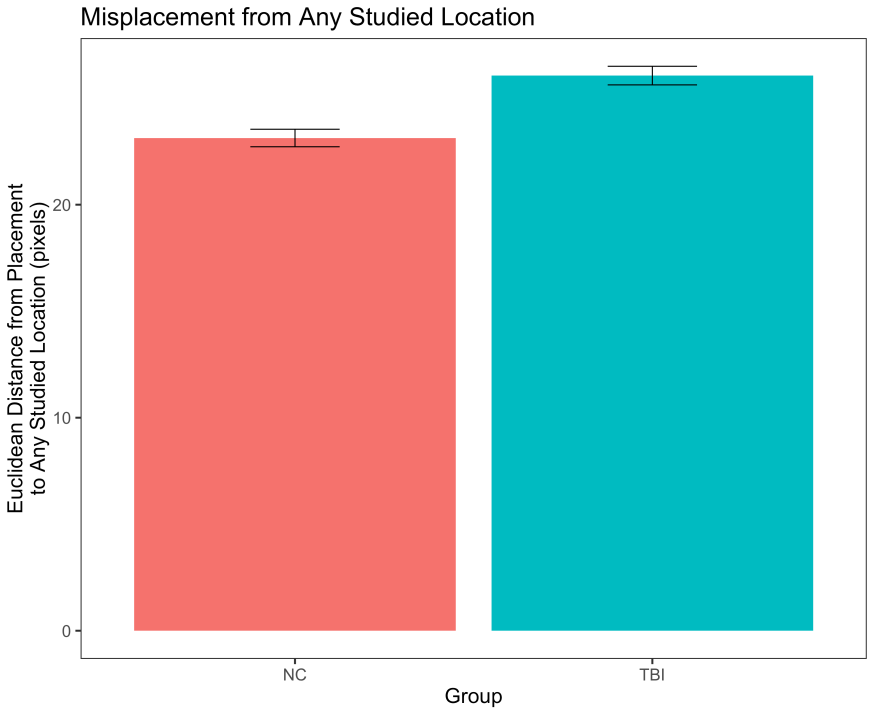

Figure 7. Average Misplacement Distance from any Valid Location

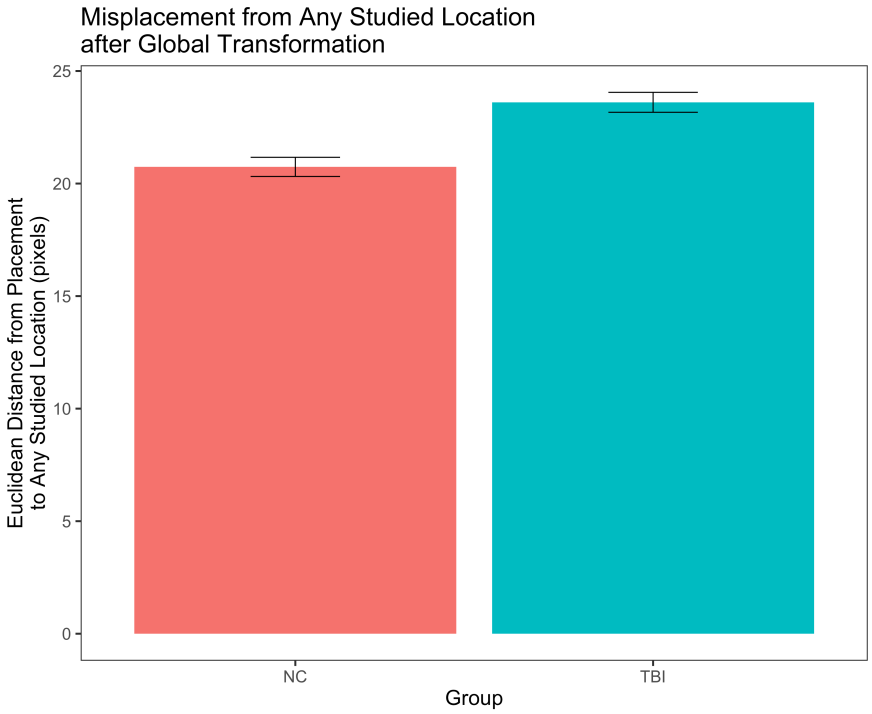

Figure 8. Average Misplacement Distance from Any Valid Location (Post-Transform)

The number of exact hits differed between groups, with patients with TBI making an average of 3.98 placements of the correct object within the accuracy window of its studied location, and healthy comparison participants making an average of 4.19 exact hits ($t(47.97) = -2.35$, $p = 0.02$). Patients with TBI made more placements of objects in another object's studied location than healthy comparison participants, even after all global transformations were applied ($t(42.29) = 3.04$, $p = 0.004$). There was no group difference in memory for locations (that is, the number of objects placed in any valid location) ($t(46.067) = -1.00$, $p = 0.32$).

The effect of set size on performance was analyzed with mixed effects models. Models were fit using the *lme4* package in R, with p-values obtained using the *lmerTest* package. The first set size analysis examined the effects of set size and group on participants' proportion of

exact hits on pre-transformed data (that is, the number of accurate placements divided by the set size for that trial, see Figure 9). An initial model included fixed effects of group and set size, and their interaction, random by-participant slopes for set size, and random by-participant and by-trial intercepts. The interaction of group and set size was not significant ($p = 0.68$) so the interaction term was removed from the model. All other fixed and random effects were maintained. There was a significant main effect of set size ($b = -0.05$, $p < 0.001$), indicating decreasing accuracy as set size increased and a significant main effect of group ($b = -0.04$, $p = 0.01$), with patients with TBI performing more poorly relative to comparison participants. A model with the same effects structure run on post-transform data included a significant main effect of set size ($b = -0.07$, $p < 0.001$) and a significant main effect of group ($b = -0.09$, $p = 0.003$, see Figure 10). When placements were considered without regard for object identity (that is, when accuracy was defined as a placement within a valid location, irrespective of object identity), the two groups did not differ significantly ($b = -0.005$, $p = 0.71$). There was a significant main effect of set size ($b = -0.03$, $p < 0.001$, see Figure 11). Full model results for all set sizes analyses are presented in Appendix A.

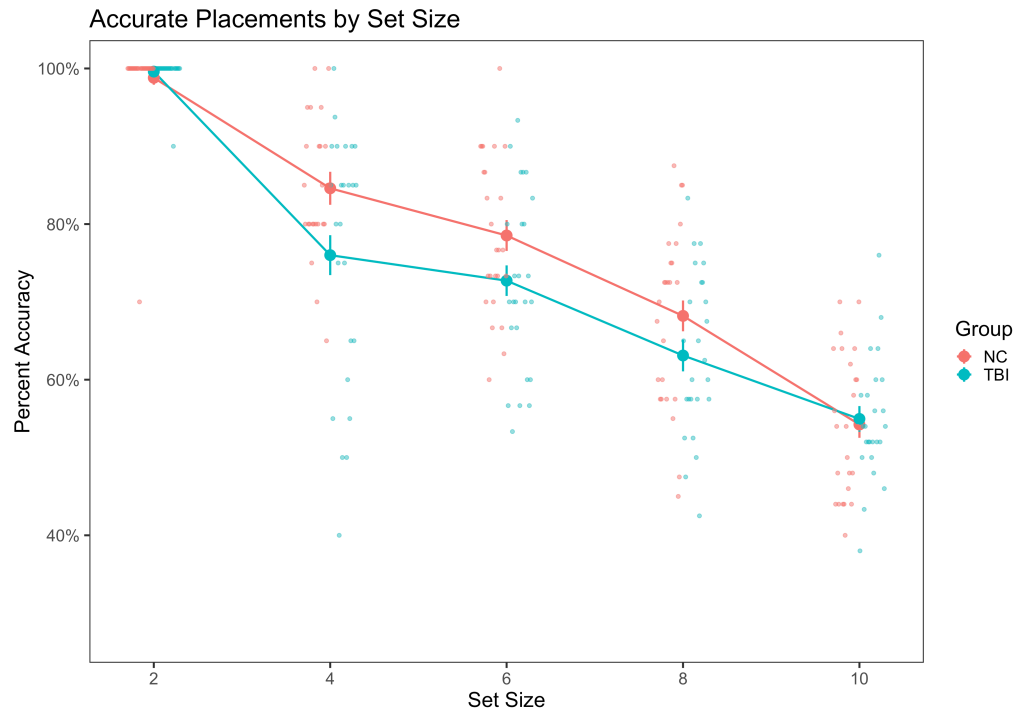

Figure 9. Average Proportion of Exact Hits by Set Size (Pre-Transform)

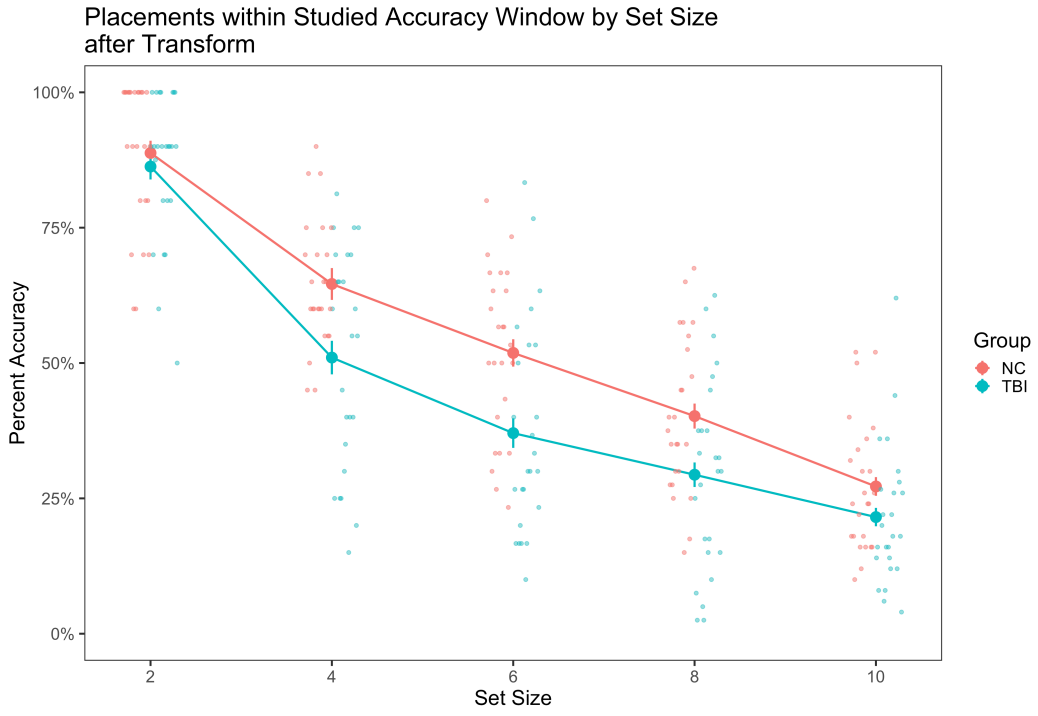

Figure 10. Average Proportion of Exact Hits by Set Size (Post-Transform)

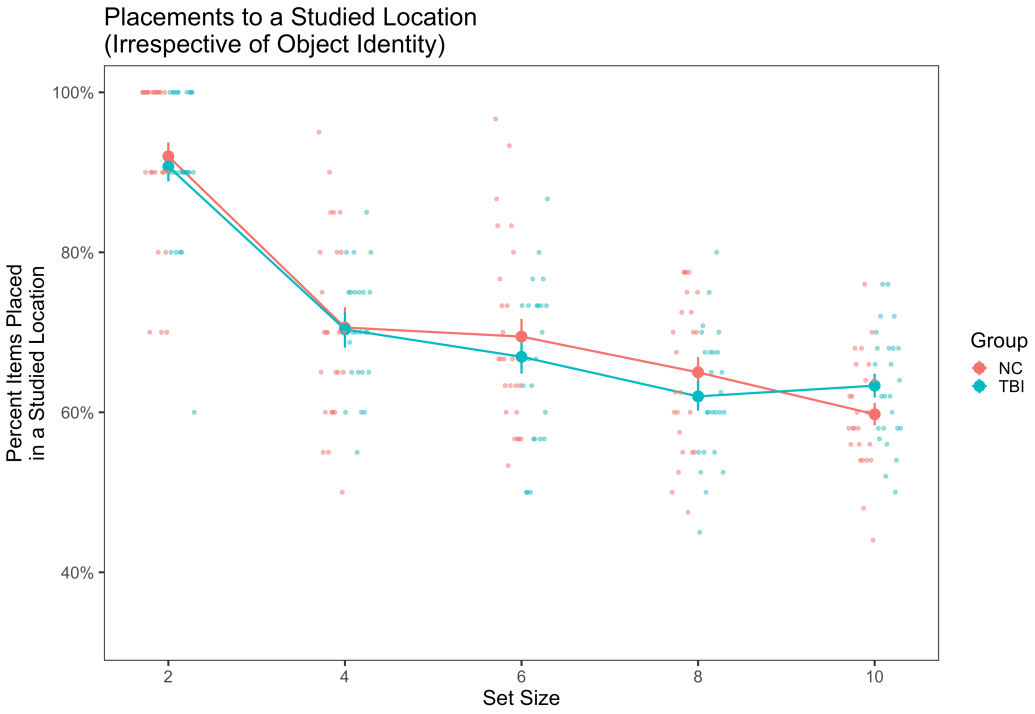

Figure 11. Average Proportion of Placements to any Valid Location (Post-Transform)

Consistent with predictions for Research Question 1a., patients with TBI were impaired on the spatial reconstruction task, relative to healthy comparison participants. This group difference is in contrast to patients' performance on the declarative memory subtest of the NIH Toolbox (see Overall Cognitive Profile, above), where patients did not significantly differ from healthy comparison participants, and all but two patients scored within a standard deviation of the normative mean. Notably, the two patients who did score more than a standard deviation below the mean did so by a single point in standard scores (both achieved a standard score of 84; 85 is one SD below the mean). Results support predictions for Research Question 1b., suggesting that the spatial reconstruction task may be a sensitive measure of declarative memory performance.

Procedural Memory

Photoelectric Rotor Pursuit.

For the photoelectric rotor pursuit task, data from 48 participants were analyzed. One participant, 5034, bumped a component off of the rotor pursuit machine during the task, causing loss of half of her data, and therefore her data (as well as data from her matched comparison participant, 1030) were excluded from analysis. As a group, patients with TBI completed the rotor pursuit task at a significantly slower baseline speed compared to healthy comparison participants ($t(41.34) = -2.27$, $p = 0.03$). After completing baseline trials, 12 (50%) of participants with TBI completed the task at the 60 rpm speed, 11 (45.83%) at 45 rpm, and 1 (4.17%) at 30 rpm. 19 (79.17%) healthy comparison participants completed the task at the 60 rpm speed, 5 (20.83%) completed the task at 45 rpm (see Figure 12). There was also an effect of sex, with male participants completing the task at significantly faster baselines than female participants ($t(45.91) = 3.27$, $p = 0.002$, Figure 13).

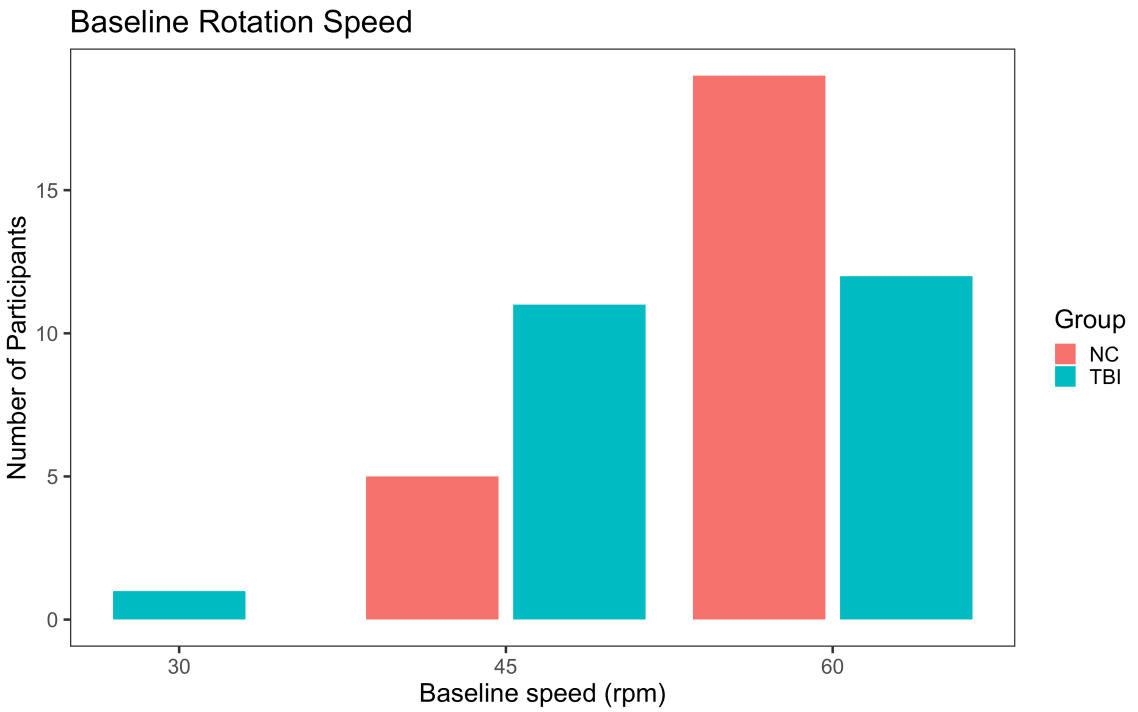

Figure 12. Baseline Rotation Speed by Group

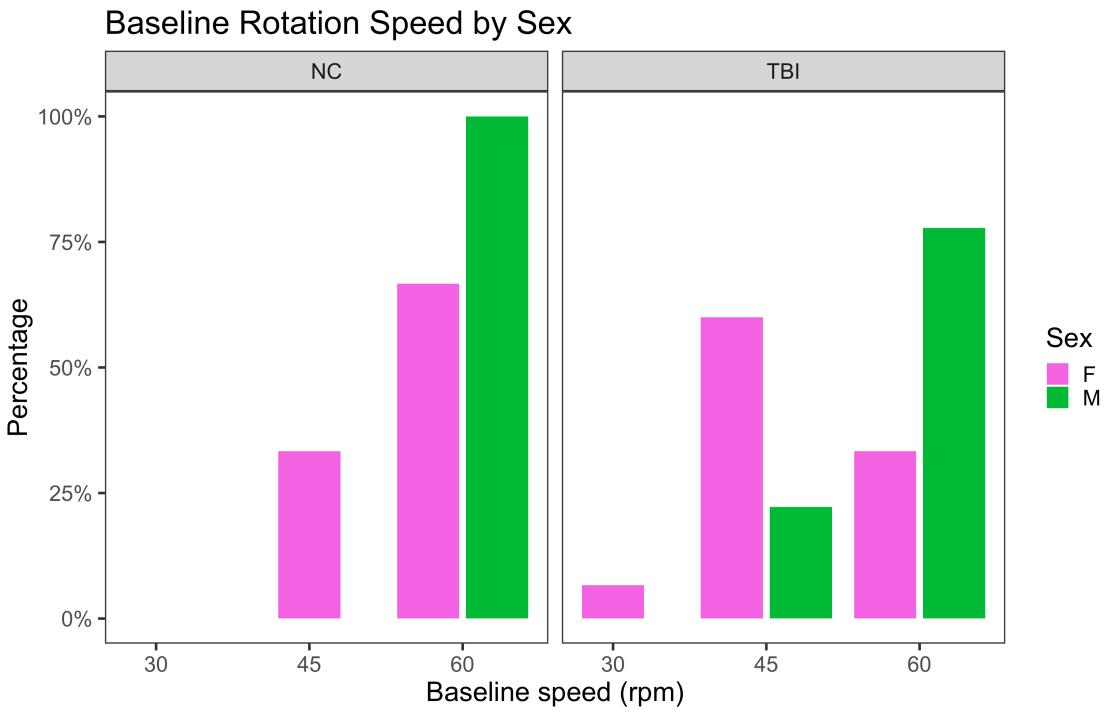

Figure 13. Baseline Rotation Speed by Sex

Average time-on-target for each trial for each group is presented in Figure 14. To facilitate comparisons with previous rotor pursuit experiments, the data are also plotted by experimental block (4 trials per block) in Figure 15. The two groups did not differ in their initial time-on-target performance during Block 1 (TBI mean = 6.32, SD = 2.24, NC mean = 6.55, SD = 2.14, $t(45.01) = -0.41$, $p = 0.68$). To examine group differences in procedural learning, a conditional growth curve model was fit to the trial-by-trial time-on-target data. Models were fit using the *lme4* package in R, with p-values obtained using the *lmerTest* package. Trials were re-coded such that the intercept of the model could be interpreted as the grand average initial time-on-target (trial 1 coded as 0, trial 2 as 1, etc.). A reduced model with fixed effects of trial, group, and their interaction and random by-participant intercepts revealed a significant trial-by-group interaction ($b = -0.06$, $p = 0.02$), in which patients with TBI demonstrated a reduced effect of trial (i.e. less growth over time) compared to healthy comparison participants.

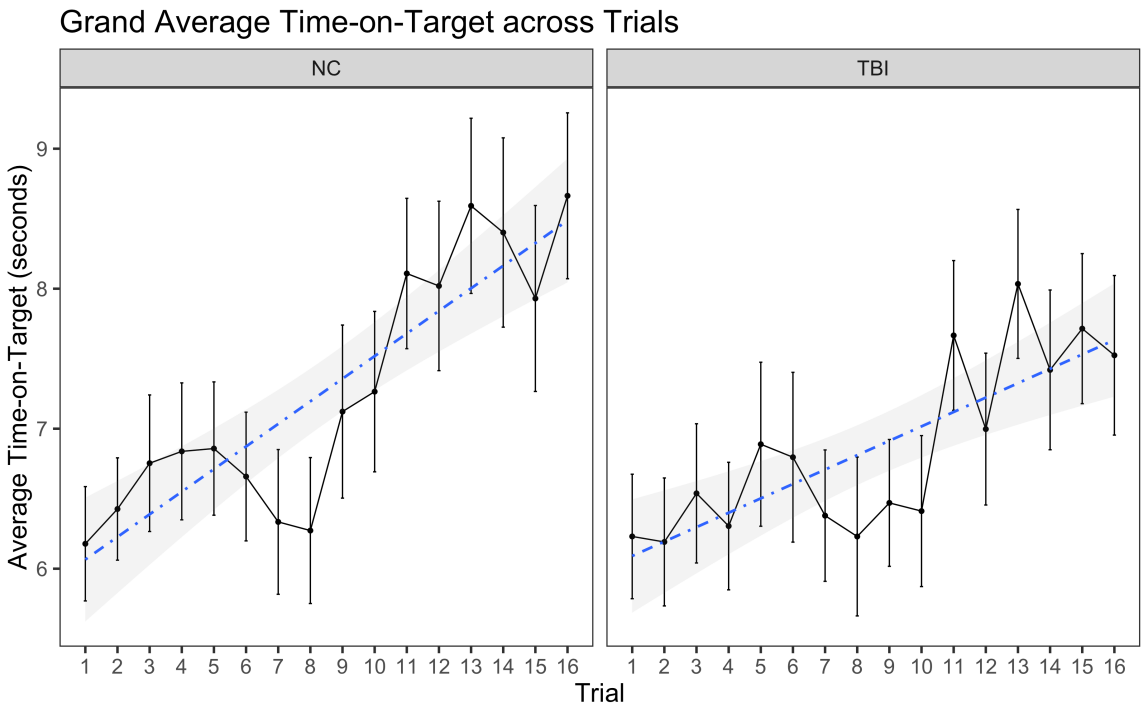

Figure 14. Average Time-on-Target across Trials.

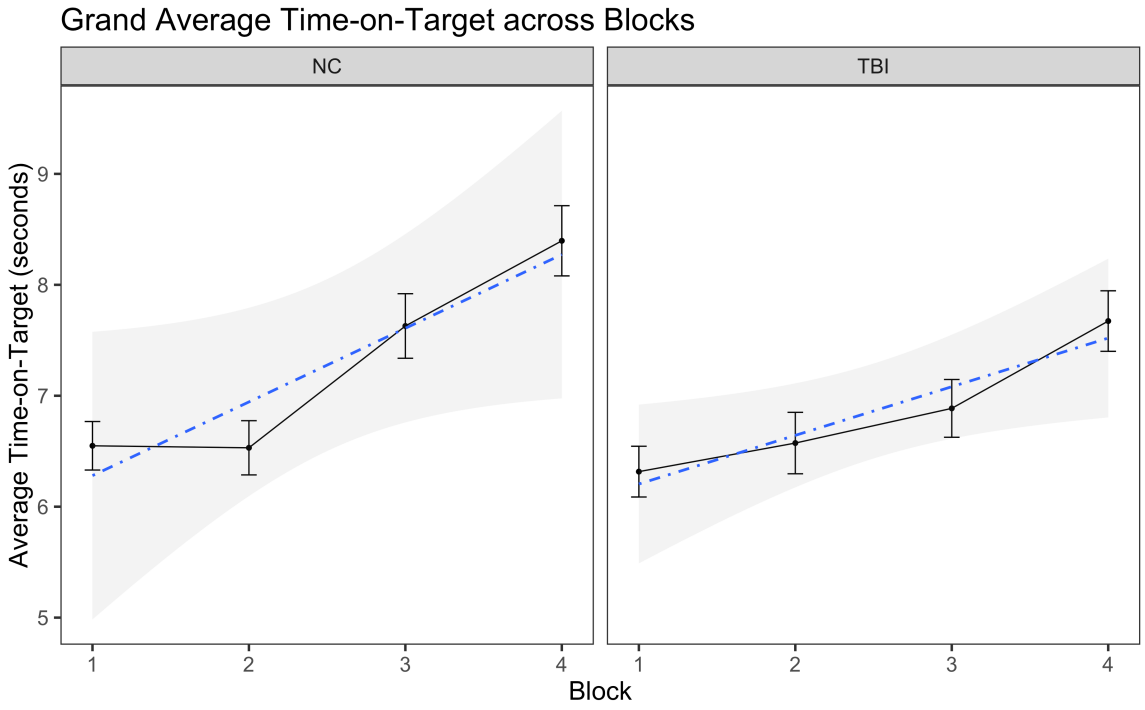

Figure 15. Average Time-on-Target by Block

However, fitting a full model with the addition of by-participant random slopes significantly improved model fit ($\chi^2(98) = 2$, $p < 0.001$) and eliminated the significant trial-by-group interaction ($b = -0.06$, $p = 0.24$). A significant effect of trial ($b = 0.16$, $p < 0.001$) indicated procedural learning of the rotor pursuit skill over time. By-participant random effects indicated significant inter-individual variability in initial performance (intercepts) and in degree of learning (slopes). Figures 16 and 17 display raw time-on-target data by participant. Figure 18 displays group and individual growth curves as predicted by the full model. Intercepts and slopes had a weak, non-significant negative correlation ($r = -0.09$), indicating that participants' initial time-on-target was not associated with their degree of learning across the task. In contrast, when the model was re-parameterized with reference to the final trial (trial 16 coded as 0, trial 15 as -1, etc.), there was a strong, significant positive correlation ($r = 0.75$) between by-participant intercepts and slopes, indicating that participants' degree of learning (slope) was associated with their final trial performance (re-parameterized intercept). Full model results for the reduced, full, and re-parameterized models are presented in Appendix B.

