CLASSIFICATION ACCURACY OF THE WISCONSIN CARD SORTING TEST (WCST) IN DETECTING NONCREDIBLE COGNITIVE PERFORMANCE IN NEUROPSYCHOLOGICAL OUTPATIENTS

A DISSERTATION SUBMITTED TO THE FACULTY OF

THE ADLER SCHOOL OF PROFESSIONAL PSYCHOLOGY

BY

LYDIA WARDIN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PSYCHOLOGY

CHICAGO, IL July 3, 2014

Committee Page

Dissertation Chair:

Douglas Whiteside, PhD, ABPP-CN
Professor of Psychology
Neuropsychology Concentration Coordinator
Board Certified Clinical Neuropsychologist
Adler School of Professional Psychology
17 North Dearborn
Chicago, Illinois 60602
(312) 662 – 4346

Dissertation Committee:

Linda Rice, PhD, ABPP-CN
Assistant Professor of Psychology
Board Certified Clinical Neuropsychologist
Adler School of Professional Psychology
17 North Dearborn
Chicago, Illinois 60602
(312) 662 – 4361

Lauren Nichols, PsyD
Professor of Psychology
Child & Adolescent Psychology Track Coordinator
Adler School of Professional Psychology
17 N Dearborn
Chicago, Illinois 60602
(312) 662 – 4356

Abstract

The current study sought to examine the classification accuracy of the Wisconsin Card Sorting Test (WCST) as an embedded performance symptom validity test (PVT) among three different samples. The sample of interest included 110 participants with mild traumatic brain injury (mTBI). For comparison, the study included 69 participants with moderate to severe traumatic brain injury (TBI) (STBI) and 155 non-neurological patients with mixed depression/anxiety psychiatric diagnoses (PSYCH). Furthermore, a logistically derived combination of Number Correct, Perseverative Responses, and Trials to First Category was created as an additional predictor variable named WCSTCOMB. Results indicated significant group differences between the credible performance (PASS) and non-credible performance (FAIL) groups for the PSYCH sample in the following variables: Number of Trials, Total Errors, Perseverative Errors, Perseverative Responses, Number of Categories Completed, and WCSTCOMB. Significant group differences in the STBI sample were found in the following variables: Number of Trials, Total Errors, Perseverative Errors, Perseverative Responses, Number of Categories Completed, and WCSTCOMB. The study found no significant differences in the WCST variables between the PASS and FAIL groups in the mTBI sample. Receiver operating characteristic (ROC) analysis found that Perseverative Responses and WCSTCOMB had acceptable classification accuracy of at least .70 in the PSCYH group. In the

STBI group, ROC analysis found that the following WCST variables had at least acceptable classification accuracy of at least .70 for the following: Number of Trials, Total Errors, Perseverative Errors, Perseverative Responses, Number of Categories Completed, and WCSTCOMB. Sensitivity, specificity, positive predictive power, negative predictive power, and recommended raw cutoff scores were provided for WCST variables with acceptable classification accuracy. Acceptable classification accuracy was not found among any WCST variables in the mTBI sample. Results do not provide support for the WCST as an embedded PVT among those with mTBI. However, results did provide support for the WCST as embedded PVTs with populations with moderate to severe TBI and depressed/anxious outpatients.

Acknowledgements

I would like to thank several people for their assistance in the completion of this study. I would like to first thank the individuals who assisted in planning, conceptualizing, and editing this research project. These individuals include Dr. Douglas Whiteside, Dr. Linda Rice, and Dr. Lauren Nichols.

Furthermore, I would like to thank Dr. Douglas Whiteside for inspiring my interest in performance validity testing in neuropsychological evaluations and for all the support and assistance in this research project. I would also like to thank him for allowing the use of the study's data from his clinical database. Without his mentorship and guidance this dissertation would not have been possible.

Lydia Wardin

<u>EDUCATION</u>

Adler School of Professional Psychology: Chicago, IL September 2009–Present
Doctoral Program in Clinical Psychology-Student
APA Accredited Program
Concentration in Neuropsychology

Adler School of Professional Psychology: Chicago, IL
Dissertation Title: ***The Wisconsin Card Sorting Test as an Embedded Measure of Effort***
Chair: Douglas Whiteside, PhD, ABPP-CN

University of Iowa: Iowa City, IA August 2005–May 2009
Bachelor of Sciences in Psychology
Minor in Human Relations
The University of Iowa Deans List (2006–2009)

Member of Gamma Sigma Alpha, National Greek Academic Honor Society (2008–2009)

<u>SUPERVISED PRACTICUM EXPERIENCE</u>

Adler Community Health Services (ACHS), Chicago, IL
September 2013–June 2014

<u>Neuropsychology Extern</u>
Supervising Neuropsychologists: Douglas Whiteside, PhD, ABPP; Linda Rice, PhD, ABPP

- Conducted outpatient neuropsychological evaluations for children, adolescents, adults, and geriatric patients from varied socioeconomic, ethnic, and cultural backgrounds.
- Answered referrals that include the assessment of differential diagnosis, treatment recommendations for traumatic brain injury, learning disorders/ADHD, seizure disorders, psychiatric disorders, neurological disorders, and dementia.
- Administered, scored, and interpreted neuropsychological test batteries. Conducted feedback sessions, and wrote integrated neuropsychological reports under supervision with a one-week turnaround time.

- Maintained administration responsibilities including scheduling, coordination of care, and managing client financial issues.
- Contacted, networked, and developed relationships with ACHS referral sources.
- Researched and developed a referral source list for potential neuropsychological referral sources in the Chicagoland area.
- Responsible for collecting and entering research data into database from neuropsychological evaluations.
- Assisted in drafting HIPAA forms and inform consents for the ACHS neuropsychology program.
- Created the neuropsychology program brochure for ACHS.

University of Illinois at Chicago (UIC) May 2012–June 2013
Neurology and Rehabilitation, Chicago IL
Advanced Doctoral Clinical Neuropsychology Extern
Supervising Neuropsychologist: Linda Laatsch, PhD

- Provided outpatient neuropsychological evaluations for children, adolescents, adults, and geriatric patients.
- Provided psychological and neuropsychological evaluations for patients on a 20-bed acute inpatient rehabilitation unit.
- Received referral sources from physical medicine, neurology, psychiatry, neurosurgery, somnology, oncology, social work, speech therapy, and occupational therapy. Common referral issues included assessment of dementia, cerebrovascular accidents, traumatic brain injuries, mental retardation, seizure disorders, psychiatric issues, brain tumors, attention deficit disorder, and learning disorders.
- Conducted psychotherapy and taught psychoeducation, relaxation techniques, and compensation strategies to patients on inpatient rehabilitation unit. Typical treatment issues included: stroke rehabilitation, spinal cord injury rehabilitation, traumatic brain injury, multiple sclerosis, arthritis, loss of limbs, cardiac and pulmonary patients, status post brain tumor resections, and psychiatric issues.
- Worked as a part of a multidisciplinary team, consisting of nurses, physicians, physical therapists, occupational therapists, nutritionists, social workers, and speech pathologists.

- Attended biweekly multidisciplinary team meetings.
- Provided cognitive rehabilitation to adult and child patients. Helped patients cope and strategize ways to adapt, improve, or compensate for their cognitive changes. Served as a motivator and coach to encourage patients during their rehabilitation process. Worked weekly with patients to improve areas of weaknesses such as attention, processing speed, memory, executive functioning, and social comportment. Used manualized treatment and cognitive rehabilitation software to teach clients skills.
- Participated in weekly neuroanatomy seminars and neuropsychology grand rounds.
- Engaged in extensive supervision in interviewing clients and their family members, providing feedback sessions to clients and their family members, report writing, conceptualization, and providing cognitive rehabilitation.

Neuropsychology Practice, Chicago, IL March 2012–August 2013
Neuropsychology Extern
Supervising Neuropsychologists: Douglas Whiteside, PhD, ABPP-CN; Linda Rice, PhD, ABPP-CN

- Provided outpatient neuropsychological evaluations in a private practice setting for children, adolescents, and adults.
- Received diverse referral issues including head injury, learning disorder, seizure disorder, attention deficit disorder, and cerebrovascular accidents.
- Provided peer supervision to students conducting neuropsychological evaluations.
- Conducted disability evaluations for the social security administration. Typical referral questions included assessment of level of functioning, ability to return to work, differential diagnosis, and ability to care for self.
- Engaged in supervision in interviewing, providing feedback, report writing, and conceptualization.
- Provided peer supervision to less advanced neuropsychology students and assisted them with administering tests, interpreting data, and writing neuropsychological reports.
- Assisted in getting the private practice organized by orchestrating the organization of testing materials, and printing and filing all paperwork for

the practice (informed consents, interview forms, fax cover sheets, HIPAA notice, authorization for release of records, testimony and deposition agreement, and intake forms).

Will County Health Department, Joliet, IL September 2011–May 2012
Doctoral Psychotherapy Extern
Supervising Psychologist: Rita Gray, PsyD

- Provided weekly individual therapy, family therapy, case management, and medication monitoring for children, adolescents, and adults in a community mental health setting.
- Received referral sources from parents, probation officers, psychiatrists, social workers, case managers, and primary physicians.
- Worked with clients with presenting problems that included anger, gang involvement, learning disorders, substance abuse, depression, anxiety, bipolar disorder, attentional difficulties, impulsivity, family discord, conduct disorder, trauma, and abuse or neglect.
- Collaborated with teachers, psychiatrists, and probation officers regarding treatment planning and diagnostic clarification.
- Administered mental health assessments, lifestyle assessments, and mental status examinations.
- Attended and presented at weekly case presentations regarding assessment and therapy cases.
- Created a seven-week anger management program for adolescent males on probation.

John J. Madden Mental Health Center, Hines, IL July 2010–June 2011
Doctoral Diagnostic Extern
Supervising Psychologist: Teresa Neeliyara, PhD

- Worked in fast-paced inpatient psychiatric hospital setting, which treated those who were in acute crisis and those with chronic mental illness.
- Completed assessment batteries for personality, cognitive, neuropsychological, and intellectual disability referrals. Collected, interpreted, and integrated all data into a comprehensive report with a week turnover deadline.

- Worked on an inpatient psychiatric unit with acutely disturbed patients suffering from a variety of disorders such as: psychosis, personality disorders, bipolar disorder, posttraumatic stress disorder, anxiety, depression, schizophrenia, substance abuse, attention deficit disorder, learning disorders, and dementia.
- Facilitated daily community meetings with patients to inform and encourage them to attend daily groups, to discuss issues they had regarding living on the unit, to teach coping skills, and to provide psychoeducation regarding their treatment options and psychiatric disorder.
- Collaborated and consulted at morning treatment meetings with a team of psychiatrists, psychologists, medical students, nurses, and social workers to discuss patient's needs, priorities, discharge status, and diagnoses.
- Attended grand rounds regarding psychiatric diagnoses, co-morbid diagnoses, psychological interventions, psychopharmacology, and neurological conditions.
- Participated in weekly trainings to further build upon diagnostic, professional, and clinical skills.

RESEARCH EXPERIENCE

Adler School of Professional Psychology, Chicago, IL Sept. 2012–May 2013
Manuscript Reviewer for the *Journal of International Neuropsychological Society* and *The Clinical Neuropsychologist*
Supervisor: Douglas Whiteside, PhD, ABPP-CN

- Reviewed research manuscripts under supervision.
- Provided written critical reviews on the grammar, design, methodology, statistics, and conclusions of the study.

Adler School of Professional Psychology, Chicago, IL October 2011–Present
Neuropsychology Research Lab
Lead Research Assistant
Supervisor: Douglas Whiteside, PhD, ABPP-CN

- Assist in managing and expanding a large neuropsychological database comprised of children and adult patients.

- Delegate research tasks to the research team and assist in determining who will have access to the database.
- Supervise student access to the database, organize times for students to do data entry, and provide requested datasets for student's research projects.
- Recruit new students to become involved in the database, and assign them to current research projects based on their research interests.
- Organize, coordinate, and prepare research meetings.
- Work to merge two large Excel databases into one functional comprehensive SPSS database.
- Search, identify, and correct errors and incorrect entries in the database.

