PHYSIOLOGICAL PREDICTORS OF BEHAVIORAL DYSREGULATION IN ADULTS WITH TRAUMATIC BRAIN INJURY: A NOVEL ECOLOGICAL MOMENTARY ASSESSMENT METHOD

by

Ashlee Brooke McKeon

B.Phil., University of Pittsburgh, 2010

M.S., University of Pittsburgh, 2013

Submitted to the Graduate Faculty of

School of Health & Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH & REHABILITATION SCIENCES

This dissertation was presented

by

Ashlee Brooke McKeon

It was defended on

May 31, 2016

and approved by

Rory Cooper, PhD, Distinguished Professor, Department of Rehabilitation Science & Technology

Elizabeth Skidmore, PhD, Associate Professor, Department of Occupational Therapy

Lauren Terhorst, PhD, Associate Professor, Department of Occupational Therapy

Dan Ding, PhD, Associate Professor, Department of Rehabilitation Science & Technology

Dissertation Advisor: Michael McCue, PhD, Professor, Department of Rehabilitation Science & Technology

PHYSIOLOGICAL PREDICTORS OF BEHAVIORAL DYSREGULATION IN ADULTS WITH TRAUMATIC BRAIN INJURY: A NOVEL ECOLOGICAL MOMENTARY ASSESSMENT METHOD

Ashlee Brooke McKeon, PhD

University of Pittsburgh, 2016

Each year in the United States 2.5 million people sustain a traumatic brain injury (TBI), making TBI one of the leading causes of death and disability. Difficulty self-regulating behavior is considered a hallmark deficit of TBI and can impede positive rehabilitation outcomes. Traditional clinic-based assessments with TBI cohorts are often compromised due to inaccurate self-reporting caused by self-awareness and/or memory challenges. This project developed out of the need for more adequate and objective assessment methods of behavior after TBI. The aim of this study was to identify physiological states that may serve as real world, real-time prodromes of behavioral dysregulation when compared to direct behavioral observation during task performance. Adults with TBI (N=14) were asked to complete a protocol of challenging, problem-solving tasks selected due to their potential to elicit negative responses believed to be indicative of dysregulation and result in task breakdown behaviors. Participants wore a portable biosensor band throughout the completion of all research tasks, which allowed for continuous physiological recordings of heart rate, breathing rate, and heart rate variability. Physiology was then synchronized with post hoc behavioral ratings conducted by trained clinicians from video recordings and analyzed through hierarchical linear mixed model methods to assess both between- and within-subject contribution related to the outcome. Testing was conducted at ReMed of Pittsburgh, a community-based brain

injury rehabilitation program. This study was funded in part through the National Science Foundation's Integrative Graduate Education & Research Traineeship (IGERT) Program in the Department of Rehabilitation Science & Technology and the School of Health & Rehabilitation Sciences Dean's Research Development Fund, all at the University of Pittsburgh.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

The current research was funded in part through the Integrative Graduate Education & Research Traineeship (IGERT) grant from the National Science Foundation, grant number 0333420 and the University of Pittsburgh School of Health & Rehabilitation Sciences Dean's Research Development Fund.

First, I would like to thank my committee chair and advisor Dr. Michael McCue. Your guidance and support these past 5 years have been invaluable to me. I am forever grateful to you for accepting me as your doctoral student and believing I am capable of great things. I know having you as my mentor has prepared me for success in my future.

Thank you to my committee: Dr. Rory Cooper, for accepting and supporting me as a student in the IGERT program, Dr. Dan Ding, for serving as my technical mentor, Dr. Lauren Terhorst, for guiding me as I navigated a challenging dataset and analytic plan and, Dr. Elizabeth Skidmore, for always encouraging me to challenge myself, think critically, and not be afraid to ask tough questions. I would not have gotten through this experience without your support and I am grateful for the time and effort you all have invested into my professional development.

I would like to extend my gratitude to the Department of Rehabilitation Science & Technology, especially the Clinical Rehabilitation and Mental Health Counseling program for providing me with a support system and home throughout both my Master's and doctoral education. Finally, without the IGERT program, which encourages graduate students to engage in cross-disciplinary training and research, this unique project would have not been possible.

I owe a special thanks to ReMed of Pittsburgh—Regina Lesako, Eduardo Jorge, and Gina Moon, without your belief in the impact my research can have on the lives of individuals with

traumatic brain injury and your willingness to open ReMed's doors to not only assist with recruitment but to also provide me with space to run my research this project would not have been as positive of an experience as it was.

Thank you to my clinicians, Evan Knutson, Lindsay Martin, Cameron LeViere, and Susan Longley for your tireless effort behavioral coding participant videos and valuable feedback on my rating scale and coding scheme. Thank you to my three expert professional raters, Dr. Jamie Kulzer, Dr. Michael Pramuka and, Susan Chase—your unique expertise and perspective made my assessment tool a much better version of itself than I could have ever produced alone.

It means a great deal to me to be able to thank my personal support system. My friends, who may not even fully understand the support they have provided me, but their friendship and belief in my abilities is something that carried me through the stressful times that are the dissertation process. My grandparents, Angela and Sam—my drive for success stems from the positive influence you both have had on my life, and this doctoral degree is just as much mine as it is yours. Your unconditional love and support of me in whatever I choose to do is one of the most meaningful memories I will take away from this process. Acknowledgment of the people who have supported me along the way would be incomplete without thanking my mother and father in-law, Becky and Donald McKeon. You both have understood since the beginning how important this journey was to me. Your support and understanding when sometimes work had to cancel or shorten planned visits and time together made things just a little easier when times became challenging.

Finally, to my husband Donald, who deserves an entire chapter dedicated to the support and partnership he has provided me during my 10-year academic career—from the beginning of this journey when I decided my goal was to earn my doctoral degree, to this moment where that

goal has finally come to fruition. Your selflessness throughout this dissertation process will never be forgotten—a debt I will spend a lifetime repaying. Watching you achieve your own professional and education goals while I moved through the doctoral process gave me motivation to continue down my own carved path. I can say with complete confidence that I could not have gotten through the hard days or remember to appreciate the good days without you believing in me and keeping me focused.

No one earns a doctoral degree alone. On the outside, it may appear to be a process driven by one person, but behind that person is a team of individuals who while living their own lives make time to lift that person up toward achieving their goal. I will be forever grateful for each of you and the influence you have had on my doctoral experience.

1.0 INTRODUCTION

1.1 PROBLEM BACKGROUND

Each year in the United States it is estimated 2.5 million people sustain a traumatic brain injury (TBI), making TBI one of the leading causes of death and disability (Centers for Disease Control and Prevention, 2010). Difficulty self-regulating, or exercising control over one's thoughts, emotions, and behaviors to adaptively fit across environment contexts is a hallmark deficit experienced by anywhere between 76-97% of adults with TBI (Sherer et al., 1998). This inability to self-manage, along with other executive functioning deficits serve as barriers to independence and functioning in major life areas including social and vocational (Ylvisaker et al., 2007). The resulting negative outcomes have been shown to contribute to decreases in both participation and quality of life (QoL). Therefore, it is not difficult to see why improved evidence-based practice for assessment and intervention for the care of the full continuum of TBI has been labeled a priority focus area by The United States Department of Defense (Helmick, Baugh, Lattimore, and Goldman, 2012).

1.1.1 Statement of the Problem

Despite the significant impact of self-regulation (SR) deficits on daily living, the majority of research on SR for persons with TBI is typically conducted in clinic settings. Due to the well-

documented poor self-awareness and difficulty generalizing learned skills across contexts (e.g., particularly from the clinic to the real world), clinical settings may not be the most appropriate delivery environment to facilitate long-term skill retention in TBI cohorts (McCue, Fairman, & Pramuka, 2010). Individuals with TBI tend to function poorly in situations that are unstructured, where a high degree of SR is needed (Levine, Deirdre, Schwartz, Boutet, & Stuss, 2000). Additionally, intervention delivery within the easily controlled and highly structured clinic may not allow for natural representation of functioning or tap into the true environment where dysfunction manifests. The challenges faced when selecting appropriate interventions for the target group are also present when attempting to understand the shortcomings of traditional assessment methods within the clinic. Highly controlled approaches may paint an unclear picture of functioning and subsequently, lead to ineffective methods for collecting data. Research has shown rehabilitation interventions can be effective and even superior in yielding positive functional outcomes when delivered directly in the settings new learned skills will be applied in (McCue, Fairman, & Pramuka, 2010). It can be assumed the same thought process can be applied to assessment.

In an overview of recent literature, McCue and colleagues provide supporting evidence for having clients engage in the rehabilitation process within an environmental context familiar to them (2010). This support stems from the theory that home or community settings provide unique contextual factors that would otherwise not be present in a controlled environment, which may positively influence rehabilitation outcomes (e.g., reduced feelings of anxiety about treatment due to overall environment familiarity and comfortability, which may lead to buy in and increased participation in the rehabilitation process; McCue et al., 2010).

Biosensor technology has been shown to be highly efficacious across chronic illness and disability for increasing self-regulatory ability through increased self-awareness and training of physiological states (Association for Applied Psychophysiology and Biofeedback, 2011; Yucha & Gilbert, 2004). More specifically, respiration rate (RR), heart rate (HR) and electric muscle activity (EMG) all appear frequently in the literature as responsive signals to target this deficit (Yucha & Gilbert, 2004). However, literature focused on the utilization of biomeasures for SR deficits, specifically for persons with TBI is scarce, especially for applications in real world settings. Due to the limitations of current assessment methods and interventions, it is suggested rehabilitation professionals consider novel methods that are ecologically valid, and shift focus to delivering rehabilitation where deficits naturally occur and where assistance is most needed.

1.2 RESEARCH PURPOSE

The main purpose of the current research project was to explore physiological states that may serve as prodromes for behavioral dysregulation (BD) onset when compared to post hoc behavior observation during task completion. This project is the initial and a critical stage in determining if a TBI-specific naturalistic rehabilitation intervention is feasible based on the prediction of SR physiologically in real-time and in the real world. With this ecologically valid intervention, biosensor technology may be used for physiological measurement and prediction of moments leading up to maladaptive behaviors that then allow for prospective cognitive rehabilitation intervention aimed at increasing SR ability. Ultimately, physiological states that accurately predict clinician judgment of behavior within the clinic may provide insight into the ability for these same biomeasures to be predictive of BD in the natural environment.

1.2.1 Research Aims

Both a development phase and an initial feasibility study were conducted prior to two full pilot studies. This was done in order to gain a greater understanding of the construct of interest and determine the feasibility of conducting research in this area. Figure 1 can be referred to for a complete research graphic of the trajectory of all studies in the greater research project.

1. Development of the Overt Behavioral Dysregulation Rating Scale (OBDRS), a novel clinician-administered assessment tool for naturalistic observation of BD in adults with TBI (Tool development)
2. Evaluation of the feasibility of measuring physiological states continuously during task completion via portable biosensor technology for synchronization with observational behavior data from the OBDRS (Study 1: Initial feasibility)
    i. Feasibility testing also assessed several secondary aims:
        1. To evaluate changes in perceived cognition and mood during task completion
        2. Exploration of physiological change patterns throughout the behavioral observation period
        3. Examine correlations between physiological signals to provide insight into factors that may have increased value or redundancy in a predictive model
        4. Assess the relationship between physiological change and BD event frequencies across time periods
2. Evaluation of physiological responsiveness (Study 1: Initial feasibility)
3. Selection of research tasks (Study 1: Initial feasibility)

4. Reliability and validity of the OBDRS (Study 2: OBDRS psychometrics)
5. Development of an EMA method utilizing biosensor technology for exploration of physiological signals that may predict BD events in adults with TBI (Study 3: EMA model)

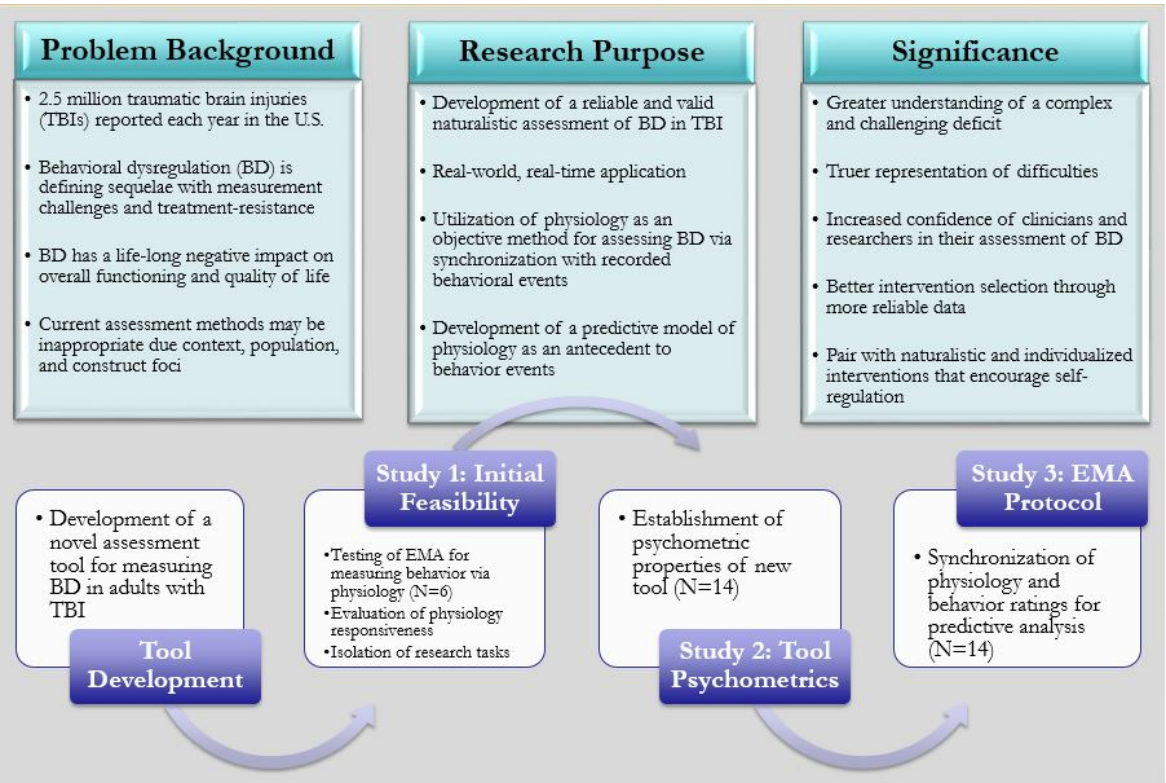


Figure 1. Research graphic

1.2.2 Research Questions

The feasibility protocol examined three research questions. The first question was, is it feasible to continuously measure physiological profiles via a portable biosensor device during the completion of challenging problem-solving tasks for synchronization with recorded behavior events? Secondly, are specific physiological signals more responsive than others are when measured

during a challenging task protocol? Thirdly, can brainteaser puzzles serve as an appropriate and challenging problem-solving task to elicit the construct of interest?

Study 2 was a more comprehensive assessment of the psychometric properties of the OBDRS. The research questions were, (1) does the OBDRS fully and solely measure the construct of BD as it presents in TBI and as defined for the purpose of this study? (2) does the OBDRS correlate with existing validated measures of BD and present as distinct from tools measuring different constructs? and (3) can the OBDRS be established as a reliable measure of BD through rater agreement of recorded behavior events?

Study 3 was an exploratory study focused on examining physiological change during the completion of challenging tasks in adults with TBI, exploring physiological variability both between and within individuals and, evaluating physiological variables as potential predictors for BD. The main research question was, in adults with TBI, can HR, HRV, and RR taken together with age and sex accurately predict BD events?

It was hypothesized that it would be feasible to continuously measure physiological profiles via a portable biosensor device during a behavioral observation period where individuals will be asked to complete several challenging tasks. For the exploratory psychometric study, it was hypothesized that examination of the OBDRSs psychometric properties will provide evidence for both adequate reliability and validity. Finally, for the EMA protocol, it was believed that HR, HRV, and RR when evaluated together with age and sex would serve as a reliable suite of prediction variables for BD.

1.2.3 Definitions of Important Terms

A lack of consensus can be found in the head injury literature for understanding TBI severity, especially for separating mild TBI (mTBI) from concussion. The current research focused solely on closed TBIs, excluding all other head injuries and including mTBI only when references to concussion were not made. This was due to the belief that concussion is a head injury separate in nature to TBI. Additionally, due to the heterogeneity in the SR literature, it was a critical task to operationally define behavioral SR as it pertains to this project. Through consolidation of the literature, as well as informal discussions with rehabilitation professionals working with the target population, behavior SR will be defined as, '*a person's ability to purposefully identify, monitor, adapt, and execute control over his/her behavior to fit in congruence with his/her environment*.'

1.2.4 Basic Limitations

Studying a multi-faceted concept such as SR can be a complex task. This comes as a result of competing expert opinion on what subdomains truly make up SR, differences in terminology across professions, and a highly heterogeneous body of literature—all of which make researching this construct challenging at best. More specifically, research on behavioral SR after TBI within the context of rehabilitation is scarce, and even more difficult to find when evaluating ecologically valid forms of assessing problematic behavior in this cohort. This project sought to pull together the available literature and analyze common themes to isolate a conceptual framework upon which to build the following research studies.

1.2.5 Study Significance

The inability to self-regulate behavior can lead to a variety of negative rehabilitation outcomes for persons with TBI (Braden et al., 2012; Jumisko, Lexell, & Soderberg, 2009; Tate & Broe, 1999). After TBI, behavioral challenges can often become defining sequelae and directly influence a myriad of negative outcomes leading to poor QoL and low participation (Liu, Chan, Lee, Li, & Hui-Chan, 2002; Ylvisaker et al., 2007). Selecting optimal interventions for increasing self-regulatory ability can be a difficult task due to the limitations of current, widely used behavior assessment methods that most likely have led to broad estimates of incidence and prevalence and consequently, an inability to understand the true functional impact of BD (Sabez et al., 2014).

Naturalistic assessment may be an answer to known assessment challenges when attempting to target problematic behavior after TBI. However, despite increasing support for implementing more naturalistic research protocols to infer clinical practice, novel EMA methods are still needed due to more traditional EMA approaches that require the ability to accurately self-report often being inappropriate for disability cohorts who experience co-morbid self-awareness and memory difficulty along with behavior challenges. Due to the connection between human physiology and behavior, continuous measurement of objective physiological parameters may be a viable option for not only detecting the presence of momentary behavior, but to also identify antecedents or even consequences of behavior through unique physiological presentation in the moments leading up to or following an event (Ali et al., 2012). This predictive ability, along with the inherent heterogeneity of physiology across persons, may subsequently lead to the development of not only ecologically valid rehabilitation approaches, but also more individualized assessment methods and intervention delivery models focused on intervening prior to the occurrence of maladaptive behaviors.

2.0 CHALLENGING BEHAVIOR PROFILES IN ADULTS WITH TRAUMATIC BRAIN INJURY: FUNCTIONAL IMPACT, REHABILITATION SHORTCOMINGS, AND NOVEL ASSESSMENT APPROACHES

2.1 TRAUMATIC BRAIN INJURY

A TBI is a type of acquired brain injury where external forces cause sudden trauma to the head and can lead to significant damage to the brain (Brain Injury Association of America [BIAA], 2011). It is currently the leading cause of death and disability for children and young adults in the U.S. (BIAA, 2011). These injuries are different from congenital or hereditary brain damage; all of which occur at birth. TBI is also distinguished from acquired, non-traumatic brain injuries where internal illness is the cause, not force (e.g., anoxia, aneurysm, or stroke). With a TBI, trauma often results from direct impact through skull fractures, penetration, or gunshot wounds, but can also occur as a result of acceleration-deceleration forces, such as in a motor vehicle accident.

Both primary and secondary trauma can occur with a TBI. Primary injuries occur directly from trauma at the moment of injury. Skull fractures, contusions, hematomas, lacerations, and diffuse axonal injury are all examples of primary injuries. Primary injuries are preventable, but not reversible. Secondary injuries can develop anywhere from the seconds immediately following initial force to even days after, sometimes unknowing to the individual—making these symptoms especially dangerous when a head injury is not addressed appropriately following impact.

Common secondary injuries include but are not limited to, hematomas, brain swelling or edema, intracranial pressure, necrosis, ischemia, metabolic changes, and cardiac changes—some of which can lead to death. These injuries are viewed as the evolution of direct impact to the brain. TBIs can be classified into two categories, closed head injury and penetrating injury. Due to variability existing between the two diagnostic categories, closed TBIs will serve as the focal point of this research. Table 1 can be referred to for a detailed breakdown of the differences between closed and open head injuries.

Table 1. Breakdown of closed and open traumatic brain injuries

Type	Characteristics	Common Forces	Primary Injury	Secondary Symptoms
Closed head injury	Skull and Dura Mater intact Brain swells with no place to expand Damage can often be global	Motor vehicle accidents Falls Violence Sports injuries Blasts	Contusion Hematoma Diffuse axonal injury Coup-contrecoup	Concussion Hematoma Cerebral contusion Intracranial pressure Intracranial infection Edema Ischemia Anoxia Epilepsy Metabolic changes Free radical discharge Hyponatremia Personality changes Cardiac changes Hypotension
Penetrating injury	Skull fracture or penetration Focal, localized damage	High velocity projectile Weapon wounds	Lacerations Skull fracture	

2.1.1 Risk Factors

Research has provided evidence for several risk factors associated with TBI. Age has shown to play a significant role when defining risk groups and attributing causal factors of injury. Three age points stand out as being at a higher risk for experiencing a TBI: small children due to falls

and abuse (statistics may be underestimated due to inaccurate reporting or non-reporting; Roehr, 2012); adolescents between 15 and 19 years of age, due to motor vehicle accidents and risk seeking behaviors; and with the elderly there is a spike in TBI rates seen due primarily to falls (Bruns & Hauser, 2003; CDC, 2006; Nolan 2005). The elderly contribute to around 22% of TBI-related hospitalizations. In addition, males experience TBI at a higher rate (59%) than females; making sex a predictive factor as well (CDC, 2006; Bruns & Hauser, 2003). Figure 1 displays a casual breakdown of TBI from statistics reported by the CDC (2010).

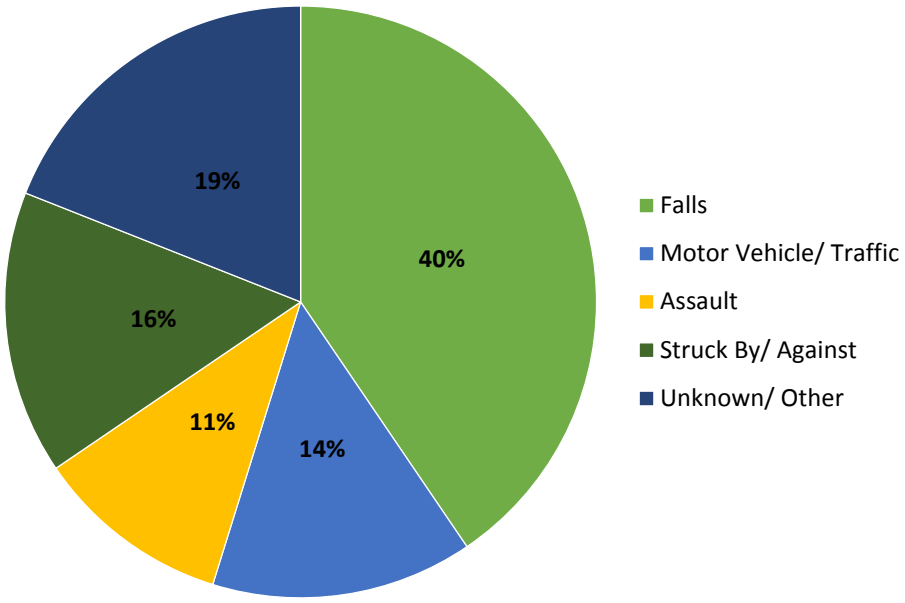


Figure 2. Breakdown of most commonly reported causes of TBI

Studies have also looked at race and ethnicity as potential risk factors for TBI and although there is evidence that young, African-American males are more likely to experience a TBI, this data may be confounded by socioeconomic status (Bruns & Hauser, 2003). It is also important to note that TBI related injuries are costly, with the U.S. spending over 76.5 billion dollars on medical costs related to head injury in 2000 (CDC, 2006; Leibson et al., 2012). These numbers are especially unfortunate when considering many TBI-related deaths can be prevented all together

from wearing seatbelts and using helmets (Nolan, 2005). Just having experienced a TBI significantly increases an individual's risk of sustaining a second or even a third TBI. In addition to what is known about risk factors, when examining TBI over the past decade, although rates of TBI-related emergency room visits have increased by 70%, only 17% of those individuals were admitted or experienced death (CDC, 2010). Statistics like these make it clear that more individuals are living with TBI than ever before, potentially resulting from increased education and/or earlier seeking of an initial diagnosis—with rates only expected to increase. Rising incidence rates speak to a critical need for rehabilitation services not only be available for these individuals, but also for the state of the science to advance along with it.

2.1.2 Diagnosis

TBIs are typically rated on a three-category scale of severity: mild, moderate, and severe. Classification is usually done immediately following trauma by medical professionals through an assessment comprised of multiple sources of information. This assessment serves as a baseline for the individual following trauma and oftentimes, subsequent administrations of specific assessment components (e.g., objective measures) will occur at multiple time points after injury to track progress.

Several criteria are often considered when assessing TBI severity. These indicators include, loss of consciousness (LOC), presence and duration of coma, post-traumatic amnesia (PTA) or confusion, disorientation, and memory loss pertaining to events occurring after the injury, objective assessments, as well as neuroradiological evidence (e.g., computer tomography scanning or magnetic resonance imaging), a positive neurological exam, and neuropsychological

testing (Mayo Clinic, 2012). Rehabilitation professionals use these informational sources as a guide for the development of appropriate and individualized rehabilitation plans for clients.

It is important to note that although these indicators are often used as a general estimate of an individual's injury severity and prognosis, the high level of variability associated with TBIs means this diagnostic approach should be viewed critically. Brain injury is not a homogeneous or highly predictable disability.

2.1.2.1 Novel classification approach

Within the field of rehabilitation, a shift has been seen in recent years on how professionals classify and approach treating TBI. Due to the inaccuracies associated with commonly used classification methods labeling injury solely from medical factors, a novel approach has emerged where focus is placed on the functional impact manifested as a result of injury, and not the characteristics of the injury itself at the time of or shortly after occurrence. For example, a rehabilitation professional may look to the severity of the consequences of injury and their impact on daily living (e.g., extreme behavioral problems or poor working memory), regardless of the injury itself being medically documented as mild, moderate, or severe.

The current research project supports a classification approach centered around the functional impact associated with injury symptomology when conceptualizing behavioral challenges after TBI. Although TBI is one of the more commonly known disabilities where behavioral problems manifest, it is only one of many cohorts in need of more optimal rehabilitation assessment methods and rehabilitation services and supports to help facilitate more adaptable behavioral outcomes.