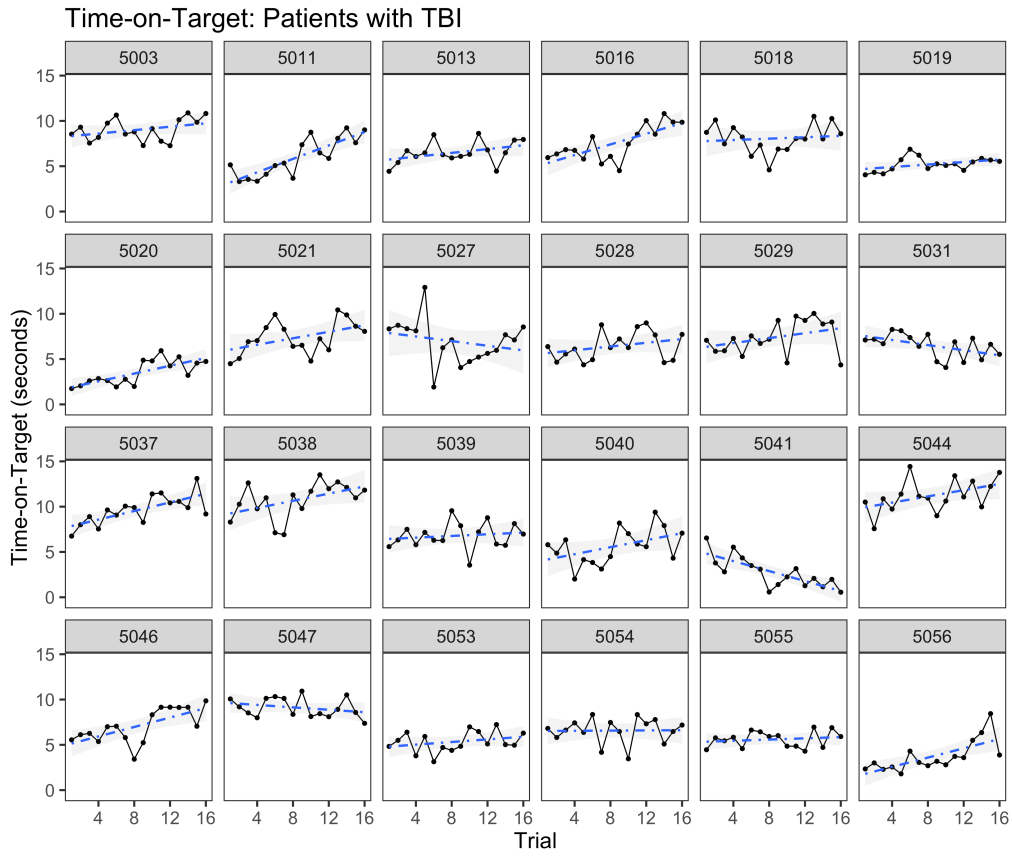

Figure 16. Time-on-Target by Participant: Patients with TBI

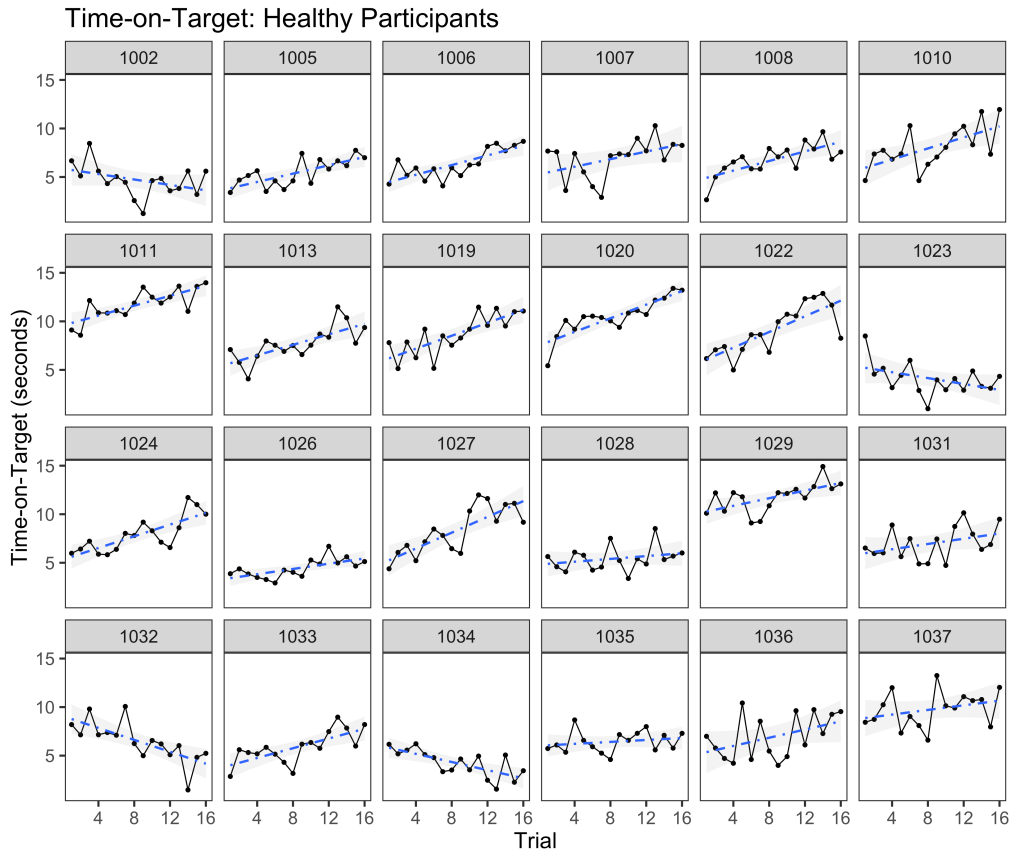

Figure 17. Time-on-Target by Participant: Healthy Adults

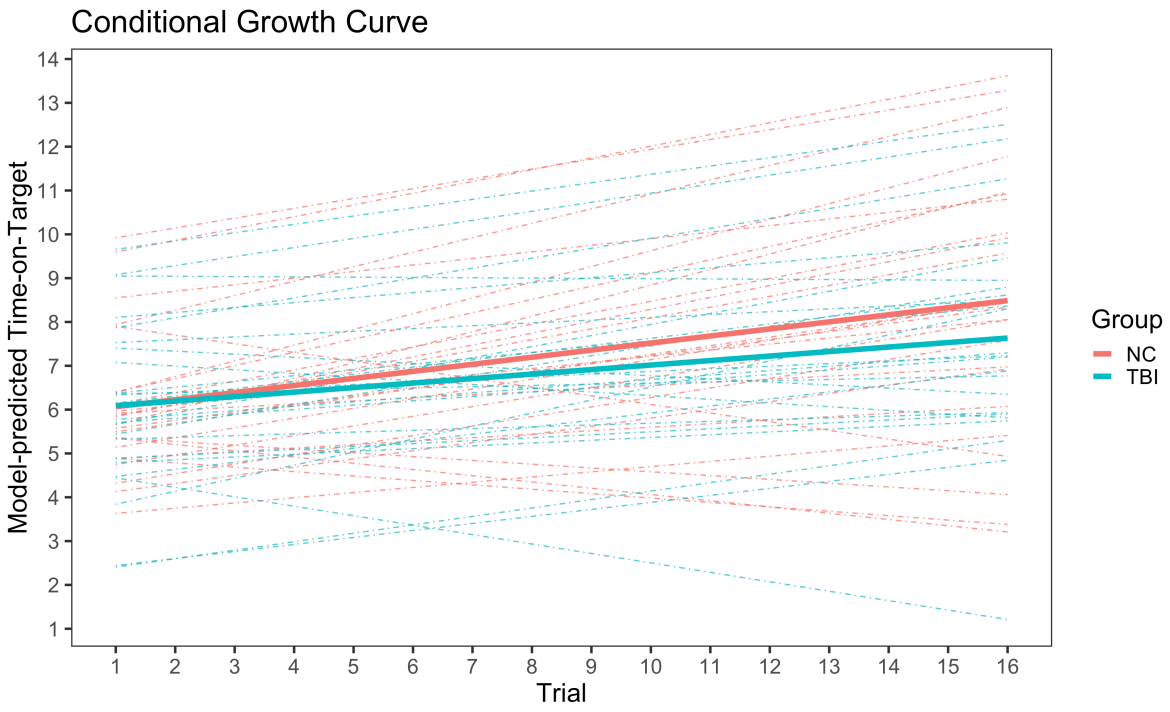

Figure 18. Photoelectric Rotor Pursuit Model-Predicted Time-on-Target

Participants' retention of the learned skill across sets was assessed by examining time-on-target for the final trial of the first set (trial 8) and the first trial of the second set (trial 9). The time elapsed between the end of the first 8-trial set and the beginning of the second was recorded for a majority of participants (22 TBI, 23 NC). There was an average of 68.51 minutes (SD = 23.14) between completion of the first set (trials 1-8) and the beginning of the second set (trials 9-16). Patients with TBI completed the intervening tasks between sets 1 and 2 significantly more slowly, resulting in a longer intervening interval (mean = 78.68 min, SD = 27.85) compared to healthy comparison participants (mean = 58.78 min, SD = 11.32, $t(27.51) = 3.11$, $p = 0.004$). On trial 8, the groups had similar average time-on-target (TBI mean = 6.23, SD = 2.78; NC mean = 6.27, SD = 2.55). Following the break, time-on-target for healthy participants was 7.12 seconds, on average (SD = 3.03) compared to 6.47 seconds (SD = 2.22) for patients with TBI (Figure 19).

There was a significant main effect of trial (b = 0.85, t(48) = 2.09, p = 0.04), indicating a small increase in time-on-target following the break. The difference in retention across groups was not significant (b = -0.61, t(48) = -1.06, p = 0.29, full model results presented in Appendix B). Differences in time-on-target between blocks 8 and 9 by participant are presented in Figure 20. There was no significant relationship between the amount of time elapsed between sets and participants' degree of retention (Figure 21, r = 0.17, t(43) = 1.11, p = 0.27).

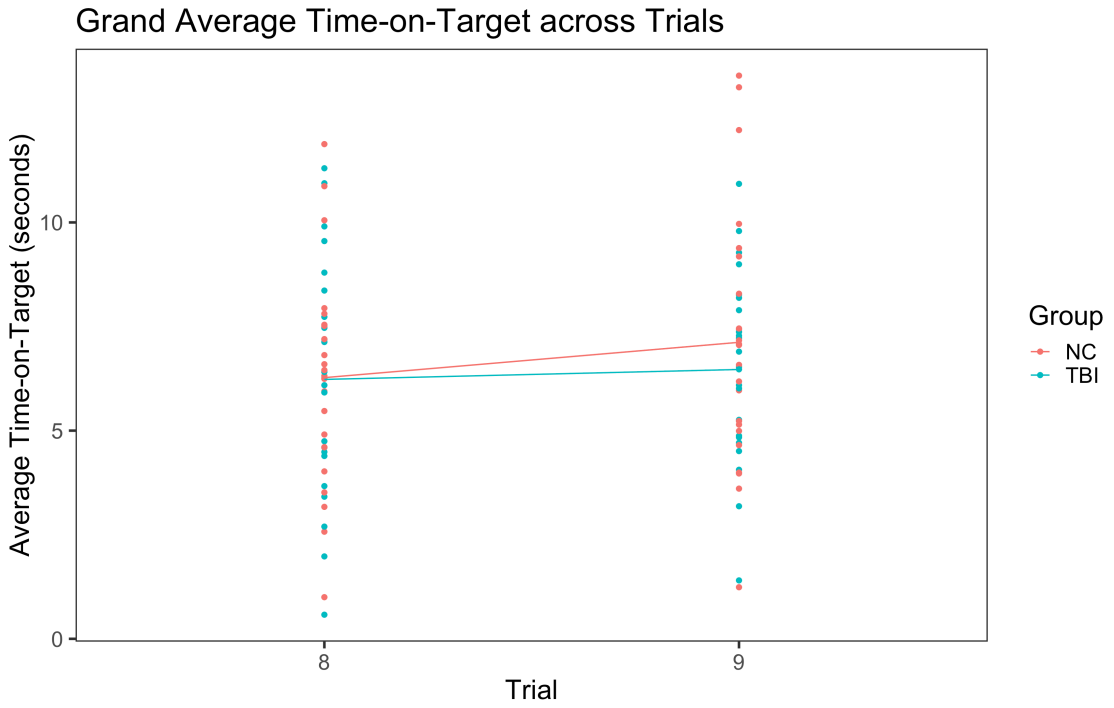

Figure 19. Time-on-Target across the Break.

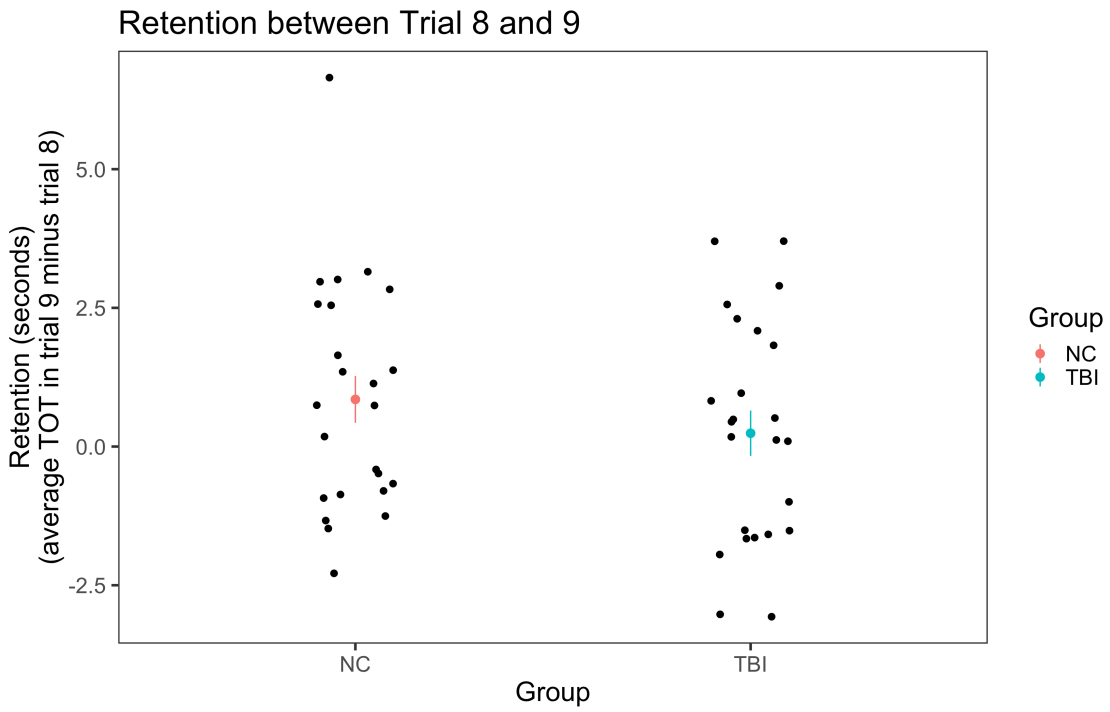

Figure 20. Retention by Group

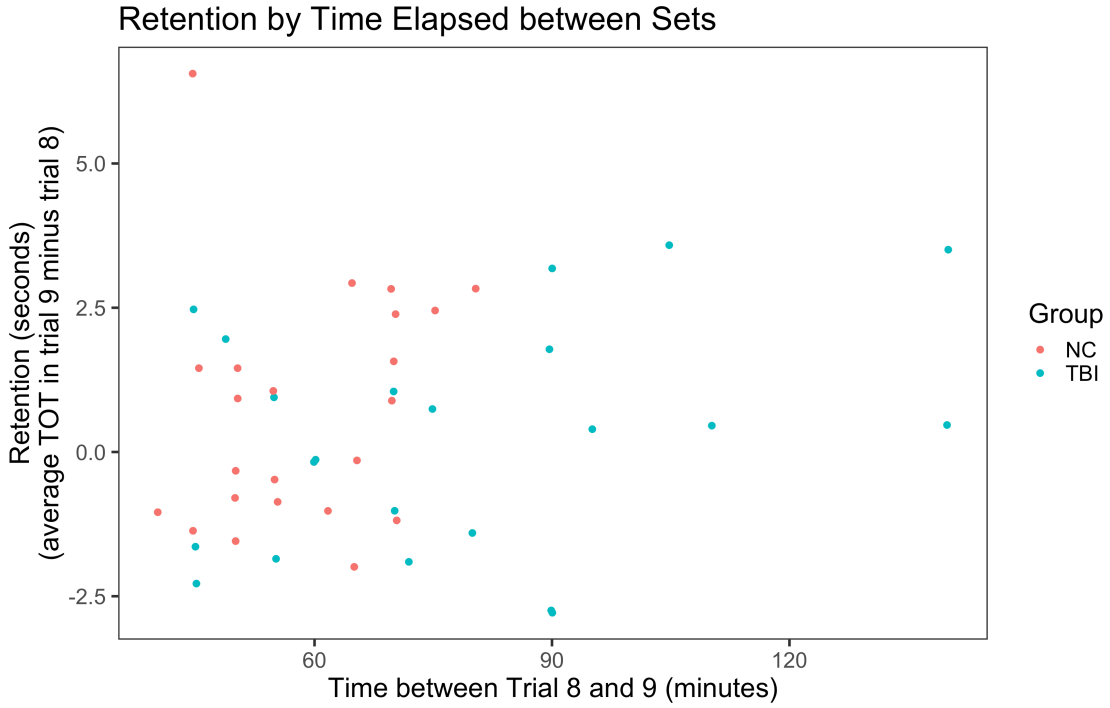

Figure 21. Relationship between Retention and Time Elapsed between Sets

Mouse-tracking Rotor Pursuit.

For the mouse-tracking rotor pursuit task, data from 48 participants were analyzed. One patient with TBI and one healthy comparison participant did not complete the task due to time constraints. Average time-on-target for each trial for each group, by feedback condition, is presented in Figure 22. To facilitate comparisons with previous rotor pursuit experiments, the data are also plotted by experimental block (2 trials per block) in Figure 23.

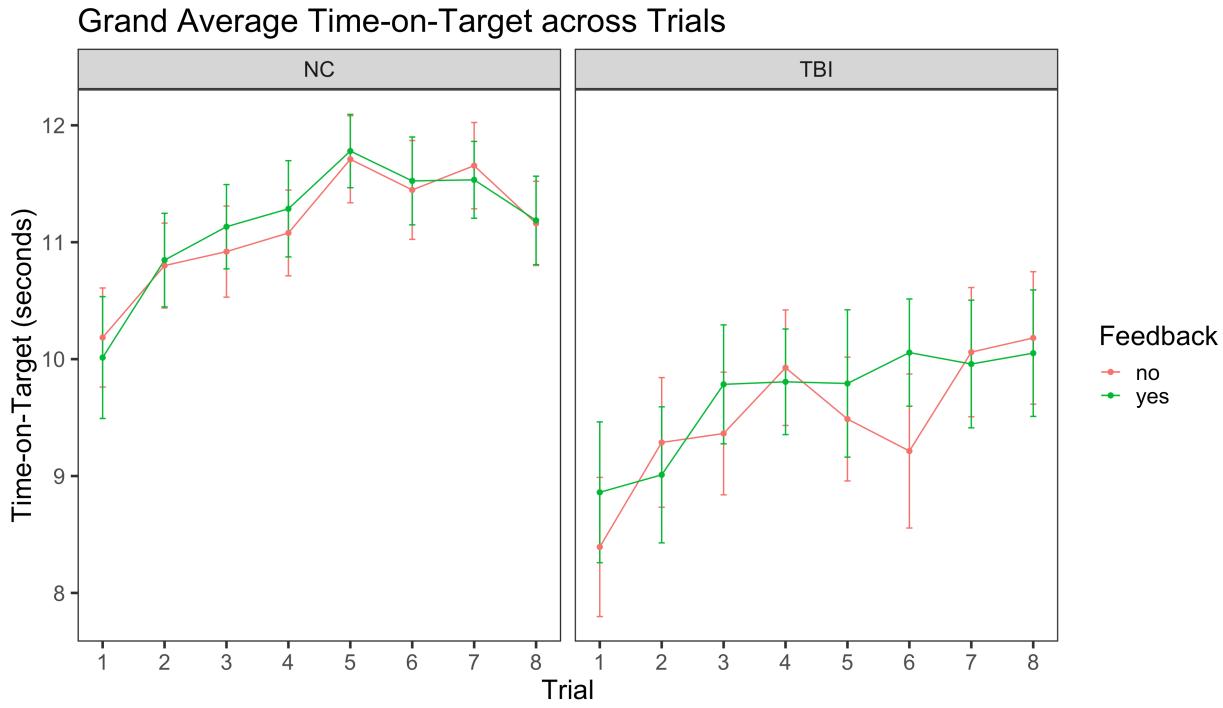

Figure 22. Time-on-Target across Trials

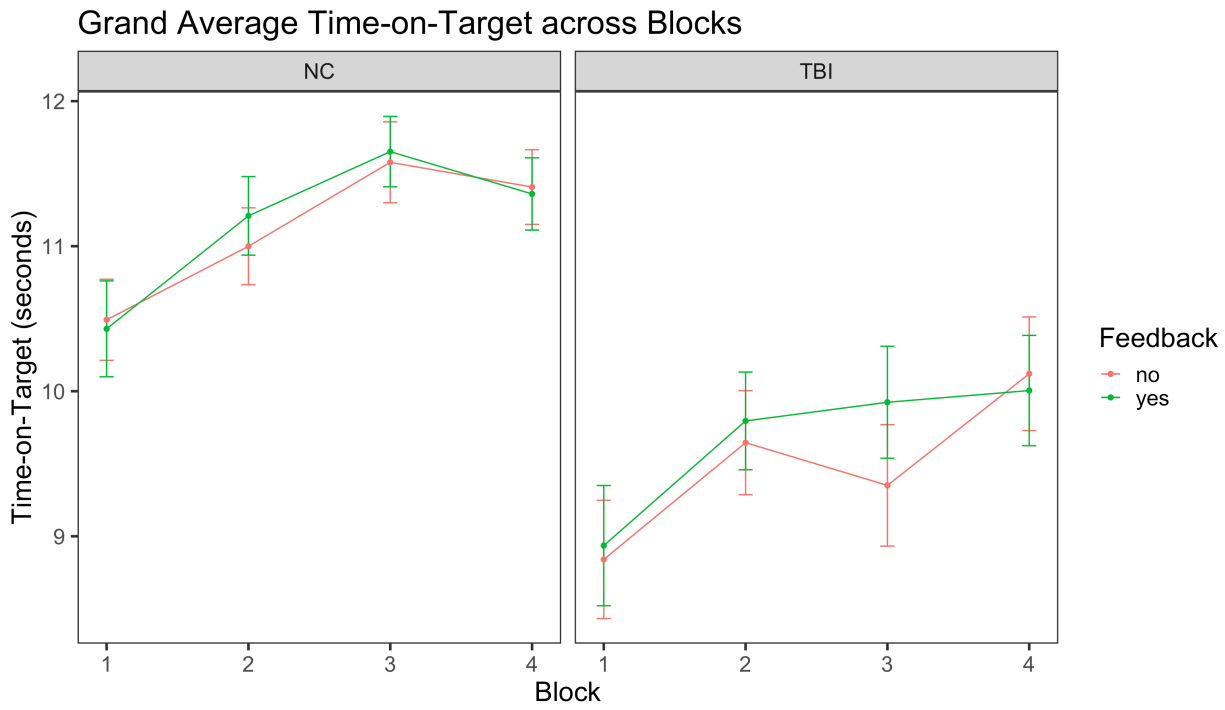

Figure 23. Time-on-Target by Block

To examine group differences in procedural learning, a conditional growth curve model was fit to the trial-by-trial time-on-target data. Models were fit using the *lme4* package in R, with p-values obtained using the *lmerTest* package. Trials were re-coded such that the intercept of the model could be interpreted as the grand average initial time-on-target (trial 1 coded as 0, trial 2 as 1, etc.). Visual inspection of the data suggested both linear and quadratic terms. An initial model was fit with fixed effects of group, feedback, and their interaction. The linear (trial) and quadratic ($trial^2$) terms were allowed to vary across participants and feedback conditions. There were no significant two- or three-way interactions between group, feedback, and either of the growth terms. There was a significant main effect of group ($b = -1.52$, $p = 0.04$), with TBI patients demonstrating significantly poorer time-on-target compared to healthy comparisons. There was a significant positive linear effect of trial ($b = 0.54$, $p < 0.001$) and a significant

negative quadratic effect of trial (b = -0.05, p = 0.004), demonstrating a positive trend with increasing downturn over time. Figure 24 displays group and individual growth curves as predicted by the model. Figures 25 and 26 display by-participant data. Full model results are presented in Appendix C.