The University of Iowa, Iowa City, IA September 2008–May 2009
Research Assistant
Structure of Psychopathology and Personality, and Assessment Construction
Supervisor: David Watson, PhD

- Performed literature searches on various personality theories and assessments and completed data entry on scale development projects.
- Obtained consent, proctored studies, and debriefed participants on a bi-weekly basis.
- Maintained confidentiality of data and anonymity of participants.
- Reported, recorded, and solved complications during experimental sessions.
- Recruited and answered questions for participants in a clinical community sample.

Publications

- Whiteside, D., Kogan, J., **Wardin, L.**, Phillips, D., Franzwa, M.G., Basso, M., & Roper, B. (under review). Boston Naming Test and Controlled Oral Word Association Test as Embedded Performance Validity Measures in a Mild Traumatic Brain Injury Sample. Manuscript submitted to *The Clinical Neuropsychologist.*
- **Wardin, L.** (2012). Homeless Women's Utilization of Health Care. *Praxis*. (Student publication of Adler School of Professional Psychology).

Presentations

- **Wardin, L**., Whiteside, D.M., Kogan, J., Widmann, G., Luu, H., Roper, B., & Basso, M. (2014, February*). Digit span total raw score as an embedded performance validity test.* Poster presented at the 42nd annual meeting of the International Neuropsychological Society (INS): Seattle, WA.

- Franzwa, M.G., Whiteside, D.M., Kogan, J., **Wardin, L**., & Phillips, D. (2014, June). *Boston Naming Test and Controlled Oral Word Association Test as performance validity tests in psychiatric and severe TBI samples.* Poster presented at the 12th annual conference of American Academy of Clinical Neuropsychology (AACN): New York, NY.

- Luu, H., Whiteside, D.M., Kerce, V., Granite, L., **Wardin, L**., Carns, D., Rice, L., Roper, B., & Basso, M. (2014, June). *Motor functioning tests as embedded performance validity measures in females.* Poster presented at the 12th annual conference of American Academy of Clinical Neuropsychology (AACN): New York, NY

- Luu, H., Whiteside, D., Granite, **L., Wardin**, L., Kerce, V., Rice, L., Roper, B., & Basso, M. (2014, February). *Motor functioning tests as embedded performance validity measure in males.* Poster presented at the 42nd annual meeting of the International Neuropsychological Society (INS): Seattle, WA.

- **Wardin, L**., Whiteside, D., Phillips, D., Franzwa, G., Kogan, J., Backus, A., & Luu, A. (2013, June). *Controlled Oral Word Association Test as an embedded symptom validity test.* Poster session presented at the American Academy of Clinical Neuropsychology (AACN) conference, Chicago, IL.

- Kogan, J., Whiteside, D., Franzwa, G., Phillips, D., **Wardin, L.,** Parikh, S., Vogler, M. (2013, June). *The Boston Naming Test as an embedded symptom validity measure*. Poster session presented at the American Academy of Clinical Neuropsychology (AACN) conference, Chicago, IL.

- **Wardin, L.**, Whiteside, D., Widmann, G., Granite, L., Ma, T., Kogan, J., & Basso, M. (2013, February). *Delayed recall trials of the California Verbal Learning Test II as embedded measures of effort.* Poster session presented at the International Neuropsychological Society (INS) conference, Waikoloa, HI.

- Hahn-Ketter, A., **Wardin, L.**, Kogan, J., & Whiteside, D. (2013, February). *Analysis of performance on test effort in dementia population.* Poster session presented at the International Neuropsychological Society (INS) conference, Waikoloa, HI.

- **Wardin, L.**, Whiteside, D., Widmann, G., Ma, T., & Granite, L. (2012). More Than Verbal Learning and Memory: An Embedded Effort Evaluation of the California Verbal Learning Test II (Abstract). *The Clinical Neuropsychologist, 26,* 438.

- **Wardin, L.**, Whiteside, D., Widmann, G., & Ma, T. (2012, June). *More than verbal learning and memory: An embedded effort evaluation of the California Verbal Learning Test II.* Poster session presented at the American Academy of Clinical Neuropsychology (AACN) conference, Seattle, WA.

- **Wardin, L.** (2010, June). *Community Health.* Poster session presented at the Adler School of Professional Psychology Community Service Practicum Fair, Chicago, IL.

TEACHING EXPERIENCE

Adler School of Professional Psychology, Chicago, IL — Fall 2013
Teaching Assistant, Cognitive and Intellectual Assessment
Supervisor: Linda Laatsch, PhD

- Co-teach class lectures on cognitive assessment, report writing, interviewing, and administration of cognitive measures.
- Teach selected lectures on cognitive assessment.

- Grade student's reports and examinations.
- Assess student's proficiency examinations on the WAIS-IV, WISC-IV, WRAT4, and WIAT-III.

Adler School of Professional Psychology, Chicago, IL Fall 2012 and Fall 2013
Teaching Assistant, Pediatric Neuropsychological Assessment
Supervisor: Linda Laatsch, PhD

- Prepare and organize class readings and pediatric neuropsychological tests for each class.
- Teach selected lectures on pediatric neuropsychology.
- Consult with the course professor regarding course pacing, which includes discussing load of assignments and weekly readings, and addressing any concerns related to student's progress.
- Collaborate with students and develop with the professor effective teaching strategies for presenting the course material to ensure optimal learning.

Adler School of Professional Psychology, Chicago, IL Spring 2013
Teaching Assistant, Biological Bases of Brain and Behavior
Supervisor: Linda Laatsch, PhD

- Consulted with course professor regarding course pacing, readings, and construction of examinations.
- Graded student's examinations and provided study review sessions.

Adler School of Professional Psychology, Chicago, IL
Fall 2011 and Spring 2012
Teaching Assistant, Intermediate and Advanced Neuropsychological Assessment
Supervisor: Douglas Whiteside, PhD, ABPP-CN

- Conducted weekly hour lab sessions with second and third year doctoral students.
- Taught and led discussions on aspects of neuropsychological assessment including assessment tools, administration, and interpretation of neuropsychological batteries.

- Created weekly PowerPoint presentations on neuropsychological case examples including presenting problem, background, interview data, testing data, results, diagnoses, and recommendations.
- Reviewed and provided feedback on neuropsychological reports written by students, and graded examinations.
- Met with students individually to assist with interpretation of assessment batteries, conceptualization of class cases, selecting research projects, and familiarizing students with the neuropsychological database at Adler.
- Held optional study sessions for midterm and final exam preparation.
- Consulted with the course instructor regarding effective ways to present course content to students.

COMMUNITY ENGAGEMENT EXPERIENCE

Adler School of Professional Psychology, Chicago, IL May 2011–June 2011
Parenting Teacher
Supervisor: Christine Beecroft, PsyD

- Created a six-week parenting course for parents of school-aged children, which included selecting and developing reading material for parents, publicizing a parent educational program, recruiting parent participants, and designing an outcome evaluation.
- Co-led parenting group in the community and served as a parent educator to answer parents' questions about their children's behavior and corrective strategies.
- Educated and consulted with parents on Adlerian parenting with focus on the following: understanding goals of children's misbehaviors, steps to redirect misbehavior, encouragements, use of natural and logical consequences, and utility of punishment.

Community Health, Chicago, IL January 2010–May 2010
Community Service Practicum Extern
Supervisor: Ornella Razetto, LCSW

- Worked on building resources and referral systems for medical practitioners and patients diagnosed with chronic illnesses such as:

epilepsy, multiple sclerosis, fibromyalgia, hypothyroidism, osteoarthritis, cancer, and hepatitis C.

- Researched multicultural services that provided free to low-cost care for the uninsured.
- Created information guides that explained the steps and requirements needed in order to receive Social Security Income, Medicare, and Social Security Disability Insurance.
- Built and strengthened relationships between Community Health and outside organizations and agencies.
- Developed and organized a CBT oriented 10-week module of group therapy for Latino males with anger management problems.
- Constructed an anger management screening tool as well as pre and post assessments to measure participant's progress.

Cumulative hours: 250

LEADERSHIP EXPERIENCE

Adler School of Professional Psychology, Chicago, IL

September 2012–January 2013

Student Representative for the PsyD Curriculum Committee

Supervisor: David Katz, PhD, ABPP-CN

- Nominated by faculty for the student representative position for the Curriculum Committee for the clinical psychology doctoral program.
- Assisted in program development and determined mandatory core courses for future PsyD students.
- Met weekly with clinical PsyD program director and faculty to provide an ongoing evaluation and development of curriculum, and provided recommendations for updating core courses.
- Advocated for students' needs and concerns regarding proposals for new courses, policies, development and revision of course sequence, and curriculum innovations.
- Communicated changes in curriculum to the Adler School of Professional Psychology Student Government.

PROFESSIONAL MEMBERSHIPS

American Psychological Association (APA) Division 40, Student Member
American Academy of Clinical Neuropsychology (AACN), Student Member
International Neuropsychological Society (INS), Student Member
Association of Neuropsychology Students in Training (ANST), Member

Table of Contents

List of Tables

Chapter 1: Introduction

An integral aspect of neuropsychological evaluations is assessing the level of effort a client puts forth in an examination. This is crucial as effort directly impacts the validity of an assessment regardless of the reason for the lack of effort (Boone, 2009). Unfortunately, traditional neuropsychological assessments are susceptible to misinterpretation since these measures rely on the patient's willingness to put forth their best effort (Heubrock & Petermann, 1998). Green, Rohling, Lees-Haley, & Allen (2001). Therefore, there has been a push for neuropsychologists to use standardized and validated quantifiable measures in order to better detect malingering and or poor effort (Mittenburg, Rotholc, Russell, & Heilbronner, 1996). In particular, referrals for neuropsychological assessment often involved personal injury litigation, social security disability, academic accommodations, worker's compensation, or other compensation seeking situations. Mittenberg, Patton, Canyock, and Condit (2002) surveyed the American Board of Clinical Neuropsychology (AACN) membership about symptom exaggeration and possible malingering and found that 29% of personal injury, 30% of disability, 19% of criminal, and 8% of medical cases involved probable malingering and symptom exaggeration. As a result, neuropsychologists frequently use stand-alone forced-choice recognition PVTs with standard cutoff scores to detect non-credible performance. For example, the Test of Memory Malingering (TOMM) (Tombaugh, 1996) is the second most commonly used malingering instrument among forensic psychologists right after the Structured Interview of Reported Symptoms (SIRS) (Archer, Buffington-Vollum, Stredny, & Handel, 2006).

However, there are downsides to using freestanding forced choice performance validity tests (PVTs). This is because freestanding PVTs can be lengthy and adds to the overall assessment time and cost of the evaluation (Suhr & Gunstad, 2000). In addition, a survey (Wetter & Corrigan, 1995) discovered that 50% of lawyers thought it was appropriate to inform their clients about the nature of testing and validity scales on particular measures. Lawyers are familiar with many of these forced choice recognition tasks and some lawyers coach their clients to ensure that their clients pass these freestanding PVTs (Ashendorf, O'Bryant, & McCaffrey, 2003). In addition, Ruiz, Drake, Glass, Marcotte, and van Gorp (2002), found websites that could negatively impact the validity of PVTs in clinical practice. For example, the authors found websites that published administration instructions for instruments with information on how to alter test performance. Therefore, a client wishing to intentionally distort his or her results could have access to information on how to successfully do so.

Boone (2009) discussed the importance of continuous assessment of effort throughout the entire neuropsychological examination involving multiple forced choice and non-forced choice PVTs. Specifically, Boone (2009) reported that an individual's effort tends to vary throughout testing and stated, "response bias typically fluctuates across an exam due to differences in feigning strategies, and that without continuous monitoring of effort and use of effort indices tapping multiple skills, some non-credible performances will not be identified" (p. 730). Therefore, multiple PVTs are needed throughout multiple neuropsychological domains. Having multiple measures of effort can also allow the examiner to detect a pattern of performance that is more difficult to feign (Boone, 2009; Larrabee, & Berry, 2007; Suhr & Gunstad, 2000).