2.1.3 Prognosis

In addition to risk factors and injury severity, the course of a TBI is also influenced by additional factors including, whether trauma results in global/diffuse (affecting the entire brain) or focal (specific areas) damage, the location of brain exposed to trauma, as well as person-centered factors such as age, pre-injury functioning, motivation for rehabilitation, and support systems. In terms of re-injury, individuals who experience a TBI are 3-times more likely to experience another throughout their lifetime and after a second TBI they are at an even more elevated risk, being 8-times more likely to sustain a third injury.

The most significant gains in terms of restoring functioning are made early on during the acute rehabilitation process, generally within six months post injury—supporting the importance of early intervention. Some researchers believe this critical window to be more generous, extending as far as two years after injury or longer. Regardless of level of functioning regained and the timeline it occurs in, effects from damage to the brain will be long lasting to some degree, especially with moderate to severe TBIs. It is unlikely for symptomology to resolve completely over time without residual effects.

2.1.4 Outcomes

Trauma to the brain can cause individuals to experience deficits in major functional domains including, physical, cognitive, behavioral, and emotional (Jumisko, Lexell, & Soderberg, 2009). Onset of a TBI is sudden, most of the time unexpected, and with that comes an adjustment period to disability that individuals often struggle with. Many of these deficits have been shown to persist even decades after injury (Anderson, Brown, Newitt, & Hoile, 2009). Short and long-term changes

to memory, affect, behavior, personality, concentration ability, language skills, insight, problem solving ability, arousal, and attention are often seen as well as chronic pain, cognitive fatigue, and decreased physical strength (Braden et al., 2012; Tate & Broe, 1999). These changes not only affect the individual, but also have a significant impact on their family, friends, and loved ones—a TBI's impact can be global (Nolan, 2005).

Functional limitations experienced by individuals with TBI often result in educational, social, personal, and vocational problems (Anderson, Brown, Newitt, & Hoile, 2009). Education difficulties can result in academic underachievement, which can often have a direct negative effect on vocational success. Social difficulties can lead to isolation and feelings of loneliness, which restricts full participation in society and reinforces antisocial behavior from the person with the brain injury (Gordon et al., 2006). Many living with a brain injury display poor time management skills and lack punctuality, both of which can serve as barriers for gaining and/or maintaining employment. Extensive research has shown low levels of functioning in multiple major life areas decreases QoL and overall life satisfaction (Jumisko et al., 2009).

SR problems are a common deficit associated with TBI that can have both short and long-term effects (Liu, Chan, Lee, Li, & Hui-Chan, 2002). A person with difficulty self-regulating will often experience negative functional and rehabilitation outcomes. Specifically, this could include, trouble making and keeping friendships, understanding and respecting the perspectives of others (Henry et al., 2006), lack of disability-related awareness, and engaging in risk seeking behaviors. Due to the pervasiveness, complexity, and impact of self-regulatory challenges, this deficit serves as the focal point for the following research.

2.2 SELF-REGULATION

Research on human behavior over the years has approached SR with a great deal of variability and comes with diverging expert opinion on what SR is truly comprised of when broken down into smaller sub constructs (Siegert, McPherson, & Taylor, 2004; Vohs & Baumeister, 2010). Terms such as self-control, self-monitoring, and self-management are often used interchangeably with SR, while other research explicitly distinguishes the terms from one another—claiming SR is a broader concept referring to feedback loops and goal-directed behavior, self-management is reserved for medication management, while self-control refers to impulse control (Vohs & Baumeister, 2010). In the field of psychology, SR definitions have evolved to not simply mean control *of* the self, but a more active human process of control *by* the self. Additionally, debate exists on whether unconscious processes should be included in an all-encompassing definition of SR and not only the conscious processes more commonly focused on (Vohs & Baumeister, 2010). Experts have come to a consensus that both conscious and automatic processes are equally important in human ability to exercise control over oneself; however, some still use the term to refer solely to conscious processes (Vohs & Baumeister, 2010). One definition of SR provided by Vohs and Baumeister (2004) includes the regulation of thoughts, emotions, attentional processes, impulses, appetites, and task performance (Vohs & Baumeister, 2010).

The lack of universally accepted terminology to describe human ability to control the self has made establishing evidence-based practice (EBP) for SR interventions within the field of rehabilitation challenging, specifically in terms of assessment, measurement, and tracking progress. Vohs and Baumeister provide an explanation for this in the *Handbook of Self-Regulation: Research, Theory, and Applications* stating, “The search for a general understanding of SR has not been coherent given the diversity of the field (Vohs & Baumeister, 2010, p.2).”

Despite definitional issues and competing expert opinion, common themes have emerged across interpretations (Vohs & Baumeister, 2010). General agreement exists that SR involves exercising control over the self to maintain regularity and keep the self in line with preferred standard processes (Vohs & Baumeister, 2010). What is also clear is that our ability as humans to exercise control over our external environmental influences and internal processes is vital to our existence. It is likely many rehabilitation and health sciences professionals would agree that cognitive, behavioral, and emotional processes all play a role in and interact together to make up the construct of SR; however, some may still argue these domains in fact represent separate deficits all together.

The provided overview on self-regulation, although necessary to highlight the conceptual and methodological challenges inherent when researching this complex construct, was determined to be too broad of a construct to focus novel pilot work on. Due to the overt, objective nature of behavior, as well as the known direct impact of maladaptive behaviors (e.g., behavioral outbursts in the workplace) on rehabilitation outcomes when compared to more covert cognitions and emotions, behavior was determined to be an appropriate scope and logical place to begin pilot research—especially for application to TBI (Ylvisaker et al., 2007). Therefore, taking influence from the consolidated definition of SR developed from existing literature, for the purposes of the current research behavioral self-regulation was defined as 'a person's ability to purposefully identify, monitor, adapt and control his/her behavior to fit within the context of his/her environment.' Figure 2 depicts the complex integration between the subdomains involved in SR and as it commonly presents in TBI. This figure conveys that although this research focused only on the behavioral subdomain, there is an underlying overlap of sub-constructs that cannot be ignored.

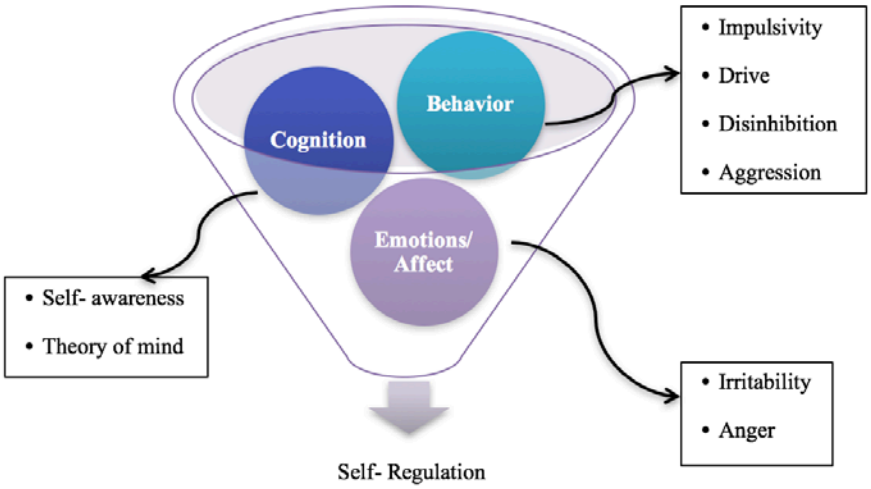


Figure 3. Interaction between the sub-domains making up self-regulation

2.3 BEHAVIORAL SELF-REGULATION

It is estimated that 40 percent of all deaths can be attributed to some form of poor SR (Schroeder, 2007). The most common examples supporting this percentage are situations involving poor decision-making skills or risk-related choices. SR involves exercising control of the self, by the self, to maintain homeostasis, define our existence as human beings, and keep in line with preferred standard processes (Baumeister, Heatherton, & Tice, p.6, 1994; Vohs & Baumeister, 2010). Three main components are believed to make up behavioral SR, (1) monitoring behaviors, (2) motivation to reduce discrepancies between societal standards and actual states, and (3) the capacity to change these discrepancies when barriers are presented. Hofmann, Schmeichel, and Baddeley state that when one or more of these components are not functioning properly, 'self-regulation failure'

occurs (2012). Regulation of the self is a particularly difficult skill for individuals with TBI due to poor self-awareness and other executive functioning abilities.

Many situational factors (e.g., temptation, impaired judgment and control, direct structural damage, or resource depletion) can cause an individual to experience behavioral SR failure (Baumeister & Heatherton, 1996; Heatherton, 2011). Failure often manifests as aggression, impulsivity, and inappropriate sexual and social behavior and leads to poor outcomes related to work, socialization, education, and QoL, but also threats to safety through risk seeking behavior such as criminal activity (Quinn & Fromme, 2010).

Impulsivity, lack of drive or self-initiation, and disinhibition, are all behaviors commonly seen in brain-injured populations who fail to self-regulate at a high level (Ylvisaker et al., 2007). Behavior problems after a TBI are believed to contribute to a much higher degree to low QoL than cognitive or emotional deficits (Ylvisaker et al., 2007). An individual with a brain injury who displays behavior problems will also most likely experience negative social and vocational outcomes as consequences of a disconnect between their behavior and societal standards and expectations.

2.3.1 Self-Regulatory Systems

Executive functioning is believed to either support self-regulatory goal pursuits or lead to SR failure (Hofman, Schmeichel, and Baddeley, 2012; Miyake, 2000). This ‘top-down control’ from the pre-frontal cortex (PFC) over sub-cortical mesolimbic dopamine system regions of the brain is responsible for control over primitive impulses and reward systems (Heatherton & Wagner, 2011; Vohs & Baumeister, 2011). Due to the numerous connections between the PFC and other areas of the brain, direct damage to the PFC has widespread consequences, and reciprocally, damage to

other areas of the brain can easily disrupt normal PFC functioning (Vohs & Baumeister, 2011). Theories of self-regulatory balance state SR failure can occur both when the balance shifts in favor of the sub-cortical areas (known as bottom-up control) or when damage is experienced directly to the PFC. When this balance tips in an unfavorable direction, dopamine receptors are activated and serve as antagonists for impulsive or aggressive behavior (Heatherton, 2011). Extensive research has shown that even when individuals are instructed to specifically inhibit an impulsive desire, a shift in activity can be observed from sub-cortical brain areas to the PFC. When the frontal lobe is directly impaired, individuals show an immediate preference for instant reward over delayed gratification (Brass & Haggard, 2007). This suggests the mechanisms involved with SR are domain-general, but the sub-cortical regions involved may vary based on the stimulus and situation (Heatherton & Wagner, 2011). Therefore, interventions focused on increasing executive functioning may yield improvements in an individual's self-regulatory ability.

2.3.2 Incidence and Prevalence

When consolidating the few known studies on the incidence of challenging co-occurring behavior profiles after TBI**,** Johnson and Balleny estimated 78% of individuals display one or more behavioral challenges in the home 18 months post injury (1996). Verbal and physical aggressive behavior towards others or objects have been estimated to be anywhere from 11-84% (Baguley, Cooper, & Felmingham, 2006; Johnson and Balleny, 1996; Tateno, Jorge, & Robinson, 2014). Additionally, research shows that 50% experience reduced social skills, 39% are disinhibited, 30-46% engage in sexually inappropriate behaviors, more than 50% abuse substances, 80% display inappropriate social behavior, 60% show a lack of initiation and/ or perseveration, and physical

aggression against the self and wandering are experienced in less than 15% of samples (Feeney, Ylvisaker, Rosen, & Greene, 2001; Johnson and Balleny, 1996; Kelly, Brown, Todd, & Kremer, 2008).

In an attempt to analyze the prevalence of behavioral SR deficits after TBI, Sabez and colleagues estimated both physical and verbal aggression as well as irritability can range anywhere from 11-96% (2014). Additionally, 40-71% lack initiation, 6-30% display inappropriate sexual behavior, wandering has been seen in 6-14%, and preservation is estimated at up to 25%. These same researchers then conducted a comprehensive, multicenter study hoping to mitigate some of the weaknesses in the literature by analyzing behavior systematically in over 500 community-dwelling adults with TBI (Sabez et al., 2014). Results showed an overall prevalence of challenging behaviors of 54%. When breaking down behaviors individually, inappropriate social behavior (33.3%) and aggression (31.9%) were the most prevalent, and 35.5% displayed co-occurring behavioral challenges. These broad and varying statistics make understanding the impact of behavior profiles post-TBI difficult at best and suggest a greater need for more systematic assessment methods.

2.3.3 Functional Impact

Individuals who self-regulate have the ability to implement what Baumeister and colleagues refer to as 'self-stopping,' or the ability to change behavior patterns by generating alternatives or halting the process all together (p. 7, 1994). Individuals with TBI often lack this self-stopping ability.

When studying correlations between challenging behavioral profiles and functional outcomes, Sabaz and colleagues found as aggression increased, so did the need for support, the level of functional disability, and mental health concerns, while participation declined (2014). The

limited literature makes it difficult to establish temporal relationships, as research has yet to provide evidence for whether behavioral challenges lead to functional decline after injury, or if the early stages of functional decline exacerbate BD (Arciniegas & Wortzel, 2014). Research has shown these problems to be extremely persistent, still being present several years post injury and often times increasing in severity when left untreated (Ylvisaker et al., 2007). A study conducted by Winkler, Unsworth, and Sloan showed that lack of emotional control (e.g., irritability and anger) led to maladaptive behaviors and community isolation an average of 8.8 years after sustaining a TBI (2006).

2.3.4 Summary

The inability to self-regulate behavior can lead to a variety of negative rehabilitation outcomes. Behavioral SR deficits are believed to stem from impaired top-down control of the PFC cortex over sub cortical areas of the brain associated with impulsivity and reward seeking. After TBI, behavioral challenges can be become defining sequelae and directly influence a myriad of negative outcomes and lead to poor QoL. Selecting optimal interventions for increasing this self-regulatory ability is a difficult task due to weaknesses in current behavior assessment methods that have most likely led to broad estimates of incidence and prevalence and consequently, an inability to understand the true functional impact of BD.

2.4 ECOLOGICAL MOMENTARY ASSESSMENT

Human beings participate in ecological momentary assessment or EMA every day most likely without even realizing it. EMA involves the delivery of assessment methods in the real world, with real-time or momentary data collection (Heron & Smyth, 2010; Shiffman, Stone, & Hufford, 2008; Stone & Shiffman, 1994). It is an alternative to more traditional assessment approaches conducted within the clinic or laboratory setting that is rooted in the importance of the relationship between the individual and the environment. Although EMA research to date has focused primarily on smoking cessation and mood, its use is expanding. Advances in technology have now made it possible to assess difficult to measure constructs through more objective and reliable data collection methods. Behavior tracking through mobile technologies or wearable physical activity monitors are just two examples of EMA application. Due to the well-known measurement difficulty associated with challenging behaviors and their impact on rehabilitation outcomes in TBI, EMA is suggested as a viable assessment method for this cohort.

EMA (also known as ambulatory assessment, naturalistic assessment, or experience sampling; Runyan et al., 2013) can signal individuals to report on their current state relevant to pre-determined outcomes of interest (e.g., mood or physical symptoms), observe behavior, or conduct contextual assessments of the environment (Intille et al., 2003). This often occurs through portable technology devices, such as smartphones or wristwatches, where data is most commonly collected at multiple randomized time points throughout the day for several continuous days. Alternatively, EMA can be non-disruptive and occur during daily living, without any need for conscious effort or input from the user (e.g., ambulatory physical activity monitors; Heron & Smyth, 2010). This method of momentary data collection eliminates dependency on recall ability and reduces concerns over inaccurate self-reporting by only asking individuals to report on

experiences at an exact moment in time. Earlier uses of EMA in the literature commonly used paper diaries, but weaknesses from this technology (e.g., information integrity and compliance rates) led to the development of more reliable sources of data collection less susceptible to biases (e.g., behavior observation or wearable self-monitoring systems; Fahrenberg, Myrtek, Pawlik, & Perrez, 2007). EMA can be used independently for data collection alone, or to infer deployment of a full community-based ecological momentary intervention or EMI. EMA delivery models highlight the importance of true representation of difficulties experienced in the natural environment, where it is believed assistance is most needed (McCue, Fairman, & Pramuka, 2010). This is in opposition to approaches requiring skill learning or insight development to occur in a highly controlled environment where it is believed challenges may be presented or reported in a biased and unnatural manner, which limits rehabilitation professionals' ability to intervene effectively when unreliable data is a concern (Smyth & Stone, 2003).

2.4.1 Applicability in Rehabilitation

In rehabilitation, naturalistic delivery models present an opportunity for more ecologically valid assessment methods; from data collection done by a clinician directly in the environment of the individual (e.g., functional based assessments) to remote monitoring facilitated through technologies. This is potentially a more methodologically sound process due to the elimination of the possibility of finding false relationships in the laboratory or clinic that do not similarly occur in real life. EMA also proves to be a highly useful approach when it may be difficult or unethical to simulate change in a controlled setting without causing unnecessary harm, but the event may happen naturally in the real world (e.g., stress inducing situations; National Commission for the Protection of Human Subjects of Biomedical and Behavioural Research, 1978). Real world

assessment methods provide the opportunity for adaptation in novel settings due to the inability to predict pseudo-randomly occurring environmental stimuli.

The real-time data collection provided by EMA lends itself to additional therapeutic advantages. A clinicians' presence can be extended into the individual's natural environment as opposed to only during finite clinic hours. This continuous opportunity for data collection allows for more effective use of clinician time, due to potentially less traveling to see clients and overall time needed for gathering reliable information (Heron & Smyth, 2010). Difficulties often experienced by clinicians when attempting to get clients to collect data between sessions to further facilitate the rehabilitation process are lessened due to events of interest and recording occurring directly in the real world as challenges naturally manifest. As a result, positive effects may be seen in terms of compliance rates, which may lead to less services and supports needed over time and less follow-up or 'booster services.' By freeing up clinician availability and increasing compliance, EMA may prove to be a cost-effective and timely assessment approach that supports a more efficient rehabilitation process through reliable data to build rehabilitation plans from that yields long-term positive outcomes.

Momentary collection of objective data also means less reliance on self-report data, which proves to be problematic when working with disability groups who, due to cognitive challenges or mental health concerns, have difficulty reporting accurately. Furthermore, spontaneous data collection eliminates the effect of diminished recall ability, or what Smyth and Stone refer to as 'retroactive reconstruction,' where the individual may recall an event differently due to the outcome (2003). This can be especially true when long periods of time separate an event and data collection. Retroactive reconstruction, which can truly invalidate recall data, can often happen when the individual's current state or mood is negative during the moment when they are asked to

remember an event or situation (Smyth & Stone, 2003; Yoshiuchi, Yamamoto, & Akabayashi, 2008). A less biased data collection method can result in greater confidence in assessment results and consequently, the interventions inferred from the data.

Additionally, by delivering assessments during daily living, EMA provides the ability for clinicians to reach greater numbers of people, including those who may be unable to consistently make it into a therapeutic setting (e.g., persons living in rural areas). A method that reaches clients more easily may also positively influence compliance rates and limit 'fake compliance' (or disguising compliance through recording of data at other times aside from the designated times). EMA also allows for examining time-sensitive phenomenon and rapidly changing states (e.g., physiology), while also being able to account for data patterns that may vary depending on the time of day (e.g., the influence of cognitive fatigue on task performance) and examine temporal relationships between variables to potentially infer cause and effect.

Arguably, the most important factor when evaluating the utility of EMA for rehabilitation is this delivery model's focus on self-management. The use of technology as an awareness tool is intended to gradually fade over time to encourage independent SR. By this time, it is expected new naturalistic skills have been learned and implemented effectively and the user now begins to drive the rehabilitation process, not the technology, until its presence as an assessment tool is no longer necessary. Skills are then reinforced long-term through positive outcomes of their use once the technology is no longer an active component of the rehabilitation plan.

2.5 BIOSENSORS AS AN OBJECTIVE EMA APPROACH

Biosensor facilitated EMA can provide ambulatory monitoring of individuals who are considered at-risk for challenging behavior (Fahrenberg et al., 2007). The physiologic system has long been directly linked to our emotions, cognitions, and behaviors (Lang, 1978) and researchers have theorized covert physiology may actually directly influence the development or maintenance of challenging behavior (Cohen, Yoo, Goodwin, & Moskowitz, 2011). This theory suggests that physiological events may serve as antecedents, concomitants and/or consequences of challenging behavior, and that certain physiological patterns (e.g., increased HR or decreased HRV) may increase the likelihood an individual engages in a particular behavior. Cohen and colleagues provide support for this, showing that interventions are twice as likely to yield successful outcomes if they target the underlying functions of behavior (2011). Physiology is believed to be one of those underlying functions. Although much of the EMA and physiology literature examines the relationship between physiology and mood regulation, it is possible there is a greater ability to answer research questions related to SR from a behavioral approach due to the clear visual presence or absence of behavior in a given moment when compared to more subjective emotions or cognitions.

Physiology is a multifaceted and highly individualized system both across and within human beings (Wilhelm, Pfaltz, & Grossman, 2006). This characteristic allows for rehabilitation assessment methods using biosensors to be tailored to the individual and dynamic as changes occur over time. However, due to this inherent variability, physiological responsiveness patterns may vary across persons. Therefore, it is recommended that several physiologic parameters be assessed to give a broader assessment of behavioral states (Wilhelm, Pfaltz, & Grossman, 2006). Naturalistic physiological monitoring can be highly useful as one component of a multimodal

assessment process that provides a comprehensive picture of the challenges experienced by an individual (Fahrenberg et al., 2007).

Behaviors are considered more spontaneous than emotions or cognitions; therefore, existing EMA questionnaires focused on mental processes and attitudes are not suitable for measuring behavior (Fahrenberg et al., 2007). Whereas a survey-based EMA delivered through a smartphone application can alert an individual at randomized time points during the day to self-report on how they feel in that exact moment, this same approach would be inefficient for measuring behavior due to the inability to simultaneously engage in a behavior while reporting on the act. Therefore, measuring behavior through more traditional EMA approaches still relies on retrospective recall to some degree, which is not appropriate for application with individuals with disabilities who experience impaired self-awareness or limitations in other executive functions such as motor skills, processing speed, and memory (e.g., TBI; Fahrenberg et al., 2007). Using biosensors as an EMA approach eliminates the need for retrospective recall or self-awareness, and reduces client burden, as physiology is monitored effortlessly without any input from the user.

Naturalistic monitoring also provides a variety of data-sampling strategies for analyzing the relationship between physiology and behavior that make it possible to answer a diverse set of research questions ranging from, continuous monitoring without further intervention to, interactive monitoring where detected physiological change prompts the user to engage in momentary self-reporting and symptom self-monitoring where the user practices self-management skills to take control of a behavioral problem and implement or seek out adaptive responses (Fahrenberg et al., 2007). These strategies also allow for the evaluation of temporal relationships to be explored between behavior occurrences and physiology.

Finally, the reliable assessment of physiological correlates of behavior is possible due to technological advancements allowing for non-disruptive physiological monitoring during daily living. When assessing physiology through wearable technology, the technology itself can become reinforcing to the individual due to the personal learning gained about unconscious bodily states that would otherwise go unnoticed. This awareness training may supports SR skill building through monitoring and adapting physiological states independently over time, as technology is eventually faded out. It is important to note that research on the utilization of biosensors as an objective EMA approach for measuring behavior is still in an exploratory state, and research is needed to empirically connect measured physiology with maladaptive behavior in the target group. The following studies sought to support this research direction.

2.5.1 Naturalistic Assessment of Behavioral Dysregulation Using Biosensors

The current gold standard for assessing behavior is clinician behavioral observation (Fahrenberg et al., 2007). But what happens when the individual goes into the real world without the clinician by his/her side? Often times, assessment in-between sessions halts because services are not directly provided by a professional or the ability to generalize learning across contexts is impaired. Additionally, the majority of EMA research utilizes self-report questionnaires aimed at asking participants to recall experiences from periods of time that are often loosely defined (e.g., 'over the past few days'), despite multiple reports providing evidence that this methodology may not accurately record experience, but instead record subjective mental representations of experience (Fahrenberg et al., 2007). This speaks to a greater need for reliable ambulatory assessment methods for measuring behavior that utilize more objective data measurement approaches.

2.5.2 Sensor Identification

Recently, physiological states have been explored as an objective EMA data collection method for measuring behavior when traditional self-reporting EMA approaches may be unreliable or inappropriate. This reliance on self-reporting is a limitation of many current assessment methods used with disability cohorts who have difficulty reporting accurately (e.g., TBI). Research questions on naturalistic behavior that were once difficult to answer due to methodological limitations may now be possible to systematically study through EMA models using biosensor technology. However, due to a lack of evidence on physiological correlates of behavior for the target population, an overview of psychological parameters of impulsivity and aggression was conducted due to both behaviors being characteristic of BD after TBI and the extensiveness of literature dedicated to this area of study.

Cardiovascular and electrodermal activity (EDA) physiological measures make up the majority of research on physiological correlates of both aggression and impulsivity. Results from a study evaluating physiological correlates of impulsivity in a non-clinical sample found low HRV during task completion similar to previous research findings on impulsive individuals (Patrick, 2008). In an extensive review of the literature on psychophysiological correlates of aggression, a trend was shown between individuals who display hostility (often through verbal and physical aggression) and increased cardiovascular reactivity when in the presence of stress or provocation (Patrick, 2008). Additionally, confirmation for studying both HR and EDA when measuring aggression also comes from a meta-analysis conducted by Lorber where the relationship between aggression and both HR and EDA was replicated; however, larger, moderate effect sizes were seen for HR when compared to EDA (2004).

Although this body of evidence is characterized by conflicting outcomes due to methodological and theoretical differences, non-replication, and heterogeneity of the behavior constructs measured, overall, the research supports the responsiveness of both HR and EDA in the presence of executive functioning deficits, decreased inhibitory control, and increased aggression (Fowles, 2000; Iacono, Carlson, & Malone, 2000). Impulsive individuals seem to display an under arousal of autonomic functioning, while aggressive behavior is linked to the inverse relationship—autonomic hyperarousal (Knyazev, Slobodskaya, & Wilson, 2002). Therefore, there seems to be a true association worth exploring between behavior challenges and physiology. Directional differences in the relationship between physiological effects and behavioral challenges suggest there may be value in analyzing target behaviors individually; however, the commonly co-occurring nature of challenging behaviors makes exploring the overall construct of behavioral SR a logical starting point.

When consolidating the evidence on physiology and behaviors of dysregulation, HR, HRV, and electrodermal skin response all emerge as physiological states worth exploring in an EMA model for measuring behavioral SR deficits in adults with TBI. Additionally, due to the known influential relationship breathing rate has on overall autonomic functioning, especially its direct connection to HRV, and documented differences in physiological presentation across sexes (de Geus, Willemsen, Klaver, & van Doornen, 1995; Kring & Gordon, 1998), both sex and RR are also important factors to consider.