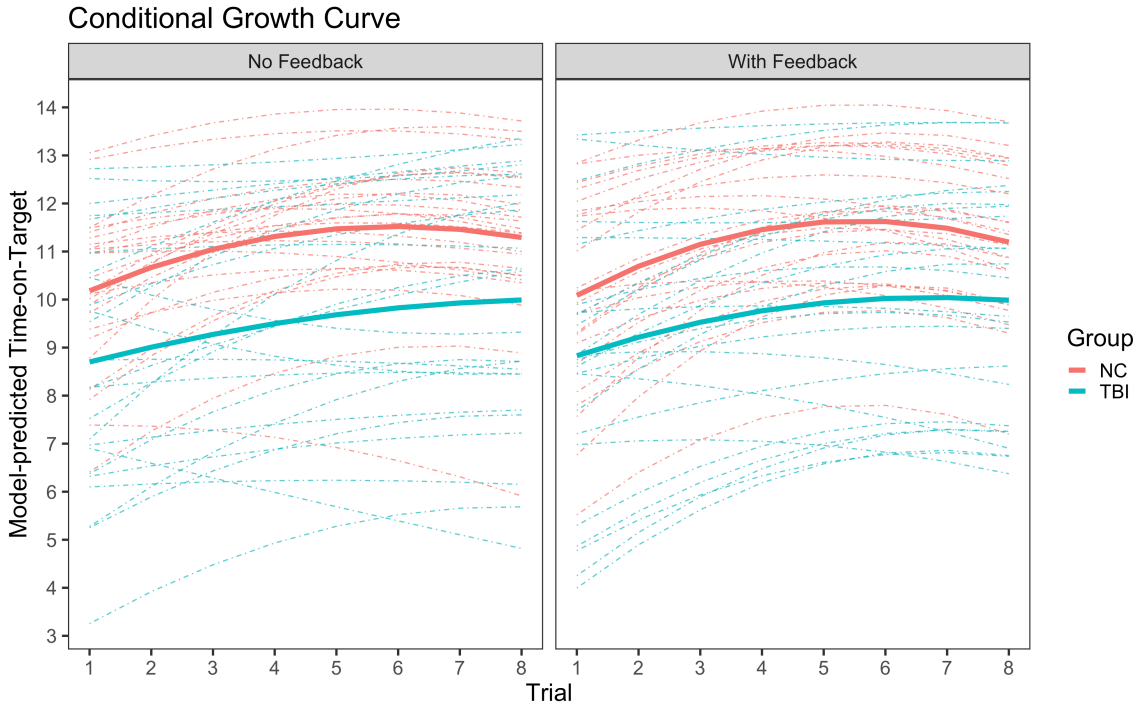

Figure 24. Mouse-tracking Rotor Pursuit Model-Predicted Time-on-Target

Neither the Feedback term nor any of its interactions with growth terms nor group were significant; therefore, Feedback was removed from the model. A model with a fixed effect of group and linear and quadratic growth terms that were allowed to vary across participants revealed the same significant effects: a significant main effect of group (b = -1.44, p 0.03) and significant positive linear (b = 0.61, p < 0.001) and negative quadratic effects of trial (b = -0.06, p < 0.001). There were significant individual differences in initial ability (intercepts) and

learning (linear and quadratic growth). There were no significant correlations among the random effects.

Figure 25. By-participant Time-on-Target: Patients with TBI

Figure 26. By-participant Time-on-Target: Healthy Adults

Mirror-Reversed Reading.

Removal of outliers

Out of the full dataset of 10,000 trials (50 participants x 200 trials), a small number of trials (n = 32, 0.32%) were removed from analysis. Some trials were removed based on contemporaneous notes taken during testing sessions (e.g. due to fire alarm going off, phone call). In addition, in a small number of instances, participants accidentally hit the space bar to advance to the next trial, prior to attempting the stimulus words.

In addition to trials identified during testing, trials were dropped prior to analysis based on an outlier detection rule established a priori based on pilot data (separate from the data reported on here). Because reading times were expected to decrease across experimental blocks, and to differ based on experimental condition (repeat vs. novel), outliers were assessed on a by-block and by-condition basis, relative to the individual participant's distribution of reading times within each condition and block. Outlier trials were defined as trials for which reading time was more than three standard deviations from the mean for that block and condition, for that particular participant. Using this outlier rule, 31 additional trials were dropped (0.31%). Thus, from the initial 10,000 trials, a total of 63 were excluded from analysis (0.63%).

Reading times across trials

In block 1 (prior to any repetitions of the "repeat" stimuli), there was significant variability in average reading times across participants, with by-participant average reading times ranging from 6.69 seconds to 46.84 seconds. However, the two groups did not significantly differ in average by-trial reading times during Block 1 (TBI mean = 26.74, SD = 17.40; NC mean = 22.79, SD = 13.81; $t(46.60) = 1.38$, $p = 0.17$). The distribution of by-participant average reading times during block one is presented in Figure 27.

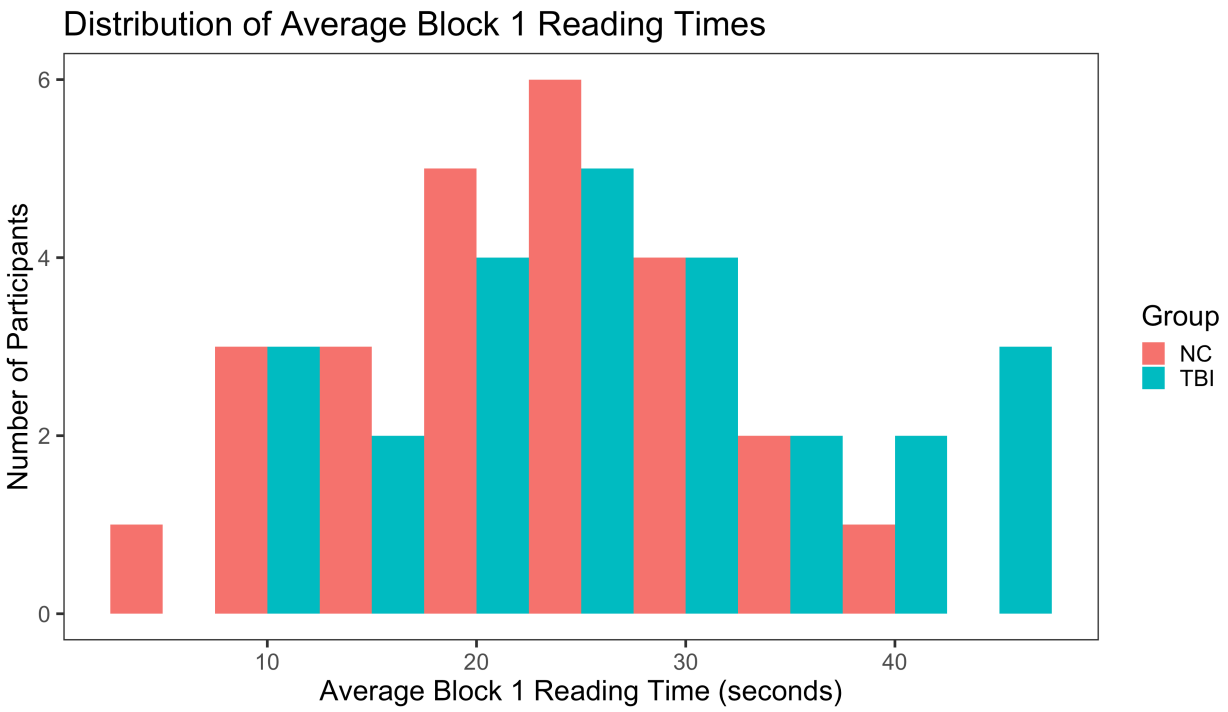

Figure 27. Distribution of Block 1 Reading Times

Reading accuracy

Reading accuracy was assessed and recorded contemporaneously during testing sessions. Overall accuracy across the task was high, with patients with TBI reading 92.03% of words accurately (SD = 9.82%) and healthy comparison participants reading 95.81% accurately (SD = 3.94%). Two patients with TBI demonstrated poor overall accuracy compared to the rest of the group (patient 5034: 58.62% accuracy, patient 5011: 68.84% accuracy; all other participants > 80% accuracy). During the first block of trials, there were no significant differences in accuracy across groups ($p = 0.17$) or conditions ($p = 0.45$). Average accuracy by block is presented in Figure 28; by-participant data is presented in Figure 29 and Figure 30.

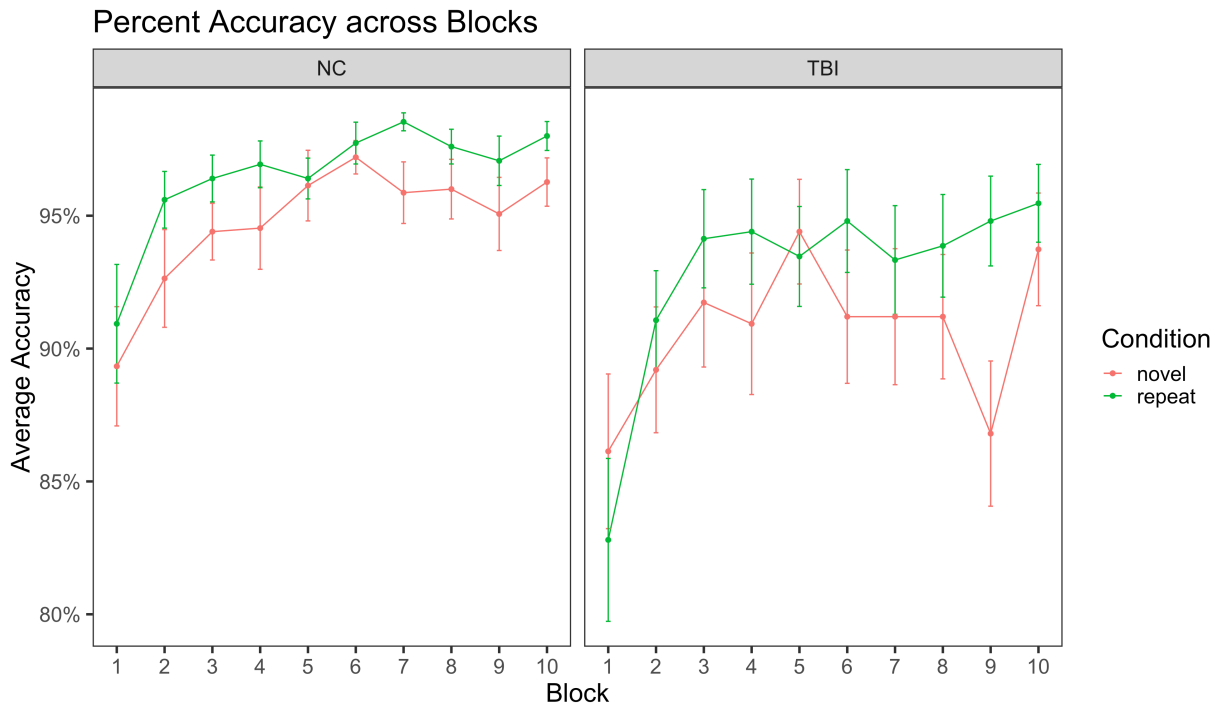

Figure 28. Mirror-Reversed Reading Accuracy by Block.

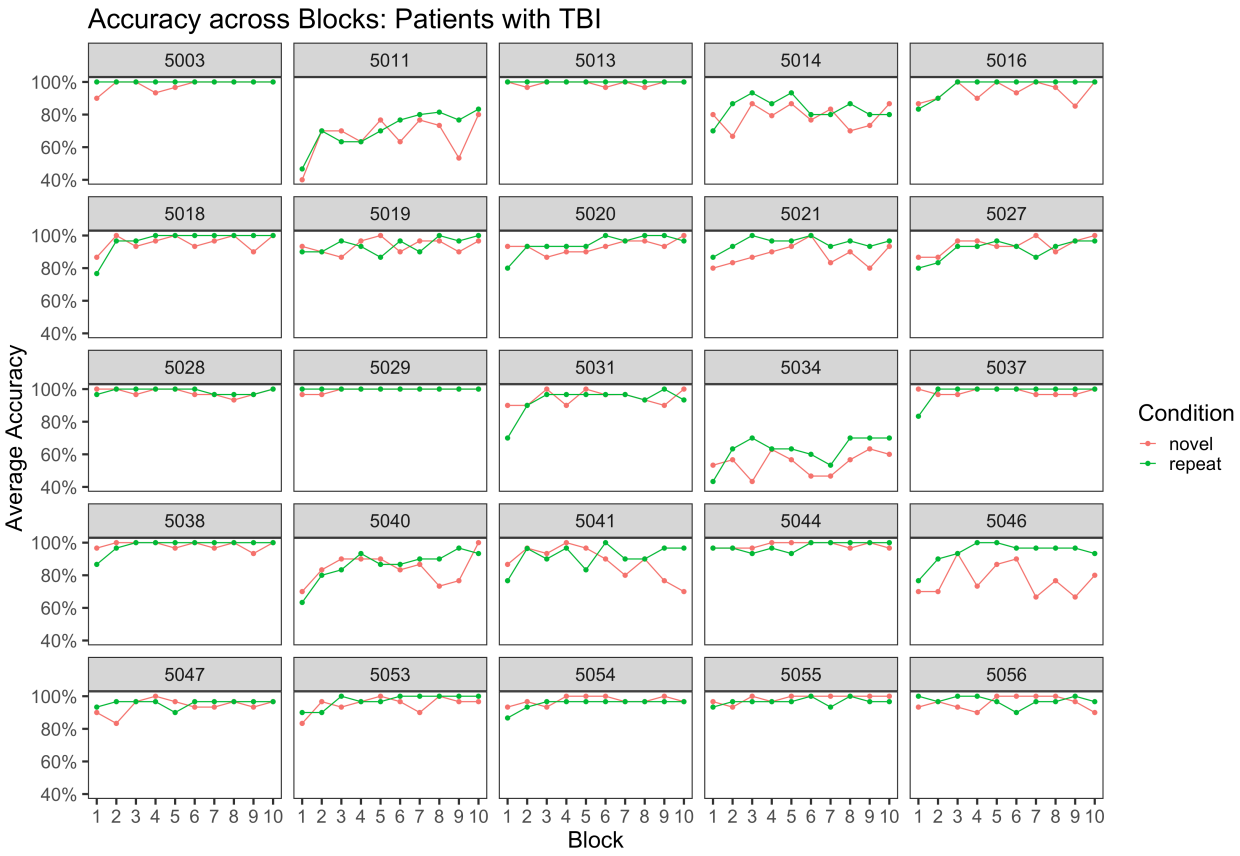

Figure 29. By-participant Accuracy: Patients with TBI

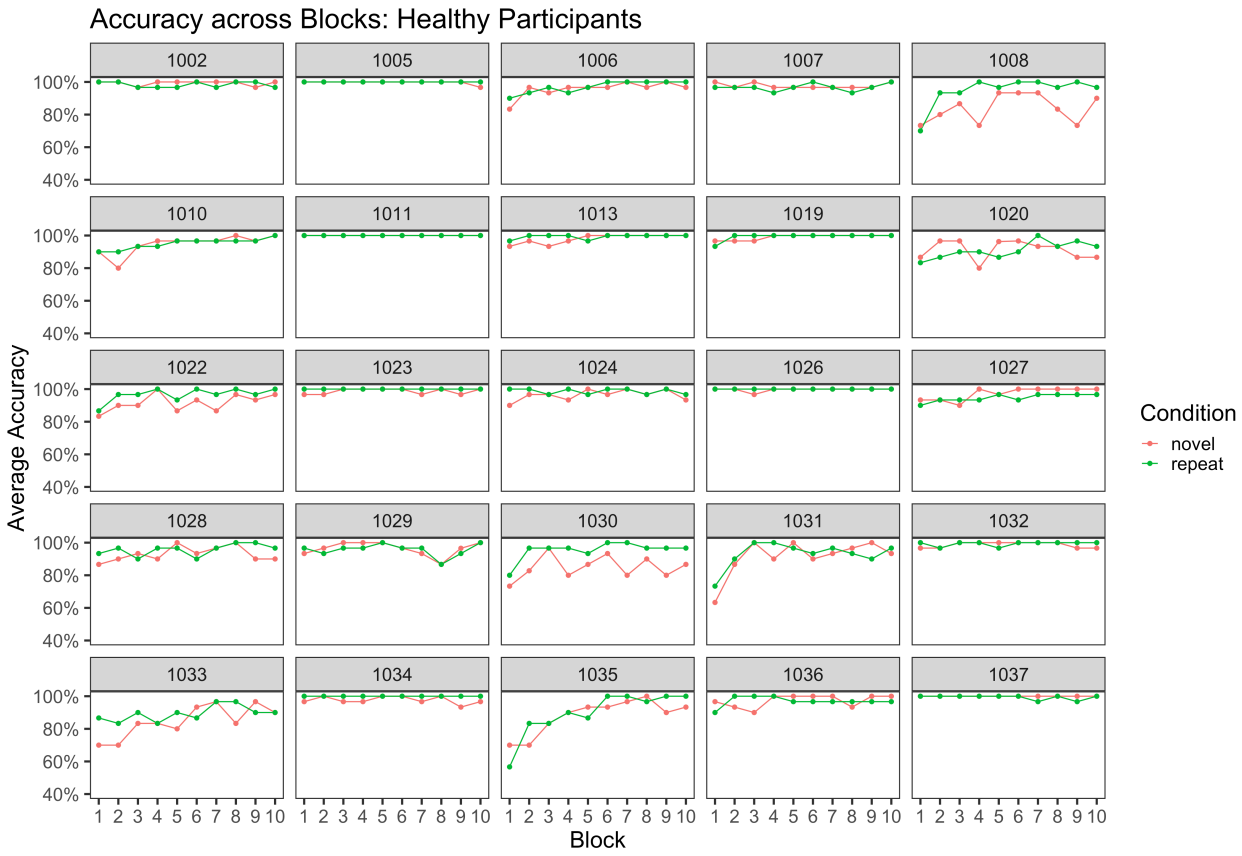

Figure 30. By-participant Accuracy: Healthy Adults

A logistic mixed effects model was fit to the accuracy data, with fixed effects of group, condition, and block, and their 2- and 3-way interactions, and random by-participant intercepts. Non-significant higher-order interaction terms were removed from the model in a step-wise fashion. The final model included fixed effects of group, block, condition, group*block, and condition*block, and random by-participant intercepts. A significant group*block interaction ($b = -0.06$, $p = 0.001$) indicated that patients with TBI demonstrated a smaller increase in log-odds of an accurate response across blocks relative to healthy comparison participants. A significant condition*block interaction ($b = 0.09$, $p < 0.001$) indicated larger increases in the log-odds of an accurate response for repeating words compared to novel words. Full model results are presented

in Appendix D. Participant's overall accuracy was significantly positively correlated (r = 0.44) with their overall reading ability, as measured by the WRAT Reading subtest (t(46) = 3.28, p = 0.002, see Figure 31).

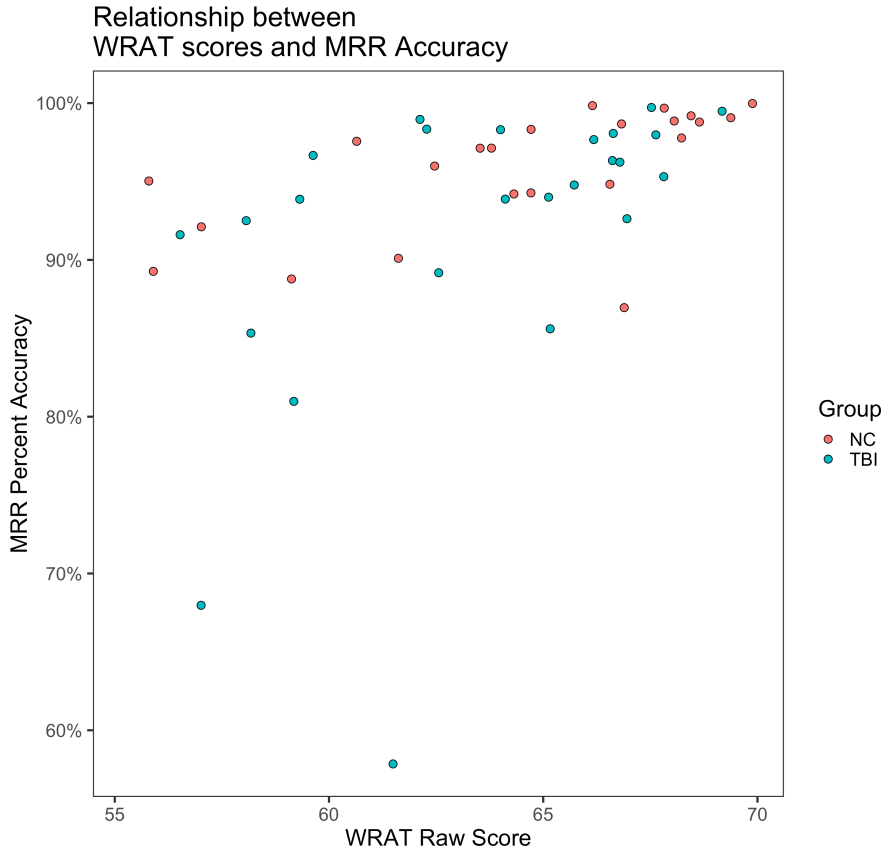

Figure 31. Overall Reading Ability and Mirror-Reversed Reading Accuracy

Mirror-reversed reading learning over time

Average reading times by group, block, and condition are presented in Figure 32. By-participant reading times for individuals with TBI are presented in Figure 33 and in Figure 34 for healthy comparison participants.