The use of multiple embedded PVTs is a way to address instances of coaching that threatens test security from information on the Internet and clients whose level of motivation and effort waivers throughout the evaluation. Validated embedded PVTs are likely more difficult to coach and therefore more difficult for the client to feign (Greve, Bianchini, Mathias, Houston, & Crouch, 2002). This is because it is more difficult to feign a pattern across testing than rather on one specific measure. Therefore, there has been an increased need for valid and reliable embedded PVT's and consequently more research is needed.

Statement of Problem

More research is needed to find additional PVTs that accurately identify credible versus non-credible performance. Specifically, it is important to find PVTs that are empirically validated among those with mild traumatic brain injury (mTBI). This is because mTBI was the most common diagnosis associated with non-credible performance (Mittenberg et al., 2002). In addition, Mittenberg et al.

(2002) found that approximately 40% of mild traumatic brain injured individuals in litigation "demonstrate malingered impairment and 5% are actually impaired, the base rate of probable malingering in cases with apparent cognitive impairment is approximately 88%" (pp. 1099). This is interesting because a meta-analysis by Binder, Rohling, and Larrabee (1997) found that 95% of individuals with mTBI showed no significant cognitive impairments three months and after the injury. Moreover, Binder (1997) found little evidence for neurological causes of most persisting complaints after an mTBI. Therefore, there is a paradoxical effect of those with mTBI involved with litigation.

Paradoxical effects of mTBI. Youngjohn, Davis, & Wolfe (1997) coined the term, "paradoxical effects," in relation to mTBI after discovering the general tendency for individuals with mTBI who exaggerate or embellish psychological or emotional functioning more appear more severely impaired compared to individuals with moderate to severe TBI. Furthermore, a meta-analysis by Binder and Rohling (1996) examined the associated between financial compensation and the experience and treatment of chronic pain. Binder and Rohing (1996) examined 157 studies totaling 3,802 pain patients and 3,849 controls. The results indicated that compensation is related to increased reports of pain and decreased reports of treatment efficacy. This same paradoxical severity can also be found in chronic mTBI individuals when testing credible performance (Carone & Bush, 2013; Greiffenstein & Baker, 2006; Thomas & Youngjohn, 2009).

Greiffenstein and Baker (2006) examined the paradoxical effect in mTBI and TOMM performance. The authors used archival data from 759 head-injured claimants who were divided into symptomatic mTBI (n = 607) and moderate to severe TBI (n = 152) groups. The results found this paradoxical or inverse dose response effect where individuals with milder brain injuries performed worse on motor and effort measures. Specifically, individuals with mTBI performed significantly lower on grip strength and the TOMM compared to individuals with moderate to severe TBI. Therefore, non-credible performance should be thoroughly assessed, especially in individuals with mTBI,

Sample of interest. This present study's primary sample is mild traumatic brain injury (mTBI) because of the reported high frequency of referrals in practice and high base rate of non-credible performance. Also, there has been extensive research on PVTs in this population. For comparison, a moderate to severe TBI (STBI) sample and a mixed depression/anxiety psychiatric diagnosis sample (PSYCH) will be used. These comparison groups will help determine the generalizability of the WCST as an embedded PVT in both neurological and psychiatric populations.

The literature has advocated for the need of continuous assessment of non-credible performance throughout an evaluation in multiple neuropsychological domains (Boone, 2009; Larrabee & Berry, 2007). Much of the current literature on PVTs has focused on memory measures (Suhr & Boyer, 1999). In contrast to memory PVTs, little research in comparison has explored executive functioning PVTs (Greve & Bianchini, 2002). Ideally an embedded executive functioning PVT would include a frequently used measure, such as the Wisconsin Card Sorting Test (Heaton, Chelune, Talley, Kay, & Curtiss, 1993).

The Wisconsin Card Sorting Test (WCST) is a standardized and validated test of executive functioning (Eling, Derckx, & Maes, 2008). A comprehensive test usage survey of neuropsychologists found that the WCST is the most commonly used test of executive functioning in the United States and Canada (Rabin, Barr, & Burton, 2005). There have been inconsistent findings in the literature regarding the utility, accuracy, and validity of using measures on the WCST as a PVT and the literature of the WCST as an embedded PVT will later be discussed (Greve, Heinly, Bianchini, & Love, 2009; King, Sweet, Sherer, Curtiss, & Vanderploeg, 2002). Therefore, more research is needed on the WCST as a PVT among specific neuropsychological populations.

Malingering versus non-credible performance. Much of the current literature conceptualizes non-credible performance as Malingered Neurocognitive Deficits (MND) articulated by Slick, Sherman, & Iverson (1999). Slick et al. (1999) operationalized MND as follows:

> Malingering of Neurocognitive Dysfunction (MND) is the volitional exaggeration or fabrication of cognitive dysfunction for the purpose of obtaining substantial material gain, or avoiding or escaping formal duty of

> responsibility. Substantial material gain includes money, goods, or services of nontrivial value (e.g., financial compensation for personal injury). Formal duties are actions that people are legally obligated to perform (e.g., prison, military, or public service, or child support payments or other financial obligations). Formal responsibilities are those that involve accountability or liability in legal proceedings (e.g., competency to stand trial). (p. 552).

MND includes several criterion including one substantial external incentive, definite response bias (below chance performance), and the behavior cannot fully be accounted for by psychological, neurological, or developmental factors. This approach implies that malingering is needed for non-credible performance to occur during testing. However, this is not always the case because non-credible performance does not always occur in the case of malingering. For example, Boone (2009) explained that non-credible performance can occur in the absence of external incentives and that cognitive effort fluctuates over the course of an evaluation. Boone (2009) advocated for a less strict evaluation of credible performance rather than the strict diagnosis of malingering. Therefore, this study will use non-credible performance to conceptualize the validity of performance rather than malingering due to issues raised by Boone (2009). For research purposes, non-credible performance will be defined as failing two or more empirically validated PVTs. Therefore, non-credible performance refers to failing

PVTs rather than malingering, which is consistent with Boone's 2009 recommendations.

Research Question. Based on the limitations in the above discussed literature, the following question was addressed: Are measures from the WCST effective as PVTs in detecting non-credible performance in both non-neurological (mTBI and PSYCH) and neurological (STBI) populations? The following hypotheses were generated to address the research question of whether the WCST is an effective embedded PVT among non-neurological and neurological populations.

Hypotheses

1. It was hypothesized that those who fail two or more previously established embedded or freestanding PVT's will perform significantly worse on the WCST.
2. It was hypothesized that the WCST had acceptable classification accuracy for detecting non-credible performance among adults with mTBI, and had sensitivity of at least 50% when specificity was set at 90%.
3. It was hypothesized that the WCST had acceptable classification accuracy for detecting non-credible performance among adults with depression/anxiety, and had sensitivity of at least 50% when specificity was set at 90%.

4. It was hypothesized that the WCST had classification accuracy for detecting non-credible performance among adults with moderate or severe traumatic brain injury, and had sensitivity of at least 50% when specificity was set at 90%.

Statement of Purpose

There has been relatively little research conducted on various measures of the WCST as embedded measures of effort (Suhr & Boyer, 1999). In fact, the research that has been conducted has found inconsistent results (Greve et al., 2009). As mentioned, non-credible performance is a concern and validated PVTs need to be utilized in clinical practice in order to accurately identify those who exhibit non-credible performance (Boone, 2009; Larrabee & Berry, 2007). Much of the current literature used a simulator design where college students were asked to simulate as if they were trying to exaggerate cognitive symptoms for compensation. Known-group design has also been utilized and will be used in this study. Known-group designs are considered to be the best approach because the sample utilizes clinical participants and consists of two criterion samples, a non-credible performance and credible performance group (Larrabee & Berry, 2007). Therefore, the present study provided a sample of middle-aged clinical adult participants (opposed to college-aged students). Finally, much of the current literature on the WCST as a PVT utilized only head trauma participants. This study will add to the literature by exploring the generalizability of the WCST as a PVT to other non-neurological populations with psychiatric diagnoses. The findings of this study will either provide support for or against using the WCST as a PVT among populations with mTBI, moderate/severe TBI, or anxiety/depression. Finally, this study will continue to develop new directions for related research on the WCST as an embedded PVT.

Assumptions and Limitations

It was an assumption of this study that all neuropsychological tests were administered in a standardized fashion by a trained and experienced examiner and that all participants understood the task directions. In addition, it was assumed that failing two or more PVTs indicated an individual's effort is non-credible.

The limitations of this study included the use of archival data and the inability to randomly assign participants into groups. The generalizability of the findings may have been limited to the demographics used in this study. For example, this study's sample consisted primarily of Caucasian participants.

Chapter II: Review of Literature

Heilbronner et al. (2009) highlighted that including performance validity testing (PVT) in neuropsychological evaluations is now the standard of care. Furthermore, Boone (2009) has advocated for continuous assessment of effort throughout a neuropsychological evaluation. Embedded PVTs are important, as they are less susceptible to lawyer coaching, less time consuming, and allow the examiner to interpret a pattern of effort throughout the neuropsychological evaluation (Ashendorf, O'Bryant, & McCaffery, 2003; Suhr & Gunstad, 2000; Wetter & Corrigan, 1995). Therefore, empirically validated measures of effort are needed in every domain of neuropsychological functioning (Boone, 2009). Specifically, there has been relatively little research on PVTs in the domain of executive functioning compared to other cognitive domains such as memory (Suhr & Boyer, 1999). Executive functioning deficits are well documented in traumatic brain injury and therefore performance in this area could determine compensation and accommodations (Larrabee, 2012). Therefore, there could be motive to feign or exaggerate deficits related to executive functioning (Larrabee, 2012). The following is a literature review related to the WCST as a PVT in neuropsychological evaluations.

Executive Functioning

Executive functioning has become an "umbrella term" and difficult to operationalize due to different definitions and theories (Goldstein, Naglieri, Princiotta, & Otero, 2014). Early theorists of executive functioning included Posner and Snyder (1975) who described executive functioning as "cognitive control". Posner and Snyder (1975) described that this process was part of the attentional system that focused and directed attention in order to guide behavior and control thought processes. Baddeley, Sala, & Robbins (1996) similarly described executive function as the "central executive" of the working memory system. The authors hypothesized that the "central executive" was a supervisory system that controls cognitive processes such as the phonological loop (auditory working memory), visuospatial sketchpad (visual working memory), and an episodic buffer (integrates information). The literature continues to associate executive functioning as a "supervisory" process (Alvarez & Emory, 2006).

More recently, Alvarez & Emory (2006) conceptualized executive functioning as a "supervisory" or "higher-level" cognitive functioning. Otero & Barker (2014) described executive function "as an umbrella term used for a diversity of hypothesized cognitive processes...they include planning, working memory, attention, inhibition, self-monitoring, self-regulation, and initiation" (p. 32). Lezak, Howieson, Bigler, and Tranel (2012) described executive functioning as "the intrinsic ability to respond in an adaptive manner to novel situations and are also the basis of many cognitive, emotional, and social skills" (p. 661). Therefore, executive functioning includes planning, organization, problem solving, flexible thinking, inhibiting behavior, and multitasking. Lezak et al. (2012)

conceptualized executive functioning into the four components of volition, planning, purposive action, and effective performance. Executive functioning coordinates multiple cognitive processes and therefore there are various neuroanatomical structures associated with executive functioning.