2.6 SUMMARY

Naturalistic assessment may be an answer to known assessment challenges when attempting to accurately measure problematic behavior after TBI. However, the development of novel EMA methods that do not rely on self-report or retrospective recall are needed due to these components being inappropriate for disability cohorts who experience low self-awareness and poor recall ability along with behavior challenges. Due to the connection between human physiology and behavior, continuous measurement of objective physiological parameters is proposed to be a viable option for the ability to detect the presence of momentary behavior, and to identify antecedents or consequences of behavior through unique physiological presentations in the moments leading up to or following an event (Ali et al., 2012). This potential for prediction, along with the inherent heterogeneity of physiology across persons may subsequently lead to the development of ecologically valid and individualized intervention delivery models focused on intervening prior to the occurrence of maladaptive behaviors. Given preliminary yet promising results from the literature, this is an area worthy of research.

The link between covert physiological states and overt behaviors may lead to more reliable assessment approaches for measuring behavior SR deficits in adults with TBI through naturalistic delivery models. Research on impulsivity and aggression suggests EDA, HR, and HRV may be the most responsive physiological correlates of behavioral SR challenges. The current research provided insight into the ability to examine temporal relationships between physiology and behavior, and therefore, explore predictive factors. Based on the literature, known physiological covariates, and methodological considerations, it was determined HR, HRV, RR, age and sex would be utilized in this research as predictors to explore in a hierarchical linear mixed model analysis to determine the predictive utility of physiology on BD. Promising results from this pilot

research provide support for replication testing of a novel EMA model for measuring behavior outcomes historically characterized by methodological challenges in similar cohorts. The current research is an initial and critical stage in developing a reliable EMA method for behavioral challenges experienced after TBI that may lead to better identification of successful interventions for increasing self-regulatory ability and ultimately support positive long-term rehabilitation outcomes.

3.0 TOOL DEVELOPMENT

3.1 RESEARCH PURPOSE AND AIMS

The aim of this study was to develop a novel clinician-administered assessment tool for measuring BD naturalistically in adults with TBI. This tool was developed for comparison with the current gold standard of behavioral observation by a clinician in hopes the OBDRS tool may serve as an improved assessment method for BD.

3.2 METHODS

3.2.1 Tool Development

The OBDRS is a clinician-administered assessment tool for naturalistic observation of BD in adults with TBI. It was developed due to a clear gap in the literature and lack of availability of tools for measuring BD naturalistically in the real world, in real-time in adults with TBI. The final version of the tool was developed through an iterative process from consolidating the literature on BD in TBI, analyzing the strengths and shortcomings of similar existing assessment tools, and receiving informal input from rehabilitation clinicians on what behaviors they typically see in their own clients that may be indicative of BD and lead to subsequent task breakdown (see Appendix A

for the OBDRS tool). It is important to note that the OBDRS was designed as a research tool and not a clinical tool to be used in real-time.

This scale was designed to categorize types of behaviors that are believed to be indicative of BD during task completion specifically in the target group. The OBDRS measures behavior type, frequency, duration, off-task behavior, and task disengagement/engagement. The OBDRS measures six target behaviors: disengagement, opposition, non-verbal (non-violent) dysregulation, physical aggression/violence, verbal dysregulation, and perseveration. Each behavior was operationally defined for the purpose of this research. Definitions of each observation period, as well as guidelines for properly recording both off-task and task engagement behaviors are also provided in the appendix. The tool is to be completed in a post hoc manner, through clinician ratings of behavior based on the defined coding scheme used while watching video recordings of behavioral observation sessions.

Using the original version of the tool, which was formatted through an embedded excel worksheet, throughout each observation period from start to finish if an individual engaged in a target behavior a clinician began working across a row in the record sheet recording: the observation period, time of behavior onset, the type of target behavior, if the individual was off-task during this behavior, a brief description of the behavior, whether or not the individual re-engaged in the task, and the time of re-engagement (if applicable). Each row represented a single behavior incidence. Each behavior incidence was recorded even if the individual redirected independently or if the onset or continuation of the behavior did not disrupt task engagement. The OBDRS can also account for events that fall in more than one target behavior category (e.g., throwing an object while also refusing to continue the task), as well as behavior incidences whose start-stop times overlap. The tool produces a total score, representing the overall frequency of all

behaviors of dysregulation during the observation period, a sub-score for measuring the frequency of each of the six target behaviors individually and, total time in seconds spent engaging in behaviors of interest. Sub-scores allow for the examination of differences across behavior type. The OBDRS is intended to be synchronized with objective physiological data recorded from external biosensor technology.

The original format of the tool proved to be labor intensive for clinicians when conducting behavior ratings. In order to increase usability, the tool was modified to be embedded into a behavioral coding software program as a behavioral coding scheme. This allowed for review of video recordings and behavior ratings to occur within one integrated system and make coding a more streamlined experience for the clinician raters.

4.0 MEASURING BEHAVIORAL DYSREGULATION VIA BIOSENSOR TECHNOLOGY: INTITIAL FEASIBILITY

4.1 RESEARCH PURPOSE AND AIMS

The primary aim of this study was to (1) examine the feasibility of a biomeasure device to accurately record changes in physiological activity that correlate with behavior during task completion, (2) evaluate physiological responsiveness, and (3) select research tasks. The following four secondary aims were also examined to provide a wide breath of data to gain a greater understanding the constructs measured and infer future pilot research.

1. Evaluate self-reported changes in cognition and mood over time during task completion
2. Explore trend patterns in physiological signals across time periods
3. Assess correlations between physiological signals that may provide insight into factors that may have the greatest utility in a predictive model
4. Examine the correlation between physiological change and clinician ratings of behavioral (gold-standard)

This phase of the research project was intended to be exploratory in nature due to the limited amount of available research to extract meaningful effect sizes to guide future research efforts. Therefore, no formal hypotheses were developed prior to running the protocol.

4.2 METHODS

4.2.1 Variables

Several physiological measures served as continuous independent variables in this study. Skin temperature (ST) was recorded using a temperature sensor taped to the body using medical self-adhesive tape. Skin conductance or GSR was recorded from two surface electrodes placed around the wrist. HR was obtained from an electrocardiogram (ECG) signal on the wrist. HRV or the difference between the maximum and minimum HR recordings within an observation period was extracted from the HR data. RR was recorded using a respiration sensor band worn over light clothing around the abdominal area with the central part of the sensor just above the navel. Alpha and beta waves were derived from overall brain activity recorded through an electroencephalogram (EEG) sensor. All signals were sampled up to 2048 hz per second. Behavior dysregulation served as the binary categorical dependent variable (0= absent; 1= present). Perceived cognition and mood, impulsivity, and daily stress were also examined to gain greater insight into whether any participants were experiencing higher levels of impulsive behavior and stress prior to completing the study. The current study also sought to examine the relationship between subjective cognitive and emotions states and objective physiological change and behavior.

4.2.2 Assessments

4.2.2.1 *Overt behavioral dysregulation rating scale*

The OBDRS is checklist developed for the purposes of this research that produces ratings on whether specific behaviors indicative of dysregulation are present or absent during a task

observation period. Data from this tool allowed for associations to be evaluated between behavior and physiological measures due to the external time stamp on the checklist and the internal time stamp embedded within the biomeasure software. Two trained clinicians each completed post hoc behavioral analyses using the tool through video recordings of each participant. See Appendix A for a copy of the scale.

4.2.2.2 *PROMIS cognition and mood scale*

A scale for measuring momentary cognition and mood scale was developed for the purposes of this research from the Patient Reported Outcomes Measurement Information System (PROMIS) database. PROMIS is a databank of highly reliable and precise patient-reported measurements on physical, mental and, social health and well-being. Ten items with established individual reliability through item response theory were selected due to their ability to measure cognitive and mood states and were merged into one questionnaire. All items are self-scored on a Likert scale ranging from 0 (never) and 4 (always). The total score of all items was used for analyses, where higher scores indicated greater difficulty with cognition and mood.

4.2.2.3 *Barrett impulsiveness scale*

The Barrett Impulsiveness Scale (BIS) is a widely used 30-item questionnaire measuring impulsivity. It is a self-administered tool where individuals are asked to rate each item on a 4-pt Likert scale ranging from 1 (rarely or never) to 4 (almost always/always). A total score is calculated from summing the ratings from the individual items, where higher scores indicate greater impulsivity. Three second-order factors, and six first order factors or subscales can also be derived from the BIS, attentional (further deduced into attention and cognitive instability), motor (further deduced into motor and perseveration), and non-planning (further deduced into self-

control and cognitive complexity). It is recommended researchers report the total score and at least the second order factors when using the BIS.

4.2.2.4 *Daily event stress scale*

The Daily Events Stress Scale is a single-item tool developed specifically for this study to assess the level of stress participants are experiencing in daily living around the time they complete the study. Individuals are asked to rate on a Likert scale from 1 (not stressed at all) to 10 (the most stressed) on how much life stress they have experienced during the past seven days. The item score was used for analyses, where higher scores indicated greater stress levels.

4.2.3 Participants and Recruitment

4.2.3.1 *Healthy volunteers*

Six healthy University of Pittsburgh undergraduate students severed as volunteers for this study. Despite this research being focused on behavioral challenges after TBI, it was determined that due to the limited available evidence to guide research efforts, it was appropriate and ethical conscious to run feasibility research on a neurotypical population to examine feasibility before exposing a more vulnerable population to the protocol. Aside from being enrolled in an undergraduate rehabilitation psychology course, no additional inclusion/exclusion criteria were used for the selection of the healthy volunteers in this study. Volunteers were recruited through an informal flyer read aloud during class and passed around. Interested individuals provided their contact information and consented for the research coordinator to contact them.

4.2.3.2 *Recruitment*

A convenience sampling method was used for identifying both the healthy volunteers and clinician raters for this study. Master's students were enrolled in the Clinical Mental Health and Rehabilitation Counseling program and participants were sampled from a class taught by the primary investigator on the project. No formal sample size estimations were conducted prior to recruitment due to the exploratory nature of this first study.

4.2.4 Clinician Training

Two Master's level students in the University of Pittsburgh Clinical Mental Health and Rehabilitation Counseling program were formally trained to serve as clinician raters for this study. Aside from enrollment in their academic program, no additional inclusion/exclusion criteria were used for sample selection.

Clinician raters were trained through a standardized protocol occurring over a 3-week period prior to rating participant videos. Training included reviewing articles on relevant content, and learning how to administer and score the OBDRS for four mock participant videos. Each video scored was reviewed in detail by the research coordinator and the raters individually. Successful completion of training was determined once both raters establish inter-rater reliability via a kappa statistic of at least .80.

4.2.5 Instrumentation

4.2.5.1 *Video recording system*

The Rehabilitation Engineering Research Center on Telerehabilitation (RERC-TR) developed the Versatile and Integrated System for Telerehabilitation (VISyTER; Parmanto et al., 2010). VISyTER is an integrated system that provides both real-time and asynchronous communication channels to support the collaboration and delivery of rehabilitation services remotely. VISyTER was used for video recordings of all participants and as the database for archiving sessions for later review. Participants were continuously recorded during all research activities, and when a recording was completed, the entire session was archived by the secure portal. This archive was treated as private data in the same secure fashion as each participant's identifiable information and was not removed from the secure TR server. A back-up video was also recorded using a standard external video camera, and these files were stored securely.

4.2.5.2 *The observer xt software*

The Observer-XT is a widely used software package developed by Noldus that allows for the collection, presentation, and analysis of behavior data. Noldus allows for behavior coding tools, like the OBDRS to be embedded into the software program and allows for clinician ratings to be easily compared, analyzed and synced with other external data (e.g., continuous physiological data; Zimmerman, Bolhuis, Willemsen, Meyer, & Noldus, 2009). Videos were uploaded temporary into the software for rating purposes, but were not stored in the Noldus system. All videos were stored on the secure server at the University of Pittsburgh as mentioned above.

4.2.5.3 *Nexus-10 biofeedback system*

Sten's Corporation in San Rafael, California produced the Nexus-10 Biofeedback System. It is a 10-channel biofeedback system with the portability to be carried on the person using the device; therefore, allowing for continuous measurement of physiological data in any context. "*The system can monitor physiological responses (e.g., brain activity, blood pressure, muscle tension, HR, respiration, skin conductance, temperature, and etc.), with a 10-channel interface device that communicates with any computer using Bluetooth. The portable system can be used to monitor emotions, attitudes, perceptions, and mental processes that can be helpful for professionals developing innovative devices, systems, and interventions for people with behavioral limitations* (Sten's Corporation, 2015)." In this project, the device was worn by study participants throughout all research activities (including pre and post testing measures). Figure 4 depicts a participant wearing the device in the testing room.

Figure 4. Participant connected to Nexus-10 device.

4.2.6 Identification of Research Tasks

Due to the potential for this research to yield results translational to rehabilitation clinical practice when working with individuals with TBI to achieve independence and employment goals through increased self-regulatory skills, the researchers felt it was necessary for the selected tasks to be functional in nature, where individuals would be required to implement problem-solving and decision making strategies—both necessary skills for living independently and gaining and maintaining employment, but common functional deficits experienced by persons with TBI. After research and informal testing of several potential functional tasks, off-the-self challenging brainteaser puzzles were selected as the research task for the current protocol (Figure 5).

Figure 5. Off-the-shelf problem solving puzzles

4.2.6.1 *Research tasks*

LEGO blocks were selected as the baseline task. This task was intended to be non-stressful, and provide baseline physiological measures to be utilized as thresholds during completion of the

problem-solving task later on in the protocol. The baseline task required participants to build a structure out of LEGO blocks. A LEGO set and an easy-to-follow instruction guide was provided. The duration of the baseline task was 10 minutes long, which was believed to be enough time to gather accurate and stable physiological data representative of each individuals' true resting state.

For the problem-solving task, three different off-the-shelf challenging puzzles series were presented to each participant (Rubik's cubes, Perplexus epic globes, and hanayama metal puzzles). Each task series consisted of three tasks each with a different level of difficulty (easy, medium, and hard) to allow for task completion to get increasingly more challenging. Participants were told they had 20 minutes to work on solving the puzzles and that compensation would be awarded for each puzzle they solved correctly. This choice was given to participants as to how they worked through attempting to solve the puzzles to ensure effort.

The recovery task was 10 minutes long and required participants to read freely through a stack of provided magazines. This task was intended to serve as a time period where participants can engage in a non-stressful task before leaving the testing room.

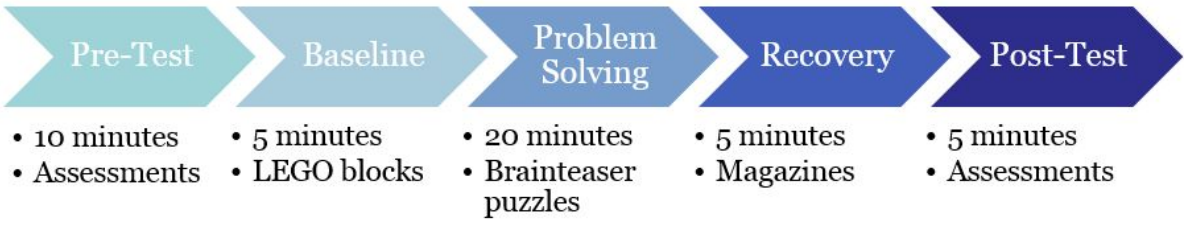


Figure 6. Feasibility protocol graphic

4.2.7 Procedure

This study followed a descriptive cross-sectional cohort design. It was non-experimental and therefore, did not utilize randomization, blinding, or allocation processes. Participants completed all research activities over one session that took no longer than 1.5 hours to complete. Participation involved having participants complete a series of challenging problem-solving tasks, as well as filling out questionnaires prior to, during, and after task completion. The research tasks were administered face-to-face, one time to all participants. The research coordinator was responsible for running all participants through the study protocol as well as administering and scoring all assessments.

After setting up the device, participants completed all pre-test assessments. Afterwards, participants were presented with the baseline task. After baseline, participants were introduced to the problem-solving task. Participants were given 20 minutes to attempt to work through the puzzles. Completion of the easiest level earned them $5; completion of the medium level earned them $10; completion of the highest difficulty level earned them $15. Participants were informed that they could choose to work on the series however they wish (e.g., working with one difficulty level for the entire time period or attempting several different levels of difficulty or puzzles). A difficulty level (e.g., easy) did not need to be completed before moving on to another—moving between tasks and returning to a previously attempted task was allowed. At six time points throughout the 20-minute duration (beginning, 5^{th} minute, 10^{th} minute, 15^{th} minute, 20^{th} minute, and immediately following task completion), the research coordinator checked in with the participant to ask them to stop what they were doing and complete the same PROMIS scale they completed as part of their pre-testing assessment. These ratings were recorded on a participant self-report rating form. Immediately following the problem-solving period participants completed

the recovery task as their final research task of the day. After the participants completed the recovery task, he/she was asked to complete the PROMIS scale a final time. Figure 6 displays a graphic of the study protocol.

After completion of all research tasks, the associated compensation was provided to the participant for each task successfully completed. Time was kept using a countdown timer visible to both the research coordinator and participants during the entire observation period. At any point during the task participants could choose to not continue working. In the event that this happened, the research task would have ended before the time cap expired—this did not occur though. All research efforts were approved by the University of Pittsburgh Institutional Review Board (IRB) prior to running any research activities (#PRO13090443).

4.2.8 Facilities and Personnel

All research activities were performed in the Rehabilitation Science & Technology (RST) Department in Forbes Tower at the University of Pittsburgh in Pittsburgh, Pennsylvania. All members of the research team were affiliated with the RST department and possessed the appropriate qualifications to run a study of this nature.

4.2.9 Data Collection

All data collected was kept confidential by giving each participant an arbitrary identification number, which served as a link between raw data and the individual providing it. This confidential data was entered into a spreadsheet (using the identification number for coding purposes) by the research coordinator for statistical analyses and stored separately (in a data safekeeping room

within the Pitt-RST facilities) from any identifiable information. This process allowed raw data to not be identified, even by other members of the research team who had access to it.

4.2.10 Data Source and Analyses

The data set analyzed contained data for six individuals on variables related to behavior, mood, cognition, stress, and physiology. Physiology served as the sole continuous independent variable with eight levels (representing the different signals measured), while mood, cognition, and stress level were analyzed as nominal independent variables. Behavioral dysregulation was the dependent binary variable. Statistical analyses examined the ability for a biomeasure device to accurately record changes in physiological activity that correlate with behavior events of interest (e.g., task breakdown) during task completion. In addition, changes in cognition and mood over time, physiological signal patterns, inter-correlations between physiological signals and synchronization between physiology and BD events were explored. All statistical analyses were conducted using the Statistical Package for the Social Sciences Software (SPSS, v. 22) and the Observer XT Software.

This exploratory data analysis focused on descriptives and correlations only. Means, standard deviations, effect sizes (Cohen's *d*), and frequencies were used for all descriptive analyses, and Pearson product moment correlations and the kappa statistic were used for measuring inter-rater reliability.

4.3 RESULTS

4.3.1 Participant Demographics

Six individuals participated in this study, four females and two males. Age (M=21, SD=1.37) was normally distributed. Most students were female (66.7%), equally distributed across sophomore, junior, and senior education levels (33.3% each), and the majority were psychology majors (83.3%) and right-handed (83.3%).

4.3.2 Pre-Testing Assessments

All assessments showed non-normal distributions ($p<.05$). These distributions along with the spread shown in the data (BIS [M=69.8, SD=7.06]; PROMIS [M=23, SD=8.22]; Daily events stress scale [M=4.8, SD=.41]), were believed to be a result of both the small sample size and the inherent variability of these constructs between college students. Average total scores on the BIS seemed to be slightly elevated (M=67, SD=7.06) from expected norms, but this result was not significant. Elevated scores on the BIS may be easily explained by the students preparing to enter their midterm exam week at the time of testing, which most likely had an impact on their performance on these measures. Although the overall BIS score was not normally distributed (*p*-value = .002), all three BIS subscales displayed normal distributions.

4.3.3 Cognition and Mood over Time

PROMIS cognition and mood mean scores proved to be stable throughout the majority of the problem solving period, with only the difference between the final five minutes (M=19.17, SD=7.36) and the first five minutes (M=17, SD=5.87), showing a clinically meaningful, small effect (*d* = .24). This result suggests it may take more time during a stressful task for participants to self-report changes in mood or cognition. Clinically significant differences were seen in mean change scores when comparing the overall problem solving period (M=18.04, SD=7.09) to baseline (M=13.83, SD= 5.95) (*d* = .64), indicating a medium effect, and a small effect was observed when comparing the recovery period (M=15.17, SD= 5.31) to baseline (M=13.83, SD= 5.95) (*d* = .24). This data shows that participants experienced greater perceived difficulty controlling their cognitive and emotional state during the problem-solving period when compared to baseline. This difficulty was seen to a much greater degree during the problem-solving period, but the small effect size during the recovery period is most likely a result of stress responses still being elevated so soon after task completion.

Table 2. Mean physiology compared to baseline and behavior frequency across time

	Alpha Amplitude	Beta Amplitude	RSP- Rate	HR	HRV	Behavior
Baseline	14.4 (3.9)	8.1 (3.1)	21.7 (2.7)	93.7 (13.5)	61.5 (18.1)	2
Problem-Solving	10.5 (1.7) *d*= 1.30***	6.0 (1.2) *d*= 0.89***	21.0 (2.0) *d*= 0.29*	88.9 (11.3) *d*= 0.39*	56.9 (16.1) *d*= 0.27*	22
Recovery	10.7 (1.4) *d*= 1.26***	6.1 (1.3) *d*= 0.84***	18.9 (2.0) *d*= 1.18***	82.8 (14.6) *d*= 0.78**	53.9 (17.3) *d*= 0.43*	7

-Mean, standard deviation and Cohen's d effect size reported
**small effect*
***moderate effect*
****large to extra-large effect*

4.3.4 Physiology across Time

Continuous recording of physiological data across all observation periods showed several signals to be more responsive than others. GSR, ST, and EMG did not show any significant change between observation periods. However, Table 2 shows a range of small to extra-large Cohen's *d* effect sizes across observation periods for EEG, RR, HR, and HRV when compared to baseline. These findings suggest that further research should consider focusing on these physiological signals.

Table 3. Inter-correlations for physiological variables

	SC/ GSR	Skin Temp	Alpha Amplitude	Beta Amplitude	EMG Amplitude	RSP- Rate	HR	HRV
SC/ GSR	X	-	-	-	-	-	-	-
Skin Temp	.13*	X	-	-	-	-	-	-
Alpha Amplitude	.05	.01	X	-	-	-	-	-
Beta Amplitude	.05	.01	.96***	X	-	-	-	-
EMG Amplitude	.01	.03	.10*	.12*	X	-	-	-
RSP- Rate	.03	.01	.02	.02	.04	X	-	-
HR	.10*	.16*	.01	.01	.04	.16*	X	-
HRV	.03	.00	.01	.01	.03	.00	.02	X

**small effect*
***moderate effect*
****large to extra-large effect*

4.3.5 Physiological Correlations

Correlations between physiological measures were examined using Pearson product moment correlations by averaging physiological recordings across the entire observation period for all participants. Results showed only minimal correlations between pairs of physiological signals at

best (r = .00 - r = .16), with the exception of the very strong positive correlation found between alpha and beta waves (r = .96), which was interpreted to be a confounding effect due to the direct influential reverse relationship the two states share (Table 3).

4.3.6 Physiological Change and Behavior Patterns

Frequency of BD incidences was recorded for each observation period through post hoc ratings of behavior from video recordings. These frequencies were then compared to physiological data. Incidences of behavioral events rose from the baseline to the problem-solving period, and began to fall back toward baseline levels during the recovery period—an expected pattern (Table 2). This data, along with physiological change data suggest that alpha amplitude, beta amplitude, HR, HRV, and RR data may have the greatest value in providing insight into BD that occurs during task completion in adults with TBI.

4.3.7 Preliminary Inter-Rater Reliability

Clinicians reached excellent inter-rater reliability on the coding of mock participant videos (N=4) during the training phase, calculated using the kappa statistic prior to rating the six study participants (κ = .83). Excellent inter-rater reliability was again achieved when analyzing the six study participants (κ = .87). This reliability, although preliminary and on a very small sample size, provides promise that the OBDRS has the ability to become a reliable behavioral assessment tool and is worthy of further exploration on larger samples.

4.4 DISCUSSION

The present study explored the relationship between BD and physiology across observation periods and found several physiological signals showed clinically meaningful differences during task completion when compared to baseline: alpha amplitude, beta amplitude, RR, HR, and HRV—suggesting these signals be analyzed further and that these specific signals may have more predictive utility than others when entered into models of BD. Non-responsiveness was seen for GSR, ST, and EMG, which is contradictory to research suggesting EMG and GSR are highly predictive measures of stress. A potential explanation for these results is the nature of the tasks participants were asked to complete and sensor placement. Participants were asked to physically manipulate objects with their hands, while also having sensors for several signals (e.g., ST, EMG, and GSR) placed on the fingers and palms. Additionally, the pressure placed on some of the sensors from simply holding onto the objects had the potential to introduce bias into the data (e.g., GSR). It is highly possible the location of the sensors had an impact on participants' ability to maintain full range of motion and move freely throughout the research tasks (e.g., EEG sensors placed on the head with short leads connecting back to the device).

An additional explanation for consideration is that the research tasks selected did appropriately capture the construct of interest (i.e. behavior dysregulation), but physiological expectations based on stress literature may not carry over to protocols where BD and not stress is the outcome of interest. The assumed link between the two constructs may be directly influenced by the tasks selected. No noteworthy correlations were found between pairs of physiological signals, with the exception of alpha and beta waves, indicating that physiology is highly individualized when averaged across persons. This supports physiological theory and previous research highlighting the challenges of analyzing physiology across individuals (Collins, 1999;

Davies & Morris, 1993). Finally, physiological signals did show to change in parallel with behavioral incidence frequencies across all participants, supporting the need for further research examining the temporal relationship between BD and physiology (Table 2).

Secondary analysis of the data showed that when compared to baseline, clinically meaningful differences were seen in perceived cognition and mood during all other observation periods. This magnitude of difference remained stable during the four checkpoint intervals within the problem-solving observation period when analyzed separately. These results suggest that perceived cognition and mood do in fact change in parallel with internal physiological states and correlate with behavior. This PROMIS measure may yield meaningful data as a more conscious assessment method along with unconscious bodily states.

Overall, the present exploratory study provided evidence that it is feasible for biomeasure technology to reliability detect changes in physiological activity during task completion that correlates with clinician behavioral observation. These findings, if replicated in larger-scaled, predictive modeling protocols have the potential to infer the development of novel assessment methods that may ultimately allow for rehabilitation service delivery prior to the onset of potentially damaging behaviors. Portable technology measuring physiological states can then extend rehabilitation services into the community, where technology will be able to not only measure, detect, and predict behavior, but also trigger full interventions through integrated mobile applications. This novel rehabilitation capability can lead to increased service delivery efficiency, through maximizing clinician time and decreasing costs. This study supports continued research in this area, specifically on establishing the psychometric properties of the OBDRS, exploring physiological states that may have the greatest value in a predictive model for BD, and the

development of a rehabilitation mobile application for pairing with portable biomeasure technology.