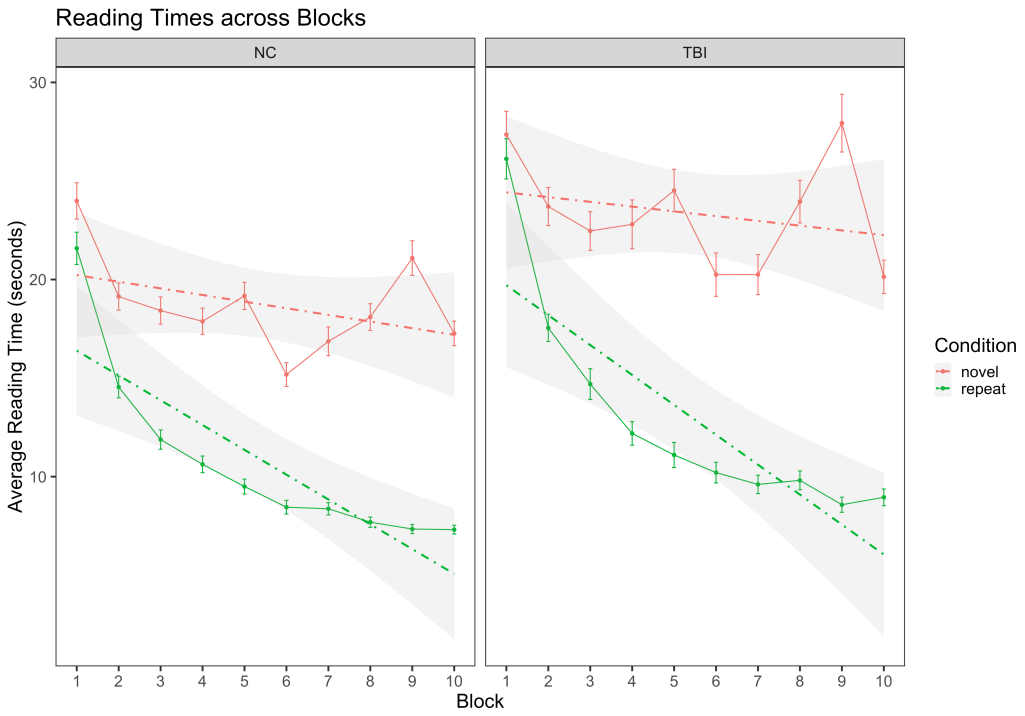

Figure 32. Reading Time by Block

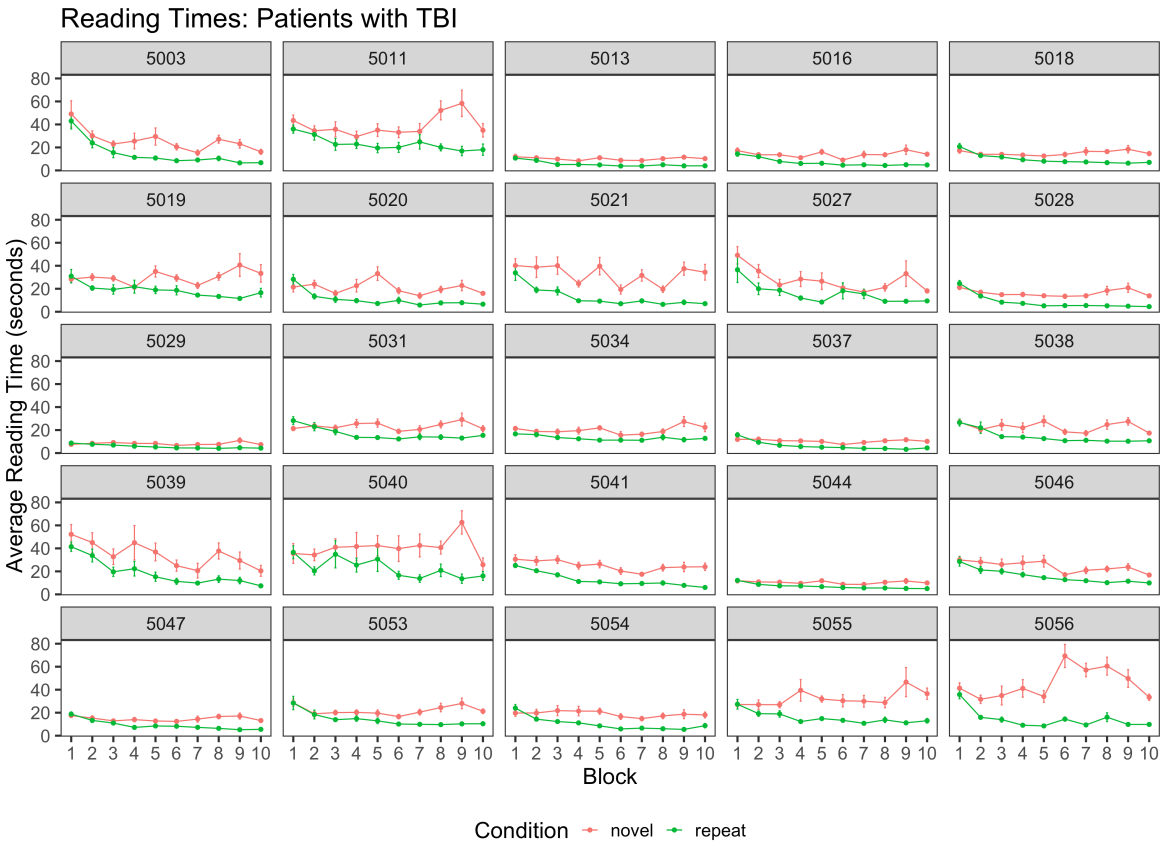

Figure 33. By-participant Reading Times: Patients with TBI

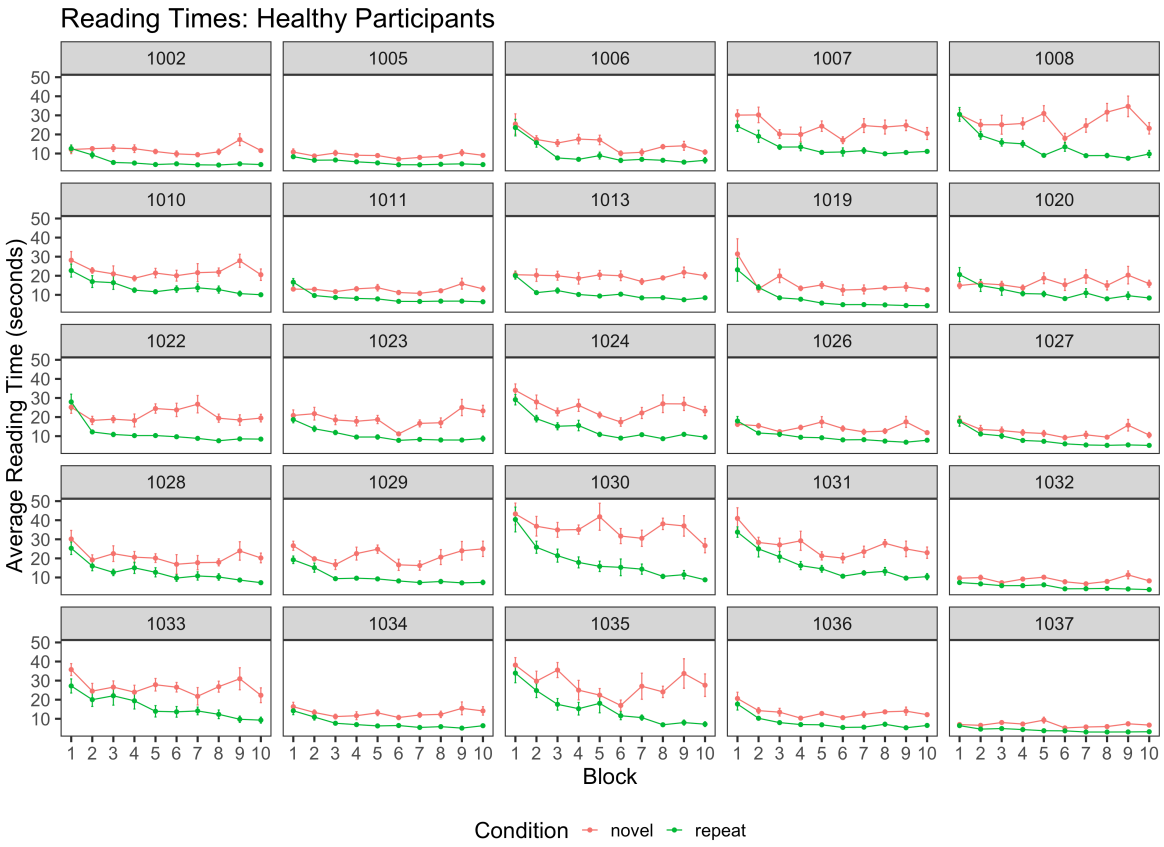

Figure 34. By-participant Reading Times: Healthy Adults

To examine group differences in procedural learning, a conditional growth curve model was fit to the reading time data. Models were fit using the *lme4* package in R, with p-values obtained using the *lmerTest* package. Blocks were re-coded such that the intercept of the model could be interpreted as the grand average initial reading time (block 1 coded as 0, block 2 as 1, etc.). Visual inspection of the data suggested both linear and quadratic terms. An initial model was fit with fixed effects of group, condition, and their interaction. The linear (block) and quadratic ($block^2$) terms were allowed to vary across participants and conditions. There were no significant three-way interactions between group, condition, and either of the growth terms. Thus, three-way interaction terms were eliminated from the model. The final model included

fixed effects of group, condition, the growth terms, and their two-way interactions. The linear (block) and quadratic (block^2) terms were allowed to vary across participants and conditions. Contrary to predictions for Research Question 2a., there was no significant interaction between group and either of the growth terms (both ps > 0.7). There was no significant interaction between group and condition (b = -2.93, p = 0.27). There were significant two-way interactions between condition and both of the growth terms, such that reading times decreased more rapidly across blocks for repeated relative to novel words (b = -6.74, p < 0.001). There was a significant main effect of group, with patients with TBI reading more slowly over all compared to healthy participants (b = 4.87, p = 0.01). Figure 35 displays group and individual growth curves as predicted by the model. Full model results are presented in Appendix D.

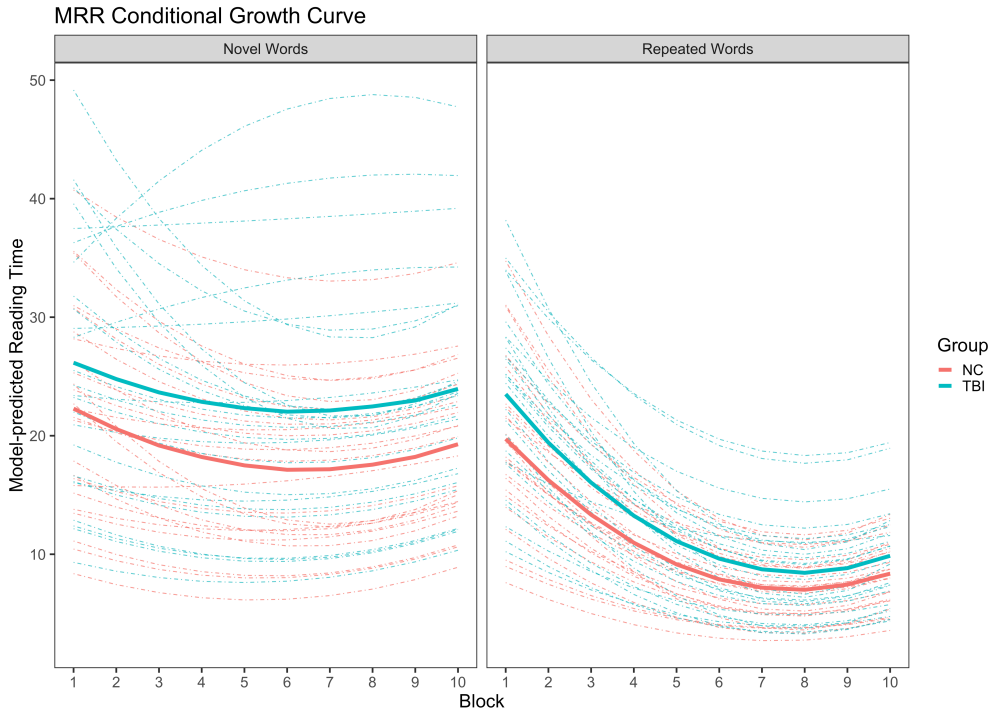

Figure 35. Model-Predicted Reading Times

A model with the same effects structure was run on data excluding all trials in which participants made an error (see Figure 36). Removing trials with at least one error (out of the word triad) resulted in the exclusion of 1,535 trials (15.44% of trials). On average, trials in which participants read at least one word inaccurately were read significantly more slowly than trials for which they read all three words accurately ($t(1723.8) = 25.75$, $p < 0.001$, Figure 37). All of the models terms remained significant following removal of errored trials with the exception of the main effect of group, which was no longer significant ($b = 3.04$, $p = 0.06$). Full model results are presented in Appendix D.

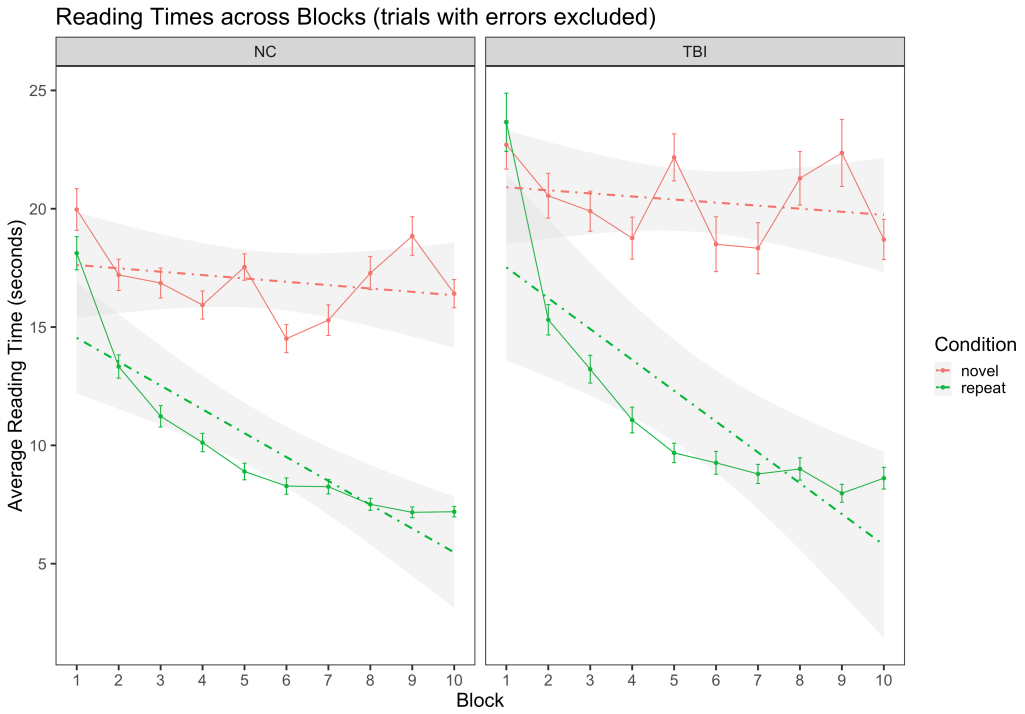

Figure 36. Reading Time by Block (errored trials excluded)

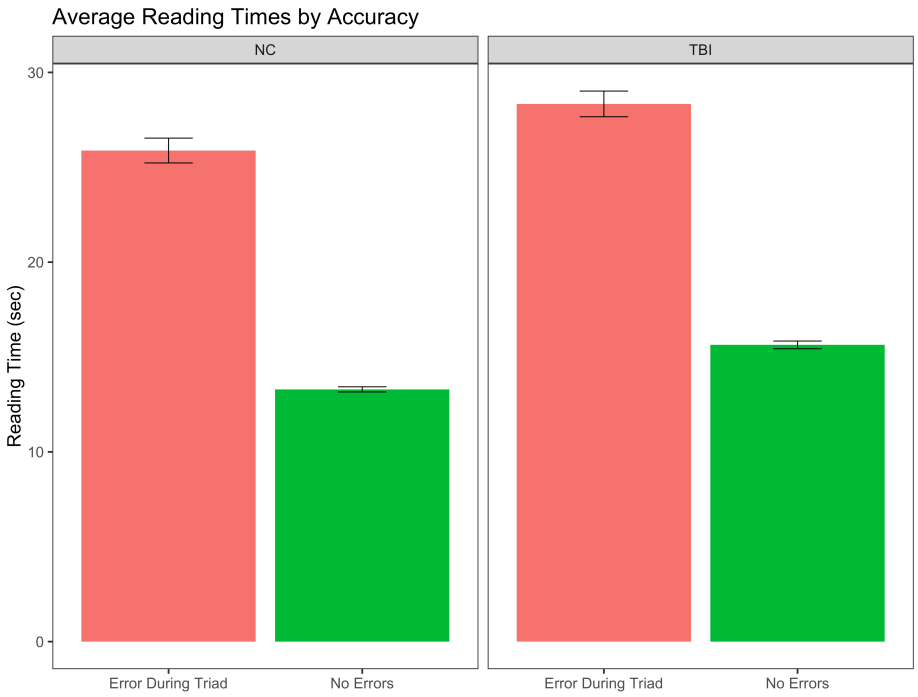

Figure 37. Reading Times for Triads with and without Errors

Serial Interception Sequence Learning.

Across training trials, participants in both groups performed with similar levels of overall accuracy. Patients with TBI hit the correct button at the correct time an average of 79.02% of the time (SD = 1.51%) during the first five blocks, and healthy comparison participants were an average of 79.89% accurate (SD = 0.08%). Changes in cue speed over the course of the task are presented in Figure 38.

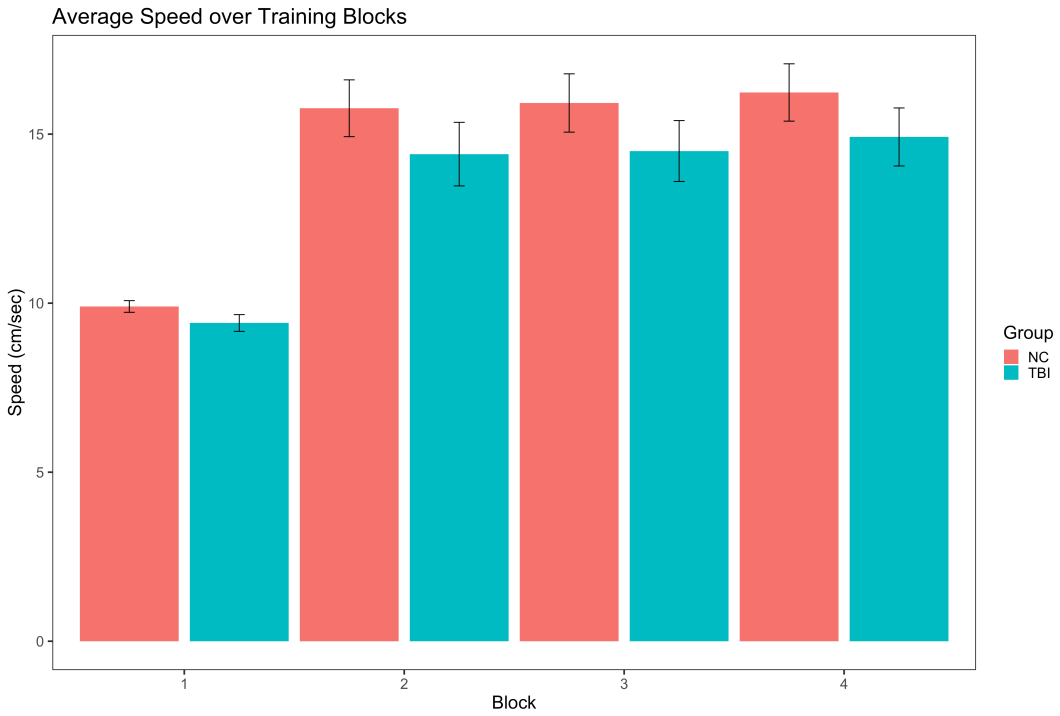

Figure 38. Cue Speed by Block

There was no significant group difference in cue speed in blocks 5 and 6 (the implicit test phase; Figure X) ($t(47.93) = -0.84$, $p = 0.41$, Figure 39). There was a significant sex effect, with male participants reaching an average cue speed of 17.35 cm/sec, compared to 14.10 for female participants ($t(32.09) = 2.55$, $p = 0.02$; Figure 40). There was also a significant effect of age, with younger participants reaching a faster average cue speed (17.22 cm/sec) compared to older participants (13.47 cm/sec; Figure 41).

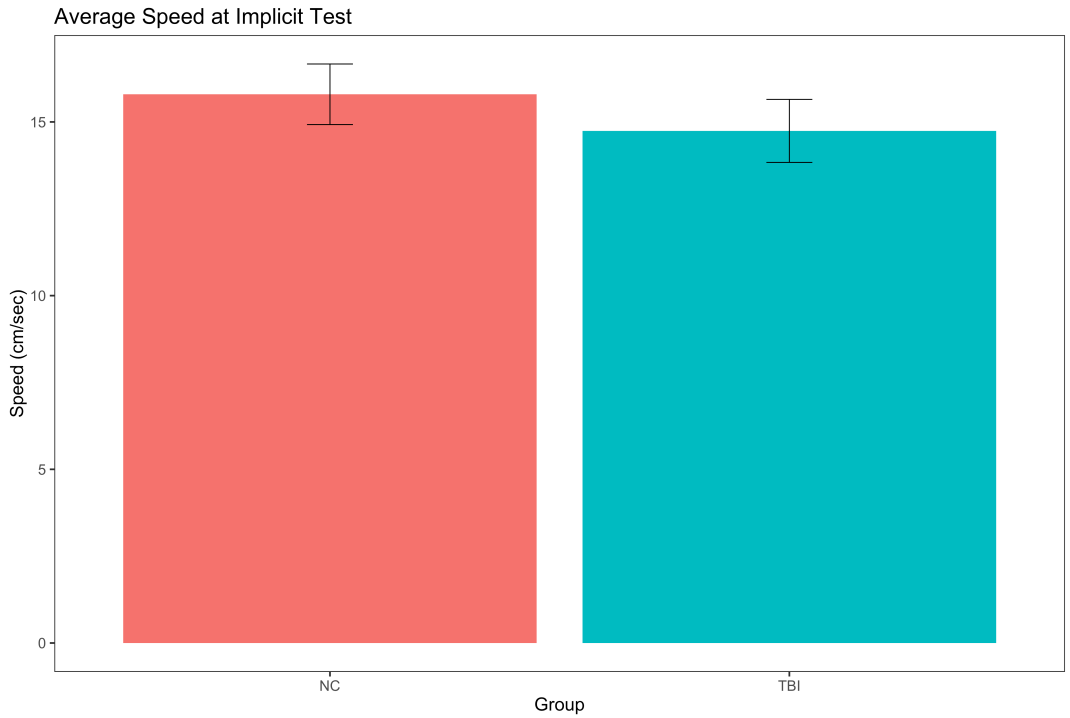

Figure 39. Average Cue Speed by Group during Implicit Test

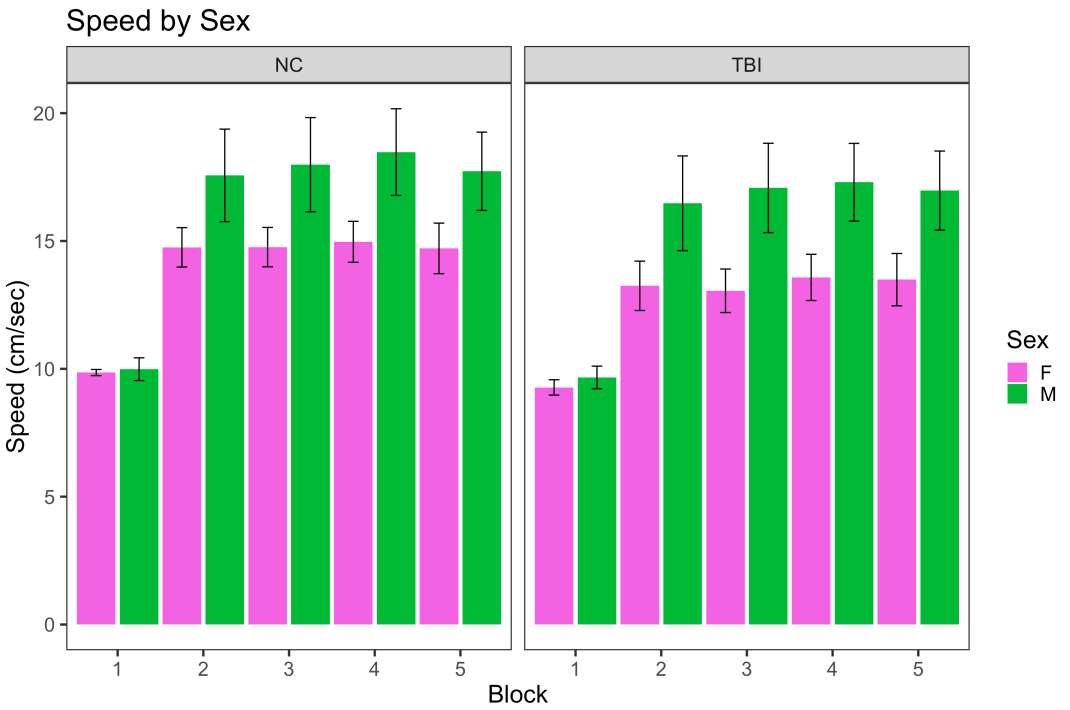

Figure 40. Cue Speed by Sex

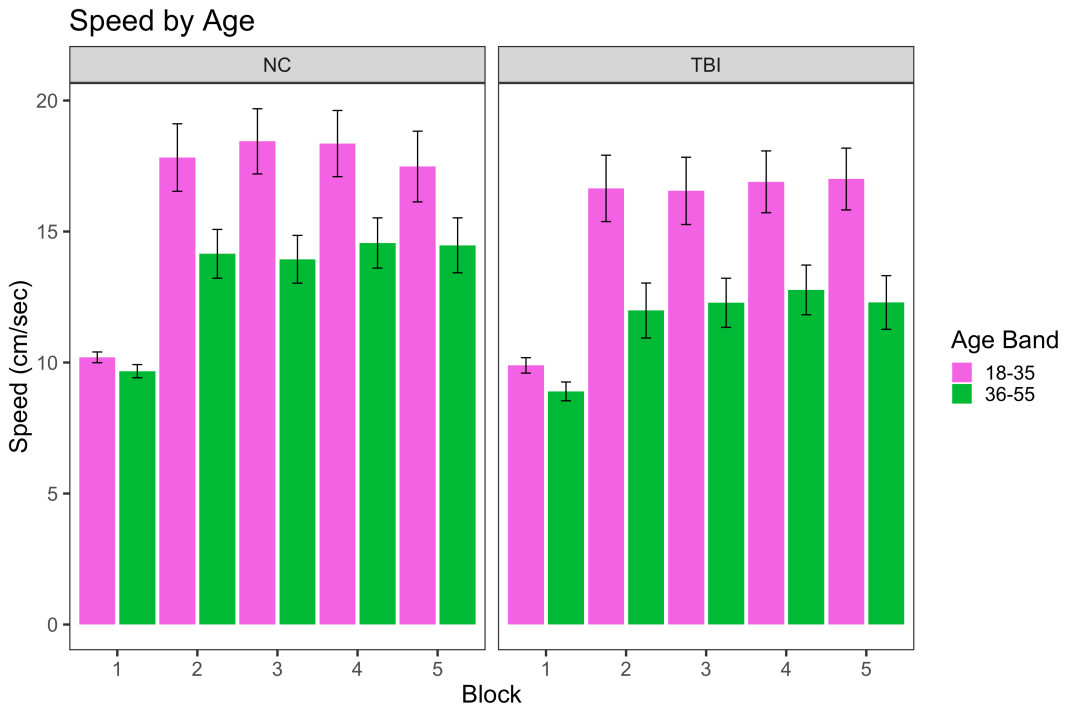

Figure 41. Cue Speed by Age

Sequence-specific performance advantage (SSPA; average percent accuracy for sequence trials minus average percent accuracy for foil trials), by group is presented in Figure 42. By-participant SSPA is presented in Figure 43 and Figure 44. Participants' procedural learning of the underlying sequence was assessed during the learning phase over blocks two through four. Models were fit using the *lme4* package in R, with p-values obtained using the *lmerTest* package. Blocks were re-coded (block 1 coded as 0, block 2 as 1, etc.) such that the intercept of the model could be interpreted as the grand average sequence-specific performance advantage (SSPA). An initial model with fixed effects of block, group, and their interaction and by-participant random intercepts and slopes indicated that, as a group, participants' SSPA was significantly greater than zero ($t(58.40) = 2.65$, $p = 0.01$). There was no significant difference in change in SSPA across blocks between the two groups ($t(91.71) = -0.96$, $p = 0.34$). Contrary to predictions for Research

Question 2a., there was no significant main effect of group (t(58.40) = 0.53, p = 0.60). There was no significant main effect of block (t(91.71) = 1.88, p = 0.06). Full model results are presented in Appendix E. During the implicit test blocks, there were no group differences (t(43.17) = 0.55, p = 0.58).