Physiological features of executive functioning. Executive functioning is traditionally conceptualized as resulting from frontal lobe damage (Lezak et al., 2012). The frontal lobes are involved with executive functioning tasks, particularly with integrating emotions and cognitions (Otero & Baker, 2014; Stuss & Alexander, 2007). This is made possible due to the multiple connections the frontal lobes have throughout the brain (Otero & Baker, 2014). Otero & Baker (2014) described the frontal lobes role in executive functioning as a "driver of a car" and stated, "the complex action of driving a car cannot be done without many different components interacting together…it is through a complex interaction of all of the parts of the car being controlled by the driver that the action of driving the car can occur" (p. 30). Therefore, the frontal lobes are not solely responsible for executive functioning. The literature has also reported that executive dysfunction can result from other parts of the brain such as damage from subcortical structures (Alvarez & Emory, 2006; Chung, Weyandt, & Swentosky, 2014; Lezak et al., 2012). The literature has reasoned that this is due to the fontal-subcortical circuitry so that the frontal lobes have many connections between the cortical and subcortical regions (Alvarez & Emory, 2006). This has been supported by neuroimaging studies using magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and positron emission tomography (PET) technology (Chung et al., 2014). Many studies have supported that the prefrontal and parietal cortices are involved with executive functioning (Chung et al., 2014). In the next section, neuroimaging findings regarding planning, working memory, responses inhibition, and set shifting will be discussed.

Neuroimaging studies on executive functioning. Planning can be conceptualized as judgment and decision-making based of ones own behaviors and the behaviors of others (Chung et al., 2014). Research has consistently found increased activation in the dorsolateral prefrontal cortex during planning tasks (Chunget et al., 2014). For example, a study by Newman, Carpenter, Varma, & Just (2003) used fMRI technology to examine the physiology of planning. Participants were compared using the Tower of London (TOL), which is a neuropsychological measure known to measure planning ability. The study found activation in the dorsolateral prefrontal cortex. Specifically, the fMRI results suggested that right prefrontal area might be involved with generating a plan. Second, the left prefrontal cortex may be involved in plan execution.

Working memory. Working memory is often conceptualized as the ability to temporarily store and manipulate information (Baddeley, 1992). Furthermore, Bledowski, Kaiser, & Rahm (2010) described working memory as " the ability to keep information active for further use, while allowing it to be prioritized, modified and protected from interference" (p. 1). Bledowski et al., 2010 explored the functional neuroimaging of working memory. The authors reviewed fMRI literature in working memory tasks and found that the dorsolateral and parietal regions activated during working memory tasks. The neuroanatomical structures of inhibition in executive functioning has also been explored.

***Inhibition*.** Inhibition refers to an executive functioning process of rejecting an automatic response in a situation (Goldman-Rakic, Thierry, Glowinski, & Christen, 1994). The literature has found different regions of the brain responsible for inhibition (Chung et al., 2014). Chung et al. (2014) speculated that this is because inhibition can be conceptualized as motor, cognitive, motivational, or automatic inhibition. Therefore, more research is needed in order to outline the different brain regions associated with the different types of inhibitory processes. Aron, Robins, & Poldrack (2004) explained that inhibition is typically associated with the dorsolateral prefrontal cortex, inferior frontal cortex, or orbital frontal cortex. The authors also proposed that the region of the right inferior frontal cortex is central for inhibiting response tendencies. In addition to inhibition, set-shifting neuroanatomical structures have been researched.

***Set-Shifting*.** Set-shifting is typically referred to the ability to flexibly switch back and forth between tasks (Miyake et al., 2000). Chung et al. (2014)

indicated that several brain regions are responsible for set-shifting tasks. In one study, Zakzanis, Mraz, & Graham (2005) explored set-shifting and brain related regions by using the Trail Making Test (TMT) and recording their brain activity using fMRI. Findings suggested that set-shifting abilities on the TMT are not exclusive to the frontal lobe. In addition to frontal regions in the left hemisphere, the left middle and superior temporal gyrus were also activated.

In conclusion, executive functioning is responsible for coordinating multiple cognitive processes. Generally, executive functioning is an umbrella term that includes several processes such as set-shifting (cognitive flexibility), problem solving, planning, working memory, inhibition, and emotional regulation (Goldstein et al., 2014; Lezak, et al., 2012). Moreover, executive functioning is not exclusively associated with the frontal lobes, but rather multiple regions of the brain are required to carryout related executive functioning tasks (Chung et al., 2014; Newman et al., 2003). Neuropsychological measures are currently used to measure executive functioning abilities. The literature has explored various neuropsychological tests that are designed to measure executive functioning. The WCST is a well-established measure of executive functioning that has been known to measure cognitive flexibility, problem-solving, and response maintenance (Greve, Bianchini, et al., 2002).

The Wisconsin Card Sorting Test (WCST)

The Wisconsin Card Sorting Test (WCST) is a commonly used clinical measure of executive functioning in neuropsychological assessment (Heaton et al., 1993). Strauss, Sherman, and Spreen (2006) described the purpose of the WCST as " to assess the ability to form abstract concepts, to shift and maintain set, and to utilize feedback" (p. 526). The WCST consists of 128 cards and four stimulus cards of various forms (crosses, circles, triangles, or stars), colors (blue, red, yellow, or green) and number of figures (one, two, three, or four). The client is instructed to match the deck of cards to the target cards. The examiner provides feedback of "correct" or "incorrect". Without warning, the examiner changes the sorting principle and the client must learn from the examiners feedback. The test is discontinued when either the examinee completes all 6 matching principles or finishes the deck of cards. The total administration time is around 10 to 15 minutes (Strauss et al., 2006). The normative sample for the WCST was created from 899 participants of six distinct samples (453 children and adolescents, 49 adults, 150 adults, 50 older adults, 124 commercial airline pilots, and 73 adults) (Heaton et al., 1993). The adult sample consisted of 384 participants that were of 20 years of age and older (M = 49.89, SD = 17.94). The mean education level was 14.95 (SD = 2.97). Overall, studies have confirmed that the WCST is a valid measure of executive functioning in the neuropsychological population (Heaton et al., 1993). This test is known to measure cognitive flexibility, problem solving, strategy, planning, and utilizing environmental feedback to shift cognitive sets

(Heaton et al., 1993). Factor analysis results indicated evidence that the WCST consists of the following three factors: cognitive flexibility, problem-solving, and response maintenance (Greve, Bianchini, et al., 2002).

Neuroanatomical structures associated with the WCST. According to Heaton et al. (1993), "the WCST is often referred to as a measure of "frontal" or "prefrontal" functioning" (pg. 1). However, Heaton et al. (1993) note that this is an "oversimplification" and stated "the frontal lobes are highly complex structures and subserve a far wider variety of cognitive functioning than those assessed by the WCST alone...and that any medical or psychological disorder that disrupts executive functions, in part or in whole, can results in impaired performance on the WCST" (p. 1). Also, Alvarez and Emory (2006) found adults with frontal lobe lesions performed significantly worse compared to healthy controls. However, the remaining 6 studies, found that participants with frontal lesions did not perform significantly different on the WCST compared to either healthy controls or diffuse or basal ganglia regions. Therefore, the authors speculated that the WCST is sensitive to frontal lobe lesions. However, the WCST performance is not specific marker of frontal lobe damage. Therefore, when one is administered the WCST, both frontal and non-frontal brain regions are activated. Nyhus and Barceló (2009) also echoed that the WCST fails to discriminate between frontal and non-frontal lesions in their critical review of the literature.

While the WCST is not specific to frontal and non-frontal brain regions, neuroimaging studies have found prefrontal cortex activation during WCST performance (Barceló & Knight, 2002; Hashimoto, Uruma, & Abo, 2008; Nyhus & Barceló, 2009; Sumitani et al., 2006). Milner's (1963) study had a large influence on viewing the WCST as a test of fontal lobe functioning (Larrabee & Berry, 2007). Milner (1963) explored the differences in WCST performance between adults with dorsolateral frontal excisions compared to individuals with orbitofrontal and posterior lesions. The study found that patients with frontal excisions performed significantly worse on the WCST compared to individuals with other brain excisions. Milner (1963) concluded that the ability to shift categories as a frontal lobe ability. The author concluded that frontal brain damage impairs the ability to shift and results in higher perseverations on the WCST. More recent studies have also explored brain regions associated with WCST performance.

A study by Sumitani et al. (2006) examined the blood oxygenation changes of 32 adult participants during the administration of the WCST. The study supported that the prefrontal cortex is activated during the WCST. Furthermore, the authors found that 20 healthy adult participants had a bilateral increase in oxygenated hemoglobin and 7 participants showed unilateral oxygenated hemoglobin (5 left sided and 3 right sided) in the prefrontal cortex.

Furthermore, Alvarez and Emory (2006) performed a qualitative review of brain activation during the WCST. Several studies supported evidence of activation in the dorsolateral prefrontal cortex during WCST performance (Berman et al., 1995; Haines et al., 1994; Kawaski et al., 1993; Konishi et al., 1998; Marenco, Coppola, Daniel, Zigun & Weinberger, 1993; Mentzel, et al., 1998; Nagahama et al., 1997; Nagahama et al., 1996; Nagahama et al., 1998; Parellada et al., 1998; Tien, Schlaepfer, Orr, & Pearlson, 1998; Volz et al., 1997; Weinberger, Berman, & Zec, 1986). Studies also found several non-frontal regions activated during the WCST including the parietal cortex, basal ganglia, temporal lobe, and occipital lobe (Alvarez & Emory, 2006). However, Alvarez & Emory concluded, "the prefrontal cortex, especially, the dorsolateral prefrontal cortex, is necessary for "normal" WCST performance" (p. 21). Alvarez and Emory (2006) concluded that the dorsolateral prefrontal cortex is central to performing set-shifting tasks like the WCST.

Set-shifting, or cognitive flexibility has been associated with successful WCST performance (Alvarez & Emory, 2006; Rubenstien, Meyer, & Evans, 2001). Specifically, Stemme, Deco, and Busch, (2007) defined cognitive flexibility as "the ability to switch attention from one aspect of an object to another" (p. 313). Stemme, Deco, and Busch (2007) speculated that perseverative errors measure an individual's cognitive flexibility. Overall, while prefrontal function is found to be a crucial process involved in the WCST, it should not be assumed as the only brain region involved and should not be considered a specific marker of prefrontal lobe function (Barceló, 2001).

Interscorer and intrascorer reliability. Axelrod, Goldman, and Woodard (1992) performed two studies on the reliability of scoring the WCST. The studies consisted of 30 adult psychiatric inpatients. The first study recruited three experienced clinicians in neuropsychological assessment to score the WCST. Excellent interscorer agreement was found with .93 for Perseverative Responses, .92 for Perseverative Errors, and .88 for Nonperseverative Errors. Additionally, excellent intrascorer reliability was found with .96 for Perseverative Responses, .94 for Perseverative Errors, and .94 for Nonperseverative Errors. The second study used six novice scorers without prior experience in scoring the WCST. The scorers were randomly assigned into two groups. One group used the standard scoring procedures given by Heaton (1981) and the other group received a set of supplemental scoring instructions in addition to the standard scoring procedures. Excellent interscorer reliability was found for both groups. For the standard scoring group the following reliabilities were found: .88 for Perseverative Responses, .97 for Perseverative Errors, and .75 for Nonperseverative errors. The standard instructions plus supplemental instructions had the following interscorer reliability: .95 for Perseverative Responses, .93 for Perseverative Errors, and .83 for Nonperseverative Errors. The authors found that the use of supplemental scoring saved scorers time when learning to score the WCST.

Standard error of measurement (SEM). SEM is an estimate of a standard deviation if an individual is given the same test multiple times (Heaton et al., 1993). Whenever an individual is administered the WCST there is a 95% chance that his or her true score with be within 1.96 standard error of measurement of their obtained score (Heaton et al., 1993).

Validity. Shute and Huertas' (1990) study supported the validity of the WCST as a measure of executive functioning. The study consisted of 58 participants who were administered a neuropsychological battery including a measure of Piagetian formal operational reasoning ability. Their factor analysis found that Perseverative Errors on the WCST was associated with Piagetian lower formal operational reasoning processes. The validity of the WCST has been also explored in other clinical groups have consisting of focal and diffuse brain damage and psychiatric groups.