4.4.1 Limitations

This study, as anticipated had several limitations. Small sample size and the convenience sampling method used as well as methodological restrictions made for a small, skewed sample on a population different from the target group. Due to these limitations, the study experienced limited statistical rigor through only running basic descriptives and correlations. However, despite these restrictions, this study was a critical exploratory phase in the greater research project—a phase that needed to happen to infer feasibility and the potential for future research.

Additionally, after running six participants through the research protocol it became clear that it might be possible to shorten the length of each research task and still see a clinically meaningful physiological effect. Due to limited research on frustrating task completion and behavior in TBI, tasks had to be selected that were believed to be challenging enough to elicit a frustration response and physiological change. Due to the neurotypical sample used and minimal number of BD incidences observed, it is difficult to answer questions regarding the appropriateness of the tasks in the protocol.

Finally, when using new technology instrumentation can be a concern. The research team took care to deliver the research protocol equally across all participants, including the use of calibrated biosensor technology and an automated computer program with an embedded protocol. However, it was determined that the Nexus-10 device, due to its sensor placement and cord lengths would be inappropriate moving forward with protocols requiring participants to move around, manipulate objects with their hands, and play an active role in research tasks. It is possible

physiological responsiveness was falsely represented in the data collected from this device due to these instrumentation challenges experienced. Without this limitation and the use of a device that better fit the needs of the protocol, responsiveness may have resulted differently. Exploration of alternative portable devices on the market should be conducted before moving forward with future protocols.

4.4.2 Protocol Modifications

Given the knowledge gained from the initial feasibility protocol, several changes were proposed moving forward to both the psychometrics and EMA protocols to strengthen study methodology.

4.4.2.1 *Tasks*

Due to the limited number of outcomes observed during task completion, it was determined that brainteaser puzzles may not be appropriately eliciting the construct of interest. It was proposed that new tasks be selected moving forward that were more supported in the literature as research tasks that may lead to the desired response.

The new research tasks chosen were selected through an extensive review of the literature. Due to a lack of a strong body of evidence isolating appropriate tasks for eliciting a negative behavioral response, a literature search was conducted from well-established stress, health, and psychology journals where studies were reviewed that used reliable and valid tasks to simulate stress responses in laboratory or clinic settings. The method of researching was based on the assumption that a direct link exists between internal stress responses, physiology, and behaviors. Therefore, selecting tasks with establish psychometric properties for eliciting a stress response was

a logical place to begin when attempting to isolate tasks that may also lead to frustration and task breakdown.

Several common themes emerged from condensing the literature on stress responses during task completion. Serial subtraction tasks, where individuals are asked to count backwards in increments starting from a large number and are corrected by a researcher and required to start over each time an error is made appeared often in the literature. Additionally, stress research protocols commonly used (as an independent task, but oftentimes along with a subtraction task) impromptu speech preparation and delivery, where an individual is told they will be given a short amount of time to plan a speech on something believed to be a salient topic (such as career goals). With this task, individuals are also told they will deliver their speech in front of a committee who will grade them on their performance while also being videotaped. The committee is always made up of confederates, who are actually members of the research team. With this type of task, deception is used and as a result carefully planned debriefing periods are conducted immediately following the completion of the research protocol to ensure the participant is informed of the reason for the deception and why it was necessary, as well as give them time to ask questions and return to a non-elevated state.

A well-established, commonly used, and standardized battery of stress tests called the Trier Social Stress Test (TSST), emerged as the most commonly implemented protocol in this literature review (Birkett, 2011; Kirschbaum, Pirke, & Hellhammer, 1993). Several components of the TSST were determined to be highly beneficial for use in this research; however, some aspects of the protocol (e.g., planned deception) were determined to be inappropriate and unnecessary for application with a vulnerable disability cohort such as TBI. This along with other considerations (e.g., the use of a confederate panel and the inability to take cortisol samples due to limited

resources), made the TSST as a standardized battery not meet the criteria for identifying appropriate research tasks for this research protocols. However, the tasks themselves were still determined to have value, and as a result, a modified version of the TSST was selected as the best approach. The serial subtraction task and a version of the speech preparation and delivery task that did not involve planned deception and simply videotaped participants giving their speech after preparing it, without the use of a confederate panel was anticipated to be enough to elicit the desired response in a controlled environment in the target population given judgements made by clinicians working on this project.

4.4.2.2 *Technology and Signals*

After extensive searching, the Bioharness-3 device was selected as an alternative to the Nexus-10 device used during initial feasibly testing. Keeping in mind the instrumentation challenges experienced during the feasibility study, it was critical for researchers to select a more user-friendly device for the current protocol. Given the needs of this project, (e.g., object manipulation and need for participant movement) and the availability of devices on the market, the Bioharness-3 was selected due to its comfortability, support of the user's full range of motion, and capability of measuring all responsive signals from Study 1, with the exception of EEG. Since user-friendliness was the main priority in selecting a new device, this left limited options to find a device that would measure all signals collected during initial feasibility testing (the ideal scenario). However, the Bioharness-3 did have the capability of measuring all responsiveness signals from Study 1, with the exception of EEG, which was determined to be an appropriate device moving forward that was also supported by the findings from Study 1.

Portable biosensor technology measuring EEG is highly limited and the devices that were available did not meet the needs of this project. Therefore, it was determined moving forward that

this signal be removed from data collection in support of a stronger methodology. Removal of the EEG signal from the current research is not reflective of a lack of utility this signal may provide to predicting BD, but rather a result of instrumentation restrictions (e.g., need for a portable, wearable device) of the current research and limited options to support those efforts. Future research examining physiological predictors of BD should still consider incorporating EEG if study methodologies allow.

5.0 OBDRS PSYCHOMETRICS[1]

5.1 RESEARCH PURPOSE AND AIMS

The aim of this study was to determine the reliability and validity of the OBDRS tool. This study was intended to serve as a follow-up to the preliminary reliability established in the initial feasibility protocol (Study 1). In addition to inter-rater reliability, content validity was assessed through a rehabilitation expert survey group and construct validity was assessed through correlating the OBDRS tool with existing 'like' and 'unlike' measures. The research questions for this study were (1) does the OBDRS fully and solely measure the construct of BD as it presents in TBI and as defined for the purpose of this study? (2) does the OBDRS correlate with existing validated measures of BD and present as distinct from tools measuring different constructs? and (3) can the OBDRS be established as a reliable measure of BD through rater agreement of recorded behavior events?

Despite the significant impact of behavior SR deficits on daily living, the most commonly used assessments for measuring behavior in persons with TBI are delivered in clinic settings, and may not be appropriate due to the easily controlled and highly structured clinic not allowing for natural representation of functioning or tap into the true environment where dysfunction manifests.

[1] Multiple sections within this chapter contain repeated content from the feasibility research discussed in Chapter 4. An editorial decision was made by the author to retain this repeated content with the goal in mind of each chapter being able to be read as standalone content methodologically.

The OBDRS is unique in that no reliable and valid research tool currently exists to assess challenging behavior profiles in post-acute TBI in naturalistic settings. Of the assessment tools that have been widely used, most have psychometric properties established through inpatient or therapeutic samples in highly controlled environments, or require retrospective recall. Additionally, existing tools often do not measure the full construct of behavior SR—only targeting a specific behavior, such as impulsivity, or even measuring a different construct all together. This is attributed to the heterogeneity of the available literature and competing expert opinion on the construct of interest. There is a need for a tool that can assess post-TBI behavior in the moment in the real world during the completion of tasks that are similar to those experienced during daily living. The OBDRS hopes to bridge this evidence gap.

5.1.1 Hypotheses

It was hypothesized that the OBDRS will prove to be a reliable and valid assessment tool for the measurement of BD in adults with TBI. More specifically:

1. Clinicians in the expert feedback group will be in perfect agreement that the content of the OBDRS represents the construct of BD.
2. The OBDRS will prove to be highly correlated ($r \geq .50$) to similar assessment tools and weakly correlated ($r < .50$) with tools measuring constructs unlike BD.
3. The OBDRS will have adequate inter-rater reliability of $\geq .80$ between the two trained clinicians.

5.2 METHODS

5.2.1 Participants and Recruitment

Recruitment and testing of all participants occurred from September 2015-March 2016. Rehabilitation professionals and study participants were all be sampled through a non-experimental convenience sampling method.

5.2.1.1 *Rehabilitation experts*

Three rehabilitation experts, consisting of two doctorate-level brain injury rehabilitation researchers and one former head injury project leader and behavioral clinician, with extensive experience working with individuals with TBI, were hand-selected by the principal investigator of this study based on their experience and ability to provide meaningful feedback on the OBDRS for the target population and for the purpose of developing a reliable tool.

5.2.1.2 *Participants*

Participants in the challenging task research protocol were adults with TBI who experience BD challenges during daily living. ReMed of Pittsburgh, a brain injury rehabilitation program in Pittsburgh, Pennsylvania served as both the recruitment and testing site due the ability to sample from a representative cohort of the target group. Eligible participants must have been currently receiving either residential or outpatient rehabilitation services at ReMed, have a primary diagnosis of TBI and subsequent behavioral challenges, and not currently receiving treatment for a psychiatric diagnosis. ReMed case managers working directly with program clients identified and

referred all potential participants by reading a provided recruitment script and scheduling interested participants to be consented and tested by the research coordinator on site.

5.2.2 Data Measurement and Collection

Pre-testing assessments, as well as behavioral events data from the OBDRS post hoc clinician ratings were used to examine the psychometric properties of the developed tool. This study identified BD as the sole binary dependent variable ('0'= behavior dysregulation absent; '1'= behavior dysregulation present) based upon researchers wanting to know if any behavior of interest was present or absent and not which target behavior specifically. BD was measured by clinician raters post hoc during each second for the entire recording period and the total score of BD frequency for each participant was used for validity analysis. This small sampling window will allow for capturing spontaneous or brief changes in the variables of interest.

The research tasks, including all pre-testing assessments took no longer than two hours to complete. All assessments were administered and scored by the primary research coordinator, while all behavior data was recorded by clinician's post hoc ratings from video recordings of participant research sessions.

All data collected was kept confidential by giving each participant an arbitrary identification number, which served as a link between raw data and the individual who provided it. This confidential data was entered into a spreadsheet (using the identification number for coding purposes) by a research team member for statistical analyses and stored separately (in a data safekeeping room within Pitt-RST facilities) from any identifiable information. This process allowed raw data to remain unidentified, even by the research team members who had access.

5.2.3 Assessments

Several pre-testing assessments were given to all participants that served as measures for evaluating both the convergent and discriminant validity of the OBDRS. All assessments were well established, widely used, and psychometrically valid measures for assessing either SR or constructs believed to be unlike the construct of interest (e.g., mindfulness).

5.2.3.1 *OBDRS rating scale*

The OBDRS is checklist developed for the purpose of this research that produces ratings on whether specific behaviors believed to be indicative of BD are present or absent during an observation period. Data from this tool allowed for associations to be evaluated between behavior and physiological measures due to the external time stamp on the checklist and the internal time stamp embedded within the biomeasure software. Two trained clinicians each completed a post hoc behavioral analyses through video recordings of each participant (see *tool development* section of Chapter 3 for detailed information on the scale and Appendix A for actual tool).

5.2.3.2 *Behavioral assessment of the dysexecutive syndrome dysexecutive questionnaire*

The self-administered version of the Behavioral Assessment of the Dysexecutive Syndrome Dysexecutive Questionnaire or DEX is a 20-item survey of common problems experienced with executive functioning deficits. It measures three factors, behavior, cognition, and emotion. Individuals self-report on their own experiences by answering on a Likert scale, ranging from 0 (never) to 4 (very often). The total sum of all items is used for scoring, where higher scores indicate greater degrees of dysregulation, and oftentimes the self-rating questionnaire is compared to an independent rater version of the DEX questionnaire—completed by a member of the

individual's support system or therapeutic team. This study used only the self-administered form due to interest in measuring only the participants; perception of the self. The DEX questionnaire has been shown to have high inter-rater reliability (κ = .88-1.00, p = .001; Emmanouel, Mouza, Kessels, and Fasotti, 2014). Research has also shown the DEX to be a valid instrument for measuring discrepancies between self-reported and support-system versions of the tool, indicating impaired perception of self-regulatory ability from the individual of focus (Emmanouel, Mouza, Kessels, and Fasotti, 2014).

5.2.3.3 *Go/no-go test*

The Go-/No-Go test is a neuropsychological measure of response inhibition and impulsivity—both components of SR (Fillmore, 2003). A version of the test is part of the Luria-Nebraska Neuropsychological Battery, a widely used standardized test battery focused on the screening and evaluation of neuropsychological difficulties (Golden, 1980). With this test, participants are given instructions to perform an action in response to two different stimuli, one representing the 'go' condition and the other representing the 'no-go' condition (e.g., "When I knock once you knock twice, when I knock twice you knock once.") The test administrator conducts six trials of four different stimuli tests, delivering the two "go/no-go" conditions in a pseudo-random format and allowing the participant time to respond. The four different tests are hand knocking on table, hand squeezing in response to colors, hand raising in response to knocking, and hard vs. soft hand knocking. The number of errors made in each trial is recorded and transferred to a scale that produces a quantitative score ranging from 0 to 8 errors. The Go/No-Go test is a validated tool for measuring impulse control in children with Attention Deficit Hyperactivity Disorder or ADHD (Weafer, Milich, & Fillmore, 2011), as well as substance abuse populations (Fillmore & Rush, 2006) and other cohorts known to display poor inhibitory control.

5.2.3.4 *Mindfulness attention awareness scale*

The Mindfulness Attention Awareness Scale or MAAS is a 15-item self-administered tool where participants are asked to rate each item on a Likert scale ranging from 1 (almost always) to 6 (almost never), for items all pertaining to everyday experiences. The MAAS measures an individuals' awareness and attention to situations taking place in the present, a core component of the greater mindfulness construct. The average of all items is used for scoring, where higher scores are a reflection of greater levels of dispositional mindfulness. The MAAS has shown to be a highly reliable and valid tool for measuring mindfulness in diverse cohorts, including community dwelling adults, college students, and individuals with cancer (Brown & Ryan, 2003).

5.2.3.5 *Satisfaction with life scale*

The Satisfaction with Life Scale (SLS) is a brief 5-item self-administered tool measuring an individual's cognitive perception of their satisfaction with life. Responses are rated on a Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). The total sum of all items is used for scoring, where higher total scores indicate greater satisfaction with life. The SLS is a well-established, reliable, and valid assessment for measuring the construct of life satisfaction (Pavot, Diener, Colvin, & Sandvik, 1991).

5.2.4 Clinician Training

Two clinicians (Master's level students from the RST department) were hand selected and trained by the research coordinator on how to complete post hoc behavior ratings using the OBDRS tool through the behavior coding software. Both raters reviewed videos of all participants completing

the research protocol. These raters did not interact with study participants, were unaware of assessment results, and completed all ratings independently of each other.

Clinician raters were trained through a standardized protocol occurring over a 3-week period prior to rating participant videos. Training included reviewing articles on relevant content, introduction to the OBDRS including how to administer the tool, and the scoring of four mock participant videos not included in the sample of videos clinicians rated for the research project. Each mock video scored was reviewed in detail by the clinician and the research coordinator to address coding concerns and provide feedback. Successful completion of training was defined as when both raters established inter-rater reliability via a kappa statistic of at least .80 on the four mock participant videos. At this point, clinicians were cleared to begin coding real participant videos.

5.2.5 Instrumentation

5.2.5.1 *Video recordings*

Video recordings were taken using a standard video recording camera. Participant videos were stored on a secure server at the University of Pittsburgh and only accessed by the researchers associated with this project. All participants were asked to give consent to be videotaped and have their videos analyzed after completion of the study before any research tasks occurred.

5.2.5.2 *The observer xt software*

The Observer-XT is a widely used software package from Noldus that allows for the collection, presentation, and analysis of behavior data. Noldus allows for behavior coding tools, like the OBDRS to be embedded into a software program that allows for clinician ratings that can be easily

compared, analyzed and synchronized with other external data sources (e.g., physiological monitors). To complete a behavior rating, a video recording is temporarily uploaded from the secure server into the software, but videos are not stored in the Noldus system permanently. All videos remained stored on a secure server after clinician ratings were completed.

5.2.6 Procedure

The current research design followed a single group, cross-sectional design. This study was non-experimental and therefore, did not utilize randomization, blinding, or allocation processes. This project was approved by the University of Pittsburgh IRB prior to conducting any research activities (#PRO15100307).

5.2.6.1 *Electronic survey group*

A small electronic survey group (N=3) was held with a group of rehabilitation professionals who have extensive experience working with individuals with TBI. This group occurred via email and required no more than one hour of each clinicians' time. The goal of this session was to have rehabilitation professionals critically review the OBDRS, specifically focusing on the tools' ability to fully and reliably capture the construct of BD as (1) it pertains to adults with TBI and (2) how the researchers on this project have chosen to operationalize the construct. Quantitative and qualitative feedback gained from the experts was utilized for tool modifications and establishing the content validity of the OBDRS.

5.2.6.2 *Challenging task protocol*

The research protocol was administered face-to-face, one time to all participants. The research coordinator was responsible for running all participants through the protocol as well as administering and scoring all pre-testing assessments. Participants were video recorded and permission to record was obtained during the written consent process. The consent process occurred immediately prior to the administration of all research tasks.

The study protocol began with a pre-testing period where three self-rating questionnaires were administered along with one clinician-administered performance measure. Following pre-testing, participants were escorted to a relaxation room where they were left alone and asked to simply relax for ten minutes. This allowed for participants to re-stabilize from any potentially evaluated states resulting from being asked to fill out potentially stressful self-assessments.

Immediately following the acclimation period, participants were introduced to the first research task. Participants completed serial subtraction where they were asked to count down from 6233 by 13s. The task ended when either the participant reached 0 or five minutes had elapsed. Participants were informed prior to beginning the task that if at any point they made an error they would be asked to stop and start the task over from the beginning. A visual timer was provided letting participants know how much time they had left to work on the task. The subtraction task used in this study was adapted from the TSST identified in development and feasibility work (Study 1).

Following the serial subtraction task, participants immediately moved into the speech preparation and delivery task. Here participants were told they would be given 10 minutes to prepare a 5-minute speech on how their injury has affected their life. They were given scrap paper and a pencil and pen to write ideas on while they planned their speech. At the end of the five-

minute preparatory period participants were then asked to stand and present their speech to the camera. What they were not told before completing the task was that they would be required to present their speech without using their notes. This element was introduced to add a spontaneous unplanned stressor into a life situation. This speech task was also adapted from the TSST.

Immediately following the speech task, participants were asked to complete a third and more functional task. The hamburger turning task is a functional task developed and used clinically for over the past 12 years by clinicians and researchers at the University of Pittsburgh (Shugars, 2007). This task requires participants to prepare 16 fake hamburger patties on a faux grill based on a 30-step procedure modeled after actual hamburger cooking process used in Wendy's restaurants. This task has been shown be a clinically useful measure of executive functioning (Shugars, 2007) and was selected to serve as a hands-on task that may be more representative of meaningful experiences during daily living that result in challenges for the target group when compared to the other two (more cognitive based) tasks in the protocol. Completion of this task required participants to first watch a brief, narrated step-by-step video of a person completing the task correctly from start to finish. Next, participants were asked if they had any questions about the task and then were introduced to task props (fake hamburger patties, faux grill, real saltshaker, and spatula). Participants were also given a list of task steps to use as a cognitive aid if they chose to do so while completing the task. Participants were told they would have 15 minutes to correctly cook all the hamburger patties and that similarly to the serial subtraction task, if they made an error, they would be alerted to the error and then asked to start the task over from the beginning.

Finally, after the completion of all research tasks, there was a 20-minute post-testing recovery period where participants were again left alone in a quiet room and asked to simply relax. This period was set aside to allow participants to return to baseline levels from any potentially

elevated states before testing concluded. After the recovery period ended, the research coordinator escorted the participant back to the main testing room, stopped the video recording, answered any remaining questions participants may still have had about the study, and provided compensation. ReMed staff remained on-call during testing in the event a participant experienced distress levels beyond what was expected and considered appropriate during a common stressful event. This type of elevated effect was not anticipated and did not occur. Figure 7 shows the full OBDRS psychometric protocol.

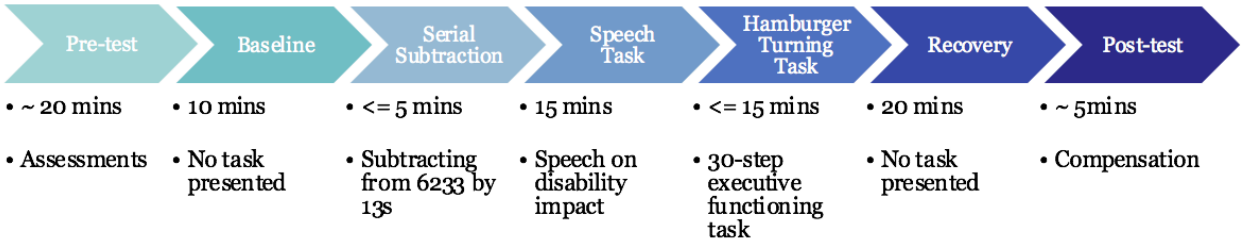


Figure 7. Tool psychometrics protocol

5.2.7 Facilities and Personnel

All research activities were conducted at ReMed of Pittsburgh's clinical facilities in Pittsburgh, Pennsylvania. Members of the research team were all affiliated with the RST department at the University of Pittsburgh and possessed the appropriate qualifications to run a study of this nature.

5.2.8 Data Source

The data set contained repeated measures data for each participant on the dependent dichotomous variable BD, where '0' = absent and '1'= present, based on post hoc clinician ratings. Therefore, many records were nested within each participant and each record represented one second of data collection. This nested setup allowed researchers to examine the ability of the OBDRS to be a reliable and valid measure of BD both within- and between-subjects. Fixed, non-nested assessment data on SR, impulsivity, life satisfaction, and mindfulness, as well as demographic data was also included. For analyses of fixed, non-nested variables (i.e. variables collected at only one time point that remain unchanging, such as age), only the first record for each participant's data was selected due to all remaining records having repeated information in the corresponding cells (i.e. if a participant was 23 years old, each record in the age column for the entire dataset for that particular individual read '23'). BD analyses used the full nested dataset, where variables had the ability to change over time (e.g., present or absent). All statistical analyses were conducted using the Statistical Package for the Social Sciences Software (SPSS, v. 22), and Noldus the Observer XT Software. All data was evaluated at the .05 alpha level.

5.2.9 Data Analyses

5.2.9.1 *Validity*

Content validity was examined from feedback gained from rehabilitation experts participating in the electronic survey group. Raters filled out a yes/no table asking if the six target behaviors within the tool are representative of the BD construct and objectively measurable. This table allowed for increased efficiency when determining rater agreement and the ability to gather quantitative

evidence for assessing content validity. Qualitative data was also collected through rater's open-ended responses to questions about the tool. Quantitative and qualitative professional feedback supporting the OBDRS as a tool that captures the entire construct of BD in the target population is a necessary stage in the process of moving toward establishing the tool's construct validity.

Several validated and widely used assessment tools were used to measure how much the OBDRS is like other tools measuring similar constructs and the extent to which the OBDRS differs from tools measuring different constructs.

The SLS and the MAAS measuring perceived life satisfaction and mindfulness ability were used to determine the discriminant validity of the OBDRS. Discriminant validity was measured through Pearson product moment correlation coefficients, evaluating the correlation between the two scales being compared. It was anticipated the OBDRS would be weakly correlated with both the SLS and the MAAS.

The DEX and the Go/No-Go Test were used to measure whether the OBDRS is highly similar to tools measuring similar constructs. Convergent validity was measured through correlations evaluating the correlation between the two scales being compared. It was anticipated the OBDRS would have high correlations with both the DEX and the Go/No-Go Test through evaluation of Pearson product moment correlation coefficients.

5.2.9.2 *Inter-rater reliability*

Due to the dichotomous nature of the outcome variable, inter-rater reliability was assessed through percent agreement or the agreement between raters on whether BD was present during the observation period. Percent agreement is calculated from the proportion of exact agreements over the total number of possible agreements, where percentages closer to 100% indicate greater agreement between raters (Portney & Watkins, 2009, p. 598). This measure of agreement, although

simple to understand, is typically reported along with additional measures of reliability due to its ability to overestimate agreement and not account for chance. For that reason, the kappa statistic was also used to obtain a chance-corrected measure of agreement between raters. The kappa statistic ranges from 0 (agreement equal to chance) to 1 (perfect agreement) and values above .60 indicate substantial agreement between raters (Sim, & Wright, 2005; Portney & Watkins, 2009, p. 598-600).

5.3 RESULTS

5.3.1 Descriptives

Table 4 shows demographic characteristics of the sample (N=14). All continuous demographic measures were normally distributed, with the exception of time at ReMed. This sample of adults with TBI was on average 40.5 years old, 152.2 months post injury, and a client at ReMed for 19.7 months. The majority of the sample was Caucasian, male, single or divorced, and receiving services from the outpatient side of the rehabilitation program—a sample make-up consistent with the greater population of individuals with TBI.

Table 4. Participant demographics

Demographic	N=14	*W*
Age (years), mean (SD)	40.5 (3.28)	0.938*
20-29	3 (22)	
30-39	4 (28.6)	
40-49	3 (21.4)	
50-59	3 (21.4)	
≥ 60	1 (7.1)	
Time Since Injury (months), mean (SD)	152.2 (36.0)	0.886*
Time at ReMed (months), mean (SD)	19.7 (9.3)	0.576
Sex, n (%)		
Male	10 (71.4)	
Female	4 (28.6)	
Race, n (%)		
Caucasian	14 (100)	
Ethnicity, n (%)		
Non-Hispanic or Latino	14 (100)	
Marital Status, n (%)		
Single	7 (50.0)	
Divorced	5 (35.7)	
Separated	1 (7.1)	
Relationship	1 (7.1)	
Cause of Injury, n (%)		
MVA as pedestrian	4 (28.6)	
MVA as driver	7 (50)	
Fall	3 (21.4)	
Living Status, n (%)		
Residential	9 (64.3)	
Outpatient	5 (35.7)	

**p-value ≥ .05*

5.3.2 Content Validity

Content validity of the OBDRS tool was assessed as an initial and critical phase of establishing evidence that the tool appropriately and fully measured the construct of BD. Expert raters assessing the tool's content validity were carefully selected due to their professional experience working directly with individuals with TBI in clinical, community, and/or research settings. Table 5 displays results from feedback given by these professionals on whether each individual behavior included in the tool measured a construct representative of BD as it may be experienced in adults with TBI, was a measurable construct, objectively defined, discrete, and reflective of a threshold behavior. Appendix F can be referred to for a complete overview of the feedback rating form provided to rehabilitation experts. Experts were in perfect agreement across the six target behaviors as presented measured the construct of interest across all criteria (Table 5). This level of agreement served as confirmatory data to support moving forward to examine the tool's ability to meet the next phase of validity, construct validity.