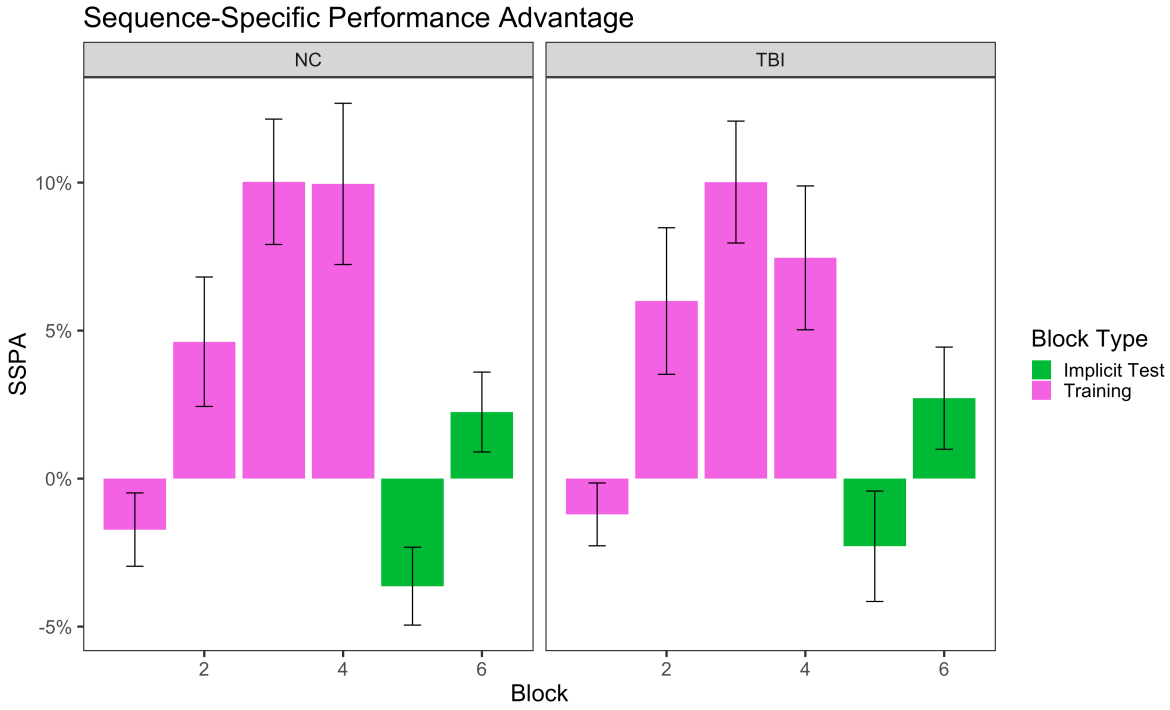

Figure 42. Sequence-Specific Performance Advantage by Block

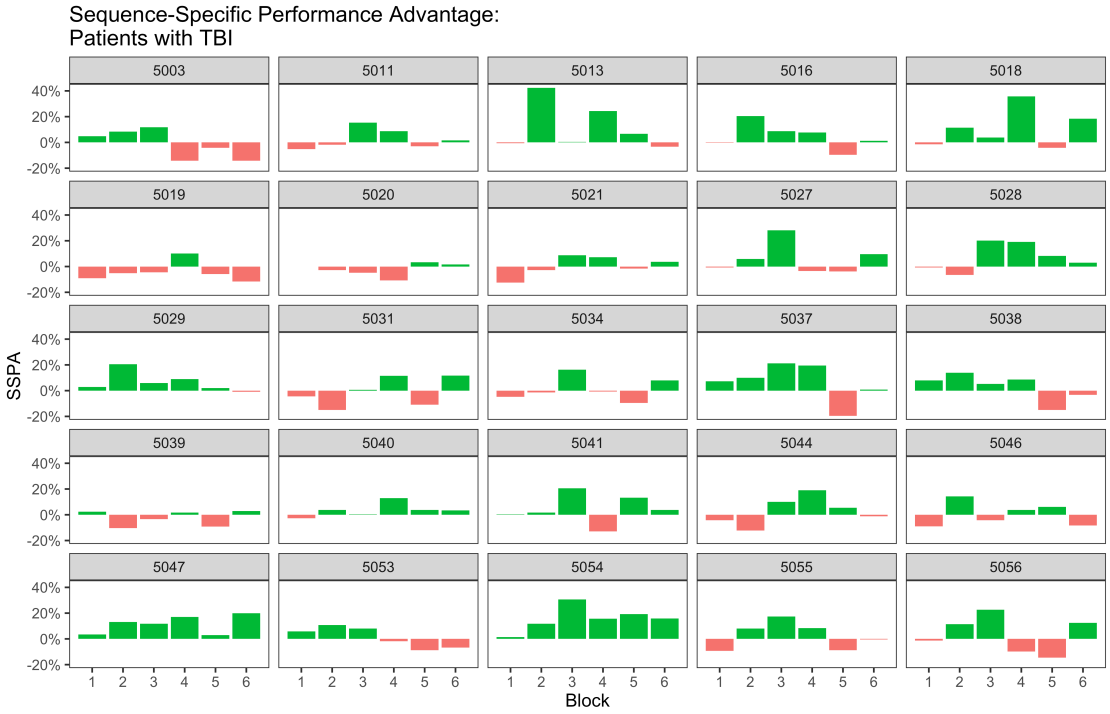

Figure 43. By-participant SSPA: Patients with TBI.
(Positive SSPA in green; negative SSPA in red)

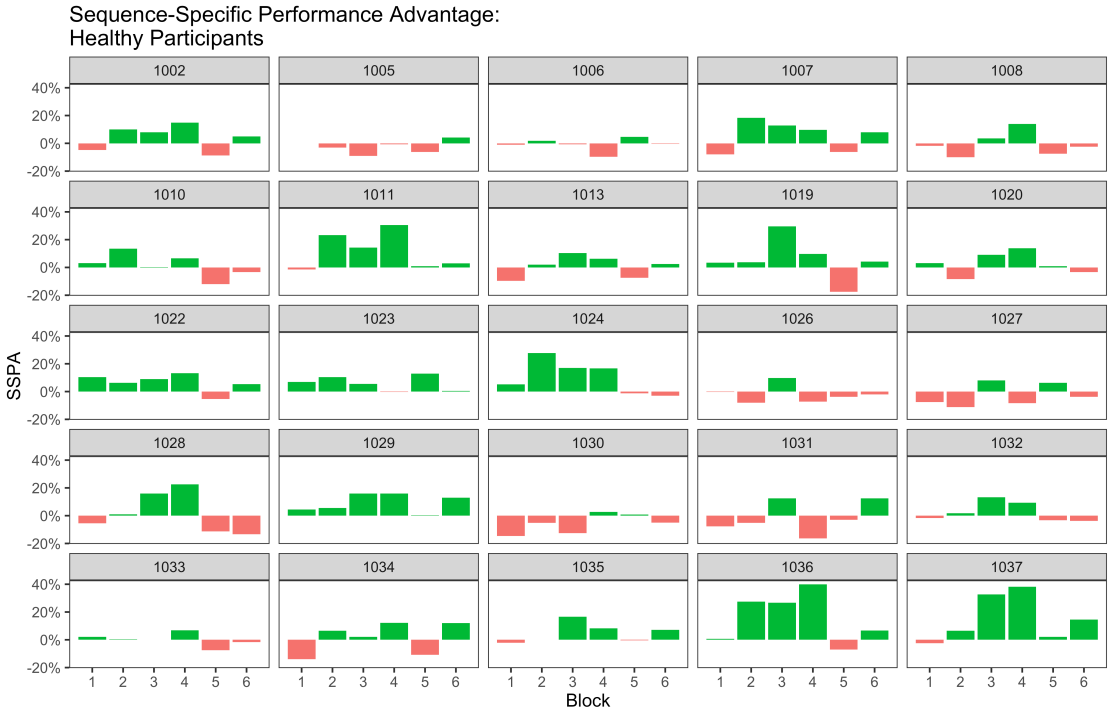

Figure 44. By-participant SSPA: Healthy Adults.
(Positive SSPA in green; negative SSPA in red)

Individual Differences

Measures.

To examine individual differences across procedural memory tasks, unconditional growth curves (i.e. no fixed effects besides the growth term) were fit to each task's data, and by-participant intercept and slope coefficients were extracted for use as individual difference measures. All unconditional growth curves had equivalent fixed and random effect structure: a linear effect of time ("trial" for both rotor pursuit tasks, and "block" for mirror-reversed reading), with random by-participant intercepts and slopes, to facilitate between-task comparisons. Two separate models were fit to the mirror-reversed reading data: one fit to the novel word triad data, and the second fit to the repeating word triad data. Participants' average SSPA during learning

blocks three and four was utilized as the individual differences measure in all subsequent analyses.

Relationships between Procedural Memory Tasks.

To address Research Question 2b., correlations between procedural memory tasks were examined. Three different metrics were assessed as individual difference measures for the rotor pursuit and mirror-reversed reading procedural memory tasks. These included by-participant slopes (participant's degree of learning), intercepts (participants' initial ability) and difference scores (participants' average performance during the final block minus their average performance in the first block). Positive slopes, intercepts, and difference scores are indicative of better performance on the rotor pursuit tasks, while negative slopes and difference scores are indicative of better performance on the mirror-reversed reading task. For SISL, participants' average sequence-specific performance advantage (SSPA) during the final two training blocks was used as an individual differences measure. Larger, positive SSPAs are indicative of better performance. Significant correlations between metrics across tasks (positive among all tasks excluding MRR; negative for all correlations involving MRR) would suggest that these experimental measures each capture the underlying construct of procedural memory and support predictions for Research Question 2b.

Assessed across all participants, irrespective of group, contrary to predictions for Research Question 2b., there were no significant correlations between any of the procedural memory tasks when by-participant slopes were used as individual difference measures (all $| r |s < 0.2$, all $ps > 0.05$), with the exception of a medium negative correlation between mouse-tracking rotor pursuit slopes and participants' average SSPA ($r = -0.35$, $p = 0.01$). Correlation matrices by

group are presented in Figures 45 and 46. When difference scores were used as the individual difference measure, there were no significant correlations among any of the tasks (all | r |s < 0.4, all ps > 0.5; Figures 47 and 48). There were significant medium and strong correlations in the expected direction for all tasks when by-participant intercepts were used as the individual difference measure, suggesting relationships between participants' initial abilities across tasks, in the absence of significant associations between their degree of learning across tasks. Correlation matrices by group are presented in Figures 49 and 50.

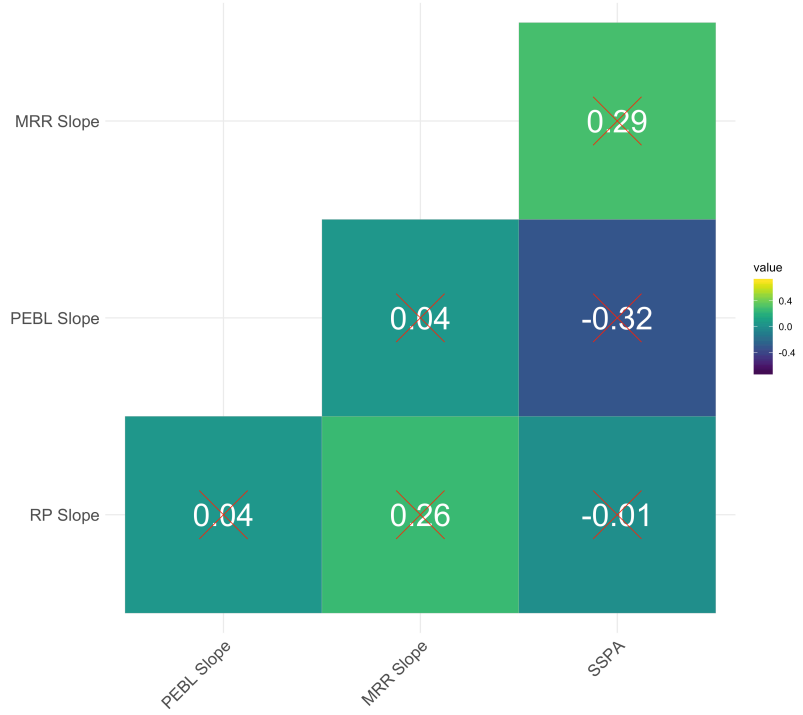

Figure 45. Correlations between By-Participant Slopes: Patients with TBI
Red Xs indicate nonsignificant correlations.

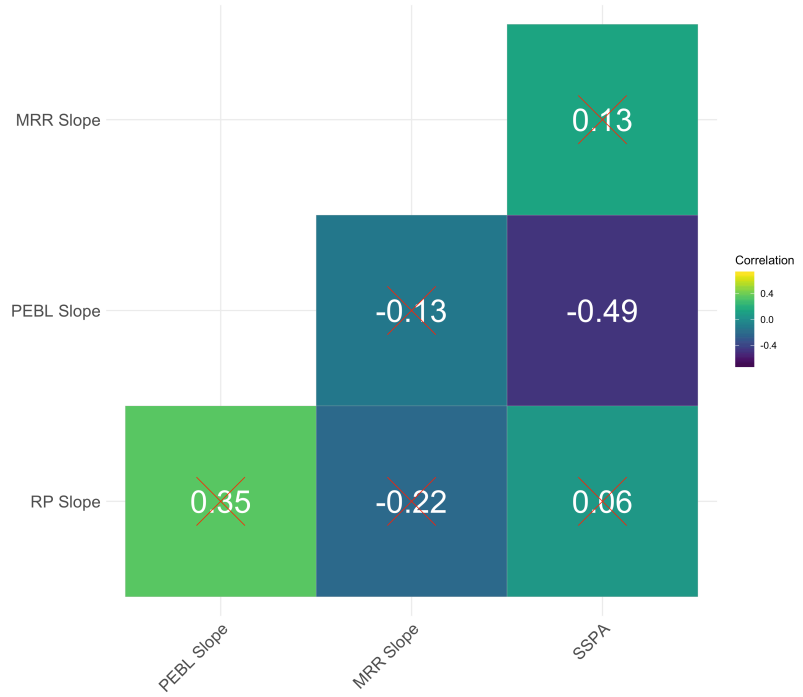

Figure 46. Correlations between By-Participant Slopes: Healthy Adults
Red Xs indicate nonsignificant correlations.

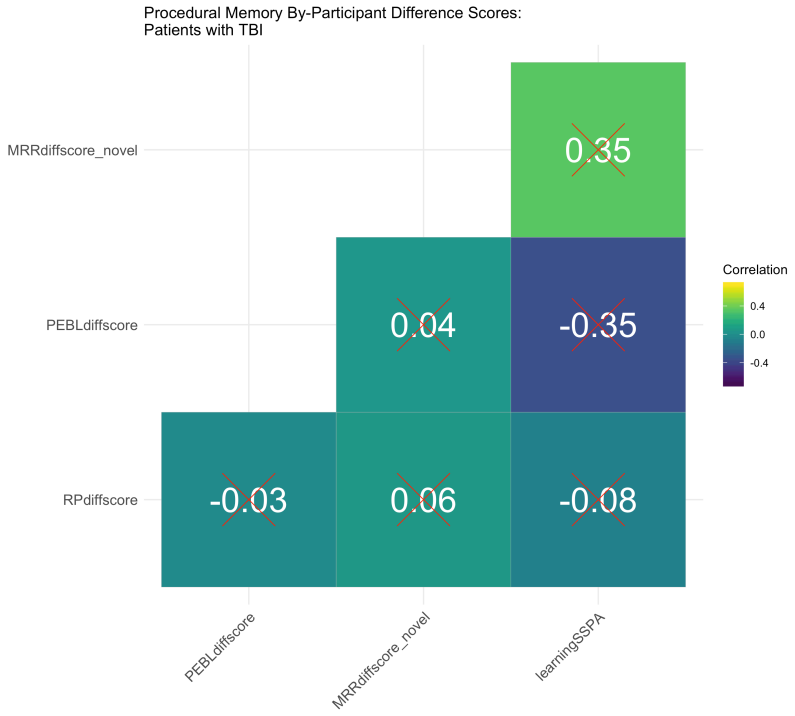

Figure 47. Correlations between Difference Scores: Patients with TBI
Red Xs indicate nonsignificant correlations.

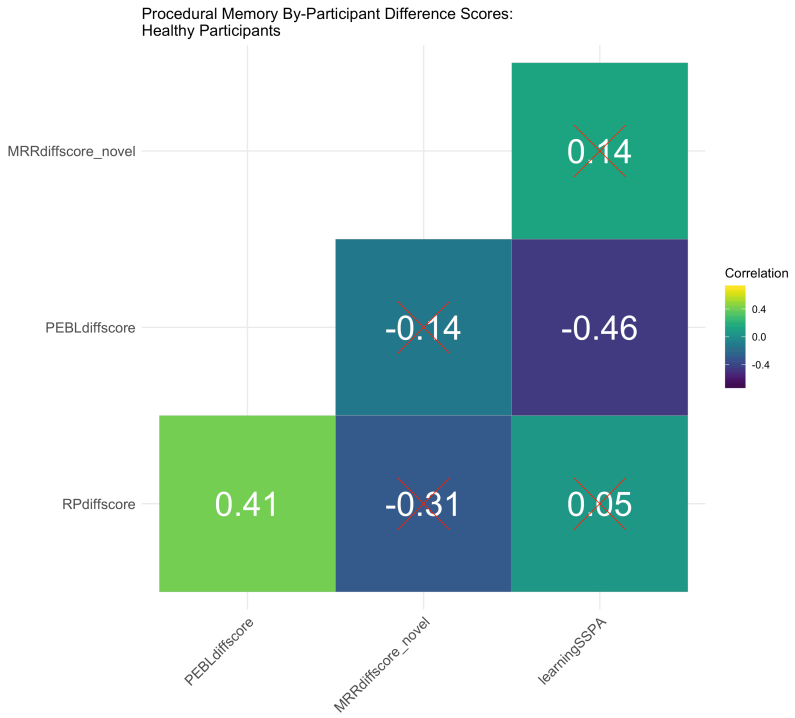

Figure 48. Correlations between Difference Scores: Healthy Adults
Red Xs indicate nonsignificant correlations.

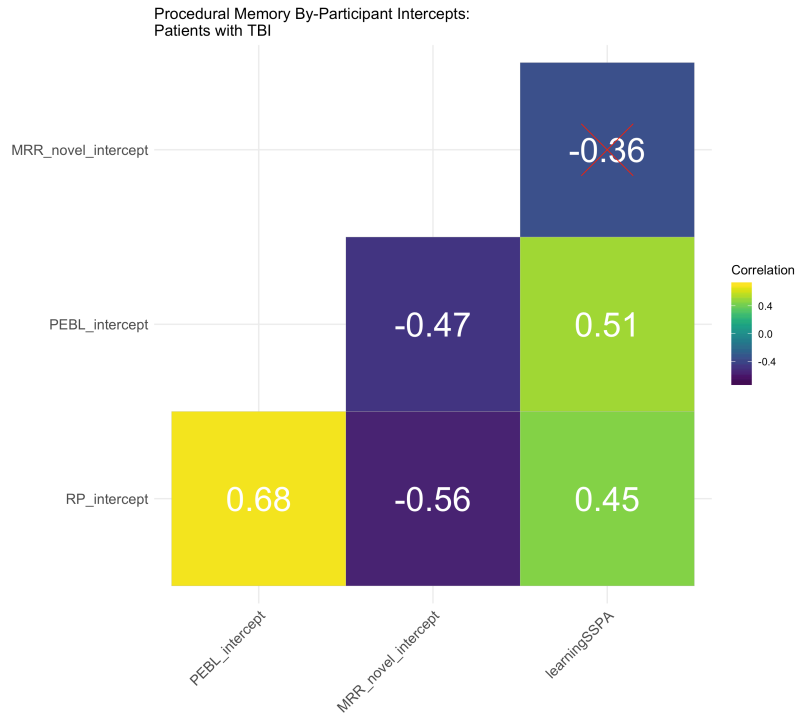

Figure 49. Correlations between Intercepts: Patients with TBI
Red Xs indicate nonsignificant correlations.

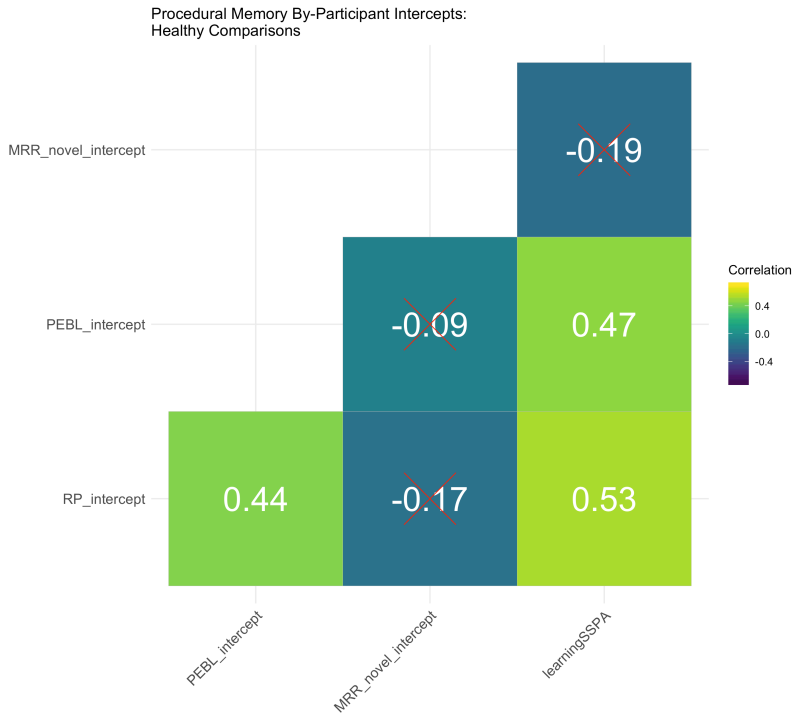

Figure 50. Correlations between Intercepts: Healthy Adults
Red Xs indicate nonsignificant correlations.

To address Research Question 2c., correlations between metrics for the photoelectric rotor pursuit task were compared to those for the mouse-tracking rotor pursuit task. Contrary to predictions, there were no significant correlations between either the slope metric or the difference score metric (the two metrics indexing amount of learning; all $|r|s < 0.2$, all $ps > 0.2$). By-participant intercepts, in contrast, were strongly correlated ($r = 0.57$; $t(44) = 4.58$, $p < 0.001$).

Spatial Reconstruction.

There were significant correlations between the NIH Toolbox Picture Sequence Memory Test (PSMT) and each of the analyzed metrics from the Spatial Reconstruction task (Figure 51). There was a strong negative correlation ($r = -0.52$) between participants' average misplacement (calculated on post-transformed data) and PSMT scores ($t(36) = -3.66$, $p = 0.001$). There was a

strong positive correlation (r = 0.59) between PSMT scores and the number of exact hits, calculated prior to global transformations (t(36) = 4.40, p < 0.001). There was a strong positive correlation (r = 0.62) between PSMT scores and the number of exact hits, calculated after all global transforms had been performed (t(36) = 4.72, p < 0.001). Participants' number of valid placements (irrespective of object identity) were also positively correlated with PSMT scores (r = 0.51; t(36) = 3.57, p = 0.001).

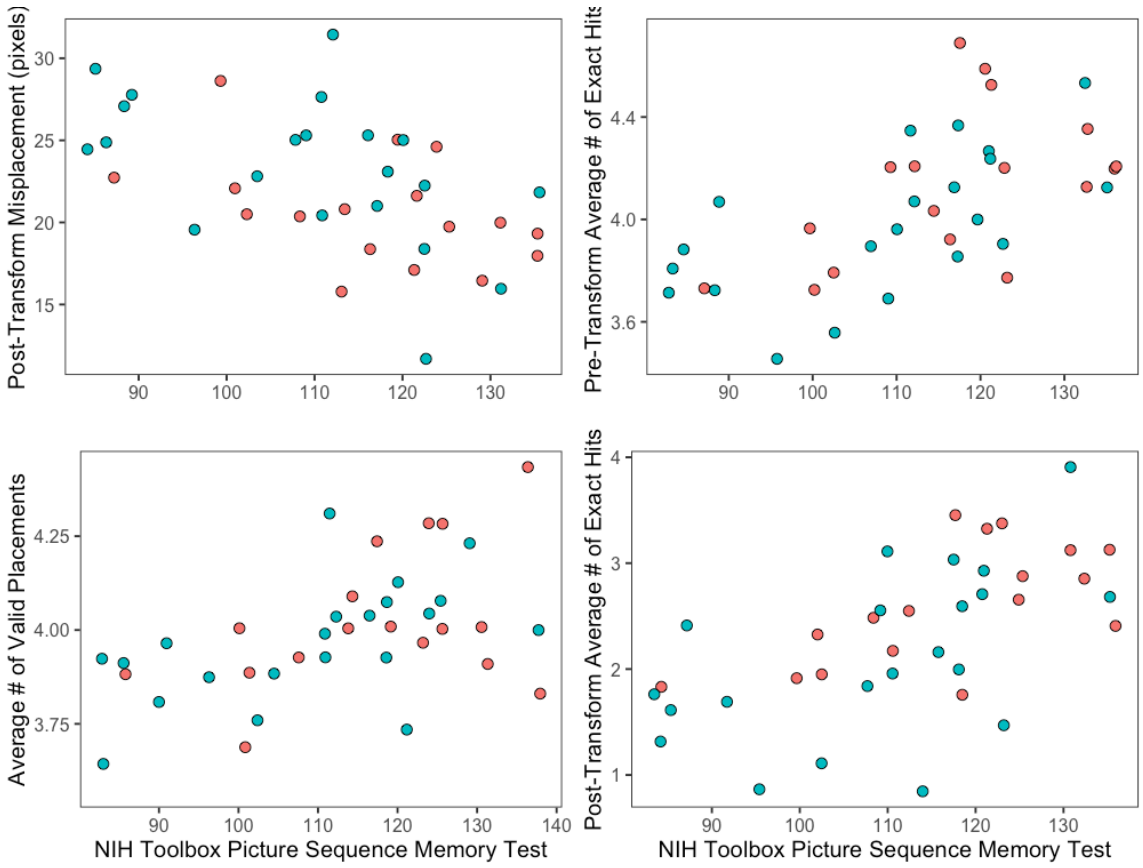

Figure 51. Spatial Reconstruction and Declarative Memory

There were significant correlations between the NIH Toolbox List Sorting Working Memory Test (LSWM) and three of the analyzed metrics from the Spatial Reconstruction task

(Figure 52). There was a strong negative correlation (r = -0.62) between participants' working memory scores and their degree of post-transform misplacement (t(36) = -4.71, p < 0.001). There was a medium positive correlation (r = 0.33) between working memory and participants' average number of exact hits, without global transforms (t(36) = 2.08, p = 0.04). There was a medium positive correlation (r = 0.38) between participants' average number of post-transform exact hits and their LSWM scores (t(36) = 2.51, p = 0.02). There was no significant correlation between participants' average number of valid placements and their LSWM scores (r = 0.22, t(36) = 1.34, p = 0.19).