A study by Weinberger et al. (1986) examined the regional cerebral blood flow (rCBF) in the dorosolateral area of the prefrontal cortex during performance on the WCST. Participants consisted of 20 adults with schizophrenia and 25 adult controls. The dorsolateral rCBF was measured while participants were resting and also during the administration of the WCST. In the control group, an increased in dorsolateral rCBF was found during the administration of the WCST compared to when resting. However, individuals with schizophrenia did not show a significant increase in rCBF during the administration of the WCST compared to when resting. The authors concluded the results provided evidence for the role of the prefrontal cortex in adequate WCST performance. WCST validity studies have also been conducted with TBI populations.

A study was also created for the Heaton et al. (1993) WCST manual. The study consisted of 502 participants with brain damage were classified into four lesion groups. These groups included 59 participants with frontal lesions, 53 participants with frontal and additional lesions, 177 participants with diffuse lesions, and 54 participants with non-frontal lesions. A group of 356 adult controls were used for comparison. An analysis of the results found that group identification accounted for 25% of the variance on WCST scores. Total Number of Errors, Percent Conceptual Responses, and Number of Categories Completed were found to be the most significant variables in differentiating the groups. For example, participants in the frontal, frontal plus, and diffuse groups made more errors and completed fewer categories on the WCST compared to participants in the control and non-frontal lesion groups. However, not all studies have found frontal brain damage to be specific to WCST performance.

A study by Anderson, Damasio, Jones, and Tranel (1991) also explored the effect of different brain lesions on WCST performance. The study consisted of 91 participants with brain lesions. Three groups were created based of location of brain lesion. The groups consisted of 49 participants in the frontal only group, 24 participants in the nonfrontal group, and 18 participants in the frontal plus group. Results demonstrated no significant differences between participants in the frontal and nonfrontal groups. Overall, lesion location and or size were not significantly related to performance on the WCST.

WCST variables. There are several scores that are derived from the administration of the WCST which are described by Strauss et al. (2006). Number of Categories Completed measures the number of correct 10 consecutive matches to a category (there are a maximum of 6 categories). Trials to First Category is another score that is the total number of trials it takes an individual to complete the first category. Perseverative Errors is the number of errors that involve perseverative on the previously completed category. Perseverative Responses includes the number of Perseverative Errors as well as responses that perseverated to the previously completed category. Failure to Maintain Set (FMS) occurs when an individual makes five or more consecutive correct matches but then makes an error before completing the category. Unique Responses are the number of times an individual fails to sort a card to any of the possible categories. Nonperseverative Errors are errors that do not involve perseveration to the previously completed category. Number of Trials Completed refers to the totals number of trials administered to the individual. The Total Number Correct is scored by the total number of correct items that were administered. The Total

Errors refers to the total errors that were performed. Percent Perseverative Errors is the number of perseverative errors in relation to the overall test performance multiplied by 100. Percent Conceptual Level Responses is the consecutive correct responses that occur in runs of three. Learning to Learn is the individual's average change in efficiency across categories that are based on the percent of error difference scores for each pair of categories.

Performance on the WCST among clinical populations

In this next section the nature of mTBI, moderate to severe TBI, and anxiety/depression populations will be discussed. Next, the expected or typical performance of individuals with mTBI, moderate to severe TBI, and anxiety/depression on the WCST will be reviewed. A TBI can be defined as a blow or jolt to the head, a penetrating wound, or a jolt to the head that is severe enough to cause functional or physical damage to the brain (Holtz, 2011). This disturbance of brain function is typically associated with normal structural neuroimaging findings (i.e., CT scan, MRI).

TBI severity and classification. TBI can be classified as mild, moderate, or severe. Furthermore, there are two broad types of TBI called closed head injury (CHI) and open head injury (OHI) (Holtz, 2011). An OHI results from an open wound such as a bullet entry and can also be referred to as a penetrating head injury (Holtz, 2011). OHI damage is usually more focal and can lead to the subsequent seizures or epilepsy (Holtz, 2011). CHI is the most common type of

TBI, which can results from falls, assaults, sport related injury, and motor vehicle accidents (Holtz, 2011). This occurs when the head comes into forceful contact with an external object (Holtz, 2011). Furthermore, a CHI is usually more diffuse and cognitive and or functional deficits are not as localized as in an OHI (Holtz, 2011). This is because CHI often results in diffuse axonal shearing (DAI), which is the breaking or stretching of myelinated axons (Holtz, 2011).

mTBI definitions. There is no universal definition of a mTBI (Holtz, 2011). However, one commonly used definition was described by the American Congress of Rehabilitation Medicine (ACRM) (Mild Traumatic Brain Injury Committee, 1993) and included the following criteria for mTBI: any period of loss of consciousness lasting less than 30 minutes, any loss of memory for events before or after the incident, any alteration in mental state at the time of the incident lasting less than 24 hours, or focal neurological deficit that may or may not be transient with a Glascow Coma Scale of 13 or higher. Another definition of mTBI by the World Health Organization (WHO) definition was derived from the ACRM definition and is defined as the following:

> Mild traumatic brain injury is an acute brain injury resulting from mechanical energy to the head from external physical forces. Operational criteria for clinical identification include: (i) 1 or more of the following: confusion or disorientation, loss of consciousness for 30 minutes or less, post-traumatic amnesia for less than 24 hours, and/or other transient

> neurological abnormalities such as focal signs, seizure, and intracranial lesion not requiring surgery; (ii) Glascow Coma score of 13-15 after 30 minutes post-injury or later upon presentation for healthcare. These manifestations of mTBI must not be due to drugs, alcohol, medications, caused by other injuries or treatment for other injuries (e.g. systemic injuries, facial injuries or intubation), caused by other problems (e.g. psychological trauma, language barrier or coexisting medical conditions) or caused by penetrating craniocerebral injury (Carroll, Cassidy, Holm, Kraus, & Corondo, 2004, p. 115).

A Mild traumatic brain injury results in a constellation of physical, cognitive, emotional, and/or sleep-related symptoms and may or may not involve a loss of consciousness (LOC). A concussion is a type of mTBI, which involves momentary loss of consciousness and can result in symptoms such as headache, dizziness, and inattention (Holtz, 2011). Duration of symptoms is highly variable and may last from several minutes to days, weeks, months, or even longer (Centers for Disease Control and Prevention, 2003). Individuals with uncomplicated mTBI are expected to recover from cognitive impairments within three months post-injury (Larrabee, 2012). A presence of a TBI and the severity is generally assessed using the Glascow Coma Scale (GCS) (Teasdale & Jennett, 1974), which measures the responsiveness in eye opening, motor movement, and verbal communication. Scores range from 3 to 15 with lower scores indicating lower functioning. Patients with scores between 13 and 15 are usually classified as having a mTBI.

Cognitive sequelae of mTBI. Binder et al. (1997) conducted a meta-analysis of 8 studies with individuals with mTBI three months post injury with persistent neuropsychological complaints. The study found that on average the effect of mTBI was undetectable in neuropsychological performance. The study found that neuropsychological domain accounted for less variability compared to injury severity. However, measures of attention had the largest effect size, but the authors speculated that this has little clinical significance. Furthermore, the authors concluded that neuropsychologists are more likely to be correct in diagnosing no brain injury than brain injury in mTBI. Binder (1997) also stated, "there is little empirical evidence that prolonged neuropsychological deficits typically are caused by mTBI" (p. 448). Binder (1997) found that of individuals with mTBI only 7 to 8% remain symptomatic. Binder (1997) believed that individuals with persistent complaints are likely best explained by psychosocial factors rather than neurological deficits from the brain injury. Binder (1997) also cautions that compensation and litigation status are also associated with reported neurological complaints even in absence of neurological injury.

A study by King and Kirwilliam (2011) explored factors that caused chronic neurological complaints after mTBI. Twenty-four participants reporting persistent post-concussive syndrome (PCS) with at least 18 months post mTBI were evaluated on a range of factors including cognition and emotional and psychosocial variables. The participants had an average post-injury time of 6.9 years. The results found that anxiety accounted for 45.9% of the variance in post-concussive complaints. The authors speculated that chronic mTBI complaints in cognition should be understood in context of a biopsychosocial model and that factors such as anxiety and depression likely account more for chronic mTBI deficits rather than neurological damage.

mTBI WCST performance. A meta-analysis reviewed 17 different studies to explore the effects of mTBI on neuropsychological functioning in adults at all stages of post-injury (Frencham, Fox, & Maybery, 2005). The study indicated "the effects of mTBI on neuropsychological functioning are small, and in general reduce to levels comparable with non-head injured individuals after the first three months" (p. 348). The study utilized the WCST and the Controlled Oral Word Association Test to measure the domain of executive functioning. The executive functioning measures failed to reach significance with respect to time post-injury. There was a small effect of mTBI on executive functioning during the acute phase (g=0.30 and SD=0.35); however, during the post-acute phase no significant effect was found. Additional meta-analyses on the neuropsychological effect of mTBI have found similar findings.

Rohling et al. (2011) performed a meta-analytic review of the neuropsychological effects of mTBI. The author's findings found no significant neurocognitive deficits by 3 months post injury. The authors found that most mTBI cases resolve neurocognitive deficits within a couple of weeks. Rohing et al. (2011) reported that their findings were consistent with the literature and other meta-analytic reviews that uncomplicated mild traumatic brain injury does not typically results in chronic neurocognitive deficits. Therefore, the authors concluded that if neuropsychological deficits occur after a mTBI that it resolves quickly.

Ord, Greve, Bianchini, & Aguerrevere (2010) explored the impact of traumatic brain injury on executive functioning and credible performance on the WCST. The study consisted of 109 patients with mTBI and 67 patients with moderate-to-severe TBI with at least one-year post injury. The authors found that individuals with credible performance on the WCST did not significantly differ from the controls. Those with moderate-to-severe TBI showed impairment on the WCST (effect of 0.09). The results indicated that effort during testing had a more significant impact on the patient's WCST performance than mild or moderate-to-severe TBI. The findings of this study highlight the need for PVT during neuropsychological evaluations of individuals with TBI.

Moderate to severe TBI. Individuals with moderate to severe traumatic brain injury usually have poorer outcomes in everyday functioning, cognition, and time to recovery from injury compared to those with mTBI (Larrabee, 2012). Furthermore, moderate to severe TBI often results in decreased personal independence and often negatively impacts an individual's ability to return to work (Larrabee, 2012). For example, individuals with GCS scores between 9 to 12 are classified as having a moderate traumatic brain injury. A severe traumatic brain injury usually results in GCS between 3 to 8. There are two types of moderate to severe TBI, closed and penetrating. Closed head injuries are caused by the movement of the brain against the skull, which can result from a fall, motor vehicle accident, or being struck by an object (Centers for Disease Control and Prevention, 2003). An open headed injury results in an injury to the brain by a foreign object entering the skull that could be caused from a bullet (Centers for Disease Control and Prevention, 2003). Larrabee (2012) highlighted that slowed processing speed, verbal learning and memory, and executive functioning are often seen in individuals with moderate to severe TBI.

Moderate to severe TBI WCST performance. The previously discussed studies from Ord et al. (2010) provided evidence that individuals with moderate to severe TBI generally perform worse on the WCST compared to controls and individuals with mTBI. Another study by Greve, Love, et al. (2002) explored the WCST in 68 individuals with chronic and severe TBI. The study found that individuals with severe TBI performed poorly on the WCST due to their impairments in cognitive flexibility, problem-solving, and response maintenance. Benge, Caroselli, and Temple (2007) studied whether the scores from the WCST among those with severe TBI could be related to functional outcomes. The study's results were counterintuitive and found that individuals with severe TBI who had more Failure to Maintain Sets (FMS) were associated with better occupational outcomes. The authors speculated this counterintuitive finding by concluding that in order to have a FMS, a participant must achieve at least five consecutive correct responses. These consecutive responses may set them apart from more impaired individuals who are unable to achieve five or more consecutive responses. Also, individuals with more Nonperseverative Errors were associated with increased supervision needs. These findings suggest that individuals with TBI's executive functioning ability can have direct functional outcomes and performance on the WCST may be used to measure ones ability to function independently. Therefore, scores of the WCST could be used as a case to substantiate a disability claim and therefore non-credible effort must be assessed on executive functioning measures such as the WCST. Larrabee (2012) discussed that individuals with moderate to severe TBI may also occasionally exaggerate symptoms and there is a need to assess for continuous effort especially among those involved in litigation.