Table 5. Expert ratings on target behaviors in the overt behavioral dysregulation rating scale

	JK		MP		SC		% Agreement
	Yes	No	Yes	No	Yes	No	
Disengagement	x		x		x		100%
Resistance	x		x		x		100%
Non-Verbal (non-violent) Dysregulation	x		x		x		100%
Physical Aggression/ Violence	x		x		x		100%
Verbal Dysregulation	x		x		x		100%
Perseveration	x		x		x		100%
Total Agreement							100%

5.3.3 Construct Validity

Construct validity of the OBDRS tool was assessed to determine if the developed tool proved to be 'like' or strongly correlated with existing validated scales measuring similar constructs, while also dissimilar or weakly correlated with existing measures assessing diverging constructs. For assessing convergent validity, two validated measures widely used for assessing self-regulatory ability were compared to the OBDRS, the DEX and the Go/No-Go. Discriminant validity of the OBDRS was assessed through comparisons between the developed tool and the SLS measuring life satisfaction, and the MAAS, measuring mindfulness. Descriptive statistics (Table 6) show all assessments were normally distributed ($p \geq .05$) providing support for running bivariate Pearson product moment correlations between all measures.

Results from a correlation matrix (Table 7) show the OBDRS displayed strong relationships with both the DEX ($r = .535$) and the Go/No-Go ($r = -.564$), indicating the measure strongly converged with validated measures of SR. When comparing the OBDRS to the DEX, the relationship between the two measures was positive, which was expected given higher scores on both instruments indicate greater dysregulation. The expected positive relationship between the OBDRS and the Go/No-Go test was confirmed as well due to higher scores on both measures also representing a greater degree of dysregulation. When examining discriminant validity, the OBDRS displayed a barely discernable, negative relationship with the SLS, which was expected given the SLS measures life satisfaction on a scale where lower numbers represent less satisfaction.

However, there was a positive, moderate correlation between the OBDRS and the MAAS (r = .469), suggesting the two measures may be assessing related constructs.

Table 6. Assessment descriptives

	Min	Max	Mean		Std. Deviation	Variance	Shapiro-Wilk Test for Normality
			Statistic	Std. Error			*W*
DEX	6	37	23.07	2.504	9.37	87.76	0.950*
MAAS	3.79	5.79	4.50	0.142	0.53	0.28	0.889*
SLS	5	25	13.93	1.357	5.08	25.76	0.966*
GO/NOGO	0	2	0.43	0.202	0.76	0.57	0.616*
OBDRS	4	30	13.64	2.032	7.60	57.79	0.930*

**p-value ≥ .05*

Table 7. Bivariate Pearson product moment correlations for evaluation of the overt behavioral dysregulation rating scale's construct validity

	DEX	MAAS	SLS	GO_NO_GO	OBDRS
DEX	1				
MAAS	-.472	1			
SLS	-.265	.26	1		
GO_NO_GO	-.428	.43	.009	1	
OBDRS	.535*	.469	-.047	.564*	1

** Correlation is significant at the 0.05 level (2-tailed).*

5.3.4 Inter-Rater Reliability

Percent agreement was calculated between two clinician raters from the proportion of behavioral agreement over the total behavioral events rated. Due to the nature of the dataset, each second of

the video recorded observation period was rated dichotomously as '1'= behavior present or '0'= behavior absent. Table 8 shows both raters were in agreement 95% of the time, with only 4682.6 seconds of total disagreement across all participant videos coded—indicating excellent reliability. Additionally, Cohen's kappa was calculated to measure the average rate of agreement while more accurately accounting for chance agreement. The resulting kappa statistic of $\kappa = .60$ falls on the boarder of moderate to substantial agreement and was significant at the .05 level.

Table 8. Inter-rater reliability analysis of the overt behavioral dysregulation rating scale

		Statistic	95% CI
Observer A and Observed B (combined results)	**Agreement (sec)**	80435.6	
	Disagreement (sec)	4682.6	
	% Agreement (/100)	.95	
	Cohen's Kappa (/1.0)	.60*	.58-.61

**Indicates significance*

Due to percent agreement and the kappa statistic measuring an overall average of agreement across all target behaviors in the OBDRS tool, an additional reliability analysis was conducted to further evaluate disagreement and whether any one specific category of behavior may be influenced by reliability results. This subjective analysis added rich information to the quantitative reliability statistics. Table 9 represents a confusion matrix of the six target behaviors of the OBDRS tool and the agreement between raters for individual behavior categories. Two important pieces of information stand out when examining this table. First, the majority of disagreements came from the disengagement/engagement behavior, where one rater coded the participant as disengaged while the other determined the individual was actually engaged at that moment in time. Secondly, taking into account all other behaviors aside from disengaged/engaged, the majority of coding discrepancies occurred when one rater coded a behavior as present and the

other rater coded that time period as having no behavior present at all. These results suggest that reliable behavior coding may be influenced by unique factors pertaining to specific behaviors.

Table 9. Confusion matrix displaying inter-rater reliability of the overt behavioral dysregulation rating scale's target behaviors

	Engaged	Disengaged	Opposition	Verbal Dysregulation	Physical Aggression/ Violence	Non-Verbal Dysregulation	Perseveration	No Records	Total
Engaged	77788.4	704.2	-	-	-	-	-	-	78492.5
Disengaged	1820.3	2205.0	-	-	-	-	-	-	4025.3
Opposition	-	-	36.4	-	-	-	-	33.6	70.0
Verbal Dysregulation	-	-	-	286.1	-	-	-	257.1	543.2
Physical Aggression/ Violence	-	-	-	-	-	-	-	0.8	0.8
Non-Verbal Dysregulation	-	-	-	-	-	119.7	-	48.2	167.9
Perseveration	-	-	-	-	-	-	-	96.4	96.4
No Records	-	-	176.4	655.9	-	889.7	-	-	1722.1
Total	79608.6	2909.2	212.8	942.0	-	1009.4	-	436.1	85118.2

**Agreements can be found on the diagonal shaded boxes and disagreements are represented in the white boxes. All data is represented in seconds during the recording period.*

5.4 DISCUSSION

The aim of this study was to establish the psychometric properties of the OBDRS tool.

Promising results were seen for both content and construct validity. Perfect agreement was shown between all three expert professional raters, indicating the six target behaviors represented in the OBDRS tool may be appropriately measuring the construct of interest as defined by the researchers. Examination of the tool's construct validity provided further support for the OBDRS's potential for measuring the construct of self-regulation and only that construct—with the tool strongly converging with validated, widely used measures of self-regulation while also showing its discriminant nature from tools measuring unlike constructs, such as mindfulness and life satisfaction. Although a moderate correlation was observed between the OBDRS and the MAAS, one potential explanation for this relationship is that several items represented on the MAAS may not only measure mindfulness, but may also cross over construct boundaries to capture aspects of self-regulation (e.g., 'I snack without being aware that I'm eating' and 'I find it difficult to stay focused on what's happening in the present'). This correlation may suggest that mindfulness and self-regulation are related constructs—at least how both tools approach measuring their respective constructs. However, the strength of the relationship was not of large concern to researchers due to the correlation coefficient being smaller in magnitude when compared to the convergent relationships measured.

When taking into account all aspects of validity measured in this study, even considering the moderate correlation between the MAAS and the OBDRS, this study provides preliminary

evidence supporting future larger-scale research examining the OBDRS's psychometric properties for measuring BD.

As expected, the OBDRS proved to be a reliable assessment of BD between two trained clinician raters. A percent agreement analysis showed the raters agreed in their post hoc behavior ratings across participants 95% of the time, showing excellent agreement. The kappa statistic, run as a more rigorous measure of reliability that also accounts for chance agreement, showed bordering substantial reliability between raters. Upon further examination of agreement, results showed the two raters disagreed the majority of the time when rating the engaged/ disengaged target behavior—suggesting there may be an inherent uniqueness when coding this particular behavior that may be negatively influencing the kappa statistic. Overall, reliability statistics from this study show promise that the OBDRS tool can be a reliable measure of BD and also warrants further exploration of the differences that may exist between different target behaviors of the tool that may impact reliability. The disengaged/engaged target behavior is the only continuous behavior coded by the OBDRS (i.e. an individual must be either engaged or disengaged at all times during the observation period—requiring raters to always code one or the other), which might speak to an increased difficulty when coding this behavior or additional training that may be needed for reliable coding of this behavior. Future exploration of the behavioral coding scheme and reliability differences across target behavior when analyzed individually could infer improvements to the coding scheme itself or clinician training protocols.

5.4.1 Potential Limitations

As is often the case with pilot rehabilitation research, this study had a small sample size that was sampled through a convenience sampling method. Although this sampling method is not

considered ideal when the goal is unbiased sampling, this method is often employed when obtaining participants in rehabilitation research due to the availability of recruitment sites, community relationships and partnerships, methodological considerations, and the need for finding participants who meet potentially difficult to find exclusion/ inclusion criteria (e.g., rare conditions). ReMed of Pittsburgh proved to be a highly valuable resource to this project by their willingness to assist with recruitment and allow the research team to conduct this project on-site. This sampling method was beneficial for an exploratory project of this nature where ensuring a cohort of individuals was selected who experienced the outcome of interest in daily living was critical.

With the goal in mind of including as many participants as possible who met inclusion/exclusion criteria, individuals from both the residential and outpatient side of the rehabilitation program were included in this study. It can be argued that individuals in the residential program may be in need of increased assistance and rehabilitation services and therefore, may not have a comparable degree of self-regulatory ability to individuals participating in the outpatient program, who mostly live independently or with family members and typically display a greater level of independence. As a result, differences between residential and outpatient samples may be worthy of future exploration with a larger sample sizes and equal individuals across groups.

As previously mentioned, when developing a new measurement tool, it is critical to ensure the tool accurately and fully measures the intended construct of interest. One of the major aims of this study was to establish psychometrics properties for the OBDRS tool. Although results were promising, showing excellent content validity between rehabilitation experts, convergence with 'like' measures and discriminant properties from tools measuring differing constructs, and

displaying sufficient inter-rater reliability between clinicians, the small-scaled nature of this project requires results to be viewed critically and establishes a need for larger-scaled projects replicating these findings. If replication is successful, researchers and clinicians can have increased confidence in the OBDRS tool as a reliable measure of BD.

Task selection is a potential limitation shared by both the feasibility and psychometric protocols. In the original study researchers questioned whether the selected tasks, although challenging, had the ability to elicit the construct of interested. It was determined at that point that moving forward new tasks be selected with more evidence to support their use with less methodological restrictions. As mentioned earlier in the text, the TSST served as a foundation for researchers to select new tasks from due to the protocol's well-established use in the health psychology literature for encouraging stress responses. Due to methodological and ethical considerations, the TSST protocol was adapted to meet the needs of this project, which also included the addition of a third, more functional task. This method of task selection gave researchers greater confidence that the tasks selected had the ability to evoke the outcome of interest. Although results seem to support the continued use of the tasks, it must be considered that these tasks were not only adapted from a standardized protocol, but also were selected based on an assumption that a direct link exists between internal stress responses and overt behaviors of frustration.

6.0 EMA PROTOCOL: BEHAVIORAL DYSREGULATION PREDICTION VIA PHYSIOLOGY[2]

6.1 RESEARCH PURPOSE AND AIMS

The aims of this research were to examine physiological change during the completion of challenging tasks in adults with TBI, explore physiological variability both between and within individuals, and evaluate physiological variables as potential predictors for BD. This project is an initial and critical stage in the process of developing a TBI-specific EMA/EMI using biosensor technology that intervenes prior to the occurrence of maladaptive behaviors and allows for prospective intervention for increasing behavioral SR ability. Ultimately, variables that correlate to a high degree with clinician judgment within the laboratory will provide insight to the potential for these same biomeasures to be predictive of BD in the natural environment where SR deficits truly manifest. This was exploratory study and the main research question was, in adults with TBI, can HR, HRV, and RR taken together with age and sex accurately predict BD events?

Given what is known from review of relevant literature on adults with TBI, BD, and physiology, it was hypothesized HR, HRV, RR, when controlling for age and sex, would be able to reliably predict the probability of BD in adults with TBI during task completion. More specifically, we believed it would be possible to isolate a predictive model that not only

[2] Multiple sections within this chapter contain repeated content from the both feasibility and psychometric research discussed in Chapters 4 and 5. An editorial decision was made by the author to retain this repeated content with the goal in mind of each chapter being able to be read as standalone content methodologically.

appropriately fits the data, but could also be applied to external datasets and guide future research efforts.

6.2 METHODS

6.2.1 Participants and Recruitment

Recruitment and testing of all participants occurred from September 2015-March 2016. Rehabilitation professionals and study participants were all be sampled through a non-experimental convenience sampling method.

6.2.1.1 *Rehabilitation experts*

Three rehabilitation experts, consisting of two doctorate-level brain injury rehabilitation researchers and one former head injury project leader and behavioral clinician, with extensive experience working with individuals with TBI, were hand-selected by the principal investigator of this study based on their experience and ability to provide meaningful feedback on the OBDRS for the target population and for the purpose of developing a reliable tool.

6.2.1.2 *Participants*

Participants in the challenging task research protocol were adults with TBI who experience BD challenges during daily living. ReMed of Pittsburgh, a brain injury rehabilitation program in Pittsburgh, Pennsylvania served as both the recruitment and testing site due the ability to sample from a representative cohort of the target group. Eligible participants must have been currently

receiving either residential or outpatient rehabilitation services at ReMed, have a primary diagnosis of TBI and subsequent behavioral challenges, and not currently receiving treatment for a psychiatric diagnosis. ReMed case managers working directly with program clients identified and referred all potential participants by reading a provided recruitment script and scheduling interested participants to be consented and tested by the research coordinator on site.

6.2.2 Sample Size Estimation

Post hoc power analyses using G*Power software were conducted at multiple time points throughout testing for updated estimations of power. Recruitment and testing remained ongoing until adequate power was met for an exploratory research project. The current study examined nested data, where thousands of observations were nested within individuals. As a starting guideline, if 30 participants were run and served as the level-2 sample size, and each participant was recorded for 90 minutes, then each participant would have approximately 5,400 observations due to data being collected each second during the recording period. Therefore, the level-1 sample size, which represents the total number of records across all individuals, would be 30 (participants) x 5,400 (observations) = 162,000 (Stone, Shiffman, Atienza, & Nebeling, p. 339). This determination was made using sample size guidelines citing the use of at least 30 participants when there are 20-40 observations per individual for conducting mixed-model (or nested data) methods (Bell, Morgan, Schoeneberger, Loudermilk, Kromrey, & Ferron, 2010; Bolger, & Laurenceau, 2013, p. 37). This study resulted far more than 20-40 observations per individual, therefore, the sample size for analyses using the nested dataset was not a concern.

6.2.3 Data Measurement

For evaluation of BD prediction via physiological states, this study identified BD as the sole dependent and dichotomous variable (‘0’= behavior dysregulation absent; ‘1’= behavior dysregulation present) based upon researchers wanting to mainly know it any behavior of interest was present or absent. Three physiological measures served as continuous independent variables and were continuously recorded via a biosensor technology device. Due to the known influence of specific demographic variables on physiology (Collins, 1999; Davies & Morris, 1993), both age and sex were entered into statistical tests as a continuous and categorical covariate respectively when appropriate. All physiological measures and the outcome variable were measured each second for the entire recording period. This small sampling window will allow for capturing spontaneous or brief changes in the variables of interest.

6.2.4 Data Collection

The research tasks took no longer than two hours to complete. Behavior data was recorded by clinician post hoc ratings from video recordings of participant research sessions and biosensor technology continuously recorded physiological states.

All data collected was kept confidential by giving each participant an arbitrary identification number, which served as a link between raw data and the individual who provided it. This confidential data was entered into a spreadsheet (using the identification number for coding purposes) by a research team member for statistical analyses and stored separately (in a data safekeeping room within Pitt-RST facilities) from any identifiable information. This process allowed for raw data to remain unidentified, even by the research team members who had access.

6.2.5 Assessments

6.2.5.1 *OBDRS rating scale*

The OBDRS is checklist developed for the purpose of this research that produces ratings on whether specific behaviors believed to be indicative of BD are present or absent during an observation period. Data from this tool allowed for associations to be evaluated between behavior and physiological measures due to the external time stamp on the checklist and the internal time stamp embedded within the biomeasure software. Two trained clinicians each completed a post hoc behavioral analyses through video recordings of each participant (see *tool development* section of Chapter 3 for detailed information on the scale and Appendix A for actual tool).

6.2.6 Clinician Training

Two clinicians (Master's level students from the RST department) were hand selected and trained by the research coordinator on how to complete post hoc behavior ratings using the OBDRS tool through the behavior coding software. Both raters reviewed videos of all participants completing the research protocol. These raters did not interact with study participants, were unaware of assessment results, and completed all ratings independently of each other.

Clinician raters were trained through a standardized protocol occurring over a 3-week period prior to rating participant videos. Training included reviewing articles on relevant content, introduction to the OBDRS including how to administer the tool, and the scoring of four mock participant videos not included in the sample of videos clinicians rated for the research project. Each mock video scored was reviewed in detail by the clinician and the research coordinator to address coding concerns and provide feedback. Successful completion of training was defined as

when both raters established inter-rater reliability via a kappa statistic of at least .80 on the four mock participant videos. At this point, clinicians were cleared to begin coding real participant videos.

6.2.7 Instrumentation

6.2.7.1 *Bioharness-3*

The Bioharness-3 from Zephyr Technology Corporation is a portable physiological monitoring system worn as a smart fabric strap around the abdominal with a biomodule disc incorporating ECG and breathing detection sensors integrated in it. The device uses Bluetooth to continuously measure and transmit physiological data in real world, real-time environmental contexts, including research labs, clinics, and general daily living contexts for users. The Bioharness-3 has the ability to measure HR, RR, HRV, activity data, posture, speed, GPS, speed and distance, and peak acceleration. It is one of few remote measurement devices on the market currently with established reliability and validity (Hailstone & Kilding, 2011; Johnstone, Ford, Hughes, Watson, Mitchell, & Garrett, 2012; Kim, Roberge, Powell, Shafer, & Williams, 2013). The Bioharness-3 comes with integrated software applications making for streamlined data transfer and analysis.

6.2.7.2 *Video recordings*

Video recordings were taken using a standard video recording camera. Participant videos were stored on a secure server at the University of Pittsburgh and only accessed by the researchers associated with this project. All participants were asked to give consent to be videotaped and have their videos analyzed after completion of the study before any research tasks occurred.

6.2.7.3 *The observer xt software*

The Observer-XT is a widely used software package from Noldus that allows for the collection, presentation, and analysis of behavior data. Noldus allows for behavior coding tools, like the OBDRS to be embedded into a software program that allows for clinician ratings that can be easily compared, analyzed and synchronized with other external data sources (e.g., physiological monitors). To complete a behavior rating, a video recording is temporarily uploaded from the secure server into the software, but videos are not stored in the Noldus system permanently. All videos still remained stored on a secure server after clinician ratings were completed.

6.2.8 Procedure

The current research design followed a single group, cross-sectional design (although data was analyzed using longitudinal methods). This study was non-experimental and therefore, did not utilize randomization, blinding, or allocation processes. This project was approved by the University of Pittsburgh IRB prior to conducting any research activities (#PRO15100307).

The research protocol was administered face-to-face, one time to all participants. The research coordinator was responsible for running all participants through the protocol. Participants were video recorded and permission to record was obtained during the written consent process. The consent process occurred immediately prior to the administration of all research tasks. Physiological signals were continuously recorded throughout the entire testing period, beginning from device set-up and extending throughout the recovery period—measuring data during any given moment for comparison with behavioral events.

The study protocol began with a device set-up and pre-testing period. During this time, participants were introduced to and asked to put on the Bioharness-3 device. Following pre-testing, participants were escorted to a relaxation room where they were left alone and asked to simply relax for ten minutes. This allowed for participants to re-stabilize from any potentially non-baseline physiological levels simply from being asked to wear a device for data collection or from filling out potentially stressful self-assessments. The first five minutes of data collection during this time period was viewed as acclimation data, and therefore, was not used in any statistical analyses. The last five minutes of collected data then served as baseline data for further analyses. Removal of the first five minutes of data increased the likelihood that the data used for baseline is a true representation of each participant's resting physiological state.

Immediately following the acclimation period, participants were introduced to the first research task. Participants completed serial subtraction where they were asked to count down from 6233 by 13s. The task ended when either the participant reached 0 or five minutes had elapsed. Participants were informed prior to beginning the task that if at any point they made an error they would be asked to stop and start the task over from the beginning. A visual timer was provided letting participants know how much time they had left to work on the task. The subtraction task used in this study was adapted from the TSST identified in development and feasibility work.

Following the serial subtraction task, participants immediately moved into the speech preparation and delivery task. Here participants were told they would be given 10 minutes to prepare a 5-minute speech on how their injury has affected their life. They were given scrap paper and a pencil and pen to write ideas on while they planned their speech. At the end of the five-minute preparatory period participants were then asked to stand and present their speech to the camera. What they were not told before completing the task was that they would be required to

present their speech without using their notes. This element was introduced to add a spontaneous unplanned stressor into a life situation. This speech task was also adapted from the TSST.

Immediately following the speech task, participants were asked to complete a third and more functional task. The hamburger turning task is a functional task developed and used clinically for over the past 12 years by clinicians and researchers at the University of Pittsburgh (Shugars, 2007). This task requires participants to prepare 16 fake hamburger patties on a faux grill based on a 30-step procedural list modeled after the actual hamburger cooking process used in Wendy's restaurants. This task has been shown be a clinically useful measure of executive functioning (Shugars, 2007) and was selected to serve as a more hands-on task when compared to the other two tasks in the protocol. Completion of this task required participants to first watch a brief, narrated step-by-step video of a person completing the task correctly from start to finish. Next, participants were asked if they had any questions about the task and then were introduced to task props (fake hamburger patties, faux grill, real saltshaker, and spatula). Participants were also given a list of task steps to use as a cognitive aid if they chose to do so while completing the task. Participants were told they would have 15 minutes to correctly cook all the hamburger patties and that similarly to the serial subtraction task, if they made an error, they would be alerted to the error and then asked to start the task over from the beginning.

Finally, after the completion of all research tasks, there was a 20-minute post-testing recovery period where participants were again left alone in a quiet room and asked to simply relax. This period was set aside to allow participants to return to baseline physiological levels before testing concluded. After the recovery period ended, the research coordinator escorted the participant back to the main testing room, stopped the video recording, disconnected the biosensor device, answered any remaining questions participants may still have had about the study, and

provided compensation. ReMed staff remained on-call during testing in the event a participant experienced distress levels beyond what was expected and considered appropriate during a common stressful event. This type of elevated effect was not anticipated and did not occur. Figure 8 depicts all phases of the research protocol.

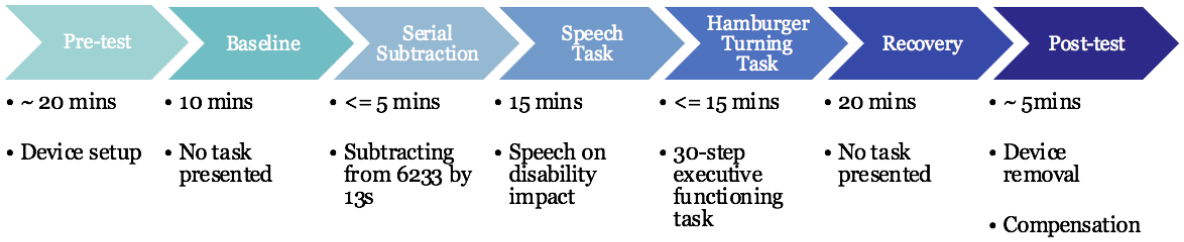


Figure 8. EMA protocol (identical to Study 2 protocol minus assessment during pre-testing period)

6.2.9 Facilities and Personnel

All research activities were conducted at ReMed of Pittsburgh's clinical facilities in Pittsburgh, Pennsylvania. Members of the research team were all affiliated with the RST department at the University of Pittsburgh and possessed the appropriate qualifications to run a study of this nature.

6.2.10 Data Source

The data set contained repeated measures data for all participants on physiology and BD. Each record represented one second of data collection. BD was the sole dependent, dichotomous variable, where '0' = absent and '1'= present based on clinician ratings. All physiological variables served as continuous independent variables and demographic data was included. Therefore, many

records were nested within each participant. This nested setup allowed researchers to examine the ability for continuous physiological data to predict BD events and change over time. All statistical analyses were conducted using the Noldus the Observer XT Software and Statistical Analysis Software (SAS, v. 9.4). All data was evaluated at the .05 alpha level.

6.2.11 Data Analysis

The present research involved examining physiological predictors of BD. The research question was, in adults with TBI who experience difficulty self-regulating behavior, do HR, HRV, age, sex, and RR significantly predict the probability that BD will occur? Due to the research question evaluating change over time in a sample of repeated measures, this study followed a longitudinal, event-contingent design, where the event is the occurrence of a target behavior.

The participants themselves served as the Level 2 or between-subjects analysis, which included person-specific variables of age and sex. Both physiological and behavioral data served as the Level 1 or within-subjects analysis.

6.2.11.1 *Data pre-processing*

Table 10 depicts the 5-step process used to process the raw dataset to fit appropriately with planned statistical analysis. Data cleaning is considered a critical stage in the analysis process when working with large, repeated measures datasets. The data was organized into spreadsheets for each participant individually, where each row represented a single observation or one second of time and each column represented one predictor variable for that record. Data was then trimmed to exclude biased time points during the observation period (e.g., device set up and acclimation), then evaluated for missing data and methods for appropriately dealing with missing data (Huck, 2012;

Portney & Watkins, 2009), aggregated when appropriate, and finally re-combined into one large dataset.

Table 10. 5-step data cleaning process

Step		Process
1	Creation of Individual Datasets	– Combined dataset used to create 14 separate datasets for each participant – This allowed for data cleaning to account for within-subject differences
2	Data File Trimming	– First 5 minutes and last 5 minutes of each participant data file was removed through trimming – This was done to ensure data was not being analyzed that captured influence from device acclimation or removal – This is a process commonly used with biosensor technology
3	Missing Values and Imputation	– Descriptives run to assess amount of missing data for each participant. – Missing data was determined to be data either outside of the device's acceptable range or system missing data. – Variables containing missing data were recoded into new variables as a '-1'. · If <=5% of data was missing--> mean imputation conducted (missing data replaced by the average of 2 nearby data points) · If >5% of data missing --> multiple imputation conducted (missing data replaced by predicted values based on available data
4	Data Aggregation	– A second data file created from the long data file containing records in seconds for all participants. – This new file aggregated all within-subject records to either a person-specific mean or total value (depending on the variable) across all time periods of interest. – This allowed for analyzing between-subject differences.