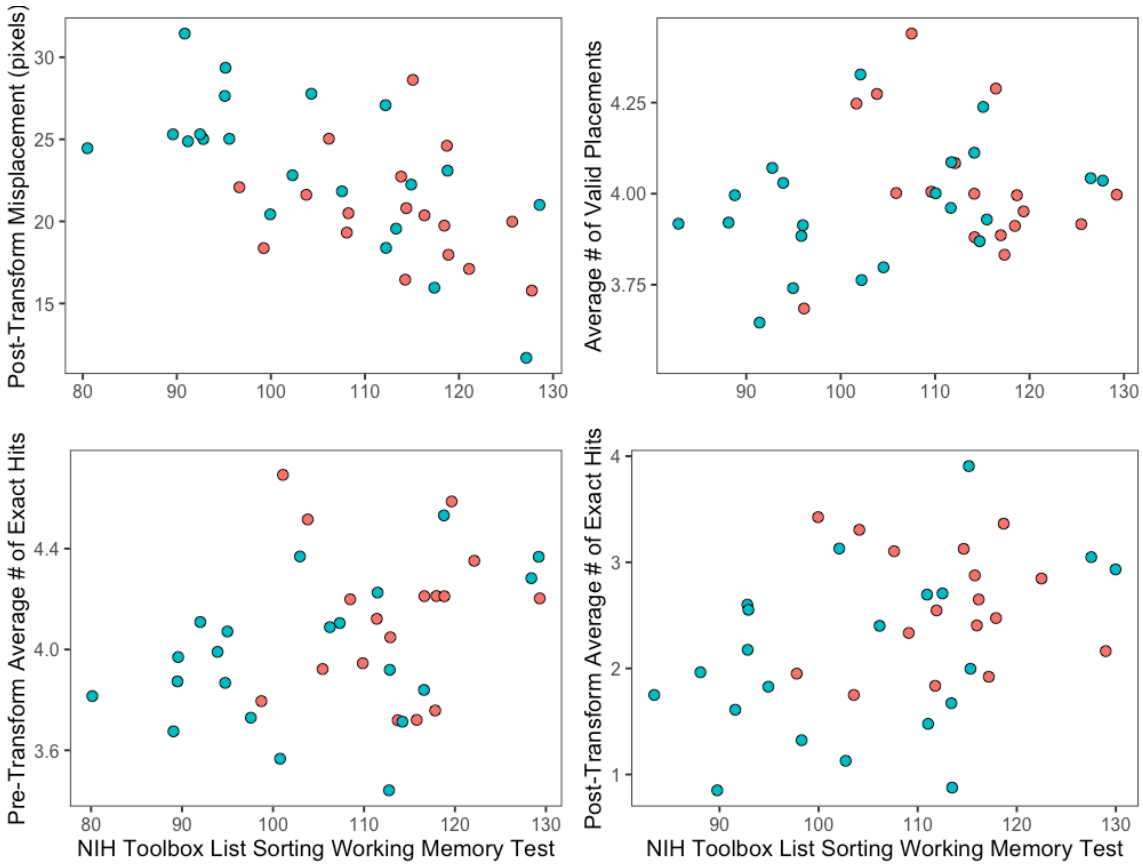

Figure 52. Spatial Reconstruction and Working Memory

There were significant correlations between the NIH Toolbox Flanker Inhibitory Control and Attention Test (ICAT) and each of the analyzed metrics from the Spatial Reconstruction task (Figure 53). There was a strong negative correlation ($r = -0.57$) between participants' average misplacement (calculated on post-transformed data) and ICAT scores ($t(36) = -4.18$ $p < 0.001$). There was a medium positive correlation ($r = 0.44$) between ICAT scores and the number of exact hits, calculated prior to global transformations ($t(36) = 2.95$, $p = 0.006$). There was a medium positive correlation ($r = 0.46$) between ICAT scores and the number of exact hits, calculated after all global transforms had been performed ($t(36) = 3.13$, $p = 0.003$). Participants' number of valid placements (irrespective of object identity) were also positively correlated with ICAT scores ($r = 0.47$; $t(36) = 3.16$, $p = 0.003$).

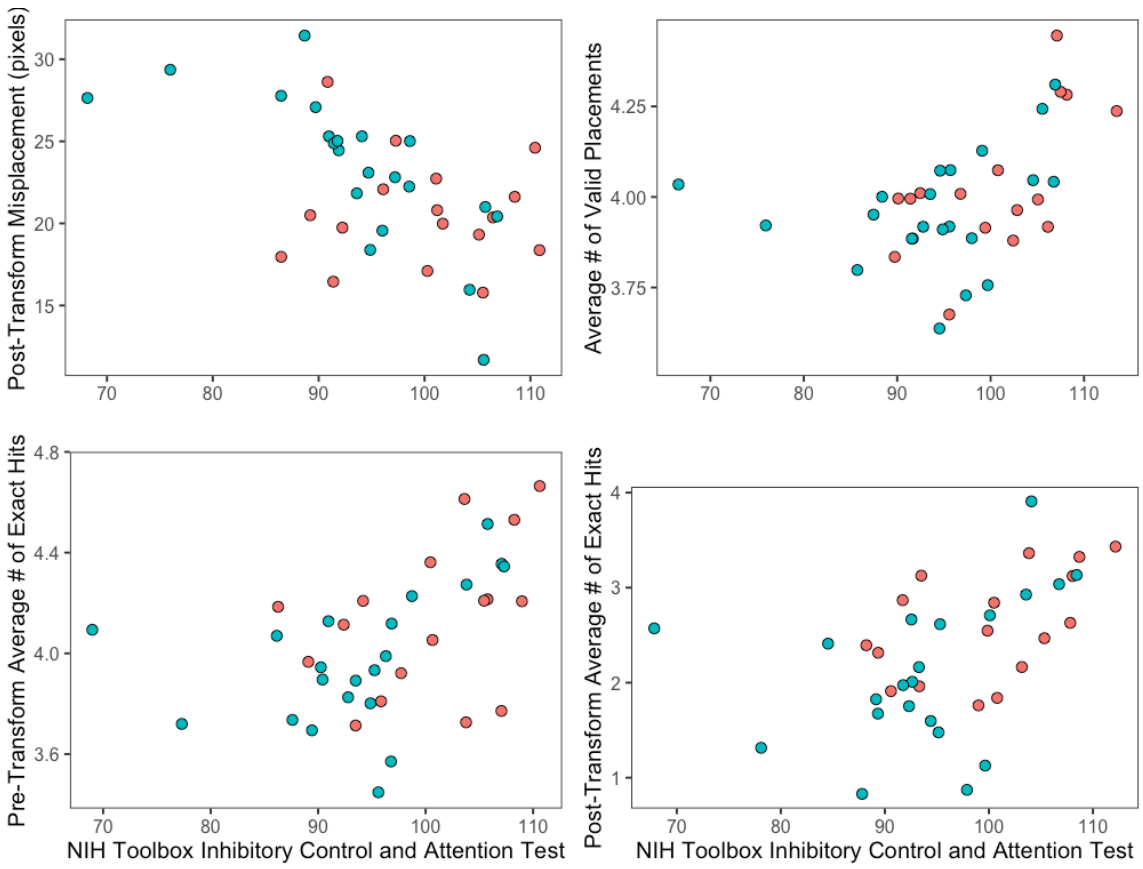

Figure 53. Spatial Reconstruction and Attention

There were no significant correlations between the NIH Toolbox Picture Vocabulary Test and any of the spatial reconstruction metrics (Figure 54, all $|r| < 0.3$, all $ps > 0.08$).

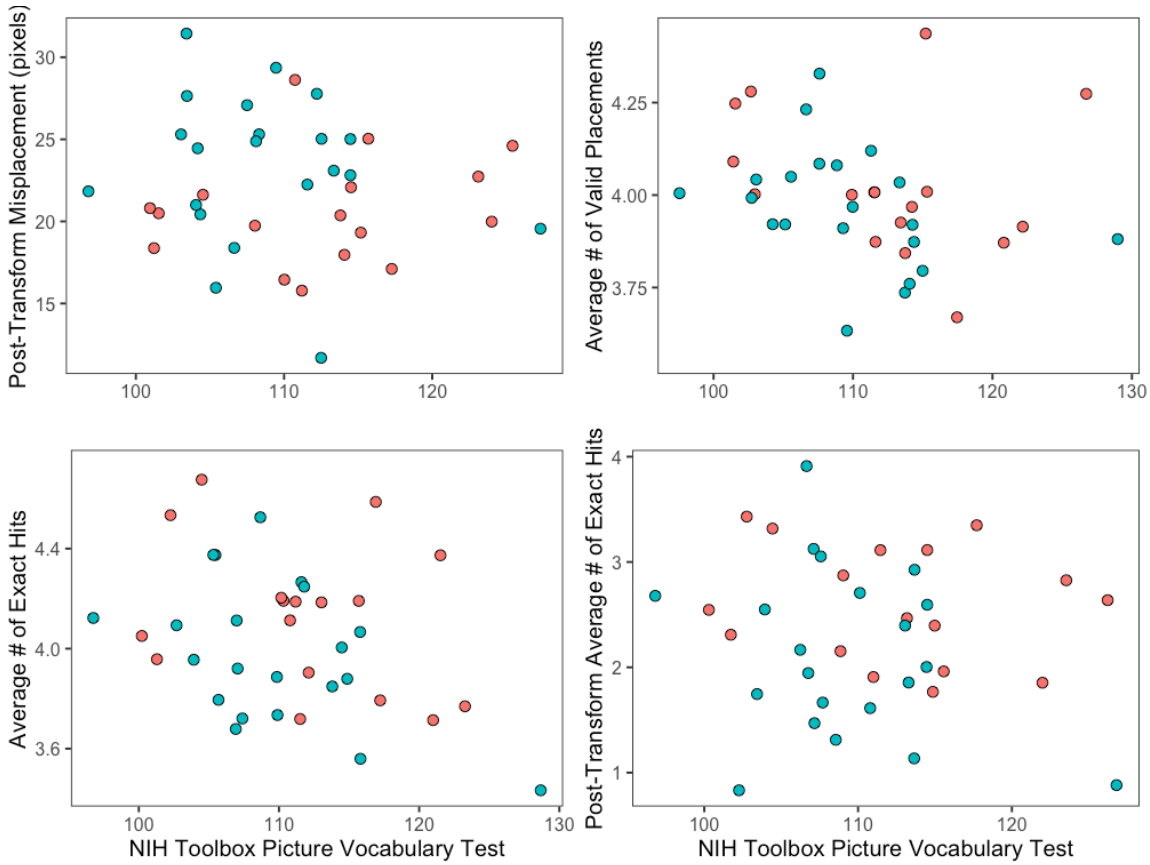

Figure 54. Spatial Reconstruction and Vocabulary

Multiple regression and partial R^2 was utilized to determine which cognitive abilities, as measured by the NIH Toolbox subtests, best predicted each spatial reconstruction metric. Each model included predictors for declarative memory, working memory, and attention. Working memory ($p = 0.009$) and attention scores ($p = 0.01$) were significant predictors of participants' amount of misplacement, calculated on post-transform data, with working memory accounting for 15.97% of the observed variance in misplacement, and attention accounting for 13.19% of the variance in misplacement. Declarative memory scores were the only significant predictor of participants' average number of exact hits, calculated on pre-transformed data ($p = 0.002$), and on post-transformed data ($p = 0.001$), accounting for 22.80% of the observed variance in exact

hits on pre-transformed data and 24.03% of the variance in exact hits calculated on post-transformed data. Declarative memory ($p = 0.008$) and attention ($p = 0.02$) were significant predictors of participants' average number of valid placements (irrespective of object identity), with declarative memory accounting for 19.20% of the observed variance and attention accounting for 14.37% of the variance.

Mirror-Reversed Reading: Repeating Triads.

Participants' by-participant slopes on the repeating word triads in the mirror-reversed reading task were positively correlated ($r = 0.30$) with their declarative memory ability, as measured by the NIH toolbox, although this correlation did not reach significance ($t(38) = 1.91$, $p = 0.06$). In contrast, participants' by-participant slopes on the novel word triads were negatively correlated ($r = 0.10$) to their declarative memory ability ($t(38) = 0.62$, $p = 0.54$). When declarative memory ability was indexed using the spatial reconstruction task, there was a significant positive correlation ($r = 0.39$) between by-participant slopes on repeating triads and participants' average accurate placements, calculated on post-transformed data ($p = 0.005$).

CHAPTER IV

DISCUSSION

A central problem for basic and clinical research in TBI is the inherent heterogeneity in the patient population. Every TBI results in a unique constellation of primary and secondary insults to neural tissue, many of which are not amenable to visualization with current clinical imaging technology. Cognitive and behavioral profiles following TBI also vary widely, even for patients who have sustained injuries with similar etiologies, with similar severity, and with similar imaging findings (Saatman et al., 2008).

In rehabilitation contexts, clinicians rely on clinical assessments across various cognitive domains to characterize these sources of heterogeneity and make decisions regarding appropriate intervention methods and goals. A primary goal of the current project is to provide a foundation for the eventual development of learning and memory phenotypes for individuals with TBI, to facilitate better clinical decision-making. The rationale behind this overarching goal is that all clinical interventions implemented by speech-language pathologists require patients to learn or relearn some set of knowledge and/or skills. Thus, consideration of a patient's particular profile of impaired and intact learning and memory ability may be crucial to ensuring that cognitive rehabilitation is as effective as possible.

The current project provides an initial dataset as a foundation for the larger goal of developing learning and memory phenotypes in TBI. The project tests learning and memory abilities across a basic distinction in the cognitive neuroscience of memory: declarative memory (memory for facts and events) and procedural memory (memory for skills and habits).

Characterizing patients' degree of impairment and ability across these memory systems is motivated by several existing intervention strategies which suggest that patients with an impairment in one memory system can compensate for these deficits using intervention methods and strategies that leverage the other (presumably intact) memory system.

A basic requirement for the development of behavioral phenotypes is the development of measures and assessments that characterize ability across the full range of heterogeneous presentations, that are valid measures of each underlying cognitive construct, and that provide scores that are reliable within a given individual and across multiple testing sessions. In the current project, I have evaluated a set of experimental measures across multiple memory systems, examining group and individual differences in these measures.

The choice of experimental measures was intentionally unbalanced: there are many standardized, nationally-normed assessments of declarative memory that are clinically useful, valid, and reliable. In contrast, there is no clinically available assessment of procedural memory, nor standardized or validated measures. In addition, while it is clear that many individuals with TBI have impairments in declarative memory ability (even following mild TBI, Geary, Kraus, Pliskin, & Little, 2010), the impact of TBI on procedural memory is less clear. A number of studies suggest intact procedural memory (Ewert, Levin, Watson, & Kalisky, 1989; Ward, Shum, Wallace, & Boon, 2002), leading to the development of intervention strategies that purport to leverage "intact" procedural memory to support learning in the context of declarative memory impairments. However, other studies (Kraus et al., 2010; Vakil et al., 2002) suggest that procedural memory is not uniformly intact across patients with TBI. Thus, the current project sought to assess procedural memory ability in patients with TBI to examine group differences as well as inter-individual variability and its association with other cognitive abilities.

Results of each domain/task are discussed in turn, below, followed by discussion of the overarching research questions and future directions.

Declarative Memory

Spatial Reconstruction.

As a group, patients with TBI demonstrated significantly poorer declarative memory for the studied locations of specific novel objects. This impairment was evidenced by larger misplacement distances for both pre- and post-transformed data, and, on average, fewer exact hits on both pre- and post-transformed data. In both groups, performance decreased as set size increased, but the effect of set size was similar across groups. Although patients with TBI were impaired on metrics that required memory for object identity-location relations, they demonstrated preserved "gestalt" memory for the relative valid locations of objects. Consistent with the findings of Horecka and colleagues (2018), these two measures appear to be related but dissociable constructs. Participants' memory for the gestalt arrangement of valid locations was predicted by their attention and working memory ability, as measured by the NIH Toolbox. Participants' ability to place objects in their particular studied location (exact hits) was predicted only by their declarative memory ability, as measured by the NIH Toolbox.

Procedural Memory

Photoelectric Rotor Pursuit.

As a group, participants demonstrated significant learning of a new procedural motor skill, improving in time-on-target across trials. Modeled with a maximal random effects structure, there were significant inter-individual differences in slopes, and no overall group difference in learning over time. When the data were modeled without the inclusion of random by-participant slopes, there was a significant group-by-trial interaction, such that patients with TBI demonstrated poorer learning of the skill than healthy comparison participants. Participants completed other experimental tasks between trials 8 and 9, resulting in an average intervening interval of about an hour. Across this interval, as a group participants demonstrated a small increase in time-on-target. There was no significant relationship between the amount of time elapsed between trials 8 and 9 and the amount of retention demonstrated by participants.

The two groups significantly differed in the speed at which they completed test trials, with patients with TBI more likely to complete the task at a slower rotation speed, based on their baseline testing. However, the baseline procedure for determining each participant's speed of rotation during testing adequately equated motor performance across participants: there was no significant difference in time-on-target between groups during the first test block, and participants' performance during the first trial (random intercept) was not significantly associated with their eventual degree of learning (random slope).

In contrast to Rigon and colleagues, there was no main effect of group on time-on-target across trials. As discussed below (see Sampling Across Heterogeneity, below), the patients included in the current sample represent relatively successful long-term outcomes. As can be

seen in Figure 18, many patients with TBI performed well within the healthy participant distribution. The difference in findings between the current study and Rigon et al., highlights the importance of significantly larger sample sizes in TBI research. The inclusionary criteria for the Rigon study and the current study were nearly identical, and the procedures for implementing the photoelectric rotor pursuit task were the same.

Mouse-tracking Rotor Pursuit.

The mouse-tracking (PEBL) version of the rotor pursuit task differed from the traditional photoelectric paradigm in two key ways: 1) there was no initial baseline testing to equate performance during the initial block across participants; and 2) the task was completed across two different conditions, one in which participants were not provided with any feedback regarding their performance (as in the photoelectric version of the task) and another in which participants received moment-by-moment feedback about their performance, as the target became illuminated when they were successfully directly on top of it. A model with maximal fixed and random effects structure indicated that there were no significant effects of Feedback on performance, so all subsequent analyses were conducted on data collapsed across feedback conditions. As a group, participants demonstrated significant learning of the new procedural motor skill, improving in their time on target across trials. In contrast to the photoelectric version of the task, there was a significant quadratic effect of trial, such that participants' relative improvement decelerated over time. Contrary to predictions, there was no significant interaction between group and either of the growth terms, indicating that there was no overall group difference in learning over time. There was a significant main effect of group, driven by worse

performance overall by patients with TBI. As in the photoelectric rotor pursuit tasks, there were significant inter-individual differences in growth terms.

Although there was significant improvement in time-on-target across trials, the lack of baseline testing for the mouse-tracking rotor pursuit task may have resulted in ceiling effects for some participants. Baseline testing in the photoelectric version of the task meant that participants were assigned a testing rotation speed based on the rotation speed for which they were closest to 5 seconds of time-on-target, which represents 25% of the total possible time-on-target for a single trial (5 / 20 sec per trial). However, participants' time-on-target during the first trial of the photoelectric version of the task was 6.18 for NCs (31% of possible time-on-target) and 6.23 for TBIs (31%). This is in part due to a number of participants whose ability to maintain contact with the target during baseline testing was well over 5 seconds, even at the fastest rotation speed, and likely because some participants began to learn the skill during the baseline testing trials. In contrast, for the mouse-tracking version of the task, participants began testing at a level that was a greater proportion of the possible time-on-target, with patients with TBI's average time-on-target 8.63 seconds (58% of 15 possible seconds), and healthy comparisons beginning at an average of 10.10 time-on-target (67% of total possible). Thus, this substantial difference in the amount of "room to grow" in the mouse-tracking version of the task, likely led to ceiling effects, particularly for healthy comparison participants.

Mirror-Reversed Reading.

As a group, participants demonstrated significant learning of a new procedural cognitive-perceptual skill, improving in reading times across trials for novel word triads. Modeled with a maximal random effects structure, there were significant inter-individual differences in slopes,

and no overall group difference in learning over time (no group-by-growth interaction). However, patients with TBI generally performed more poorly across the task compared to healthy comparison participants. There was a significant interaction between condition and the growth terms, indicating that reading times for repeating triads decreased more rapidly compared to novel word triads.

When trial types were considered separately, there were significant improvements over time across participants for both trial types (novel and repeating). Significant improvement over time on novel triads suggests learning of the procedural cognitive-perceptual skill of mirror-reversed reading. Significant improvement over time on repeating triads suggests a combination of repetition priming and new declarative learning of repeating triads. Consistent with the findings of Cohen and Squire (1980), participants' degree of learning on repeating triads was positively associated with their declarative memory ability, with a significant positive correlation between participants' repeating triad slopes and their performance on the spatial reconstruction task.

In general, participants read stimulus words accurately, with increasing accuracy over the course of the task. Accuracy increased more rapidly over time for repeating versus novel triads for healthy comparison participants compared to patients with TBI. Triads that contained one or more reading errors were generally read more slowly than triads for which participants did not make an error. Participants' overall accuracy across the task was significantly correlated with their overall word reading ability as measured by the WRAT Reading Subtest.

Visual appraisal of by-participant plots suggests that many participants may have experienced fatigue or floor effects during the second half of the task (after the break). During the first half of the task (average read time in block 5 minus block 1), participants improved on

average by 3.82 seconds (SD = 7.19) on novel triads, and by 13.55 seconds (SD = 6.96) on repeating triads. In contrast, during the second half of the task (average read time in block 10 minus block 6), participants' reading times *slowed* by an average of 0.92 seconds (SD = 7.03) on novel triads, and only improved by an average of 1.19 seconds (SD = 2.18) on repeating triads.

Serial Interception Sequence Learning.

During training trials, participants demonstrated learning of the underlying pattern, indicated by significantly more accurate responses to sequence trials compared to foil trials during blocks two through four. Neither group demonstrated a sequence-specific performance advantage during the implicit test blocks.

There was no difference between groups in the average speed at which the task was completed, although there were significant effects of age and sex. In both groups, participants' average SSPA during learning blocks was significantly correlated with their initial ability on the both rotor pursuit tasks. This suggests that individual SSPAs may be indexing a participant's underlying motor ability, rather than their procedural learning of the underlying pattern.

Research Question 1a

Are individuals with moderate-severe TBI impaired on the Spatial Reconstruction task relative to healthy comparison participants?

As predicted, individuals with TBI performed significantly more poorly on the spatial reconstruction task compared to healthy comparison participants. In particular, patients with TBI demonstrated impairments on aspects of the task that have previously been associated

specifically with hippocampal dysfunction; that is, patients with TBI demonstrated a disruption specifically in the assignment of particular objects to their studied locations, but did not demonstrate impairments in their ability to reconstruct the general, gestalt "shape" of the object-object relations in space. Across both groups, performance declined as set size increased, but the groups did not differ in the degree to which set size influenced their accuracy.

Research Question 1b

Does the Spatial Reconstruction Task have increased sensitivity in detecting and characterizing declarative memory ability, over traditional neuropsychological measures, in individuals with moderate-severe TBI?

To date, the spatial reconstruction task has been developed and tested on individuals with severe declarative memory impairments due to focal bilateral hippocampal damage (Horecka et al., 2018; Watson et al., 2013). A version of the task has also been implemented in healthy older and younger adults, with performance differentiating between the two groups (Clark et al., 2017), suggesting that this task may be a useful measure for declarative memory that might capture the range of possible memory impairments following traumatic brain injury. While neuroimaging suggests that TBI disrupts hippocampal structure and function, even following mild injuries (J. J. Chang et al., 2009; de la Plata et al., 2011; Geary et al., 2010; Leh et al., 2017), and patients with TBI often report subjective memory concerns that are not captured by neuropsychological measures, these disruptions may not be fully captured by existing standardized declarative memory assessments.

Indeed, in the current sample of individuals with TBI, patients with TBI did not significantly differ from healthy comparison participants on a standardized measure of declarative memory (the NIH Toolbox Picture Sequence Memory Test). Despite memory ability "within the normal range", patients with TBI demonstrated significantly impaired performance on the spatial reconstruction task relative to healthy comparison participants. This suggests that the spatial reconstruction task may be a more sensitive measure of declarative memory, identifying subtle disruptions in memory function that go unnoticed by commonly used assessments of declarative memory. To be clear, there are existing standardized assessments of declarative memory that likely provide nuanced characterization of patients' memory profiles (e.g. Weschler Memory Scales). However, these assessments require significant time to administer (> 60 minutes), and thus are infrequently administered in clinical settings to characterize patients' memory impairments, where clinicians must balance assessment time between multiple cognitive complaints. The timing of the spatial reconstruction task (~30 minutes) is a significant advantage for its use in clinical practice. The spatial reconstruction task may by more sensitive to subtle disruptions in declarative memory and may be useful for identifying patients for whom memory performance has changed but have remained within the "range of normal" on simpler memory tasks.

In Horecka et al., 2018, the authors chose to use healthy comparison performance to determine the size of accuracy windows for each patient with hippocampal amnesia. Here, accuracy windows were determined on each participant's data, aggregated across all trials, meaning that accuracy windows were not directly equated across groups. There are benefits and drawbacks to both approaches. In the first, the bar for "accuracy" is equivalent in both groups, but may be biased based on a particular healthy individual's ability on the task. With the

approach taken in the current study, patients may have been given more leeway in their placements if their performance is generally nosier than comparison participants. That patients' placements are generally "noisier" than healthy comparisons was evident in pre-transformed data: when accuracy windows were computed solely on participants' raw placement data, accuracy windows were significantly larger for patients with TBI compared to healthy comparison participants. However, global transformations appeared to adequately address this issue: when accuracy windows were computed based on participants' data *after* global transformations, the diameter of accuracy windows did not significantly differ between groups.

The utility of computing accuracy windows on a by-participant basis versus the development of "normative" accuracy windows should be the focus of future studies in the development of the spatial reconstruction task as a valid and reliable clinical measure of declarative memory. A large sample norming study, in which accuracy windows are computed for many individuals with similar demographic characteristics (e.g. age, sex, education) may be useful for equating scoring methods across healthy and clinical populations.

A key strength of the spatial reconstruction task is its ability to capture both group differences and meaningful individual differences. This quality of the task may be influenced by the set size manipulation, which allows for characterization of memory ability across the continuum of possible performance, without ceiling or floor effects. The version of the spatial reconstruction task implemented in the current study included larger set sizes than had been implemented in previous studies, in an attempt to capture subtle deficits in declarative memory. A post-hoc analysis of the data, excluding set sizes 8 and 10 (to more closely mirror previous implementations of the task), revealed the same significant effects, but with a larger effect size for the group effect indicating larger differences between the two groups. However, examining

correlations between the smaller set-size dataset and NIH Toolbox measures yielded weaker correlations been by-participant spatial reconstruction metrics and participants' other cognitive abilities. This likely reflects the reduced variability in performance across smaller set sizes. This difference in the effect of set size on the results highlights a key distinction between what is required for the identification of group-level effects and what is required for identifying meaningful individual differences.

Adapting the spatial reconstruction task to a computer-adaptive format may allow for probing of performance across set sizes in an efficient manner and contribute to the utility of the task as a clinical assessment. A computer adaptive version of the task would likely reduce task length, and help to avoid patient frustration or fatigue by eliminating larger set size trials if the patient demonstrates consistent impairment at smaller set sizes. In addition, allowing high-performing individuals to reach increasingly larger set sizes would allow for characterization of exceptional declarative memory ability in healthy individuals and patients of various etiologies and allow for the examination of the impact of exceptional memory ability on real-world behavior and outcomes.

Another benefit of the spatial reconstruction task for assessing declarative memory in clinical settings is that the task does not require active clinician involvement during administration of the task. This would allow for administration of the task by support personnel (e.g. speech-language pathology assistants), freeing up skilled clinician time. Alternatively, as "point of care" / "point of service" documentation becomes an increasing necessity for clinicians in medical settings with productivity requirements, the lack of clinician involvement in the administration of the spatial reconstruction task allows clinicians to simultaneously collect important clinical data while they document results of previous assessments. Finally, automated

scoring methods would eliminate clinician time spent hand-scoring standardized assessment results.