Demakis (2003) conducted two meta-analyses of the sensitivity of the WCST to frontal and lateralized frontal brain damage. The first meta-analysis compared patients with frontal damage to nonfrontal damage. The study found patients with frontal damage achieved significantly fewer categories completed and more perseverative errors. The second meta-analysis consisted of patients with frontal damage and compared left versus right damage. The results found no significant differences for the left versus right frontal brain damage.

Performance Validity Testing (PVT) in severe TBI. Tombaugh (1997) conducted a study of TOMM performance in 158 neurologically impaired adults. Of these 158 participants, 45 were diagnosed with traumatic brain injury. Participants had over 92% accuracy on Trial 2 of the TOMM. The authors concluded that the study provided evidence that the TOMM is generally insensitive to neurological impairments. Other studies in the literature have also found the TOMM to be insensitive to neurological impairments from moderate to severe TBI.

Batt, Shores, and Chekaluk (2008) examined the effect of severe TBI on the Test of Memory Malingering (TOMM) and the Word Memory Test (WMT). The authors assessed 69 adults with severe TBI under the conditions of full effort, distraction, or simulated malingering. The authors indicated that the WMT and TOMM both had excellent sensitivity. However, the WMT had significantly higher false positive rates compared to the TOMM. The authors concluded that the WMT might require more than effort especially among individuals with severe TBI. Also, the study provides evidence that the TOMM may be used as an accurate measure of effort among those with severe TBI.

Furthermore, Haber and Fichtenberg (2006) examined the TOMM in non-malingering moderate to severe TBI patients. The study replicated the Tombaugh

(1996) study and included 50 participants. The results were consistent with Tombaugh (1996) and Rees, Tombaugh, & Boulay (2001) and found the TOMM and found that the cutoff score of 5 errors produced adequate sensitivity and perfect specificity for all participants with moderate to severe TBI. Overall, the literature suggests that the TOMM is not sensitive to neurological impairments from TBI (Batt, et al., 2008; Haber, Fichtenberg, 2006; Tombaugh, 1997).

Psychiatric diagnoses. It is important to understand the impact of psychiatric disorders on executive functioning. The impact of psychiatric disorders on executive functioning in the absence of neurological impairments has also been explored in the literature. For example, depression or anxiety might negatively impact performance on the WCST and result in increased Perseverative Errors and FMS (Borkowska & Rybakowski, 2001; Moritz et al., 2002). This is because depression and anxiety have been found to decrease problem solving ability (Strauss et al., 2006).

Psychiatric disorders and WCST performance. Smitherman et al. (2007) studied the performance of the WCST on 86 adult outpatients with self-reported anxiety or depression. All participants were administered the WCST along with the Beck Depression Inventory—2nd Edition (BDI-2) and the State-Trait Anxiety Inventory (STAI) to assess for mood. The study found that variance caused by depression or anxiety after controlling for age, gender, and IQ was less than 3%. Overall, the study suggested that self-reported anxiety or depression had little effect on executive functioning measures such as the WCST. The authors noted that their findings were congruent with Martin, Oren, & Boone (1991) after controlling for IQ.

Martin et al. (1991) explored the severity and depression type on WCST performance. The sample consisted of 48 adults with the following depression diagnoses: 13 with Major Depression, 17 with dysthymia, and 18 non-psychiatric controls. All participants were interviewed using the Structured Clinical Interview for DSM-III Revised, which assigned participants to their diagnostic group. All participants were administered the Beck Depression Inventory (BDI), the Hamilton Rating Scale for Depression (HRSD), Wisconsin Card Sorting Test (WCST), and the Verbal IQ subtests of the Wechsler Adult Intelligence Test-Revised (WAIS-R). The results found that depression symptom severity and diagnosis were associated with lowered performance on certain WCST variables. For example, Percent Perseverative Errors were significantly higher and Percent Conceptual Level Responses were significantly lower among depressed participants. Furthermore, the severity of depression predicted scores on Total Errors, Failure to Maintain Set, and Percent Perseverative Responses. The authors speculated the more severe symptoms of depression might impair problem-solving abilities. However, after the participants were controlled for verbal IQ, the individuals with depression were not significantly different from the non-psychiatric controls. The authors suggested clinicians should interpret

WCST scores with caution among depressed individuals. Also, the authors urged for future research in the WCST performance among depressed inpatients.

Ilonen et al. (2000) explored WCST performance among participants with severe first-episode depression. The sample consisted of three groups: 28 patients with psychotic depression, 29 patients with nonpsychotic depression, and 30 healthy controls. A psychiatrist made all diagnoses and all participants were matched for age and education. All participants were administered the Wechsler Adult Intelligence Scale-Revised (WAIS-R), WCST, and the Rorschach Comprehensive System (CS). Results found the psychotic depression group performed significantly lower on the WAIS-R FSIQ compared to the severe nonpsychotic depression and healthy control groups. Also, participants with major depression in both the psychotic and non-psychotic group had severe impairments in the WCST. Severely depressed participants had significantly higher perseverative errors and perseverative responses compared to health controls. The authors speculated that poor performance on the WCST in severely depressed patients may be "the result of intense emotional distress and psychomotor retardation" (p. 279). Overall, the authors concluded that severely depressed individuals in the early stage of illness likely have significant executive functioning impairments.

PVT in depression/anxiety. It is also important to consider the impact of psychiatric illnesses on PVTs. O'Bryant, Finlay, & O'Jile (2007) explored

TOMM performances in 262 outpatient adults with self-reported symptoms of depression and anxiety. Overall, the results suggested that depression and anxiety did not have a significant effect on TOMM performance. Therefore, this study provided evidence supporting the use of the TOMM as a measure of cognitive effort among those with depression and anxiety. Yanez, Fremouw, Tennant, Strunk, & Coker (2006) explored the effects of severe depression on TOMM performance among 20 outpatient adults seeking disability and 20 non-depressed participants. The study suggested that depression did not significantly impact TOMM performance compared to the control group. Therefore, the study suggests that the TOMM can be used as a measure of cognitive effort among severely depressed individuals and was consistent with the previous literature (Ashendorf, Constantinou, & McCaffrey, 2004; Rees et al., 2001).

Furthermore, Iverson, Page, Koehler, Shojania, and Badii (2007) found that TOMM performance was not affected by chronic pan or depression in patients with fibromyalgia. The study examined 54 participants with mild and severe levels of depression and found not one participant failed the TOMM. No studies were found to date examining the WCST as an embedded PVT with a non-neurological psychiatric sample.

WCST as a Performance Validity Test (PVT) Studies

Over the last 15 years there have been numerous studies that explored the WCST's ability to detect non-credible performance among those with neuropsychological deficits (Greve et al., 2009). Studies have examined mild, moderate, and severe traumatic brain injuries as well as mixed neuropsychological samples. However, to date there have not been any studies that examined the utility of the WCST as a PVT among a non-neurological psychiatric populations. The designs of these studies vary in the literature. For example, there are simulators, specificity, and known groups designs. As previously discussed, known group designs are considered to be the most robust design as it allows for a control comparison group and uses clinical participants (Larrabee & Berry, 2007). The following is a review of the current and past literature in regards to the utility of the WCST as a PVT among those with traumatic brain injury and or psychiatric diagnoses. A summary of the known studies exploring the WCST as a PVT are displayed in Table 1.

Early studies. Bernard, McGrath, and Houston conducted one of the earliest studies in 1996. The authors sought to create a discriminant function formula to help determine malingering on the WCST using a simulator design. Their sample included 24 college students who were told to simulate brain damage. Participants were told to act like a "real patient" who was seeking compensation while trying not to make their poor effort obvious. The control group consisted of 21 college students who were told to perform their best. In addition, there were 89 participants with mixed neurological diagnoses without brain injury, and 70 closed head injury participants. The researchers explored the following WCST variables: Total Errors, Perseverative Errors, Categories Completed, and Unique Responses. They hypothesized that the simulators would score statistically lower on the Number or Categories Completed because they predicted that would be an obvious indicator of poor performance. Conversely, they hypothesized that the simulators would score comparatively lower on more subtle measures of Perseverative Responses and Perseverative Errors. Three discriminant analyses were conducted on the following three comparisons: simulated malingerers to controls, simulated malingers to close head injury patients, and simulated malingerers to neurological patients.

The authors used all three discriminant analyses were evaluated on half of the groups (Bernard et al., 1996). The second half of each group was used for cross-validation. In the malingerers versus controls, the number of categories completed was the only WCST variable to discriminate between the groups and the only variable to meet criteria for inclusion. This analysis accurately classified 91% of the patients in the original analysis and 96% of the patients in cross-validation. In the malingerers versus closed head injury patients, both the number of Categories Completed and Perseverative Errors met criteria for inclusion. The cross-validation was acceptable at 95%. The classification accuracy was 86% and specificity was 94%. For malingers versus mixed neurological patients without brain injury, the Number of Categories Completed and Perseverative Responses met inclusion criteria. Cross-validation was 91%. They found specificity was

100% and sensitivity was 58%. Overall, the author's findings indicated excellent classification accuracy and provide supportive evidence of the WCST as one of multiple effort measures.

The study by Bernard et al. (1996) also has several limitations. The most significant limitation is that the participants were simulators and not true malingerers. Therefore, the results may not be generalizable to individuals involved with litigation or with neurological impairments. Also, the closed head injury group consisted of both mild and moderate traumatic brain injury. This study indicated that additional research with greater external validity with real-word patients with non-credible performance is needed to increase the generalizability of these significant findings. Several studies have replicated the findings of this study.

Suhr and Boyer (1999) attempted to replicate the findings of Bernard et al. (1996). They also used undergraduate simulators to design their study. To expand the literature they used known groups design. They used the WCST to discriminate between 41 undergraduates who simulated malingering, 31 normal undergraduates, 17 probable malingering (with at least two external indicators of malingering), and 16 mild to moderate brain-injured patients. They found Number of Categories Completed and Perseverative Errors to be highly negatively correlated and therefore did not include Perseverative Errors to predict group classifications. The authors compared the groups to the variables of

Categories Completed, Total Errors, Perseverative Errors, and Failure to Maintain Set (FMS). Instead of using discriminant analysis, they created a logistic regression to Number of Categories and Failure to Maintain Set.

When comparing student simulators versus controls, the authors correctly identified 70.7% of simulators (sensitivity) and correctly identified 87.1% of students (specificity) (Suhr & Boyer, 1999). In the patient sample, their analysis accurately identified 82.4% of probable malingerers (sensitivity) and 93.3% of motivated TBI patients (specificity). The author's findings provided support for the use of the WCST as an indicator of poor cognitive effort. The limitations of this study include small sample sizes in the patient groups, the questionable generalizability of the student groups, and the authors did not provide cross-validation. Additional studies are needed in order to better understand the generalizability of these findings. Studies have also replicated the findings of this study using different patient populations.

Specificity of Bernard and Suhr formulas. Donders (1999) examined the specificity of the Bernard et al. (1996) formula. The sample included 130 patients with mild, moderate, and severe traumatic brain injury within 1 year of their injury. Participants included a wide age range of adults from 17 to 74 years of age. All participants were carefully screened to ensure that they did not have external incentives for poor performance. Those who were seeking financial compensation were excluded. Donders aimed to see if Bernard's formula could demonstrate a false-positive error rate of no greater than 10% in a TBI sample. The Bernard et al. (1999) formula yielded a false positive error rate of 5% (n = 7). Those seven patients also had severe TBI, posttraumatic stress disorder, severe pain, or substance abuse. In addition, Donders questioned the accuracy of Bernard's formula with older adults, which may lead to increased false positive rates as performance on the WCST declines with age. This study is limited in that it did not explore the sensitivity of the Bernard formula. Conclusions of this study call for more studies that explore the sensitivity Bernard et al.'s formula with those with poor effort. The next study explores the sensitivity of the Bernad et al. formula in addition to the sensitivity of the Shur formula.