6.2.11.2 *Generalized linear mixed model*

Due to the hierarchical nature of the dataset, a hierarchical linear model (HLM) was the most appropriate analysis for the nested data. However, traditional HLM models assume normally distributed, continuous outcomes as well as independence between measures—assumptions

violated by the dataset used in this study. Data that is hierarchical in nature, but with a research question concerned with both between and within-subject differences (yielding correlated data), along with a non-normal data distribution, are not appropriate for traditional HLM models (Garson, 2013). Hierarchical generalized linear mixed modeling (GLMM) served as an alternative HLM approach that could account for the correlated nature of the dataset, and non-normal categorical outcome variable. GLMMs are an extension of generalized linear models that include random effects in the linear predictor (Garson, 2013). They are more complex than generalized linear models, but are appropriate when the primary focus is on evaluating the effect of predictor variables on an outcome over time within a given individual (SAS Institute, 2005). It was determined that the SAS PROC GLIMMIX procedure was the most appropriate analysis to answer the research question (Littell, Stroup, Milliken, Wolfinger, & Schabenberger, 2006; Walls, & Schafer, 2005).

6.2.11.3 *Model building process*

A structured model building process was followed based on guidelines from Schabenberger (2005) and Bell, Ene, and Schoeneberger (2013). The process began with an unconditional model, where no predictors were included. This model was then compared to more complex models as parameters were added in and each model was evaluated for improvement of model fit from the previously tested model. This process continued until it was determined a model had been isolated to answer the research question. Due to the exploratory nature of this study, the model building process remained simple and used forward selection, where one variable was added into the model at a time. Predictor variables in the model were the three physiological variables as well as age, sex and, time since injury.

The procedure used followed a binomial distribution due to the dependent variable being dichotomous in nature and utilized a logit link function. The Laplace estimation, a commonly used estimation technique used in PROC GLIMMIX, was used to model the data through a 'quasi-likelihood' approach, which approximates maximum likelihood. Since maximum likelihood approaches are inappropriate for non-normal data, the Laplace estimation allows for models to be analyzed for best fit using the same strategies used in multi-level linear modeling (Ene, Leighton, Blue, & Bell, 2015).

6.2.11.4 *Model fit and examination*

Model fit was determined using the Bayesian Information Criterion, or BIC. The BIC is a penalized likelihood method commonly used to evaluate mixed model accuracy through maximum likelihood estimation. Given the likelihood of observed data automatically increases when additional parameters are added into a model, the BIC gives a more accurate justification for adding or removing parameters based on a model complexity penalty (Seltman, 2009). This penalty estimates the accuracy of a model while taking into account both the number of parameters and number of participants included. When comparing two or more models, the smaller BIC indicates a better fitting model. Although determination of what magnitude of difference between two BIC values should be considered significant is highly subjective (Bell, Ene, & Schoeneberger, 2013), suggested ranges provided by O'Connell and McCoach (adapted from Raftery 1995) were used along with researcher judgement to determine model fit—where a difference of 6-10 suggested strong evidence for a better fitting model and a difference of 10 or greater indicated very strong evidence in favor of the more complex model (2008).

As a secondary supporting method for determining a strong balance between the model's complexity and fit to the data, once a model with the best fit was selected according to the BIC,

the Pearson chi-square/df ratio was evaluated to confirm that the variability in the selected model had been properly modeled without concern for over dispersion of correlated data, where ratios closest to 1 indicate a good fit (Schabenberger, 2005). In addition to goodness of fit statistics individual odds ratios (ORs) were examined for each predictor while accounting for all other variables in the model (Schabenberger, 2005).

6.3 RESULTS

Three versions of the full dataset were used depending on the analysis. Descriptives were analyzed by selecting only the first record for each individual in the dataset, given the variables examined were fixed and not susceptible to change over repeated measures (e.g., age). The resulting dataset was 14 records long, where one record represented one participant's data. Physiology and behavior trends were also analyzed using this aggregated file, where variables were represented as condensed averages across participants in order to examine between-subject differences across all time periods of interest. For the PROC GLIMMIX procedure, a long data file, made up of over 88,000 uncondensed, repeated measure observations was used for exploring predictive factors. This dataset used a 2-level nest, where level-2 data represented the individual participants and level-1 data represented the records nested within each participant (Walls & Schafer, 2005). Due to this study examining physiology's ability to predict behavior only during the completion of challenging task periods, the time period of interested was extracted from the original long file—resulting in a smaller dataset of 28,327 records.

6.3.1 Descriptives

Table 4 shows demographic characteristics of the sample (N=14). All continuous demographic measures were normally distributed, with the exception of time at ReMed. This sample of adults with TBI was on average 40.5 years old, 152.2 months post injury, and a client at ReMed for 19.7 months. The majority of the sample was Caucasian, male, single or divorced, and receiving services from the outpatient side of the rehabilitation program—which is representative of the greater population of individuals with TBI.

6.3.2 Behavior Outcomes and Physiology Trends

Behavior event frequencies and the amount of time participants spent engaging in target behaviors were analyzed to determine if in fact the OBDRS tool could pick up behaviors of interest and warrant a GLMM analysis. Table 11 provides a breakdown of the six target behaviors measured during both the baseline and task period (recovery is not represented due to methodological constraints not allowing for measurement of behavior during this time period). Overall, 191 behavior event were coded, representing a total of 3,393 seconds of total recording time being represented by a '1' or present behavior. Both perseveration and physical aggression/violence were never coded by either clinician and therefore, these behaviors are not represented in the table. Verbal dysregulation was the most commonly coded behavior, representing 104/191 or 54.5% of the total number of behaviors coded. This was compared to the next most commonly coded behavior category of disengagement (18.8%), with non-verbal dysregulation following close behind (18.3%), and opposition representing only 8.4% of all coded behaviors. When analyzing

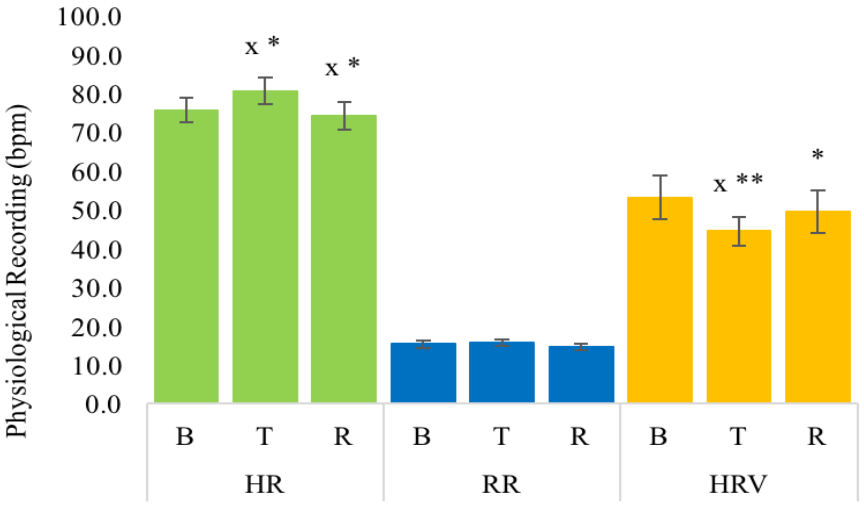
behaviors by time period as expected, the majority of coded behaviors (180/191) occurred during the task period, representing 94.2% of all behaviors coded.

Table 11. Overt behavioral dysregulation rating scale behavior count data

Observation Period	# of Behaviors Observed	Engaged Time (sec)
Baseline	**11**	**240**
Disengagement	*3*	
Opposition	*2*	
Non-verbal dysregulation	*3*	
Verbal dysregulation	*3*	
Task	**180**	**3153**
Disengagement	*33*	
Opposition	*14*	
Non-verbal dysregulation	*32*	
Verbal dysregulation	*101*	
Total	**191**	**3393**

Average between-subject physiological trends over time (baseline, task, recovery) were examined to assess change patterns. Physiological states that showed clinically meaningful change during the task period may suggest these same signals have potential to be predictive of BD during task completion. Cohen's *d*av was used to assess mean change, and is a form of the traditional Cohen's *d* often used with repeated measures data (Lakens, 2013). Figure 9 displays average change of HR, RR, and HRV over time. When examining physiological change from baseline to the task period, a clinically meaningful effect was seen for both HR (*d*av = -.405) and HRV (*d*av = .502), with HRV showing the larger effect. These changes indicate both HR and HRV were responsive to the tasks presented. Physiological change was also observed from the task period to recovery—which is anticipated when a recovery period with no stimuli is presented after the completion of challenging tasks. Here, both HR and HRV showed small, clinically meaningful effect sizes (*d*av = .415; *d*av = -.330) that were significant at the .05 level. These numbers suggest

that from the task period to recovery, HR is decreasing and HRV variability is increasing. Also worth noting was the lack of a clinically meaningful change in RR between any of the time periods. This may provide evidence that this physiological measure may not prove to be predictive of BD. This analysis provided only a method for examining between-subject change, and although valuable, results should be interpreted carefully when within-subject differences are also of importance.



Note: x represents statistical significance | number of * represents size of effect

Figure 9. Physiology by participant and time (between-subjects effect)

Figure 10 provides a visual representation of mean physiology for each time period of interest against BD (represented as total number of seconds). When focusing on the task period (the primary period of interest), this period was characterized by a very large percent of total BD, along with increased HR, decreased HRV, and stable RR when compared to baseline. With the exception of RR being unchanging, both HR and HRV moved in expected directions when individuals are exposed to stressful stimuli (increase in HR and decrease in HRV). This data suggests both behavioral and physiological change occurred during the task period when compared

to baseline and recovery, providing support for examining whether physiology can in fact be predictive of BD.

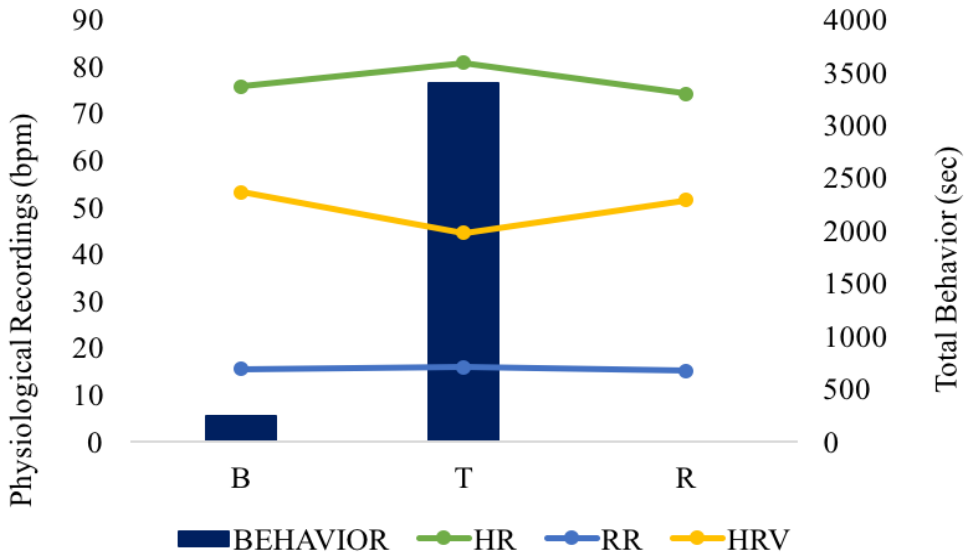


Figure 10. Average physiology and total behavior across time

Finally, in support of highlighting the inherent physiological variability between persons, Figure 11 is a visual comparison of two participants who seemly emerged as a physiological and behavioral responder (left side) and non-responder (right side). With the responder participant, the task period was characterized by increased HR and decreased HRV, and a steady change was also seen with RR—with breathing rate increasing (the expected direction). Behavior also increased substantially during the task period, especially during the speech task. Contrastingly, the non-responder saw seemingly stable measures for HR and RR, along with slow but steady increase in HRV, moving in the opposite direction as expected. One similarity the responder and non-responder share is the behavioral spike during the speech task. Although physiological and behavioral differences between individual research tasks is outside the scope of the research question for this project, it certainly is a pattern worthy of further exploration. What Figure 11

does do is further the argument for studying both within and between subject variability when researching physiology and behavior.

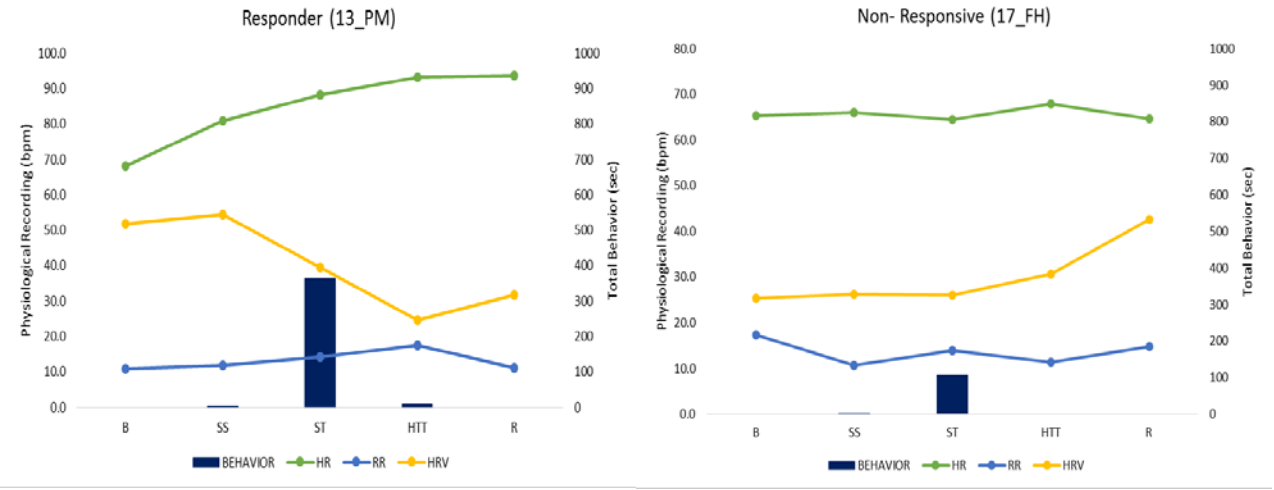


Figure 11. Example of a physiological and behavioral responder and non-responder

6.3.3 Generalized Linear Mixed Model

6.3.3.1 *Preliminary analyses*

Although GLMMs come with the flexibility of less assumptions due to their ability to account for non-normally distributed data, there is still one assumption researchers must at least be aware of, and that is multi-collinearity. Pearson product moment correlations were run between all pairs of physiological predictors in order to assess whether strong relationships existed between any two predictors and subsequently, may impact model fit by adding redundancy when entered into a model together. Table 12 displays a correlation matrix containing physiological measures of HR, RR, and HRV and as expected, all measures were moderately correlated with one another. Again, these correlations did not suggest a GLMM was not an appropriate fit for the data, but instead provided additional information that may assist with model selection and interpretation.

Table 12. Evaluation of multicollinearity of continuous predictors in a generalized mixed linear model

	HR	RR	HRV
HR	1		
RR	0.362	1	
HRV	-0.441	0.345	1

** Correlation is significant at the 0.05 level (2-tailed).*

While between-subject trends were examined for average physiology and behavior, within-subject variability may be even more important to examine due to the GLMM analysis chosen for its unique ability to capture intra-subject data through repeated measures analyses. Figure 12 represents three average growth trajectories, one for HR, RR, and HRV respectively. Each graph allows for a visual analysis of both within- and between-subject variability by the plotting light grey lines, one representing each participant, and a thicker, solid black line representing the overall sample average (Singer & Willett, 2003). While the black line is a visual for the overall average, what is most interesting about a growth trajectory is its ability to visually display both within and between subject variability across time. The x-axis in all three plots represents time (baseline, task 1, task 2, task 3, and recovery) and the y-axis displays physiological readings. These graphs show that as expected, large amounts of variability were seen both between- and within-subjects when measuring physiology. Of special note is the HRV growth trajectory, showing arguably the greatest variability from its large peaks and valleys for some participants, while others remained more stable over time. This may be an indication of the predictive value of HRV.

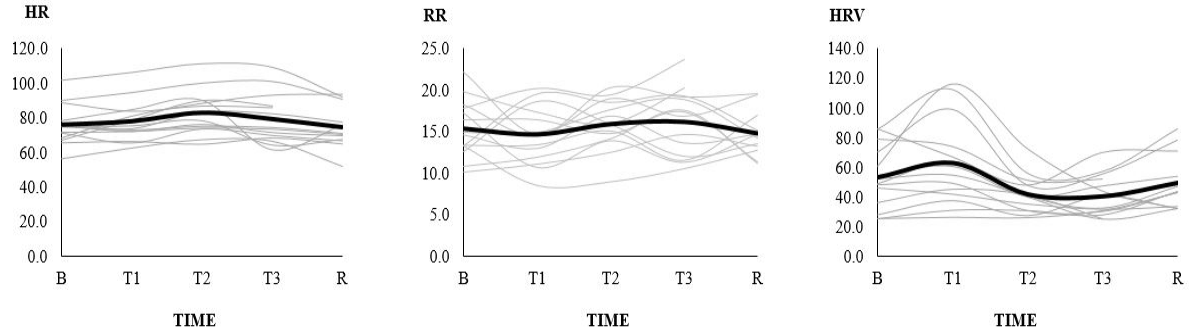


Figure 12. Average growth trajectories for physiological variables

Figure 13 provides an additional method for viewing physiological variability in the sample. Individual panels represent each participant and display within-subject change on the particular physiological variable of interest. This is in contrast to the average growth trajectories, which plot all sample data on one set of axes. Panels can also be compared to visually assess variability over time between-subjects (Singer & Willett, 2003). Unique to panel plots is the ability to more concretely identify members of the sample as responders (individuals showing physiological change over time) and non-responders (those not showing much change). Although not a formal analysis, examination of data from Figures 12 and 13 isolate change patterns worthy of further study. Of the most notable was the lack of responsiveness of RR in the majority of the sample, the high peaks and valleys associated with HR and HRV, and overall movement in the expected direction during periods of stress, and the unanticipated direction RR moved in when compared to expected RR patterns during stressful experiences.

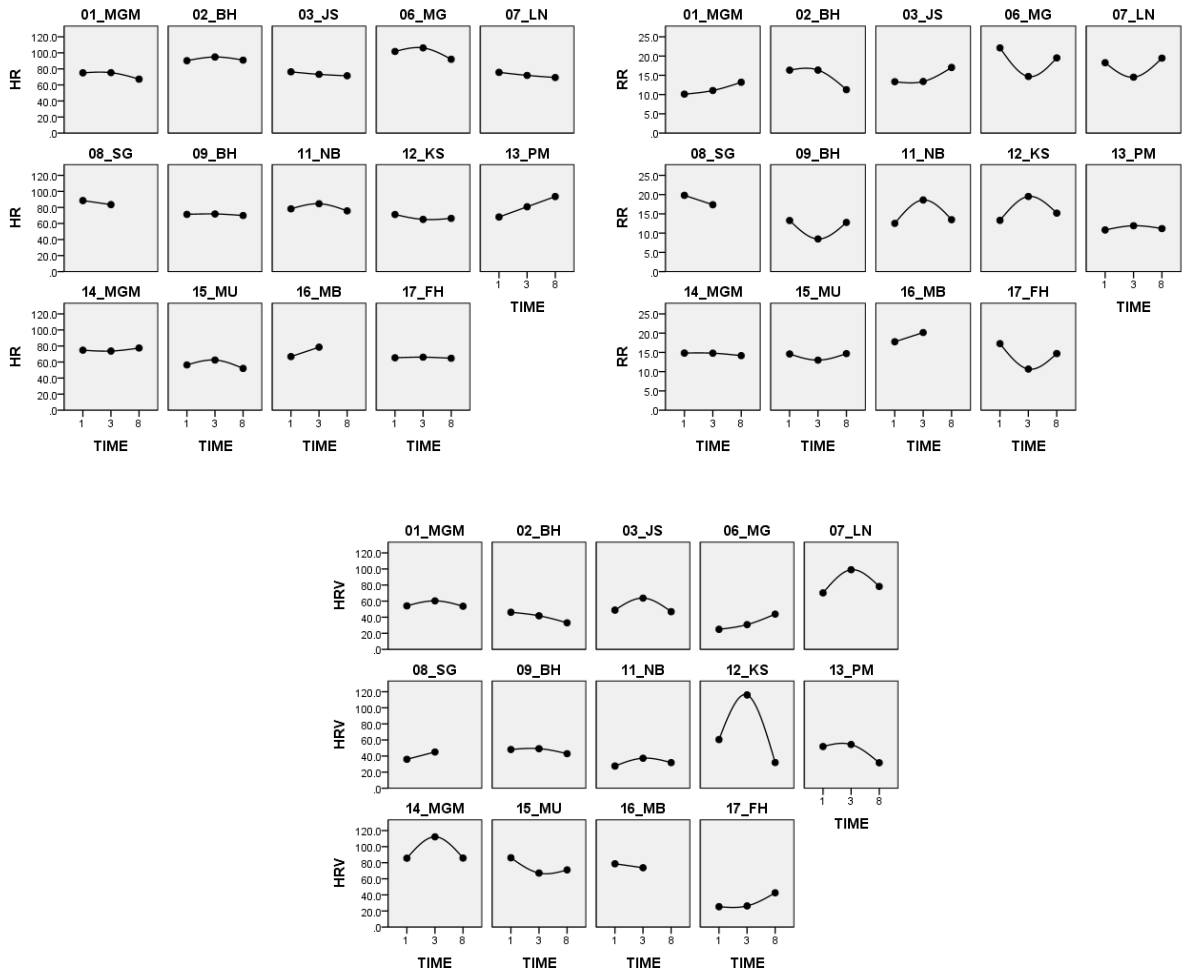


Figure 13. Empirical growth panel plots showing between- and within-subject physiological variability

6.3.3.2 *Model fit*

An exploratory GLMM was run using the SAS PROC GLIMMIX procedure to determine the probability of HR, RR, and HRV predicting BD when controlling for age, sex, and time since injury due to evidence of their influence on physiology and post-injury behavior (Buss, Larsen, Westen & Semmelroth, 1992; Sharma, 2003). Fourteen participants were included in the model, with individual records nested within each participant. This dataset contained 28,327 observations, where 2,648 or 9.3% of the data contained an event as the outcome. The syntax written for this

model was constructed to predict the probably of the outcome of interest being a '1', or that behavior was present.

An unconditional model including no predictors was first run to calculate the intraclass correlation coefficient (ICC) using the Level-2 covariance estimate for the intercept and an estimated Level-1 error variance of 3.29 suggested from Snijders & Bosker (1999) (as cited in O'Connell et al., 2008 due to dichotomous mixed models assuming no Level 1 residual). The ICC showed that the participants alone could account for 48.9% of the variability in BD, suggesting that 51.1% of the variance in the data may stem from within-subjects factors, further supporting the use of a multi-level model for this nested dataset. This effect is not surprising given the well-documented differences in physiological states both within and across individuals (Buss, Larsen, Westen & Semmelroth, 1992; Sharma, 2003).

Examination of model results overall did not yield positive findings for discussion; that is, there were negligible effects (ORs near 1) for all physiological variables. Evaluation of ORs of demographic variables did however show a large, clinically meaningful effect for sex, where females ('=1') had 4.23 greater odds than men of experiencing BD.

6.4 DISCUSSION

The aims of this study were to (1) examine physiological change during the completion of challenging tasks in adults with TBI, (2) explore physiological variability both between and within individuals and, (3) evaluate physiological variables as potential predictors of BD.

Examination of the physiological recordings from the Bioharness-3 and coded behavioral incidences showed a large increase in observed BD during the task period. Physiological change

mirrored this increased behavior, as clinically meaningful change was seen in both HR and HRV when comparing average change between baseline and the task period. Interestingly, and not expected was the lack of a meaningful change in RR during all time periods of interest, suggesting this physiological state may not be sensitive to the tasks presented in the current study. Both HR and HRV experienced changes in expected directions (i.e. increased HR and decreased HRV) indicative of a stress response during the task period. Further examination of average change trajectories displaying both within- and between-subject differences showed that although between-subject averages of physiology moved in expected directions, individual change was highly varied—moving in the intended direction for some individuals, or remaining unchanged or even moving in unexpected directions for others). Building upon this within-subject variability is the large degree of between-subject variability seen when examining physiology on a case-by-case basis. Results support well-established literature on the physiological differences that exist both across and within individuals—something researchers must be conscious of when interpreting results (Collins, 1999; Davies & Morris, 1993). This inter and intra physiological variability may speak to a need for isolating characteristics that define individuals as 'responders' and 'non responders' and also distinguish these two groups from each other for analysis. Future research should focus on (1) determining if it is possible to objectively identify factors that characterize individuals as physiological and behavioral responders and non-responders, (2) examining additional differences between these two groups (e.g., demographic differences), and (3) building a model that may better predict BD through use of a dataset consisting of only responder data. It is possible the prediction of BD through physiological change will be an assessment method most effective when tailored specifically to individuals who display a subset of defining characteristics (e.g., higher than normal baseline HR).

Results from the unconditional model showed 48.9% of the variability in the outcome was explained by between-subject factors, while the remaining 51.1% may stem from within-subject variability. These results provide further support for examining both within- and between-subject change through repeated measures data when physiology is a construct of interest. This finding sheds light on a potential shortcoming of literature only looking at between-subject changes—suggesting over half of the variability in the outcome can be lost.

It was determined that known influencing factors of physiology should be controlled for during model building due to their potential influence on the predictive utility of the physiological signals included in the model. However, results showed both age and time since injury, as well as all three physiological variables did not result in clinically meaningful prediction of the outcome when included in the model.

One positive finding that stood out and is worthy of discussion is the very large difference in odds seen between males and females experiencing the outcome—with females being at 4.23 greater odds of experiencing BD than men. Several potential and highly likely explanations exist for this sex difference. A connection might exist between sex and the specific type of BD engaged in during the protocol. Preliminary visual examination of the data and reports from the research coordinator showed that women more commonly engaged in non-verbal and verbal behaviors (e.g., crying and negative self-talk), while men seemed to be more oppositional or disengaged (e.g., requesting breaks, quitting tasks, or asking to do complete a task different than instructions stated). Therefore, it might be that sex is not necessarily the important explanatory factor in these results, but maybe it is the type of target behavior, and that a particular subset of the six target behaviors measured may increase the likelihood of predicting the outcome over others. It may also be that

there truly is a difference between men and women regardless of the type of behaviors expressed when experiencing a maladaptive behavioral response to the research task.

Looking further into the finding that men most commonly disengaged compared to women, disengagement was also the behavior clinician raters most commonly disagreed on when coding. It is possible sex differences are being misrepresented due to coding challenges that may exist with this behavior in particular when compared to the five other behaviors. Additionally, unique to the coding of disengagement is when a participant quit a task, they were required to allow the timer to reach zero before being able to move onto the next task or recovery period. In this case, clinicians were trained to code the outcome as present from behavior onset through the remaining time on the clock the participant was not working on the task. Although disengagement was in fact present during this time frame, the stimuli causing the behavior response (i.e. the task itself) was not present if the individual was not actively participating. Therefore, the coding of a '1' during this time period may be falsely influencing sex results. Building upon the uniqueness of the disengagement behavior and its potential influence on study results, is that disengaging may not simply be an example of BD. Disengagement may in fact be serving as an effective avoidance strategy or coping withdrawal for the participants who decided to discontinue working on the research task, and as a result participants reduced their physiological response due to removal of the stressor. Whereas the data shows a behavioral of dysregulation is present, those individuals may in fact have been having a positive or event neutral experience due to the task being over.