Finally, the tablet-based spatial reconstruction task is easily portable, for use as a bedside assessment, and does not require a verbal response, which increases its utility for patients with concomitant communication impairments. With the addition of a large normative dataset, computer adaptive testing, and an implementation of the current error metric pipeline on the tablet itself, the spatial reconstruction task would provide an efficient, clinically feasible measure of declarative memory that could be easily implemented across multiple clinical settings (at bedside, in outpatient settings, in schools, and in home heath contexts).

The current project has contributed to the characterization of declarative memory impairments in TBI. Although patients with TBI did not significantly differ from comparison participants on the NIH Toolbox measure of declarative memory, they significantly differed from healthy participants on all spatial reconstruction metrics associated with hippocampal pathology and declarative memory impairment. A second contribution of the current study is in extending the task to include larger set sizes (i.e. 6, 8, and 10 objects) compared to previous studies (Clark et al., 2017; Horecka et al., 2018). The inclusion of larger set sizes helped to prevent ceiling effects: following global transformations, even healthy participants were, on average, < 40% accurate in their placements. However, gestalt memory for the relative arrangement of valid locations remained high across set sizes: even at the 10 set size, both groups placed, on average, > 60% of items in a valid location. Consistent with previous findings in patients with focal hippocampal damage, assignment of a particular object to its studied location (rather than placement to any, valid location) appears to depend critically on an individual's declarative memory ability: declarative memory was the only significant predictor of exact hits, both pre-

and post-transform. Other aspects of the task, such as reconstructing the relative, gestalt "shape" of the objects (object-object relations), and the average degree of misplacement away from studied locations, were predicted by participants' declarative memory, working memory, and attention in tandem. Indeed, the spatial reconstruction task likely draws on interacting cognitive systems, but the error metrics as developed by Horecka and colleagues produce robust, individual differences that appear to index distinct memory processes. In future studies, examining associations between these error metrics and neuroanatomical data will help to clarify and confirm these relationships.

An open question is the degree to which cognitive abilities beyond declarative memory influence spatial reconstruction performance. Results of this study suggest that working memory and attention abilities influence performance, particularly for participants' memory for gestalt location information. Future studies should examine whether other factors might influence performance, influencing the validity of the task or obscuring declarative memory performance. One question is whether the use of particular cognitive strategies influences performance on the task. For example, several participants reported attempting to use verbal strategies to remember specific object locations, by assigning "names" to the novel objects during the study phase. Anecdotally, participants indicated that this strategy was significantly more difficult/less effective at larger set sizes; perhaps larger set sizes decrease the potential influence of differing cognitive strategies. It is also possible that, as set sizes increase and it becomes more difficult to remember specific object-location information, participants shift to a strategy that emphasizes reconstruction of the gestalt location information, and pay less attention to object identity information. This may explain why the proportion of exact hits consistently declines across set sizes, while the proportion of placements to valid locations remains more consistent as set sizes

increase. Another possible influence on task performance is participant motivation and willingness/ability to exert cognitive effort. Future studies should examine the influence of motivation and effort on spatial reconstruction task performance.

Research Question 2a

Are individuals with TBI impaired on a battery of experimental procedural memory tasks relative to healthy comparison participants?

Contrary to predictions, as a group, patients with TBI did not demonstrate differences in procedural learning over time on any of procedural memory tasks relative to healthy comparison participants. On the mouse-tracking rotor pursuit and mirror-reversed reading tasks, there was a main effect of group, such that patients with TBI performed more poorly than healthy comparison participants overall, but did not differ in their degree of learning over time (slope) compared to healthy participants.

Previous studies have described impairments in procedural memory in patients with TBI (Kraus et al., 2010; Vakil et al., 2002). One possibility is that differences across studies result from a lack of sampling across the full range of heterogeneity in patients with moderate-severe TBI (see Sampling Across Heterogeneity, below). Another possibility is that a small subset of patients with TBI have true procedural memory impairments, but this subset is obscured by generally intact performance of the group as a whole. To visualize whether a subset of the TBI group might be performing outside of the distribution of performance in healthy comparison participants, a composite procedural memory score was computed, based on the average of by-participant z-scores across tasks, computed on slopes for rotor pursuit and mirror-reversed

reading tasks and average SSPA for the serial interception sequence learning task (note that z-scores for the MRR task were multiplied by -1, since negative slopes on this task indicate learning, compared to positive slopes/scores for the remaining tasks). As can be seen in Figure 55, the distribution of composite scores for patients with TBI overlaps with that of healthy comparison participants, but with a greater number of scores below 0 (indicative of below-average learning), while the majority of healthy comparison participants had composite scores above 0 (indicative of better-than-average learning). Sixty-one percent of patients with TBI had composite scores below zero, compared to thirty percent of healthy comparison participants.

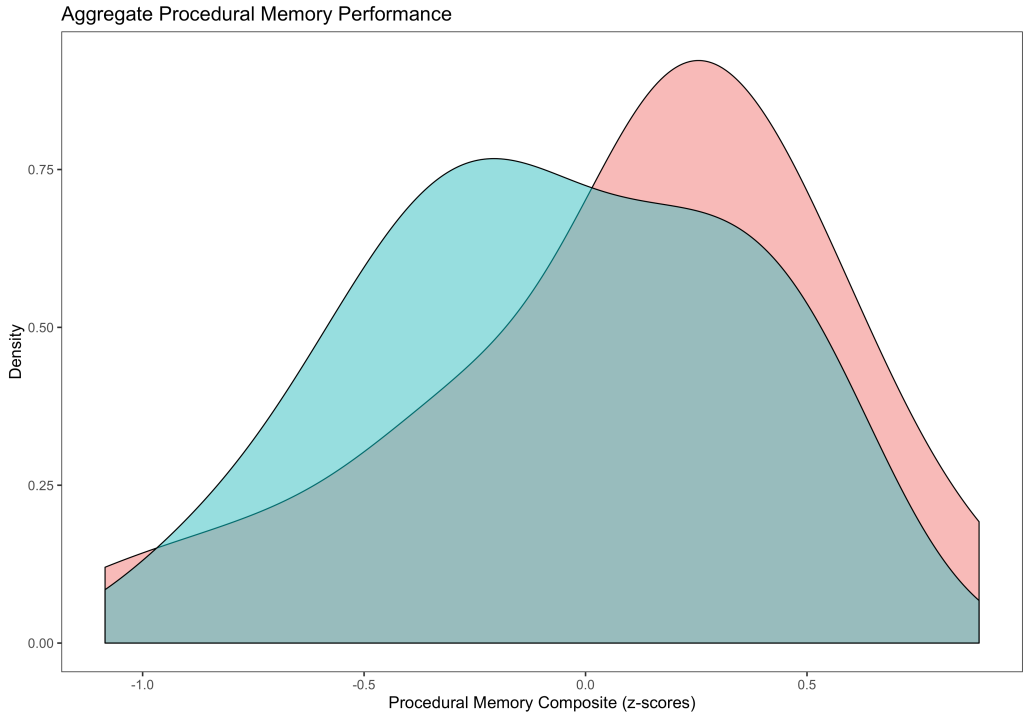

Figure 55. Procedural Memory Composite Distribution

This supports the possibility that a subset of patients with TBI may demonstrate impaired procedural learning, though the sample as a whole demonstrates similar procedural learning compared to healthy comparison participants.

While the possibility that a subset of individuals with TBI do have true procedural memory deficits should be explored, in larger samples and with better developed assessments, an encouraging interpretation is that the majority of individuals with TBI do have intact procedural memory ability. This would align with previous studies demonstrating similar procedural memory abilities in patients with TBI compared to healthy adults (Ewert et al., 1989; Ward et al., 2002), and would provide support for existing treatment strategies and programs that seek to leverage procedural memory deficits to compensate for declarative memory deficits (Glisky, 1993; Ylvisaker & Feeney, 1998). Adjudicating between these differing interpretations will require a better understanding of how procedural memory is assessed and quantified; the lack of significant correlations between procedural memory measures (discussed below), provides a significant challenge to the validity of these widely used tasks.

Finally, an open question is whether patients with TBI might demonstrate deficits in procedural learning of new skills over longer intervals. That is, would patients with TBI reach an asymptote earlier in their long-term learning of a particular skill than would healthy comparison participants, or would they show less retention or consolidation of a particular skill over multiple testing sessions. Although both the photoelectric rotor pursuit and mirror-reversed reading tasks involved an intervening interval of at least a half an hour between halves of the task, this interval is likely not sufficient to reveal potential differences in the amount of retention of a newly learned skill that may emerge with longer intervals between testing sessions. Because procedural memory is characterized by slower acquisition over time compared to declarative memory, future

studies should examine procedural learning across more prolonged time periods and with differing amounts of time between testing sessions.

Research Question 2b.

Is there a relationship in performance across procedural memory tasks?

Contrary to predictions, there were no significant associations among the different procedural memory measures, as examined using both by-participant slopes and difference scores between the final and initial blocks of each task.

Few studies have directly compared performance across procedural memory tasks. Two studies examined correlations between procedural memory tasks (Gebauer & Mackintosh, 2007; Horan et al., 2008). Horan and colleagues tested artificial grammar learning and probabilistic classification tasks in a group of patients with schizophrenia, with no significant correlation between these two tasks. Gebauer and Mackintosh examined correlations between artificial grammar learning and serial reaction time tasks, and found no significant correlation. Artificial grammar learning and probabilistic classification tasks were not selected for the current study because of known contributions of declarative memory ability to each of these tasks (Channon et al., 2002; Knowlton, Squire, & Gluck, 1994; Poldrack & Foerde, 2008). Thus, while the lack of significant correlations between procedural memory tasks in the current study mirrors that in Horan et al. (2008) and Gebauer & Mackintosh (2007), it is perhaps more surprising, given the tasks reported here are more frequently described as "pure" procedural memory tasks.

Evidence from a recent study suggests some reason to be optimistic about the possibility of developing valid and reliable measures of procedural memory (Kalra, Gabrieli, Finn,

Madison, & States, 2019). In this study, Kalra and colleagues assessed the test-retest reliability of a battery of procedural memory tasks: artificial grammar learning, serial reaction time, probabilistic classification, and an implicit category learning task. Participants were tested on two counterbalanced alternate versions of each task with 1-4 intervening weeks between testing sessions. The serial reaction time, probabilistic classification, and category learning tasks demonstrated adequate to good test-retest reliability. Learning during the category learning and probabilistic classification tasks was significantly correlated. In a factor analysis, all three tasks loaded on a common factor. Results of this study suggest that with careful development and testing, it may be possible to develop a valid and reliable measure of procedural memory.

The mouse-tracking rotor pursuit and mirror-reversed reading tasks appear to have significant ceiling and floor effects, particularly for healthy comparison participants. These effects constrain the possible variability in slopes for individuals with higher levels of initial ability. Future iterations of these tasks could begin to account for these problems by better equating participant performance at the outset of each task. For example, the addition of baseline testing to the mouse-tracking rotor pursuit task may be sufficient to increase correlations between the slope scores of the two rotor pursuit tasks. Another alternative would be to develop normative data that takes into account initial ability. For example, for the rotor pursuit and serial interception sequence learning tasks, norms could be developed at each level of sex, age, education, and motor ability.

The procedural memory measures utilized here represent the “gold standard” in empirical studies of procedural memory. However, these tasks have historically arisen from studies that sought to contrast procedural memory with better-studied declarative memory. Procedural memory studies have often involved characterization of the memory system as simply memory

which is “not declarative.” Despite a wide range of conditions for which procedural memory is purported to be disrupted (and, in several disorders, theorized to be central to the disorder itself, e.g. Ullman & Pierpont, 2005), procedural memory tasks have not been carefully tested and developed for use in assessing stable individual differences. The data collected here suggests that such an endeavor will be a significant challenge. Despite ostensibly testing the same underlying construct, there were no significant correlations among the procedural memory tasks (as measured by individual slopes or using difference scores), as would be expected if the measures were valid assessments of a shared underlying ability.

One possibility is that any stable individual differences in actual procedural “learning” ability is masked by large individual differences in participants’ initial ability for a given skill. This possibility is suggested by a number of the current project’s findings. First, although there were no significant correlations among procedural memory tasks, measured using either by-participant slopes or difference score, there were significant correlations among participant’s initial performance (intercepts). Inter-task correlations were strongest between tasks that required similar underlying abilities; for example, intercepts for the two rotor pursuit tasks were the most strongly correlated, as well as correlations between each of the rotor pursuit tasks and the other motor procedural task: SISL. In addition, visual appraisal of individual performance during the mirror-reversed reading and rotor-pursuit tasks suggests that participants who begin each task with better initial performance have substantially less “room to grow” compared to participants who have poorer initial ability. Thus, these participants may reach ceiling levels of performance, reducing the magnitude of their individual slope or difference score.

Critical next steps in the potential use of these tasks as individual difference measures for research and in the clinic will be improving task validity and assessing each task’s reliability.

Determining how best to equate performance across participants, or how to adjust participants' degree of learning relative to their initial performance will be critical. Once reliability data is obtained for each of the tasks, this information can be used to calculate the standard error of measurement for each task. One possibility would be to use a metric like the Reliable Change Index to assign individual difference scores to each participant.

Research Question 2c.

Is there a relationship between the traditional, photoelectric implementation of the rotor pursuit task and a freely-available, computerized version of the task?

The rotor pursuit task is perhaps the most widely used measure of procedural memory across disorders (Don, Schellenberg, Reber, DiGirolamo, & Wang, 2003; Heindel, Butters, & Salmon, 1988; Hsu & Bishop, 2014; Lee & Tomblin, 2015). Traditionally, this task has been implemented using a specialized piece of equipment that records time-on-target using a photo-sensitive wand. The machine is bulky, immobile, and expensive, which would impede translation of the task to real-world clinical practice. A mouse-tracking version of the task has been available since 2012, but prior to the current study had not been directly tested against the traditional photoelectric version of the task. In the current sample, learning on the photoelectric rotor pursuit task (as measured using by-participant slopes and difference scores) did not significantly correlate with learning on the mouse-tracking version of the task. When difference scores on the two tasks were assessed across all participants, scores were very weakly correlated ($r = 0.13$); however, when only healthy comparison data was included in the analysis, difference scores were moderately correlated ($r = 0.41$) although this correlation did not reach significance ($p =$

0.05). In a larger sample, and with the addition of baseline testing to the mouse-tracking version of the task, it is possible that task demands might be better equated and reveal significant correlations between learning measures.

The mouse-tracking version of the rotor pursuit task offered several advantages compared to the photoelectric version. The ability to administer the task on a computer or tablet would facilitate use in clinical practice. The fact that it is freely available, open source, and modifiable also facilitates translation to clinical practice. In addition, the mouse-tracking version was better tolerated by patients with TBI than the photoelectric version. While the photoelectric version was generally well-tolerated, a small number of patients with TBI reported that completing the task caused dizziness. In addition, the mouse-tracking version of the task allowed participant's to self-initiate each trial with a mouse click. In contrast, the photoelectric version requires an experimenter to initiate the first trial, with subsequent trials automatically beginning 8 seconds following the conclusion of the previous trial. For many participants, this led to anticipation of the beginning of each trial, or "missing" that a trial had begun, likely introducing significant noise into the photoelectric rotor pursuit data, unrelated to procedural learning ability. Taken together, the mouse-tracking rotor pursuit task represents a promising possibility for the development of a procedural memory measure that would easily translate to real-world clinical assessment, provided the task is modified to include baseline testing and is extensively normed across age, sex, and education levels.

Patient Toleration of Tasks

The current study provides evidence that patients with TBI are able to tolerate and successfully complete each of the experimental tasks examined here. This information will be

useful if these tasks are to be developed into clinically-useful assessments. Although both groups reported that the spatial reconstruction task was challenging, there were no difficulties in obtaining this data from either comparison participants or patients with TBI. A small number (n = 2) of patients with TBI reported dizziness during the photoelectric rotor pursuit task, but were able to complete the task. During the second half of the mirror-reversed reading task, patient 5056 reported significant fatigue and double vision; a significant increase in reading times for novel triads in the second half of the task is present in her data. Contemporaneous notes taken during testing sessions suggested that a small number of patients with TBI (n = 2) quickly became frustrated and put forth less effort during the task, compared to other participants. One of these patients (5034), began responding "I don't know" to word triads starting midway-through the second half of testing (around block 8). Across the entire MRR task, 5034 was a significant outlier in terms of overall accuracy and in terms of the degree of accuracy that would be predicted by her WRAT Reading score. In summary, experimental tasks were generally well-tolerated by patients with TBI, with a small number of patients demonstrating signs of frustration, fatigue, or dizziness.

Sampling Across Heterogeneity

Given the inherent variability in the population of individuals with TBI, sampling across the full range of this heterogeneity is critical for ensuring that results are generalizable to the population at large. Thus, future studies should examine declarative and procedural memory abilities in much larger samples of patients. The current sample may represent an unusually high performing group of individuals with TBI**:** of the twenty-three individuals with TBI for whom employment information is known, nineteen are currently employed (82.61%). This is

substantially higher than reports that suggest only 33-50% of individuals with moderate-severe TBI return to gainful employment (Ponsford & Spitz, 2015). In addition, the small number of patients with standard scores lower than a standard deviation from the normative mean across subtests of the NIH Toolbox suggest that, despite having sustained moderate-severe injuries as determined using the Mayo classification system, these individuals with TBI have experienced a remarkable degree of cognitive recovery. Although inclusion of participants with TBI with few demonstrable cognitive impairments may reduce the likelihood of group differences in performance, these individuals likely represent an understudied subgroup within the larger population of individuals with TBI, and should be included in larger sample investigations, that may serve to determine the factors influencing both favorable outcomes following TBI, as well as poorer outcomes.

Limitations and Future Directions

The current study is limited by its relatively small sample size. Although the sample size reported here is commensurate with that of many published studies examining cognition in patients with TBI, future studies should include substantially larger sample sizes to increase power and to increase the generalizability of study findings.

Measurement issues, particularly across procedural memory tasks, are a limitation that should be addressed systematically in future studies. Development of valid, reliable procedural memory measures would represent a significant advance, not only in TBI research and clinical practice, but across multiple clinical populations. No validated measure of procedural memory exists for use in clinical practice, despite a large body of empirical evidence demonstrating procedural memory deficits in many clinical populations commonly served by speech-language

pathologists (e.g. William's syndrome, (Vicari, Verucci, & Carlesimo, 2007), language impairment, (Evans, Saffran, & Robe-Torres, 2009; Kemeny & Lukacs, 2010; Lum, Conti-Ramsden, Page, & Ullman, 2012; Lum, Gelgic, & Conti-Ramsden, 2010; Sengottuvel & Rao, 2013; Tomblin, Mainela-Arnold, & Zhang, 2007), dyslexia, (Hedenius et al., 2013; Howard, Howard, Japikse, & Eden, 2006; Nicolson, Fawcett, Brookes, & Needle, 2010)), Broca's aphasia, (Goschke, Friederici, Kotz, & van Kampen, 2001), Huntington's disease & Parkinson disease (Heindel, Salmon, Shults, Walicke, & Butters, 1989)). This conspicuous gap in clinical assessment batteries has persisted (Lum & Conti-Ramsden, 2013) despite its acknowledgement for nearly three decades (Sunderland, 1990). Results from the current study suggest that a key issue in the development of such an assessment is the development of procedures that reduce the impact of initial ability on participants' learning.

Although the spatial reconstruction task demonstrates more promise as a future clinical measure than the procedural memory tasks as implemented here, all tasks described here would benefit from systematic evaluation of their psychometric properties. Future studies should assess the test-retest reliability of each task. For the spatial reconstruction task, future work should determine whether a computer adaptive version of the test captures individual differences to the same degree as the current task implementation and should assess differing methods for determining accuracy windows. For the procedural memory tasks, in addition to developing new ways to equate initial performance, future studies should test methods beyond by-participant slopes and difference scores for quantifying change over time. For example, once the standard error of measurement for each procedural memory task is known (following test-retest reliability studies), the Reliable Change Index may be a useful way of quantifying individual performance on these tasks.

The impact of factors beyond declarative memory on the spatial reconstruction task should also be examined. Differences in strategies deployed during the study phase, or differences in participant' motivation and effort, may contribute to performance and should be investigated.

Conclusion

The current study is part of a larger effort to develop learning and memory phenotypes to better characterize patients with moderate-severe TBI, improve treatment decision-making, and more accurately predict long-term outcomes. This study motivates the continued development of declarative and procedural memory measures that are sensitive to group- and individual-differences to characterize the influence of multiple memory systems on real-world outcomes. Use of such measures to characterize memory and learning ability in patients with TBI promises to improve both treatment decision-making and prediction of long-term outcomes.

APPENDIX

Appendix A. Spatial Reconstruction

Exact Hits Initial Model

proportion of exact hits ~ group * set_size + (1 + set_size | participant) + (1 | trial)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant	(Intercept)	3.080e-04	0.017549	
	set size	2.931e-06	0.001712	1.00
trial	(Intercept)	3.570e-03	0.059748	
Residual		3.846e-02	0.196103	

Number of observations: 1240, Number of participants: 50, Number of trials: 25

Fixed effects:

	Estimate	Std.Error	df	t value	p-value
(Intercept)	1.085307	0.033706	32.1	32.199	< 2e-16 ***
group	-0.045089	0.026552	236.	-1.698	0.0908 .
set size	-0.052760	0.005065	31.9	-10.416	8.5e-12 ***
group*set size	0.001612	0.003964	426.7	0.407	0.6844

Group coded: NC = 0, TBI = 1

Exact Hits

proportion of exact hits ~ group + set size + (1 + set size | participant) + (1 | trial)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant	(Intercept)	3.243e-04	0.018009	
	set size	2.681e-06	0.001637	1.00
trial	(Intercept)	3.571e-03	0.059758	
Residual		3.843e-02	0.196036	

Number of observations: 1240, Number of participants: 50, Number of trials: 25

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	1.080688	0.031749	25.41	34.04	< 2e-16 ***
group	-0.035799	0.013610	48.31	-2.63	0.0114 *
set size	-0.051959	0.004666	23.08	-11.13	9.28e-11 ***

Group coded: NC = 0, TBI = 1

Post-Transform

proportion of exact hits (post-transform) ~ group + set size + (1 + set size | participant) + (1 | trial)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant	(Intercept)	0.0097689	0.09884	
	set size	0.0001915	0.01384	-0.47
trial	(Intercept)	0.0082968	0.09109	
Residual		0.0589612	0.24282	

Number of observations: 1240, Number of participants: 50, Number of trials: 25

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	0.992426	0.049965	31.71	19.862	< 2e-16 ***
group	-0.091201	0.029385	47.88	-3.104	0.0032 **
set size	-0.074665	0.007159	26.35	-10.429	7.54e-11 ***

Group coded: NC = 0, TBI = 1

Identity Stripped

proportion of placements to valid locations ~ group + set size + (1 + set size | participant) + (1 | trial)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant	(Intercept)	5.618e-03	0.074954	
	set size	8.629e-05	0.009289	-1.00
trial	(Intercept)	8.028e-03	0.089598	
Residual		4.086e-02	0.202139	

Number of observations: 1240, Number of participants: 50, Number of trials: 25

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	0.912913	0.045796	26.38	19.934	< 2e-16 ***
group	-0.004550	0.012359	176.68	-0.368	0.713
set size	-0.033380	0.006781	24.67	-4.923	4.71e-05 ***

Group coded: NC = 0, TBI = 1

Appendix B. Photoelectric Rotor Pursuit

Full Model

time ~ trial * group + (1 + trial | participant)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant	(Intercept)	3.37282	1.8365	
	trial	0.02258	0.1503	-0.09
Residual		1.98452	1.4087	

Number of observations: 768, Number of participants: 48

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	6.06568	0.39923	48	15.194	< 2e-16 ***
trial	0.16142	0.03441	48	4.691	2.29e-05 ***
group	0.02532	0.56459	48	0.045	0.964
trial*group	-0.05859	0.04866	48	-1.204	0.235

Trial coded: 0-15

Group coded: NC = 0, TBI = 1

Reduced Model

time ~ trial * group + (1 | participant)

Random effects:

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	4.245	2.06
Residual		2.496	1.58

Number of observations: 768, Number of participants: 48

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	6.06568	0.44784	57.4	13.544	<2e-16 ***
trial	0.16142	0.01749	720	9.229	<2e-16 ***
group	0.02532	0.63335	57.4	0.040	0.9682
trial*group	-0.05859	0.02474	720	-2.368	0.0181 *

Trial coded: 0-15

Group coded: NC = 0, TBI = 1

Re-parameterized Model:

time ~ trial * group + (1 + trial | participant)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant	(Intercept)	7.72026	2.7785	
	trial	0.02258	0.1503	0.75
Residual		1.98452	1.4087	

Number of observations: 768, Number of participants: 48

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	8.48701	0.58355	48	14.544	< 2e-16 ***
trial	0.16142	0.03441	48	4.691	2.29e-05 ***
group	-0.85345	0.82526	48	-1.034	0.306
trial*group	-0.05859	0.04866	48	-1.204	0.235

Trial coded: -15-0

Group coded: NC = 0, TBI = 1

Retention Model

time ~ trial * group + (1 | participant)

Random effects:

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	4.803	2.192
Residual		1.987	1.410

Number of observations: 96, Number of participants: 48

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	6.27250	0.53189	63.98	11.793	<2e-16 ***
trial	0.84967	0.40690	48	2.088	0.0421 *
group	-0.04254	0.75221	63	-0.057	0.9551
trial*group	-0.60962	0.57545	48	-1.059	0.2947

Trial coded: 0-1

Group coded: NC = 0, TBI = 1

Appendix C. Mouse-tracking Rotor Pursuit

Model with Feedback Effect

time-on-target ~ (trial+ trial^2) * group * feedback + (1 + trial + trial^2 | participant*feedback)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant*feedback	(Intercept)	5.527254	2.3510	
	trial	0.220433	0.4695	-0.58
	trial^2	0.001576	0.0397	0.57 -0.93
Residual		1.150782	1.0727	

Number of observations: 784, Number of clusters (participant*feedback): 96

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	1.018e+01	5.141e-01	96.40	9.808	< 2e-16 ***
trial	5.406e-01	1.559e-01	101.9	3.467	0.000771 ***
trial^2	-5.458e-02	1.874e-02	104	-2.913	0.004383 **
group	-1.517e+00	7.264e-01	96.05	-2.088	0.039431 *
feedback	-9.607e-02	7.270e-01	96.40	-0.132	0.895149
trial*group	-2.319e-01	2.193e-01	98.86	-1.057	0.292986
trial^2*group	3.534e-02	2.630e-02	99.04	1.343	0.182184
trial*feedback	1.396e-01	2.205e-01	101.9	0.633	0.528231
trial^2*feedback	-1.994e-02	2.650e-02	104	-0.753	0.453426
group*feedback	1.441e-01	1.027e+00	96.05	0.140	0.888705
trial*group*feedback	-7.888e-03	3.102e-01	98.86	-0.025	0.979765
trial^2*group*feedback	7.493e-04	3.720e-02	99.04	0.020	0.983970

Trial coded: 0-7

Group coded: NC = 0, TBI = 1

Feedback coded: No feedback = 0, Feedback = 1

Model without Feedback Terms

time-on-target ~ (trial + trial^2) * group + (1 + trial + trial^2 | participant)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant	(Intercept)	4.2130900	2.05258	
	trial	0.0380432	0.19505	-0.45
	trial^2	0.0001034	0.01017	0.79 -0.90
Residual		1.5618673	1.24975	

Number of observations: 784, Number of participants: 48

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	10.13442	0.44564	48.3	22.741	< 2e-16 ***
trial	0.61041	0.10886	123.8	5.607	1.27e-07 ***
trial^2	-0.06455	0.01407	400.5	-4.588	6.01e-06 ***
group	-1.44470	0.62968	48.2	-2.294	0.0262 *
trial*group	-0.23665	0.15267	119.1	-1.550	0.1238
trial^2*group	0.03585	0.01970	391.9	1.819	0.0696 .