Greve and Bianchini (2002) examined the specificity of the Bernard and Suhr formulas. They utilized seven samples in their study which included: 133 normal college students with no psychiatric or neurological diagnoses, 76 unselected college students, 44 patients in an inpatient substance abuse treatment program, 69 participants with chronic TBI living in a residential treatment facility, 83 stroke patients in an inpatient physical rehabilitation facility, and a mixed outpatient and inpatient neurological samples (N = 128 and 360). The Bernard et al. formula yielded the following false-positive errors rates: 2.3% for normal college students sample 1, 15.8% for unselected college students sample 2, 20.5% for substance abusers, 41.7% of severe TBI, 12% of cerebrovascular accident, 7.8% of the first mixed neurological sample, and 13.3% of mixed the second neurological sample. The Suhr formula yielded the following false positive error rates: 0% of normal first college students sample, 1.3% of second college students sample, 18.2% of substance abusers, 26.1% of severe traumatic brain injury, 26.5% of cerebrovascular accident, 14.1% of the first mixed neurological sample, and 20.3% of the second mixed neurological sample. The authors concluded that both of these equations produced greater than 10% false-positive rates, which is unacceptably high. This study suggested that the formulas produce high false-positive rates with patients with severe pathology.

However, unlike the Suhr formula, the authors found the Bernard formula produced acceptable false positive error rates of less than 10% when used in specific diagnostic subgroups of the mixed patient samples for TBI and CVA, thus replicating Donders (1999) (Greve & Bianchini, 2002). The Greve and Bianchini (2002) stated that their findings of low specificity might be due to their underlying assumptions and operationalizations of malingering. The authors suggested that exploring non-credible performance rather than malingering with measures on the WCST may increase the classification accuracy. The authors discussed that individuals may choose alternative ways to fake bad on the WCST, and therefore homogeneity cannot be assumed from simulated malingerers. The authors suggested that future research must "empirically identify the various approaches to malingering on a given test and then apply DFA (discriminant function analysis) to derive equations which are sensitive to those strategies" (p. 53).

Ashendorf et al. (2003) designed a specificity study of the WCST and California Verbal Learning Test (CVLT) using the Bernard and Suhr formula in healthy older adults. Participants included 197 community-dwelling older adults with a mean age of 64.57 (SD = 5.52, range = 55-75). All participants scored at least a 48 out of 50 on trial 2 of the TOMM. The authors used three Bernard et al. (1996) equations in their sample. This included a more liberal mixed neurological equation, undergraduate control equation, and a TBI equation. The Bernard formula was used as suggested by authors, and a conservative 90% likelihood of malingering cutoff was applied for the Suhr formulas. When applying the Bernard formula, it was significantly related to age. Specificity for the Bernard formula ranged from 55.1% to 91.4%. Specifically, the Bernard et al. (1996) control equation had a specificity of 55.1%, the TBI equation had a specificity of 75.7, and the mixed neurological equation had a specificity of 91.4%. The Suhr formula specificity ranged from 47.6% in the undergraduate equation to 62.7% in the patient sample equation. For Unique Responses in the WCST, a cutoff of >5 produced a false positive error of 8.6%, and Perfect Matches Missed resulted in 95.1% specificity. Overall, the Suhr formula produced the highest false positive rate of 52.4%. The Bernard formula's rate was half as high but still unacceptable and had a false positive error rate of 25%. The authors stated that reduced performance in the WCST with aging negatively affects the validity of the Bernard and Suhr formulas. They also speculated that severely impaired individuals would also yield high false positive rates when using both formulas. This study urged for future research to establish the sensitivity of these formulas with different populations in order to generalize to neuropsychological practice. The literature consists additional studies with known-group designs that allow testing the generalizability of the WCST as a PVT in different populations.

Known-group studies. In addition to specificity studies, there are several known-group studies. Sweet and Nelson (2007 reviewed a study by Miller, Donders, and Suhr (2000) that examined the accuracy of the Suhr and Bernard formulas in 90 mild and severe TBI patients. Three groups of 30 participants each included severe TBI without incentive, mild TBI without incentive, and mild TBI with incentive. Of those participants, 13 individuals failed either the Warrington Recognition Memory Test (WRMT) or the Test of Memory Malingering (TOMM) and were deemed "likely malingerers". Ten of these likely malingerers came from the mild TBI with incentive group. The other 3 individuals were not from the incentive groups and therefore could not be classified as malingering, which left 10 possible malingerers out of 30 patients with incentive with a base rate of 33%. After applying both formulas, none of the formulas correctly identified any of the participants who failed the TOMM or WRMT and therefore sensitivity was unacceptably low at 0%. The specificity of the Bernard formula was 96% and the specificity of the Suhr formula was 95%. This is especially disappointing because in this study the formulas were better at identifying non-malingers than malingerers. The authors discussed that the findings for the Bernard formula may be due to the fact that their formula was derived using simulators instead of true malingerers. The authors cautioned using these formulas in clinical practice. More recent studies have followed these future directions and created a formula using real patients rather than stimulators (King, Sweet, Sherer, Curtiss, & Vanderploeg, 2002).

King et al. (2002) examined the WCST classification accuracy for detecting those patients with credible from non-credible performance. They sought to develop new variables that may be useful in detecting non-credible performance on the WCST. In this study, they looked at a new logistic regression-based indicator, which included the Number of Categories Completed, FMS, and Percent Conceptual Level Responses. In addition, they examined the utility of the Bernard and Suhr formulas. They conducted three studies. The first study consisted of a sample of 33 patients with moderate to severe chronic TBI and 27 TBI with insufficient effort (who failed malingering tests). Those with insufficient effort scored significantly poorer on the WCST indicators relative to the non-litigating TBI group. The Bernard formula had acceptable specificity of 94% with sensitivity of 63%. The Suhr formula was similar at 88% specificity with 59% sensitivity. The authors concluded that the new WCST indices were not supported in their clinical usefulness.

In the second study by King et al. (2002), the sample consisted of 75 acute moderate to severe TBI patients and aimed to assess the specificity of the WCST variables and insufficient effort in the acute stage of recovery. The acute TBI patients performed worse on all WCST variables except for FMS, Other Responses, and Nonperseverative Errors compared to chronic TBI group. The Bernard formula had specificity of 73% for the acute TBI group. The Suhr formula resulted in 75% of acute TBI patients being accurately classified. The authors discussed that the acute TBI patients were not administered symptom validity tests to assess effort, which serves as a limitation.

In the third study by King et al. (2002), the authors examined the specificity of the WCST variables insufficient effort formulas to a sample of 130 mild to severe TBI patients who post-acute compared to study 2 sample. The Bernard formula had a false positive error rate of 5% and their new formula had a false positive error rate of 1%. The Suhr formula was least accurate and had a false positive error rate of 15%.

A limitation of King et al. (2002) is that none of the patients in the sample were administered PVTs to assess for effort and rather they assumed that effort was acceptable due to the absence of litigation. The overall, findings of these studies indicate that good effort can be determined; however, they did not establish if insufficient effort could be determined using the WCST. In addition, these studies suggested variability among populations. Therefore, clinicians must be aware of individual characteristics such as type of injury, severity of injury, and time since injury. This study also is limited in that it did not demonstrate acceptable sensitivity across forensic samples, and therefore future research is needed in studying forensic samples. The next study discussed utilized participants in litigation for comparison.

Larrabee (2003) examined the detection of malingering among the FMS index of the WCST among four other standard neuropsychological tests. The sample consisted of 26 patients who performed significantly worse than chance on the Portland Digit Recognition Test (PDRT) who claim to suffer from brain damage, and 31 moderate/severe TBI including those involved in litigation. A cutoff of >1 FMS yielded a sensitivity and specificity of 48%. Among the non-malingering patients, no one had more than 3 FMS. In the malingering group, 12% had more than 3 FMS and 20% had more than 2. FMS was involved in false positives, and all of those participants had severe TBI and none of them had more than 3 FMS. Therefore, the authors suggested that for patients without a severe TBI and more than 3 FMS are likely exhibiting poor effort. Limitations of this study include sample sizes and mixed diagnoses in the samples. The authors encourage for additional replications of this study with larger sample sizes and more specific clinical groups.

Greve, Bianchini, et al. (2002) examined the classification accuracy of Bernard and Suhr formulas and two types of unique responses. The sample included 89 patients with mild and moderate-severe TBI who were then divided into groups based on Slick et al. (1999) malingering criteria. The four groups consisted of 17 no-incentive controls, 32 probable malingering, 30 suspected malingering (positive on only one Slick et al. criteria), and 10 incentive only. The authors used two cut-off levels to examine the classification accuracy of the four indicators. The first cut-offs were the following: unique responses and perfect matches-missed >0, Bernard > -3.00, and Suhr > 1.90. The second cut-off levels were based on the recommendations for the Bernard formula and 90% probability of malingering for the Suhr formula and was more conservative and as follows: Unique Responses and Perfect Matches-Missed >1, Bernard >0, and Suhr > 3.68. Classification accuracy was examined and all indicators had false positive rates of approximately 10%. The four indicators also correctly identified about one third of the probable malingers as well. The perfect match-missed marker had 100% specificity but low sensitivity of 10%. When using the more conservative cutoff scores, about 50% of the probable malingerers were correctly identified with acceptable specificity.

Greve, Bianchini, et al. (2002) discussed that their results provided evidence to three different strategies that malingerers employ on the WCST. Over 50% of the malingerers had higher FMS, which was identified through using the Suhr formula. It is likely that these individuals avoided getting too many correct responses in a row. The second most common strategy involved Unique Responses or not matching to any of the three categories on the target card. And finally one third of the probable malingerers responded validity on the WCST. The authors concluded that many individuals might not see the card-sorting task as a cognitive task that is affected by neurological deficits. This article gives light to the variability of poor cognitive effort strategies that are employed and that non-credible performance is generally not homogeneous among individuals or among an individual test battery. A notable limitation of this study is the small sample sizes, especially for the incentive only group (n = 10). This study urges future research to explore new combinations of scores that accurately distinguish good effort from poor effort. This study calls for more research to determine the limits of specificity of these examined WCST indicators. The strength of the next study discussed includes a large sample size.

Heinly, Greve, Love, & Bianchini (2006) explored the classification accuracy of the following WCST indicators: Unique Responses, FMS, Suhr formula, Bernard formula, and King formula. The sample consistent of the following: 137 mild TBI patients, 139 moderate-severe TBI patients, 101 chronic severe TBI patients that resided in a residential facility, and more the 1,000 general clinical patients. The mild and moderate-severe TBI groups were divided into malingering and non-malingering groups based on the Slick et al. criteria.

The authors then examined the sensitivity and false positive error rates. The authors found Categories Completed, Perseverative Responses, Perseverative Errors, and FMS to accurately differentiate the malingering form non-malingering groups. The reported false positive errors varied from 5% to 10%. The authors noted that the false positive error rates for the chronic and severe TBI group and those in the clinical group with more severe pathology (dementia and stroke) were unacceptably high. The authors concluded that the WCST should not be used to identify malingerers among those with severe TBI or pathology. Furthermore, they noted that clinicians should be cautious of those with mTBI who score similar to these patients. A strength of this study includes the larger sample sizes and comprehensive examination of TBI severity.

Greve et al. (2009) explored the classification accuracy of for seven indicators of the WCST in detecting malingering in mild TBI. Their sample consisted of 373 TBI patients and 766 non-TBI clinical patients. The authors used Slick et al. (1999) criteria for Malingered Neurocognitive Dysfunction (MND). After the groups were assigned into malingering and non-malingering categories, the sample sizes were as followed: mild TBI not MND (n = 55, 20% with no incentive), moderate/severe not MND (n = 92, 47.8% with no incentive), mild TBI MND (n = 38, 0% with no incentive), moderate severe MND (n =12), and indeterminate (had external incentive with suspect effort but insufficient to be classified as malingering) (n = 85).