These potential explanations for the differences in sex seen in the data suggest sex, differences between target behaviors, disengagement as a target behavior versus a form of coping withdrawal, and the coding approach itself may all be worthy of further exploration to determine if any of the identified factors may be influencing differences observed in sex when predicting BD.

This overall lack of positive findings from the GLMM, although not anticipated, suggests that the statistical approach may not have been the most appropriate for analyzing factors that may influence BD. Therefore, these results provide strong support for future research efforts to focus primarily on exploring more optimal approaches of analyzing the current dataset (e.g., machine learning techniques) to answer the research question.

6.4.1 Potential Limitations

With the goal in mind of including as many participants as possible who met inclusion/exclusion criteria, individuals from both the residential and outpatient side of the rehabilitation program were included in this study. It can be argued that individuals in the residential program may be in need of increased assistance and rehabilitation services and therefore, may not have a comparable degree of self-regulatory ability to individuals participating in the outpatient program, who mostly live independently or with family members and typically display a greater level of independence. As a result, differences between residential and outpatient samples may be worthy of future exploration with a larger sample sizes and equal individuals across groups.

Task selection is a potential limitation shared by all protocols. In the original study, researchers questioned whether the selected tasks, although challenging, had the ability to elicit the construct of interested. It was determined at that point in order to move forward new tasks be selected with more evidence to support their use with less methodological restrictions—which was done prior to conducting OBDRS psychometric testing in Study 2. As mentioned earlier in the text, the TSST served as a foundation for researchers to select new tasks from due to the protocol's well-established use in the health psychology literature for encouraging stress responses. Due to methodological and ethical considerations, the TSST protocol was adapted to meet the needs of

this project, which also included the addition of a third, more functional task. This method of task selection gave researchers greater confidence that the tasks selected had the ability to evoke the outcome of interest. Although results seem to support the continued use of the tasks, it must be considered that these tasks were not only adapted from a standardized protocol, but were selected based on an assumption that a direct link exists between internal stress responses and overt behaviors of frustration. Therefore, another potential limitation of this project is the lack of literature on BD after TBI and physiology, and TBI. Due to this lack of evidence, it was determined by the researchers that utilization of well-established literature in stress, health, and psychology could serve as a supportive foundation to guide research efforts. It can be presumed there is a logical link between internal stress responses and overt manifestations of maladaptive behavior, but the connection is assumed to a degree. Future research exploring these relationships could greatly increase the understanding of the connections between constructs.

It is important to also mention that when examining the task period, differences may exist between the three individual tasks in the predictive utility of physiology on BD. This was an area outside the scope of this research project; however, it is possible that the differing nature of the three tasks presented will cause different degrees of physiology change and behavior frequencies. A secondary analysis focused on examining whether significance differences exist between the three tasks may indicate some tasks are better at eliciting the outcome of interest, and that future protocols should be adapted to focus solely on that particular task or a collection of similar tasks as opposed to the three highly different tasks selected for this protocol.

Due to the exploratory nature of this project, the protocol was delivered within a controlled clinic setting. Although this delivery context does not support the ultimate goal of EMA research (i.e. naturalistic contexts), this initial stage proved to be necessary in determining if physiological

change had the ability to predict the outcome in the target group. To gain supporting evidence, researchers needed to have increased control over the sample and environment. It is believed results from this study do support the future development of a community-delivered protocol.

Finally, although this study chose to analyze repeated measures data through a GLMM, this approach is not the only way of examining this complex dataset. Based on the overall lack of positive research findings from the model, our statistical approach may not have been appropriate and other methods must be explored to truly answer the research question. Machine learning—a commonly used approach for determining predictive factors of an outcome using large and complex datasets, is an alternative that was considered early on when conceptualizing this project. This approach may have better matched the research question and therefore, could have resulted in more meaningful findings that led to a greater and more complex discussion on physiology's role in predicting BD. Despite the limited rigor of the GLMM, what the current study was able to accomplish was to (1) establish evidence of clinically meaningful physiological change over time during the completion of challenging tasks in adults with TBI and (2) provide support for the importance of examining physiological variability both between and within individuals. These findings establish a direct relationship between physiology and BD—an area currently under addressed in the literature. This evidence alone provides support for future research to continue examining the potential of physiology to predict BD in the target group through alternative approaches.

7.0 SUMMARY

7.1 IMPLICATIONS

7.1.1 Novel Assessment Method for Consideration

The current research aimed to develop and provide support for a needed objective assessment method of BD that had the capability of being delivered through an EMA approach. When focusing on the target group of adults with TBI who experiencing difficulty self-regulatory behavior during daily living, no naturalistic assessment method existed for accurately and reliably measuring the construct of interest. Exploration of naturalistic delivery models is important for rehabilitation researchers to consider due to the increased ecological validity that comes with this data collection method; however, it arguably comes with an even increased importance for TBI researchers and those working clinically with individuals with TBI. TBI is a disability where individuals experience alarmingly high rates of poor self-awareness, lack of self-regulatory ability, behavioral breakdowns, and memory challenges. Symptoms like these shed light on a disconnect between traditional assessment methods of behavior and the target group due to not only their delivery within highly controlled clinic settings that may be masking true representation of behavioral challenges experienced during daily living, but also the dependency on the retrospective recall and self-awareness abilities of the individual when asked to report on behavior. It is encouraged rehabilitation professionals begin exploring assessment methods that utilize objective forms of data collection that do not require self-awareness or retrospective recall abilities and move

toward methods more complimentary with and supportive of the challenges faced by individuals with TBI.

7.1.2 Reliable and Naturalistic Data Collection

By utilizing objective EMA protocols, researchers and clinicians can have confidence that the data collected is not only reliable, but representative of how phenomena of interest manifest in the real world. This increased confidence will allow for an increased understanding of constructs measured and the relationships between them. The current research serves as only one exemplar of how objective EMA protocols can be developed and tested. Results from this study can give researchers a greater understanding of physiological patterns after TBI and how those physiological states not only change during periods of time when challenging experiences are encountered, but also the potential for physiology to predict maladaptive behavioral outcomes. It gives researchers and clinicians the ability to draw conclusion about overt behavior and its relationship with internal physiological states. This EMA method can be used for research purposes or in the clinic when measuring physiology is of importance, and can also be paired with a rehabilitation intervention.

7.1.3 Clinician Efficiency

A data collection method that is not only reliable, but one that can collect data continuously outside of finite clinic hours may increase clinician efficiency. With EMA, using objective data sources such as wearable physiological monitors allow for continuous data collection during daily living without the presence of a clinician. This method of collecting data can result in larger pools of data to infer challenges from and needed services, and a more accurate and reliable representation

of how deficits truly manifest outside of the clinic walls. This rich data has the potential for clinicians to spend less time collecting data over many sessions, spend less time asking the client to recall experiences during each session, and increase the ability to select the most appropriate interventions and rehabilitation plan due to confidence in reliable data to infer decisions. Wearing a device that continuously monitors and collects data independently may also decrease client burden, given data collection occurs without any needed input from the user. This reduced burden may also positively influence client buy-in. Clinicians may also see decreased time spent obtaining client buy-in, due to the self-knowledge gained about internal physiological states while wearing the device being reinforcing in itself. Increased clinician efficiency often leads to reduced costs spent on rehabilitation as a result of maximizing clinician time though more efficiently working with clients to achieve goals.

7.1.4 Individualized Intervention Pairing

The ability to successfully use an objective EMA approach to measure constructs of interest in community settings also opens doors for opportunities to pair assessment methods with naturalistic rehabilitation interventions. Future directions of this research include the use of a mobile application delivering an audio prompt that serves as a warning sign, indicating physiology has moved over a threshold that may predict a negative behavioral event may soon follow. This is only one example of the potential EMA data collection methods may have when paired with intervention. Development of full mobile applications that provide more structured and involved rehabilitation intervention components when paired with physiological monitors may also be possible (e.g., relaxation applications), or even the deployment of mobile applications that prompt the user to enter additional information about what they are experiencing in that exact moment

when meaningful physiological change is detected. These interventions have the capability to be individually tailored across persons to deliver the best fitting intervention based on individualized needs. While one individual may respond to prompts encouraging them to engage with a mobile application focused on relaxation techniques, another may only need an auditory prompt for redirection. This pairing of assessment and intervention in the real world is another way of maximizing clinician efficiency.

7.1.5 Client-Centered Approach

An objective EMA approach fits highly in line with the rehabilitation principle of the importance of the rehabilitation process being client centered and person-driven. Using this research as an example, future assessment and intervention community-delivered protocols may see participants gain an increased awareness about their own behavior and physiology from the wearable technology itself and their ability to engage in that process. Wearing the device and being able to track change may be reinforcing in itself due to the knowledge gained about the self. Interventions paired with EMA can not only be individualized but also encourage self-management due to deployment occurring during daily living and the requirement that the individual attend to warning signs and respond adaptively to potentially harmful situations. The use of technology for EMA data collection protocols often includes covert devices already used by the general community (e.g., smartphones and wearable activity monitors), which increases the chance an individual will buy-in to using the device and not feel they are doing something different than anyone else. In many cases, individuals using EMA collect data on devices they already own, and without any knowledge of anyone around the person. This technology is intended to fade over time to encourage independent and unprompted SR skills once new behaviors are learned and a greater

sense of self-awareness has been reinforced. EMA and intervention are truly a pairing that can reinforce the client-centered and client driven principles that guide rehabilitation.

7.2 FUTURE DIRECTIONS

Results from this research shed light on multiple areas worthy of further research to establish additional support of the OBDRS as an accurate and reliable assessment of BD as well as the ability for physiological change to predict BD and serve as a novel and naturalistic assessment approach that can occur in real-time.

7.2.1 Secondary Data Analyses

Before conducting future new research protocols, the current dataset can be used to further examine promising findings in the form of secondary data analysis. This dataset allows for (1) further examination of characteristics that may separate 'responders' from 'non-responders'; (2) use of machine learning techniques to more appropriately examine physiological factors that may predict BD; (3) the creation of predictive models focused on predicting individual target behaviors to assess if any particular behaviors may be better predicted by physiology; (4) a more thorough assessment of the role sex may play in experiencing BD; (5) the evaluation of differences in physiological trends and behavioral outcomes across different tasks and; (6) the use of different data cleaning processes (e.g., handling the disengagement behavior coding differently or removing behaviors all together prior to analysis), to test sensitivity and specificity of future predictive models.

Additionally, mentioned earlier in the text was the potential for machine learning techniques to add richness to study findings and be applied to the current dataset. ORs from statistical models help to explain the magnitude of the effect a variable has on the outcome, but may have limited ability to explain how well the model predicts the outcome. Machine learning would provide more information on the sensitivity and specificity of the model selected. This can be done through cross validation where the dataset for each participant is separated into two separate datasets—one for testing and the other for validation. The testing dataset can be up to 80% of the original dataset and is used for model building. The remaining (up to 20%) data is set aside for the final model to be applied to in order to determine if the clinician rater's observation of behavior matches the outcome from the model. This is done for each participant separately to examine if a model can be built for each participant individually, and therefore, provide support for the development of individual models of BD.

7.2.2 New Research Protocols

Examination of the reliability of validity of the OBDRS support running large-scaled studies aimed at more concretely establishing the tool's utility for measuring BD. Currently no tool exists specifically for naturalistic measurement of BD in non-acute TBI cohorts, and follow-up replication protocols will be a necessary and a critical next step to generate further support for and awareness of the tool. It is intended the OBDRS become a widely used assessment tool for the construct of interest and fill an existing gap in the literature.

When focusing on future studies aimed at further examination of physiology's ability to predict BD, results from this study suggest a logical next step is to explore factors that may be influencing predictive model outcomes (e.g., task type, target behaviors, and sex), or the addition

of new signals not included in the current research that may possess predictive value (e.g., voice, motion, tone/inflection). A larger-scaled psychometric replication study can also be used as an opportunity to analyze tasks and target individual behaviors while also exploring any influential role sex may play. A larger sample will allow sufficient power to be achieved when analyzing these group differences. Findings from this study will also allow researchers to assess whether defined groups can be characterized as 'responders' and 'non responders.'

Results gained from a follow-up protocol looking at factors influencing the predictive relationship of physiology on BD will ultimately lead to a more optimal fitting model that better predicts the outcome. This increased confidence will allow researchers to then use the best fitting model to determine whether a "therapeutic window" exists in the moments prior to a behavior occurring. The identification of a therapeutic window is a critical next step in the future of this research. If it can be determined that physiology is in fact serving as a warning sign in the moments leading up to a behavior, it must then be determined if that window of time is wide enough for subsequent cognitive rehabilitation intervention to be delivered in. Machine learning with an autoregressive model could again be useful here for examining which sized therapeutic window provides the best predictive value.

The intervention selected for delivery within the window is intended initially to be simple; something an individual can attend to quickly in the moment and something that can serve as an indicator that an undesirable behavioral event may occur. This can be something as simple as a tone or vibration from a mobile device that is integrated with biosensor technology. The ideal here is that the technology continuously measures physiological states and when those measures cross over a pre-determined threshold that is individualized to the person, the intervention (e.g., tone)

will deploy from the mobile application. Each individual will be trained to attend to the tone and realize what it means when it is heard prior to intervention deployment.

If a critical therapeutic window can be established along with the development of an intervention for pairing with the assessment protocol, the next phase in this research is real world application. Here individuals will wear the biosensor device during daily living, allowing their physiology to be recorded at any given moment in time and allowing for the cognitive intervention to be deploy during daily living when life situations are experienced that cause physiological change outside of pre-determined 'normal' limits. It is the hope that this novel assessment and intervention pairing can be effectively delivered outside of the clinic and allow for a true test of a context-specific model that supports increasing self-regulatory ability. If future research aims are achieved, long-term research aims will focus on adaptation and replication of the protocol to other disability cohorts who experience difficulty self-regulating behavior (e.g., individuals with Autism Spectrum Disorder, Attention Deficit Hyperactivity Disorder, and/or anxiety).

7.3 CONCLUSION

Physiological states have been explored as an objective EMA data collection method for measuring behavior when traditional self-reporting EMA approaches may be unreliable or inappropriate. This reliance on self-reporting is a limitation of many current assessment methods used with disability cohorts who have difficulty reporting accurately (e.g., TBI). Research questions on naturalistic behavior that were once difficult to answer due to methodological limitations may now be possible to systematically study through EMA models using biosensor technology. Due to the impact of persistent and often intervention-resistant behavior challenges after TBI and the negative

impact the inability to self-regulate has on daily living, it is important for rehabilitation professionals to explore novel assessment methods in an attempt to bridge this evidence gap and support positive rehabilitation outcomes.

APPENDIX A

OVERT BEHAVIORAL DYSREGULATION RATING SCALE

Purpose: This scale has been designed to categorize types of behaviors that are believed to be indicative of behavioral dysregulation during task completion and subsequently, lead to task breakdown. The OBDRS measures behavior type, frequency, duration, off-task behavior, and task re-engagement. It is to be completed by the clinician during direct behavior observation or post hoc behavior ratings.

Measure: The OBDRS measures the following 6 target behaviors:

- Disengagement
- Opposition
- Non-Verbal (non-violent) Dysregulation
- Physical Aggression/ Violence
- Verbal Dysregulation
- Perseveration

Each behavior is operationally defined on the following page based on the purpose of the rating scale. Definitions of each observation period, as well as guidelines for properly recording task re-engagement behavior and unique coding time restrictions are also provided. Please refer to the Appendix for specific examples of each target behavior prior to completing the OBDRS.

Instructions: Throughout each observation period from start to finish, if the individual engages in 1 of the 6 target behaviors, begin working across a row in the record sheet recording: the observation period, time of behavior onset, the type of target behavior, if the individual was off-task during this behavior, a brief description of the behavior, whether or not the individual re-engaged in the task, and the time of re-engagement (if applicable). Each row represents a single behavior incidence. Each behavior incidence is recorded even if the individual redirects independently or if the onset or continuation of the behavior does not disrupt engagement.

You may observe several behaviors during a particular period or none at all. Additionally, a behavior may fall under more than one target behavior category-if so, mark all the target behaviors that apply to the behavior being rated in that particular row. Finally, it is possible that one behavior incidence can begin while another behavior is ongoing (e.g., the individual walks away from the task [disengagement] and then throws their pen against the wall [physical aggression/violence]). If this happens, make sure to record these behaviors in separate rows.

OPERATIONAL DEFINITIONS

Disengagement (D): Volitional withdraw from the task or situation. Disengagement means the person is not actively engaging in the task (e.g., either by sitting down but not working on the task or getting up and removing themselves from the task table), and not simply being inattentive or distracted (see non-verbal/ non-violent dysregulation). An individuals' disengagement cannot be due to unplanned environmental influences (e.g., a fire alarm going off). Whether the individual disengages is independent of their decision to return to the task at a later time point, or to not return. Disengagement can be dually coded along with another target behavior for the same behavior incidence. This is because being disengaged can occur simultaneously with another form of behavioral dysregulation (but can also occur alone with coding additional behaviors). For example, if the person threatens to quit, but continues to complete the task, that behavior incidence is coded solely as verbal dysregulation; however, if the individual says they are going to quit and then proceeds to get up and walk away from the task table, then that behavior occurrence is coded as both verbal dysregulation AND disengagement. It is important to make note that although disengagement may occur in the same time frame as another coded behavior, their associated start/ stop times may differ.

Opposition (O): The act of opposition to any aspect of the task or situation. This act can be in response to a current situation (e.g., demanding help with a task) or an anticipated situation (e.g., an individual stating prior to the task their refusal to fill out questionnaires afterwards).

Non-Verbal (non-violent) Dysregulation (NVD): An act that does not involve using words. These acts are non-violent in nature. Non-verbal/ non-violent cues can be represented as an overt behavior (e.g., staring out of the window) or vocal sounds (e.g., crying, moaning, changes in vocal tone). This type of behavior can be due to inattentiveness or distraction and can be either volitional (e.g., foot stomping) or involuntary (e.g., self-stimulating behavior).

Physical Aggression/ Violence (PAV): An act that causes the self or another individual to be in danger or potential danger. Physical aggression or violence can also be displayed through the individuals' interaction with an object, regardless if this act involves another person (e.g., throwing a chair at a wall or throwing it directly at a person). The presence of these acts is independent of any long-lasting damage done.

Verbal Dysregulation (VD): Changes in or manipulation of verbal processes. These acts can be directed at the self (e.g., self-deprecating comments), another individual (e.g., interrupting), or an object (e.g., yelling or making insults). A verbal dysregulation act can also be non-directive. For verbal dysregulation to be coded, word must be produced. Any other changes in speech or vocal sounds (e.g., pitch changes, crying) must be coded as non-verbal dysregulation. It is important to note that constructive self-talk/ self-guidance is not a form of verbal dysregulation, but a beneficial problem-solving strategy, and therefore, should not be coded.

Perseveration (P): Repetition or continuation of a motor or verbal act or gesture that persists even after it is no longer an appropriate response. Coding a behavior of perseveration means there is evidence to suggest the persistent behavior is unwarranted or beyond effectiveness. This act can be in the form of extended maintenance (e.g., difficulty moving on from a specific topic), verbal

repetition or recurrence (e.g., repeatedly saying the same word or phrase), or a continuous and uninterrupted motor act (e.g., constant slamming of a hand on a table). An act of perseveration continues to an excessive degree, frequency, or time point. Its presence is independent of any redirection by the individual or given from an external source. Behaviors of perseveration are not simply repetitive acts. While an individual asking several times for clarification on tasks instructions shows repetition, this behavior may be a problem-solving strategy and necessary to aid them in completing the task, and therefore goes un-coded. Preservation implies the individual is unable to generate an alternative behavior, often due to a lack of cognitive inflexibility to generate an alternative (e.g., the inability to follow changes in task instructions despite effort to do so). Only repetitive acts that appear to not be related to a problem-solving strategy should be coded, and fall under either verbal dysregulation OR non-verbal/ non-violent dysregulation.

OBSERVATION PERIODS

Baseline (B): The 10 minutes prior to the completion of pre-testing assessments where the individual is asked to simply relax and not engage in any research tasks. This serves as an acclimation period after the participant is hooked up to the technology.

Pre-Testing Assessments (PTA): Period of time when participants are asked to complete four assessment measures. This time frame occurs immediately following the BASELINE period.

Serial Subtraction (SS): Period of time when participants are asked to complete a serial subtracting task requiring them to count down from 6233 by 13s. The task will end when either (1) the participant successfully reaches 0 or (2) 5 minutes has elapsed. This time frame occurs immediately following the PRE-TESTING ASSESSMENT period.

Speech Task (ST): Period of time when participants are asked to spend 10 minutes preparing a short speech on their disability, and the immediate 5 following minutes delivering their speech without looking at their notes. This time frame occurs immediately following the SERIAL SUBSTRACTION period.

Hamburger Turning Task (HTT): Period of time when participants are asked to complete a 30-step procedural task on how to cook hamburgers. This time frame occurs immediately following SPEECH TASK period.

Recovery (R): Period of time where no research tasks are presented and the individual is asked to simply focus on relaxing. This time frame occurs immediately following the HAMBURGER TURNING TASK period.

RE-ENGAGEMENT

Task re-engagement refers to whether or not the individual returns to working on the task as instructed AFTER a period of being disengaged. Re-engagement can be viewed as the opposite of being disengaged. A period of disengagement MUST be followed by a point where the individual re-engages in the task again (unless, they choose to not continue with the task at all). Remember, a target behavior can still occur without causing the individual to be disengaged (e.g., threatening to quit the task but continuing to work). Therefore, re-engagement is only possible after the individual has a break in task completion. This behavior is coded by marking the individual as engaged, which also ends the disengagement period simultaneously.

CODING TIME RESTRICTIONS

There are two important things to consider when coding the 6 behaviors of dysregulation.

1. Behaviors falling within the category of non-verbal (non-violent) dysregulation can be challenging to code as a result of difficulty discerning whether the behavior is occurring all together (e.g., face scrunching), or if the behavior is maladaptive in nature (e.g., repeated pen tapping). Therefore, in order to code a non-verbal (non-violent) behavior, the behavior incidence must be 3 seconds in duration or longer. Behaviors in this category that do not meet this minimum duration go un-coded. It is important to note that the 3-second minimum guideline only applies to this behavior category, incidences in all other behavior categories will always be coded regardless of their duration.

2. When determining if multiple behavior incidences of the same behavior type (e.g., Opposition) are separately coded behaviors or part of one larger, longer behavior incidence, each behavior must have at least 3 seconds of time between them for them to be coded as separate behaviors independent of one another. If less than 3 seconds occurs between the end of one behavior and the beginning of the next, the incidence is coded as one, larger behavioral event. For behaviors incidences of different behavior types occurring closely in time (e.g., Opposition and physical aggression/ violence), if less than 3 seconds occurs between the two behaviors code the start and stop times for the two behaviors separately and treat as two completely separate behavior incidences. It is possible the start/ stop times for two different behaviors occurring closely in time completely or partially overlap, this is ok.