Trial coded: 0-7

Group coded: NC = 0, TBI = 1

Appendix D. Mirror-Reversed Reading

Reading Accuracy

accuracy ~ group*block + condition*block + (1 | participant), family = binomial

Random effects:

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	1.573	1.254

Number of observations: 29961, Number of participants: 50

Fixed effects:

	Estimate	Std. Error	z value	p-value
(Intercept)	2.95490	0.27212	10.859	< 2e-16 ***
group	-0.41938	0.37278	-1.125	0.26059
block	0.11966	0.01692	7.071	1.54e-12 ***
condition	-0.05318	0.10090	-0.527	0.59815
group*block	-0.06129	0.01897	-3.232	0.00123 **
block*condition	0.09046	0.01852	4.883	1.04e-06 ***

Block coded: 0-9

Group coded: NC = 0, TBI = 1

Novel = 0, Repeat = 1

Learning over time

read-time ~ (block + block^2)*group + (block + block^2)*condition + condition*group +

(1 + block + block^2| participant*condition)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr	
participant*condition	(Intercept)	45.01	6.709		
	block	27.45	5.239	-0.15	
	block^2	12.64	3.555	0.10	-1.00
Residual		97.88	9.893		

Number of observations: 9937, Number of clusters (participant*condition): 100

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	18.5873	1.3478	101.82	13.791	< 2e-16 ***
block	-4.7380	1.1073	103.05	-4.279	4.22e-05 ***
block^2	4.1809	0.8841	118.93	4.729	6.26e-06 ***
group	4.8700	1.8940	102.55	2.571	0.011568 *
condition	-7.7268	1.8940	102.55	-4.080	8.94e-05 ***
block*group	-0.3409	1.2789	103.15	-0.267	0.790360
block^2*group	0.1075	1.0213	119.14	0.105	0.916361
block*condition	-6.7425	1.2789	103.15	-5.272	7.46e-07 ***
block^2*condition	3.7269	1.0213	119.14	3.649	0.000392 ***
group*condition	-2.9335	2.6440	100.03	-1.110	0.269870

Block coded: 0-9

Group coded: NC = 0, TBI = 1

Novel = 0, Repeat = 1

Model run on data without errors

read-time ~ (block + block^2)*group + (block + block^2)*condition + condition*group +

(1 + block + block^2| participant*condition)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant:condition	(Intercept)	31.68	5.629	
	block	26.57	5.155	0.01
	block^2	12.52	3.538	-0.02 -1.00
Residual		63.39	7.962	

Number of observations: 8402, Number of clusters (participant*condition): 100

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	17.6739	1.1391	99.6	15.515	< 2e-16 ***
block	-3.4902	1.0651	96.48	-3.277	0.001460 **
block^2	3.1247	0.8340	105.86	3.746	0.000292 ***
group	3.0403	1.6140	100.42	1.884	0.062489 .
condition	-7.4471	1.6102	99.57	-4.625	1.13e-05 ***
block*group	-0.5940	1.2389	98.16	-0.480	0.632648
block^2*group	0.3327	0.9710	107.84	0.343	0.732538
block*condition	-6.0304	1.2384	97.96	-4.870	4.29e-06 ***
block^2*condition	3.4028	0.9704	107.56	3.507	0.000664 ***
group*condition	-1.2500	2.2805	100.15	-0.548	0.584834

Block coded: 0-9

Group coded: NC = 0, TBI = 1

Novel = 0, Repeat = 1

Appendix E. Serial Interception Sequence Learning

SSPA

SSPA ~ block * group + (1 + block | participant)

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
participant	(Intercept)	27.062	5.202	
	block	1.025	1.012	1.00
Residual		98.778	9.939	

Number of observations: 150, Number of participants: 50

Fixed effects:

	Estimate	Std. Error	df	t value	p-value
(Intercept)	5.537	2.092	58.4	2.647	0.0104 *
block	2.666	1.420	91.71	1.877	0.0637 .
group	1.559	2.958	58.4	0.527	0.6003
block*group	-1.936	2.008	91.71	-0.964	0.3375

Group coded: NC = 0, TBI = 1

Block coded: 0-3

REFERENCES

Adams, J. H., Graham, D. I., & Gennarelli, T. A. (1985). Contemporary neuropathological considerations regarding brain damage in head injury. In *Central Nervous System Trauma Status Report* (pp. 65–78).

Adams, J. H., Graham, D. I., Gennarelli, T. A., & Maxwell, W. L. (1991). Diffuse axonal injury in non-missile head injury. *Journal of Neurology, Neurosurgery, and Psychiatry*, *54*, 481–483.

Andelic, N., Ye, J., Tornas, S., Roe, C., Lu, J., Bautz-Holter, E., … Aas, E. (2014). Cost-Effectiveness Analysis of an Early-Initiated, Continuous Chain of Rehabilitation after Severe Traumatic Brain Injury. *Journal of Neurotrauma*, *31*(14), 1313–1320.

Arabi, Y. M., Haddad, S., Tamim, H. M., Al-Dawood, A., Al-Qahtani, S., Ferayan, A., … Rugaan, A. (2010). Mortality reduction after implementing a clinical practice guidelines-based management protocol for severe traumatic brain injury. *Journal of Critical Care*, *25*(2), 190–195.

Chang, J., Youn, T. S., Benson, D., Mattick, H., Andrade, N., Harper, C. R., … Diaz-Arrastia, R. R. (2009). Physiologic and functional outcome correlates of brain tissue hypoxia in traumatic brain injury. *Critical Care Medicine*, *37*(1), 283–90.

Channon, S., Shanks, D., Johnstone, T., Vakili, K., Chin, J., & Sinclair, E. (2002). Is implicit learning spared in amnesia? Rule abstraction and item familiarity in artificial grammar learning. *Neuropsychologia*, *40*(12), 2185–97.

Cicerone, K. D., Langenbahn, D. M., Braden, C., Malec, J. F., Kalmar, K., Fraas, M., … Ashman, T. (2011). Evidence-Based Cognitive Rehabilitation: Updated Review of the Literature From 2003 Through 2008. *ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION*, *92*(4), 519–530.

Clark, R., Tahan, A. C., Watson, P. D., Severson, J., Cohen, N. J., & Voss, M. (2017). Aging affects spatial reconstruction more than spatial pattern separation performance even after extended practice. *Hippocampus*, *27*(6), 716–725.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*(4466), 207–210.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33*(A), 497–505.

Cooper, P. R. (1985). Delayed brain injury: secondary insults. In *Central Nervous System Trauma Status Report*.

Coronado, V. G., Xu, L., Basavaraju, S. V, McGuire, L. C., Wald, M. M., Faul, M. D., … Centers for Disease Control and Prevention (CDC). (2011). Surveillance for traumatic brain injury-related deaths--United States, 1997-2007. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C. : 2002)*, *60*(5), 1–32.

Corrigan, J. D., Rust, E., & Lamb-Hart, G. L. (1995). The nature and extent of substance abuse problems in persons with traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *10*(3), 29–46.

Courville, C. B. (1942). Coup-contrecoup Mechanism of Cranio-cerebral Injuries. *Archives of Surgery*, *45*, 19–43.

Dahdah, M. N., Barnes, S., Buros, A., Dubiel, R., Dunklin, C., Callender, L., … Shafi, S. (2016). Variations in Inpatient Rehabilitation Functional Outcomes Across Centers in the Traumatic Brain Injury Model Systems Study and the Influence of Demographics and Injury Severity on Patient Outcomes. *Archives of Physical Medicine & Rehabilitation*, *97*(11), 1821–1831.

de la Plata, C. D. M., Garces, J., Kojori, E. S., Grinnan, J., Krishnan, K., Pidikiti, R., … Diaz-Arrastia, R. (2011). Deficits in functional connectivity of hippocampal and frontal lobe circuits after traumatic axonal injury. *Archives of Neurology*, *68*(1), 74–84.

Ditton, P. M. (1999). Mental Health and Treatment of Inmates and Probationers. *U.S. Department of Justice Bureau of Justice Statistics Special Report*, *July*, 1–12.

Don, A. J., Schellenberg, E. G., Reber, A. S., DiGirolamo, K. M., & Wang, P. P. (2003). Implicit learning in children and adults with Williams syndrome. *Developmental Neuropsychology*, *23*(1–2), 201–25.

El-Menyar, A., El-Hennawy, H., Al-Thani, H., Asim, M., Abdelrahman, H., Zarour, A., … Latifi, R. (2014). Traumatic injury among females: does gender matter? *Journal of Trauma Management & Outcomes*, *8*(1), 8.

Elliott, M., & Parente, F. (2014). Efficacy of memory rehabilitation therapy: A meta-analysis of TBI and stroke cognitive rehabilitation literature. *Brain Injury*, *9052*(12), 1–7.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical Learning in Children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *52*(April), 321–336.

Ewert, J., Levin, H., Watson, M., & Kalisky, Z. (1989). Procedural Memory During Posttraumatic Amnesia in Survivors of Severe Closed Head Injury. *Archives of Neurology*, *46*(8), 911–916.

Farace, E., & Alves, W. (2000). Do women fare worse: a metaanalysis of gender differences in traumatic brain injury outcome. *JOURNAL OF NEUROSURGERY*, *93*, 539–545.

Faul, M., Xu, L., Wald, M. M., & Coronado, V. G. (2010). *Traumatic Brain Injury in the United States: Emergency Department Visits, Hospitalizations and Death 2002-2006*. *Centers for Disease Control and Prevention, National Center for Injury Prevention and Control*. Atlanta, GA.

Finkelstein, E. A., Corso, P. S., & Miller, T. (2006). *The incidence and economic burden of injuries in the United States*. New York: Oxford University Press.

Frankowski, R. F. (1986). Descriptive epidemiologic studies of head injury in the United States: 1974-1984. *Advances in Psychosomatic Medicine*, *16*, 153–172.

Geary, E. K., Kraus, M. F., Pliskin, N. H., & Little, D. M. (2010). Verbal learning differences in chronic mild traumatic brain injury. *Journal of the International Nueorpsychological Society*, *16*, 506–516.

Gebauer, G. F., & Mackintosh, N. J. (2007). Psychometric Intelligence Dissociates Implicit and Explicit Learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(1), 34–54.

Gerber, L. M., Chiu, Y.-L., Carney, N., Hartl, R., Ghajar, J., Härtl, R., & Ghajar, J. (2013). Marked reduction in mortality in patients with severe traumatic brain injury. *Journal of Neurosurgery*, *119*(6), 1583–90.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, *12*(4), 395–427.

Glisky, E. (1993). Training Persons with Traumatic Brain Injury for Complex Computer Jobs: The Domain-Specific Approach. In D. Thomas, F. Menz, & D. McAlees (Eds.), *Community-Based Employment Following Traumatic Brain Injury* (pp. 3–29).

Gobel, E. W., Blomeke, K., Zadikoff, C., Simuni, T., Weintraub, S., & Reber, P. J. (2013). Implicit perceptual-motor skill learning in mild cognitive impairment and parkinson's disease. *Neuropsychology*, *27*(3), 314–321.

Goranson, T. E., Graves, R. E., Allison, D., & La Freniere, R. (2003). Community integration following multidisciplinary rehabilitation for traumatic brain injury. *Brain Injury*, *17*(9), 759–774.

Goschke, T., Friederici, A. D., Kotz, S. a., & van Kampen, A. (2001). Procedural Learning in Broca's Aphasia: Dissociation between the Implicit Acquisition of Spatio-Motor and Phoneme Sequences. *Journal of Cognitive Neuroscience*, *13*(3), 370–388.

Graham, D. I., Adams, J. H., & Doyle, D. (1978). Ischaemic Brain Damage in Fatal Non-Missile Head Injuries. *Journal of the Neurological Sciences*, *39*, 213–234.

Hannay, H. J., Howieson, D. B., Loring, D. W., Fischer, J. S., & Lezak, M. D. (2004). Neuropathology for Neuropsychologists. In M. D. Lezak, D. B. Howieson, & D. W. Loring (Eds.), *Neuropsychological Assessment* (4th ed., pp. 157–285). Oxford: Oxford University Press.

Hart, T., Kozlowski, A. J., Whyte, J., Poulsen, I., Kristensen, K., Nordenbo, A., & Heinemann, A. W. (2014). Functional Recovery After Severe Traumatic Brain Injury: An Individual Growth Curve Approach. *ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION*, *95*(11), 2103–2110.

Hedenius, M., Persson, J., Alm, P. a, Ullman, M. T., Howard, J. H., Howard, D. V, & Jennische, M. (2013). Impaired implicit sequence learning in children with developmental dyslexia. *Research in Developmental Disabilities*, *34*(11), 3924–3935.

Heindel, W. C., Butters, N., & Salmon, D. P. (1988). Impaired learning of a motor skill in patients with Huntington's disease. *Behav Neurosci*, *102*(1), 141–147.

Heindel, W. C., Salmon, D. P., Shults, W., Walicke, A., & Butters, N. (1989). Neuropsychological Evidence for Multiple Implicit Memory Systems: A Comparison of Alzheimer's, Huntington's, and Parkinson's Disease Patients. *The Journal of Neuroscience*, *9*(2), 582–587.

Horan, W. P., Green, M. F., Knowlton, B. J., Wynn, J. K., Mintz, J., & Nuechterlein, K. H. (2008). Impaired Implicit Learning in Schizophrenia. *Neuropsychology*, *22*(5), 606–617.

Horecka, K. M., Dulas, M. R., Schwarb, H., Lucas, H. D., Duff, M., & Cohen, N. J. (2018). Reconstructing relational information. *Hippocampus*, *28*(2), 164–177.

Howard, J. H., Howard, D. V, Japikse, K. C., & Eden, G. F. (2006). Dyslexics are impaired on implicit higher-order sequence learning, but not on implicit spatial context learning. *Neuropsychologia*, *44*(7), 1131–44. http://doi.org/10.1016/j.neuropsychologia.2005.10.015

Hsu, H. J., & Bishop, D. V. M. (2014). Sequence-specific procedural learning deficits in children with specific language impairment, *3*(116), 352–365.

Hyder, A. A., Wunderlich, C. A., Puvanachandra, P., Gururaj, G., & Kobusingye, O. C. (2007). The impact of traumatic brain injuries: A global perspective. *NeuroRehabilitation*, *22*, 341–353.

Kalra, P. B., Gabrieli, J. D. E., Finn, A. S., Madison, W., & States, U. (2019). Evidence of stable individual differences in implicit learning. *Cognition*, *190*(May), 199–211.

Kemeny, F., & Lukacs, A. (2010). Impaired procedural learning in language impairment: results from probabilistic categorization. *Journal of Clinical and Experimental Neuropsychology*, *32*(3), 249–58.

Knowlton, Squire, L. R., & Gluck, M. a. (1994). Probabilistic classification learning in amnesia.

Learning & Memory, *1*(2), 106–120.

Kraus, M., Little, D., Wojtowicz, S., & Sweeney, J. (2010). Procedural learning impairments identified via predictive saccades in chronic traumatic brain injury. *Cognitive and Behavioral Neurology*, *23*(4), 210–217.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Langlois, J. a, Rutland-Brown, W., & Wald, M. M. (2006). The epidemiology and impact of traumatic brain injury: a brief overview. *The Journal of Head Trauma Rehabilitation*, *21*(5), 375–8.

Lee, J. C., & Tomblin, J. B. (2015). Procedural Learning and Individual Differences in Language. *Language Learning and Development : The Official Journal of the Society for Language Development*, *11*(3), 215–236.

Leh, S. E., Schroeder, C., Chen, J.-K., Mallar Chakravarty, M., Park, M. T. M., Cheung, B., … Petrides, M. (2017). Microstructural Integrity of Hippocampal Subregions Is Impaired after Mild Traumatic Brain Injury. *Journal of Neurotrauma*, *34*(7), 1402–1411.

Levin, H. S. (1990). Memory deficit after closed-head injury. *Journal of Clinical and Experimental Neuropsychology*, *12*(1), 129–153.

Lu, J., Gary, K. W., Neimeier, J. P., Ward, J., & Lapane, K. L. (2012). Randomized controlled trials in adult traumatic brain injury. *Brain Injury*, *26*(December), 1523–48.

Lum, J.G., & Conti-Ramsden, G. (2013). Long-term memory: A review and meta-analysis of studies of declarative and procedural memory in specific language impairment. *Topics in Language Disorders*, *33*(4), 282–297.

Lum, J.G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *48*(9), 1138–54.

Lum, J.G., Gelgic, C., & Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *45*(1), 96–107.

Ma, V. Y., Chan, L., & Carruthers, K. J. (2014). Incidence, Prevalence, Costs, and Impact on Disability of Common Conditions Requiring Rehabilitation in the United States: Stroke, Spinal Cord Injury, Traumatic Brain Injury, Multiple Sclerosis, Osteoarthritis, Rheumatoid Arthritis, Limb Loss, and Back Pa. *ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION*, *95*(5), 986–995.

Malec, J. F., Brown, A. W., Leibson, C. L., Flaada, J. T., Mandrekar, J. N., Diehl, N. N., & Perkins, P. K. (2007). The mayo classification system for traumatic brain injury severity. *Journal of Neurotrauma*, *24*(9), 1417–1424.

McDonald, S., Togher, L., & Code, C. (1999). The nature of traumatic brain injury: Basic features and neuropsychological consequences. In S. McDonald, L. Togher, & C. Code (Eds.), *Communication Disorders Following Traumatic Brain Injury* (pp. 19–54). Psychology Press Ltd.

Mcintosh, T. K. (1993). Pharmacologic Therapies in the Treatment of Experimental Traumatic Brain Injury : A Review. *Journal of Neurotrauma*, *10*(3), 215–261.

Morton, M. V, & Wehman, P. (1995). Psychosocial and emotional sequelae of individuals with traumatic brain injury: a literature review and recommendations. *Brain Injury*, *9*(1), 81–92.

Mueller, S. T. (2012). The PEBL Pursuit Rotor Task.

Mueller, S. T., & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods*, *222*, 250–259.

Murray, L. L., Ramage, A. E., & Hopper, T. (2001). Memory Impairments in Adults with Neurogenic Communication Disoders. *Seminars in Speech & Language*, *22*(2), 127–136.

Nicolson, R. I., Fawcett, A. J., Brookes, R. L., & Needle, J. (2010). Procedural Learning and Dyslexia. *Dyslexia*, *16*, 194–212.

Pagulayan, K. F., Temkin, N. R., Machamer, J., & Dikmen, S. S. (2006). A Longitudinal Study of Health-Related Quality of Life After Traumatic Brain Injury. *Archives of Physical Medicine and Rehabilitation*, *87*(5), 611–618.

Paniak, C., Reynolds, S., Phillips, K., Toller-lobe, G., Melnyk, A., & Nagy, J. (2002). Patient complaints within 1 month of mild traumatic brain injury: A controlled study. *Archives of Clinical Neuropsychology*, *17*, 319–334.

Papenberg, M. (2018). minDiff: Minimize differences between groups.

Pavio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement*, *76*(3, part 2).

Poldrack, R. A., & Foerde, K. (2008). Category learning and the memory systems debate. *Neuroscience and Biobehavioral Reviews*, *32*(2), 197–205.

Ponsford, J. L., Downing, M. G., Olver, J., Ponsford, M., Acher, R., Carty, M., & Spitz, G. (2014). Longitudinal Follow-Up of Patients with Traumatic Brain Injury: Outcome at Two, Five, and Ten Years Post-Injury. *Journal of Neurotrauma*, *31*(1), 64–77.

Ponsford, J. L., Olver, J. H., & Curran, C. (1995). A profile of outcome: 2 years after traumatic brain injury. *Brain Injury*, *9*(1), 1–10.

Ponsford, J. L., & Spitz, G. (2015). Stability of Employment Over the First 3 Years Following Traumatic Brain Injury. *JOURNAL OF HEAD TRAUMA REHABILITATION*, *30*(3), E1--E11.

Rigon, A., Turkstra, L., Mutlu, B., & Duff, M. (2016). The female advantage : sex as a possible protective factor against emotion recognition impairment following traumatic brain injury. *Cognitive, Affective, & Behavioral Neuroscience*, 866–875.

Sanchez, D. J., Gobel, E. W., & Reber, P. J. (2010). Performing the unexplainable: Implicit task performance reveals individually reliable sequence learning without explicit knowledge. *Psychonomic Bulletin and Review*, *17*(6), 790–796.

Sarapata, M., Herrmann, D., Johnson, T., & Aycock, R. (1998). The role of head injury in cognitive functioning, emotional adjustment and criminal behaviour. *Brain Injury*, *12*(10), 821–42.

Sengottuvel, K., & Rao, P. K. S. (2013). Aspects of grammar sensitive to procedural memory deficits in children with specific language impairment. *Research in Developmental Disabilities*, *34*(10), 3317–31.

Shiroma, E. J., Ferguson, P. L., & Pickelsimer, E. E. (2010). Prevalence of traumatic brain injury in an offender population: a meta-analysis. *Journal of Correctional Health Care : The Official Journal of the National Commission on Correctional Health Care*, *16*(2), 147–59.

Smith, D. H., Meaney, D. F., & Shull, W. H. (2003). Diffuse axonal injury in head trauma. *Journal of Head Trauma Rehabilitation*, *18*(4), 307–319.

Sunderland, A. (1990). Clinical Memory Assessment: Matching the Method to the Aim. In D. Tupper & K. D. Cicerone (Eds.), *The Neuropsychology of Everyday Life: Assessment and Basic Competencies* (pp. 169–183). Boston, Massachusetts: Kluwer Academic Publishers.

Taylor, C. A., Bell, J. M., Breiding, M. J., & Xu, L. (2017). Traumatic Brain Injury–Related Emergency Department Visits, Hospitalizations, and Deaths — United States, 2007 and 2013. *MMWR. Surveillance Summaries*, *66*(9), 1–16.

Togher, L., McDonald, S., & Code, C. (1999). Communication problems following traumatic brain injury. In S. McDonald, L. Togher, & C. Code (Eds.), *Communication Disorders Following Traumatic Brain Injury* (pp. 1–18). Psychology Press Ltd.

Toglia, M. P., & Battig, W. R. (1978). *Handbook of Semantic Word Norms*. New York: Erlbaum.

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural Learning in Adolescents

With and Without Specific Language Impairment. *Language Learning and Development*, *3*(4), 269–293.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specfic to language: The Procedural deficit hypothesis. *Cortex*, *41*, 399–433.

Vakil, E. (2005). *The effect of moderate to severe traumatic brain injury (TBI) on different aspects of memory: a selective review. Journal of clinical and experimental neuropsychology* (Vol. 27).

Vakil, E., Biederman, Y., Liran, G., Groswasser, Z., & Aberbuch, S. (1994). Head-injured patients and control group: Implicit versus explicit measures of frequency of occurrence. *Journal of Clinical and Experimental Neuropsychology*, *16*(4), 539–546.

Vakil, E., Kraus, A., Bor, B., & Groswasser, Z. (2002). Impaired skill learning in patients with severe closed-head injury as demonstrated by the serial reaction time (SRT) task. *Brain and Cognition*, *50*(2), 304–315.

Vicari, S., Verucci, L., & Carlesimo, G. A. (2007). Implicit memory is independent from IQ and age but not from etiology: Evidence from Down and Williams syndromes. *Journal of Intellectual Disability Research*, *51*(12), 932–941.

Ward, H., Shum, D., Wallace, G., & Boon, J. (2002). Pediatric Traumatic Brain Injury and Procedural Memory. *Journal of Clinical and Experimental Neuropsychology*, *24*(4), 458–470.

Watson, P., Voss, J., Warren, D., Tranel, D., & Cohen, N. (2013). Spatial Reconstruction by Patients with Hippocampal Damage is Dominated By Relational Memory Errors. *Hippocampus*, *23*(7), 570–580. http://doi.org/10.1002/hipo.22115.Spatial

Watt, S., Shores, E. A., & Kinoshita, S. (1999). Effects of reducing attentional resources on implicit and explicit memory after severe traumatic brain injury. *Neuropsychology*, *13*(3), 338–349.

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., … Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(Issue 11, Supplement 3), S54–S64.

Wilson, B. A. (1998). Recovery of cognitive functions following nonprogressive brain injury. *Current Opinion in Neurobiology*, *8*(2), 281–287.

Wilson, B. A. (2009). *Memory Rehabilitation: Integrating Theory and Practice*. New York: Guilford Press.

Ylvisaker, M., & Feeney, T. J. (1998). *Collaborative Brain Injury Intervention: Positive Everyday Routines*. San Diego: Singular Publishing Group.

Zaloshnja, E., Miller, T., Langlois, J. A., & Selassie, A. W. (2008). Prevalence of Long-Term Disability From Traumatic Brain Injury in the Civilian Population of the United States, 2005. *Journal of Head Trauma Rehabilitation*, *23*(6), 394–400.