TARGET BEHAVIOR EXAMPLES

Disengagement
• Refusal to continue with task • Wanders away from task table • Initiates taking a break from task • Noticeable slowing of task completion pace and/ or effort

Opposition

- Uncooperative, resistant
- Overly demanding
- Inflexible to change/ redirection
- Appears to intentionally deviate from task rules/ instructions

Non-Verbal (non-violent) Dysregulation

- Jaw clenching
- Grimace
- Face scrunching
- Sighing
- Foot stomping
- Sudden quietness
- Increase in talking speed or frequency
- Cries or laughs
- Increase in heavy breaths or breathing rate
- Body shaking (e.g., leg bouncing or fidgetiness)
- Folding of arms
- Rocking, rubbing, moaning, or other self-stimulatin behavior
- Clenching fist(s)
- Pen clicking
- Finger tapping
- Makes loud noises (direct or indirect)
- Change in voice tone (e.g., volume, pitch or sound)
- Inappropriate touching/misjudgment of personal space

Verbal Dysregulation

- Difficulty articulating thoughts
- Expressions of frustration (e.g., “this is so frustrating)
- Sub-vocalization (e.g., talking or mumbling under breath)
- Self-deprecating comments/ negative self-talk
- Abrupt or dismissive verbalization (e.g., responses that suggest intentional non-responsiveness)
- Interrupts research staff
- Multiple requests for breaks or help with task
- No vocal filter (e.g., inappropriate comments, use of foul language or insults)
- Shouts angrily (direct or indirect)
- Threat making toward others (verbal or through gestures)

Physical Aggression/ Violence

- Explosive and/or unpredictable aggression toward self, others, or objects
  - Bangs head, hits fist into objects, throws self on floor or into an object
  - Throws objects or physically acts on an object with or without breaking it
  - Strikes others (with or without causing physical injury)
  - Picks or scratches skin, hits self, pulls hair (or other minor physical behaviors)

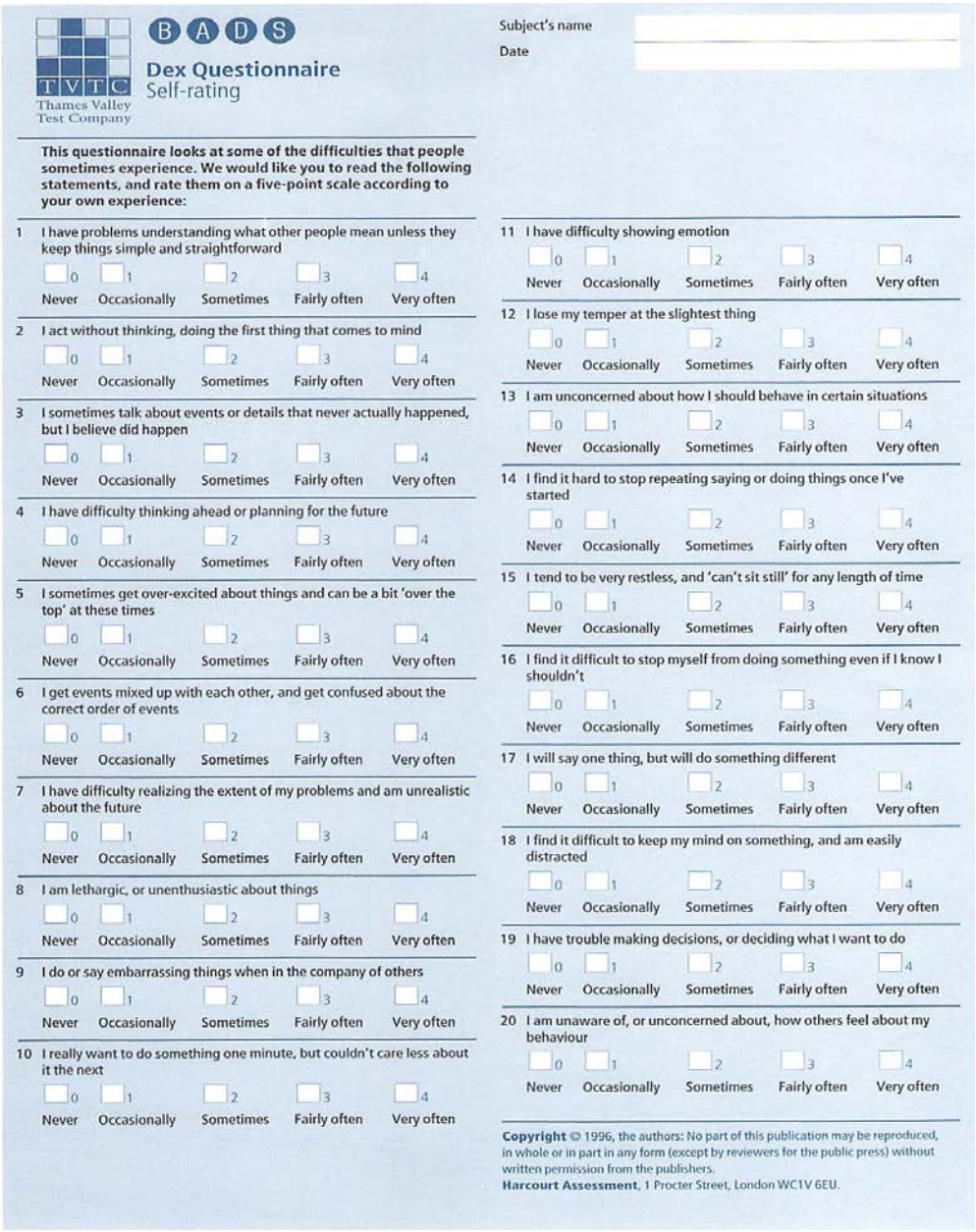
Perseveration

- Restlessness or excessive movement for a continued period of time
- Persistent behavior (motor, verbal, or gesture) that is unwarranted or beyond effectiveness

APPENDIX B

BEHAVIORAL ASSESSMENT OF THE DYSEXECUTIVE SYNDROME

DYSEXECUTIVE QUESTIONNAIRE



BADS

Dex Questionnaire
Self-rating

Thames Valley Test Company

Subject's name

Date

This questionnaire looks at some of the difficulties that people sometimes experience. We would like you to read the following statements, and rate them on a five-point scale according to your own experience:

1 I have problems understanding what other people mean unless they keep things simple and straightforward

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

2 I act without thinking, doing the first thing that comes to mind

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

3 I sometimes talk about events or details that never actually happened, but I believe did happen

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

4 I have difficulty thinking ahead or planning for the future

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

5 I sometimes get over-excited about things and can be a bit 'over the top' at these times

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

6 I get events mixed up with each other, and get confused about the correct order of events

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

7 I have difficulty realizing the extent of my problems and am unrealistic about the future

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

8 I am lethargic, or unenthusiastic about things

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

9 I do or say embarrassing things when in the company of others

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

10 I really want to do something one minute, but couldn't care less about it the next

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

11 I have difficulty showing emotion

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

12 I lose my temper at the slightest thing

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

13 I am unconcerned about how I should behave in certain situations

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

14 I find it hard to stop repeating saying or doing things once I've started

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

15 I tend to be very restless, and 'can't sit still' for any length of time

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

16 I find it difficult to stop myself from doing something even if I know I shouldn't

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

17 I will say one thing, but will do something different

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

18 I find it difficult to keep my mind on something, and am easily distracted

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

19 I have trouble making decisions, or deciding what I want to do

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

20 I am unaware of, or unconcerned about, how others feel about my behaviour

☐0 ☐1 ☐2 ☐3 ☐4

Never Occasionally Sometimes Fairly often Very often

APPENDIX C

MINDFULNESS ATTENTION AWARENESS SCALE

Day-to-Day Experiences

Instructions: Below is a collection of statements about your everyday experience. Using the 1-6 scale below, please indicate how frequently or infrequently you currently have each experience. Please answer according to what *really reflects* your experience rather than what you think your experience should be. Please treat each item separately from every other item.

1	2	3	4	5	6
Almost Always	Very Frequently	Somewhat Frequently	Somewhat Infrequently	Very Infrequently	Almost Never

I could be experiencing some emotion and not be conscious of it until some time later.	1	2	3	4	5	6
I break or spill things because of carelessness, not paying attention, or thinking of something else.	1	2	3	4	5	6
I find it difficult to stay focused on what's happening in the present.	1	2	3	4	5	6
I tend to walk quickly to get where I'm going without paying attention to what I experience along the way.	1	2	3	4	5	6
I tend not to notice feelings of physical tension or discomfort until they really grab my attention.	1	2	3	4	5	6
I forget a person's name almost as soon as I've been told it for the first time.	1	2	3	4	5	6
It seems I am "running on automatic," without much awareness of what I'm doing.	1	2	3	4	5	6
I rush through activities without being really attentive to them.	1	2	3	4	5	6
I get so focused on the goal I want to achieve that I lose touch with what I'm doing right now to get there.	1	2	3	4	5	6
I do jobs or tasks automatically, without being aware of what I'm doing.	1	2	3	4	5	6
I find myself listening to someone with one ear, doing something else at the same time.	1	2	3	4	5	6

I drive places on 'automatic pilot' and then wonder why I went there.	1	2	3	4	5	6
I find myself preoccupied with the future or the past.	1	2	3	4	5	6
I find myself doing things without paying attention.	1	2	3	4	5	6
I snack without being aware that I'm eating.	1	2	3	4	5	6

APPENDIX D

SATISFACTION WITH LIFE SCALE

The Satisfaction with Life Scale

By Ed Diener, Ph.D.

DIRECTIONS: Below are five statements with which you may agree or disagree. Using the 1-7 scale below, indicate your agreement with each item by placing the appropriate number in the line preceding that item. Please be open and honest in your responding.

1 = Strongly Disagree
2 = Disagree
3 = Slightly Disagree
4 = Neither Agree or Disagree
5 = Slightly Agree
6 = Agree
7 = Strongly Agree

_____1. In most ways my life is close to my ideal.

_____2. The conditions of my life are excellent.

_____3. I am satisfied with life.

_____4. So far I have gotten the important things I want in life.

_____5. If I could live my life over, I would change almost nothing.

APPENDIX E

GO/NO-GO TEST

Participant ID: _____

Item Number	Description	Criteria	Score	Qualitative Scores

(48)	Now I am going to knock on the table. If I knock once, I want you to knock twice; and if I knock twice, I would like you to knock once.
	1:[_____] 2:[_____] 2:[_____] 1:[_____] 2:[_____] 1:[_____]
		# errors
		0= none
		1= 1
		2= 2-6
			0 1 2	

(49)	Please take my hand. Now, if I say "red", I want you to squeeze and then relax your hand; if I say "green", do nothing.
	1) red:[_____] 2) green:[_____] 3) green:[_____] 4) red:[_____] 5) red:[_____] 6) green:[_____]
		# errors
		0= none
		1= 1
		2= 2-6
			0 1 2	

(50)	If I knock once, I want you to raise your right hand. If I knock twice, I want you to raise your left hand.
	1:[_____] 2:[_____] 1:[_____] 2:[_____] 2:[_____] 1:[_____]
		# errors
		0= 0-1
		1= 2
		2= 3-6
			0 1 2	

(51)	If I knock hard, you knock gently. If I knock gently, you knock hard. [Demonstrate hard and gentle knocks.]
	G:[_____] H:[_____] H:[_____] G:[_____] H:[_____] G:[_____]
		# errors
		0= none
		1= 1-2
		2= 3-6
			0 1 2	

APPENDIX F

EXPERT FEEDBACK EVAULATION FORM

Instructions: *Please review the attached rating scale. Your feedback is welcomed on all aspects of the tool including: format, instructions, user friendliness, clarity, operational definitions of target behaviors, research tasks selected for observation periods, and appendix of target behavior examples. As you evaluate the tool, please provide answers to the following questions:*

1) Are the 6 target behaviors related to behavioral signs of dysregulation in the target population (please select yes/no)?

	Yes	No
Disengagement		
Resistance		
Non-verbal(non-violent) dysregulation		
Physical Aggression/ Violence		
Verbal Dysregulation		
Perseveration		

2) Are the definitions for each target behavior sufficiently objective and behaviorally stated?

3) Are the behaviors being measured discrete?

4) Do the target behaviors reflect a threshold behavior (i.e. can it be dichotomously present or absent)?

5) Are any behaviors not represented on the scale that should be?

BIBLIOGRAPHY

Ali, A. A., Hossain, S. M., Hovsepian, K., Rahman, M. M., Plarre, K., & Kumar, S. (2012, April). mPuff: Automated detection of cigarette smoking puffs from respiration measurements. In *Proceedings of the 11th International Conference on Information Processing in Sensor Networks* (pp. 269-280). ACM.

Alderman, N., & Wood, R. L. (2013). Neurobehavioural approaches to the rehabilitation of challenging behaviour. *NeuroRehabilitation*, *32*(4), 761-770.

American Psychological Association. (2010). *Publication Manual of the American Psychological Association (6th ed.).* Washington, DC: American Psychological Association.

Anderson, V., Brown, S., Newitt, H., & Hoile, H. (2009). Educational, vocational, psychosocial, and quality-of-life outcomes for adult survivors of childhood traumatic brain injury. *Journal of Head Trauma Rehabilitation, 24*(5), 303-312.

Arciniegas, D. B., & Wortzel, H. S. (2014). Emotional and behavioral dyscontrol after traumatic brain injury. *Psychiatric Clinics of North America, 37*(1), 31-53.

Association for Applied Psychophysiology and Biofeedback. (2011). Disorders and treatment. Retrieved from http://www.aapb.org/i4a/pages/index.cfm?pageid=3404

Baguley, I. J., Cooper, J., & Felmingham, K. (2006). Aggressive behavior following traumatic brain injury: How common is common? *The Journal of Head Trauma Rehabilitation, 21*(1), 45-56.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry, 7*(1), 1-15.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. Academic Press.

Bell, B. A., Ene, M., & Schoeneberger, J. (2013). A multilevel model primer using SAS PROC MIXED. In SAS Global Forum (pp. 0-19).

Bell, B. A., Morgan, G. B., Schoeneberger, J. A., Loudermilk, B. L., Kromrey, J. D., & Ferron, J. M. (2010). Dancing the sample size limbo with mixed models: How low can you go? In *SAS Global Forum* (pp. 11-14).

Birkett, M. A. (2011). The trier social stress test protocol for inducing psychological stress. *Journal of Visualized Experiments, 56*.

Bolger, N., & Laurenceau, J. P. (2013). Intensive longitudinal methods: An introduction to diary and experience sampling research. Guilford Press.

Braden, C., Cuthbert, J., Brenner, L., Hawley, L., Morey, C., Newman, J., Staniszewski, K., & Harrison-Felix, C. (2012). Health and wellness characteristics of persons with traumatic brain injury. *Brain Injury, 26*(11), 1315-1327.

Brain Injury Association of America. (2011). About brain injury. Retrieved from http://www.biausa.org/about-brain-injury.htm

Brain Injury Association of America Brain Injury Association of America. (2011). Treatment. Retrieved from http://www.biausa.org/brain-injury-treatment.htm

Brass, M., & Haggard, P. (2007). To do or not to do: The neural signature of self-control. *The Journal of Neuroscience*, *27*(34), 9141-9145.

Brown, K.W. and Ryan, R.M. (2003). The benefits of being present: The role of mindfulness in psychological well-being. *Journal of Personality and Social Psychology*, *84*, 822-848.

Bruns, J., & Hauser, W. (2003). The epidemiology of traumatic brain injury: A review. *Epilepsia*, *44*(s10), 2-10.

Buss, D. M., Larsen, R. J., Westen, D., & Semmelroth, J. (1992). Sex differences in jealousy: Evolution, physiology, and psychology. *Psychological Science*, *3*(4), 251-255.

Carr, E. G., Dunlap, G., Horner, R. H., Koegel, R. L., Turnbull, A. P., Sailor, W., ... & Fox, L. (2002). Positive behavior support evolution of an applied science. *Journal of Positive Behavior Interventions*, *4*(1), 4-16.

Casals, M., Girabent-Farrés, M., & Carrasco, J. L. (2014). Methodological quality and reporting of generalized linear mixed models in clinical medicine (2000–2012): A systematic review.

Centers for Disease Control and Prevention. (2003). Glasgow coma scale. Retrieved from http://www.cdc.gov/masstrauma/resources/gcs.pdf

Centers for Disease Control and Prevention. (2006). Traumatic brain injury in the united states: Emergency department visits, hospitalizations, and deaths, 2002-2006. Retrieved from http://www.cdc.gov/traumaticbraininjury/pdf/blue_book.pdf

Centers for Disease Control and Prevention. (2010). TBI: Get the facts. Retrieved from http://www.cdc.gov/traumaticbraininjury/get_the_facts.html

Cohen, I. L., Yoo, J. H., Goodwin, M. S., & Moskowitz, L. (2011). Assessing challenging behaviors in Autism Spectrum Disorders: Prevalence, rating scales, and autonomic indicators. In *International handbook of autism and pervasive developmental disorders* (pp. 247-270). Springer: New York.

Collins, K. J. (1999). Physiological variation and adaptability in human populations. *Annals of Human Biology*, *26*(1): 19-38.

Davies, B., & Morris, T. (1993). Physiological parameters in laboratory animals and humans. *Pharmaceutical Research, 10*(7), 1093-1095.

de Geus, E. J., Willemsen, G. H., Klaver, C. H., & van Doornen, L. J. (1995). Ambulatory measurement of respiratory sinus arrhythmia and respiration rate. *Biological Psychology*, *41*(3), 205-227.

Emmanouel, A., Mouza, E., Kessels, R. P., & Fasotti, L. (2014). Validity of the Dysexecutive Questionnaire (DEX). Ratings by patients with brain injury and their therapists. *Brain Injury*, *28*(12), 1581-1589.

Ene, M., Leighton, E. A., Blue, G. L., & Bell, B. A. (2015). Multilevel Models for Categorical Data Using SAS® PROC GLIMMIX: The Basics. In SAS Global Forum 2015 Proceedings.

Fahrenberg, J., Myrtek, M., Pawlik, K., & Perrez, M. (2007). Ambulatory assessment--monitoring behavior in daily life settings: A behavioral-scientific challenge for psychology. *European Journal of Psychological Assessment*, *23*(4), 206.

Feeney, T. J., Ylvisaker, M., Rosen, B. H., & Greene, P. (2001). Community supports for individuals with challenging behavior after brain injury: An analysis of the New York State Behavioral Resource Project. *The Journal of Head Trauma Rehabilitation, 16*(1), 61-75.

Fillmore, M. T. (2003). Drug abuse as a problem of impaired control: current approaches and findings. *Behavioral and Cognitive Neuroscience Reviews*, *2*(3), 179-197.

Fillmore, M. T., & Rush, C. R. (2006). Polydrug abusers display impaired discrimination-reversal learning in a model of behavioural control. *Journal of Psychopharmacology*, *20*(1), 24-32.

Fowles, D. C. (2000). Electrodermal hyporeactivity and antisocial behavior: Does anxiety mediate the relationship? *Journal of Affective Disorders*, *61*(3), 177-189.

Garson, G. D. (2013). Fundamentals of hierarchical linear and multilevel modeling. *Hierarchical linear modeling: guide and applications. Sage Publications Inc*, 3-25.

Golden, C. J. (1980). The Luria-Nebraska neuropsychological battery: manual.

Gordon, W., Zafonte, R., Cicerone, K., Cantor, J., Brown, M., Lombard, L., … & Chandna. T. (2006). Traumatic brain injury rehabilitation: State of the science. *American Journal of Physical Medicine & Rehabilitation, 84*, 343-382.

Hailstone, J., & Kilding, A. E. (2011). Reliability and validity of the Zephyr™ BioHarness™ to measure respiratory responses to exercise. *Measurement in Physical Education and Exercise Science*, *15*(4), 293-300.

Heatherton, T. F. (2011). Neuroscience of self and self-regulation. *Annual Review of Psychology*, *62*, 363.

Heatherton, T. F., & Wagner, D. D. (2011). Cognitive neuroscience of self-regulation failure. *Trends in Cognitive Sciences, 15*(3), 132-139.

Helmick, K., Baugh, L., Lattimore, T., & Goldman, S. (2012). Traumatic brain injury: Next steps, research needed, and priority focus areas. *Military Medicine, 177*(8S), 86-92.

Henry, J., Phillips, L., Crawford, J., Ietswaart, M., & Summers, F. (2006). Theory of mind following traumatic brain injury: The role of emotion recognition and executive dysfunction. *Neuropsychologia, 44,* 1623-1628.

Heron, K., & Smyth, J. (2010). Ecologically momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology, 15*(1), 1-39.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16*(3), 174-180.

Huck, S.W. (2012). *Reading Statistics and Research* (6th Ed). Boston: Allyn & Bacon.

Iacono, W. G., Carlson, S. R., & Malone, S. M. (2000). Identifying a multivariate endophenotype for substance use disorders using psychophysiological measures. *International Journal of Psychophysiology, 38*(1), 81-96.

Intille, S. S., Tapia, E. M., Rondoni, J., Beaudin, J., Kukla, C., Agarwal, S., ... & Larson, K. (2003, January). Tools for studying behavior and technology in natural settings. In *UbiComp 2003: Ubiquitous Computing* (pp. 157-174). Springer Berlin Heidelberg.

Johnson, R., & Balleny, H. (1996). Behaviour problems after brain injury: Incidence and need for treatment. *Clinical Rehabilitation, 10*(2), 173-180.

Johnstone, J. A., Ford, P. A., Hughes, G., Watson, T., Mitchell, A. C., & Garrett, A. T. (2012). Field based reliability and validity of the BioHarness™ multivariable monitoring device. *Journal of sports science & medicine, 11*(4), 643.

Jumisko, E., Lexell, J., & Soderberg, S. (2009). The meaning of feeling well in people with moderate or severe traumatic brain injury. *Journal of Clinical Nursing, 18,* 2273-2281.

Kelly, G., Brown, S., Todd, J., & Kremer, P. (2008). Challenging behaviour profiles of people with acquired brain injury living in community settings. *Brain Injury, 22*(6), 457-470.

Kim, J. H., Roberge, R., Powell, J. B., Shafer, A. B., & Williams, W. J. (2013). Measurement accuracy of heart rate and respiratory rate during graded exercise and sustained exercise in the heat using the Zephyr Bioharness™. *International journal of sports medicine, 34*(6), 497.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1-2), 76-81.

Knyazev, G. G., Slobodskaya, H. R., & Wilson, G. D. (2002). Psychophysiological correlates of behavioural inhibition and activation. *Personality and Individual Differences*, *33*(4), 647-660.

Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality and Social Psychology*, *74*(3), 686.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863.

Lang, P. J. (1978). Anxiety: Toward a psychophysiological definition. *Psychiatric Diagnosis: Exploration of Biological Predictors*, 365-389.

Leibson, C., Brown, A., Long, K, Ransom, J., Mandrekar, J., Osler, T., & Malec, J. (2012). Medical care costs associated with traumatic brain injury over the full spectrum of disease: A controlled population-based study. *Journal of Neurotrauma, 29*, 2038-2049.

Levine, B., Deirdre, D., Schwartz, M., Boutet, I., Stuss, D. (2000). Assessment of strategic self-regulation in traumatic brain injury: Its relationship to injury severity and psychosocial outcome. *Neuropsychology, 14*(4), 491-500.

Littell, R. C., Stroup, W. W., Milliken, G. A., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models*. SAS institute.

Liu, K., Chan, C., Lee, T., Li, L., & Hui-Chan, C. (2002). Self-regulatory learning and generalization for people with brain injury. *Brain Injury, 16*(9), 817-824.

Lorber, M. F. (2004). Psychophysiology of aggression, psychopathy, and conduct problems: A meta-analysis. *Psychological Bulletin*, *130*(4), 531.

Mayo Clinic. (2012). Traumatic brain injury: Symptoms. Retrieved from http://www.mayoclinic.com/health/traumatic-brain-injury/DS00552/DSECTION=symptoms

McCue, M., Fairman, A., & Pramuka, M. (2010). Enhancing quality of life through telerehabilitation. *Physical Medicine and Rehabilitation Clinics in North America, 21,* 195-205.

National Commission for the Protection of Human Subjects of Biomedical and Behavioural Research. (1978). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. Maryland: ERIC Clearinghouse.

Nezlek, J. B. (2001). Multilevel random coefficient analyses of event-and interval-contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin*, *27*(7), 771-785.

O'Connell, A. A., & McCoach, D. B. (2008). Multilevel modeling of educational data. IAP.

Parikh, S., Koch, M., & Narayan, R. K. (2007). Traumatic brain injury. *International Anesthesiology Clinics*, *45*(3), 119-135.

Parmanto, B., Saptono, A., Pramana, G., Pulantara, W., Schein, R. M., Schmeler, M. R., & Brienza, D. M. (2010). VISYTER: Versatile and integrated system for telerehabilitation. *Telemedicine and e-Health*, *16*(9), 939-944.

Patrick, C. J. (2008). Psychophysiological correlates of aggression and violence: An integrative review. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *363*(1503), 2543-2555.

Pavot, W. G., Diener, E., Colvin, C. R., & Sandvik, E. (1991). Further validation of the satisfaction with life scale: Evidence for the cross-method convergence of well-being measures. Journal of Personality Assessment, *57*, 149-161.

Portney, L.G. & Watkins, M.P. (2009). *Foundations of clinical research: Applications to practice (3rd Ed).* Upper Saddle River, NJ: Pearson Education, Inc.

Quinn, P. D., & Fromme, K. (2010). Self-regulation as a protective factor against risky drinking and sexual behavior. *Psychology of Addictive Behaviors*, *24*(3), 376.

Roehr, B. (2012). Traumatic brain injury is more widespread than previously estimated. *British Medical Association*, 345.

Runyan, J., Steenbergh, T., Bainbridge, C., Daughert, D., Oke, L., & Fry, B. (2013). A smartphone ecological momentary assessment/ intervention 'app' for collecting real-time data and promoting self-awareness. *PLOSone, 8*(8), 1-9.

Sabaz, M., Simpson, G. K., Walker, A. J., Rogers, J. M., Gillis, I., & Strettles, B. (2014). Prevalence, comorbidities, and correlates of challenging behavior among community-dwelling adults with severe traumatic brain injury: A multicenter study. *The Journal of Head Trauma Rehabilitation*, *29*(2), E19-E30.

SAS Institute. (2005). The GLIMMIX procedure.

Schabenberger, O. (2005). Introducing the GLIMMIX procedure for generalized linear mixed models. *SUGI 30 Proceedings*, 196-130.

Schroeder, S. A. (2007). We can do better—improving the health of the American people. *New England Journal of Medicine*, *357*(12), 1221-1228.

Seltman, H. J. (2009). *Mixed models: A flexible approach to correlated data.* Experimental Design and Analysis.

Sharma, S. (2003). Athlete's heart-effect of age, sex, ethnicity and sporting discipline. *Experimental Physiology*, *88*(05), 665-669.

Sherer, M., Bergloff, P., Levin, E., High, W., Oden, K., & Nick, T. (1998). Impaired awareness and employment outcome after traumatic brain injury. *Journal of Head Trauma Rehabilitation, 13*, 52-61.

Shiffman, S., Stone, A., & Hufford, M. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*, 1-32.

Shugars, S. L. (2007). *A functional assessment of executive functioning: The hamburger turning task* (Master's thesis, University of Pittsburgh).

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Physical therapy, *85*(3), 257-268.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.

Smyth, J., & Stone, A. (2003). Ecological momentary assessment research in behavioural medicine. *Journal of Happiness Studies, 4*, 35-52.

Sten's Corporation (2015). Nexus-10 -mark II priced without sensors. Retrieved from http://stens-biofeedback.com/products/nexus-10-mark-ii-priced-without-sensors.

Stone, A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioural medicine. *Annals of Behavioural Medicine, 16*(3), 199-202.

Stone, A., Shiffman, S., Atienza, A., & Nebeling, L. (2007). *The science of real-time data capture: Self-reports in health research*. Oxford University Press.

Tate, R., & Broe, G. (1999). Psychosocial adjustment after traumatic brain injury: What are the important variables? *Psychological Medicine, 29*, 713-725.

Tateno, A., Jorge, R. E., & Robinson, R. G. (2014). Clinical correlates of aggressive behavior after traumatic brain injury. *The Journal of Neuropsychiatry and Clinical Neurosciences, 15*(2), 155–160.

The Defense and Veterans Brain Injury Center. (2010). Mild traumatic brain injury-concussion. Retrieved from http://www.publichealth.va.gov/docs/exposures/TBI-pocketcard.pdf

Thompson, L., Thompson, M., & Reid, A. (2010). Neurofeedback outcomes in clients with Asperger's syndrome. *Applied Psychophysiology & Biofeedback, 35*, 63-81.

Vohs, K. D., & Baumeister, R. F. (Eds.). (2011). *Handbook of self-regulation: Research, theory, and applications*. Guilford Press.

Walls, T. A., & Schafer, J. L. (Eds.). (2005). *Models for intensive longitudinal data*. Oxford University Press.

Weafer, J., Milich, R., & Fillmore, M. T. (2011). Behavioral components of impulsivity predict alcohol consumption in adults with ADHD and healthy controls. *Drug and Alcohol Dependence*, *113*(2), 139-146.

Wilhelm, F. H., Pfaltz, M. C., & Grossman, P. (2006). Continuous electronic data capture of physiology, behavior and experience in real life: Towards ecological momentary assessment of emotion. *Interacting with Computers*, *18*(2), 171-186.

Winkler, D., Unsworth, C., & Sloan, S. (2006). Factors that lead to successful community integration following severe traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *21*(1), 8-21.

Ylvisaker, M., Turkstra, L., Coehlo, C., Yorkston, K., Kennedy, M., Sohlberg, M., & Avery, J. (2007). Behavioural interventions for children and adults with behaviour disorders after TBI: A systematic review of the evidence. *Brain Injury*, *21*(8), 769-805.

Yoshiuchi, K., Yamamoto, Y., & Akabayashi, A. (2008). Application of ecological momentary assessment in stress-related diseases. *BioPsychoSocial Medicine*, *2*(13), 1-6.

Yucha, C., & Gilbert, C. (2004). Evidence-based practice in biofeedback and neurofeedback. Retrieved from http://www.aapb.org/files/public/Yucha-Gilbert_EvidenceBased2004.pdf

Zimmerman, P. H., Bolhuis, J. E., Willemsen, A., Meyer, E. S., & Noldus, L. P. (2009). The Observer XT: A tool for the integration and synchronization of multimodal signals. *Behavior Research Methods*, *41*(3), 731-735.