Greve et al. (2009) found that the mTBI/MND group performed significantly worse than mTBI/non-MND group and similar if not worse to the moderate-severe and severe TBI groups. Therefore, the authors suggested that the WCST is sensitive to malingering at a group level. Significant group effects were found for all WCST variables except for unique responses. Perseverative responses had poor classification accuracy and detected 16% of MND and had a false positive error rate of 11%. Similar results were seen for Categories Completed and Unique Responses. The FMS of equal to or greater than 3 detected 29% of MND and had a false positive error rate of 11%. The Suhr and King et al. formula classified malingering relatively well. They had false positive error rates at 6% or less. The Bernard et al. formula performed worse and had a likelihood ratio of 2.6. The authors concluded that more than 3 FMS is a rate event in patients with dementia, stroke, and severe pathology and therefore such scores in m TBI are likely a result of malingering. The authors concluded that the formulas were more accurate at classifying malingering from not malingering than individual WCST variables. The authors discuss that the false negative error rate is too high and individual WCST should not be used a primary method of malingering detection. However, they concluded the WCST could be a helpful tool in detecting malingering among mTBI patients.

Summary. All studies with the exception of one (King et al., 2002) used the Slick et al. (1999) criteria for malingering. However, PVTs are tests of effort or non-credible performance rather than measures of malingering. Furthermore, malingering is a clinical diagnosis that requires clinicians to infer about an individuals mental state (e.g., if one is intentionally malingering) (Boone, 2009). In most cases, PVTs do not provide enough evidence to determine if one is intentionally malingering. Also, failing PVTs may be due to reasons other than malingering (e.g, fatigue, boredom, or lack of interest in testing) (Schutte & Axelrod, 2013). Thus, research using the concept of non-credible performance would be beneficial since this concept is a directly observable one, rather than an inferential one like "malingering." Another limitation is that most of the research in the current literature consisted of small sample sizes; therefore, there is a need for future research with larger sample sizes. Also, the most current research study was published over three years ago and more up to date studies are needed. General findings in the literature found that the formulas or a combination of individual variables of the WCST had better classification accuracy than individual variables alone. Also, the literature generally found that the WCST is not useful in detecting non-credible performance among more severe TBI patients. Overall, the literature had variable findings and more research is needed to help determine a pattern reliable pattern of detection accuracy of non-credible performance on the WCST among those with mTBI.

Chapter III: Methodology

Type of Study

The current study is a retrospective quantitative study that employs a convenience sample from an archival database set. Therefore, this study will use nonrandomization and a quasi-experimental design. The study is also a known group design (Larrabee & Berry, 2007).

Variables

The independent/predictor variables include the individual WCST indices (Categories Completed, FMS, Perseverative Errors, Perseverative Responses, Nonperseverative Errors, Trials to First Category, Number Correct, Number of Trials, and Total Errors). The dependent variable is the credible (PASS) versus non-credible (FAIL) performance. All participants were administered the Test of Memory Malingering (TOMM, Tombaugh, 1996) and/or the Word Memory Test (WMT, Green, Allen, & Aster, 1996). In addition to the following embedded PVTs were administered: Brief Test of Attention (BTA, Schretlen, 1997), Trail Making Test-Part B (TMT-B, Halstead & White, 1950), Continuous Performance Test-2nd Edition (CPT-II, Conners et al., 2000) Omission Errors, Judgment of Line Orientation (JLO, Benton, Sivan, Hamsher, Varney, & Spreen, 1994), Benton Facial Recognition (Benton et al., 1994), Rey Complex Figure Test (Copy) (Meyers & Meyers, 1995), Wechsler Adult Intelligence Scales Fourth Edition, Digit Span Adjusted Scaled Score (WAIS-IV, Wechsler, 2008), and

Finger Tapping Test (dominant hand) (Halstead & White, 1950). All PVTs used are previously empirically validated with published cutoff scores and their references are displayed in Table 2. Participants were placed in either the PASS or FAIL group based off performance on PVTs. Specifically, participants who failed only one PVT were excluded from the study altogether, given the ambiguity of their performance. Participants who failed at least two PVTs including the WMT and/or TOMM were placed in the FAIL group. Participants who passed all PVTs (listed on Table 2) were placed in the PASS group.

Units of Measurement

This study aimed to identify differences between individuals, and t-tests and group comparisons were used to analyze whether the two groups responded performed significantly differently on the WCST and that demographics between groups were not significantly different. This study aimed to determine the classification accuracy of the WCST's ability to identify participant's with credible and non-credible performance.

Participants

The methodology of this study was replicated from Whiteside et al. (2014.) The current retrospective study consisted of neuropsychological assessment data from a collaborative clinical psychological assessment database. This database had Institutional Review Board (IRB) approval. Board-certified neuropsychologists working in clinical settings provided or supervised all evaluations and were responsible for the diagnoses included in this database. These neuropsychologists interpreted the WCST in this study and were not blind to any test results. However, the WCST was not used for determining performance validity at the time. All PVT's used to determine adequate cognitive effort are provided in Table 2 and include the following measures: Test of Memory Malingering (TOMM, Tombaugh, 1996), Word Memory Test (WMT, Green et al., 1996), Brief Test of Attention (BTA, Schretlen, 1997), Trail Making Test-Part B (TMT-B, Halstead & White, 1950), Continuous Performance Test-2nd edition (CPT-II, Conners et al., 2000) Omission Errors, Judgment of Line Orientation (JLO, Benton et al., 1994), Benton Facial Recognition (Benton et al., 1994), Rey Copy (Meyers & Meyers, 1995), Digit Span Adjusted Scaled Score (Wechsler, 2008), and Finger Tapping Test (dominant hand) (Halstead & White, 1950). As discussed above, these PVTs were used for group assignment. The literature that provides validated evidence for using these measures, as PVTs will be further discussed in the procedures section. Participants were consecutive neuropsychological referrals seen between 2005 and 2013 who were age 18 years or older who completed at least three of the established free standing or embedded PVTs (including the TOMM and/or WMT), as well as the WCST. As noted above, three different diagnostic samples were studied, including two non-neurological samples with individuals previously diagnosed with a possible mild traumatic brain injury (mTBI group) or a comparison group of individuals diagnosed with a depressive and/or anxiety disorder (PSYCH group). Individuals in the mTBI or PSYCH sample who presented with comorbid neurological disorders were excluded from this study. Individuals with comorbid psychiatric diagnoses were not excluded from this study. The third sample consisted of individuals who were previously diagnosed with moderate to severe traumatic brain injury (STBI group) and also was used a comparison sample.

As previously discussed, the Slick et al. (1999) MND criteria were not used to determine the credible and non-credible performance groups. Instead this study considers credible and non-credible performance rather than a diagnosis of malingering (Boone, 2009). Therefore, this study defines non-credible performance as a failure of 2 or more PVTs, not malingering, which is consistent with Boone's (2009) recommendations.

Group assignment. Group membership was determined by placing participants into the PASS group or FAIL group based on whether they failed two or more PVTs or passed all PVTs. Consequently, participants who only failed one PVT were excluded from the study. In order to be placed in the FAIL group, an individual must fail at least one standalone empirically validated measure of effort (TOMM and/or WMT). In addition, to failing a standalone measure of effort, a participant must have also failed at least one of the embedded PVTs listed in Table 2 to be placed in the FAIL group. Table 3 provides the number of participants who passed or failed PVTs in all three samples. No significant group differences on any demographic variable were found between the PASS and FAIL groups for age, gender, ethnicity and years of education. Compensation seeking status was defined as individuals who have known external incentive. Therefore, participants who were seeking a personal injury settlement, disability benefits, academic accommodation request, or were referred by attorneys, disability insurance, or educational settings were defined as compensation seeking. Specific demographic characteristics including comorbid psychiatric diagnoses for all three samples are displayed in Table 3, 4 and 5.

Sample recruitment. The study is a retrospective study and therefore there was no active participant recruitment for this study. Participants were consecutive adult outpatients referred for comprehensive neuropsychological assessment. All participants were assigned a number into the database and therefore confidentiality of the participants was maintained. No identifying information of participants was included in the database. All participants signed consent prior to testing that allowed for their results to be used for research purposes. The participants data was kept on a password protected network. Data will be kept on a password-protected computer and password-protected file for a minimum of seven years.

Exclusionary Criteria

In addition to those partieipants who failed a single PVT, other exclusionary criteria included diagnosis of a learning disability, substance abuse history, or a Full Scale IQ (FSIQ) below 71. Research has demonstrated that factors such as low IQ, learning disability, or chronic substance abuse can negatively impact WCST performance (Heaton et al., 1993). Therefore, these diagnoses were excluded in order to limit false positives in the non-credible performance group. Furthermore, IQ and education are highly correlated with performance on the WCST (Heaton et al., 1993).

Mild Traumatic Brain Injury Sample

The mTBI sample consisted of 110 consecutive outpatients. Demographic information including gender, marital status, years of education, referral source, and comorbid diagnoses can be found in Table 4 and 5. Participants were from the same dataset as Whiteside et al., 2014 with the same participant selection. Participants were included if they had a prior diagnosis of a mTBI based "on specific criteria (Bodin, Yeates, & Klamar, 2012), including evidence for injury to the head (PTA), Glascow Coma Scale (GCS) of 13 or more at time of injury, and no abnormalities on neuroimaging as determined by the original evaluating neuropsychologist" (Whiteside et al. 2014, p. 10). Participants were 18 years and older and were 6 to 24 months post injury without prior neuropsychological evaluations. Radiologists who were not affiliated with the research database performed all neuroimaging and therefore neuroimaging was not available for this study. However, the mTBI sample had negative findings on their neurological examinations.

Moderate to Severe Traumatic Brain Injury Sample

The STBI sample consisted of 69 consecutive outpatient individuals. All individuals in this sample ranged from 10 to 48 months post-injury. Table 4 and 5 includes demographic information. Participants were included in the STBI sample based on a diagnosis of moderate or severe TBI from a neurologist or physician using the severe TBI criteria (Bodin et al., 2012) of greater than 24 hours loss of consciousness, GCS of <9 and/or abnormalities on neuroimaging. The moderate TBI criteria (Bodin et al., 2012) were based of 30 minutes to 24 hours of loss of consciousness, GCS between 9 and 12, and/or evidence for abnormalities on neuroimaging. Independent radiologists who were not involved with this study conducted all neuroimaging. Therefore, neuroimaging data is not available for this study.

Mixed Mood/Anxiety Disorder Diagnosis Psychiatric Sample

For this sample, 155 consecutive outpatients were included. The PSYCH sample were individuals who were referred for a neuropsychological evaluation secondary to cognitive complaints, but had no evidence of a neurological disorder based on neuropsychological evaluation and a review of medical and neurological records. Demographics of this sample can be found in Table 4 and 5, which includes specific diagnoses using the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition, Text Revision (DSM-IV-TR, American Psychiatric Association, 2000).

Data Analysis

There was a lack of significant differences between the PASS and FAIL groups in regards to age and education as presented in Table 6 and therefore, the raw scores on the WCST measures were used in this study. The steps for the data analysis are the same for each of the three sample groups. The analysis will begin with the population of interest, the mTBI group, followed by the STBI and PSYCH comparison samples.

The first step of the data analysis consisted of analyzing the means and standard deviations of the WCST measures and then to analyze group differences. The Kolmogorov-Smirnov (K-S) test of normality found that the WCST variables did not meet the assumption for normality *($p<.05$)*, therefore non-parametric tests were used. Mann-Whitney U tests were used to analyze group differences.

Then, a logistically derived variable was created among the WCST predictor variables. In order to calculate the logistically derived variable, three WCST variables were entered into a logistic regression analysis with PASS or FAIL group membership as the outcome variable in the mTBI sample of interest. Beta values obtained from the logistic regressions were utilized to calculate the variable, called WCSTCOMB. The logistically derived WCSTCOMB variable was also used in the analysis with, the STBI, and PSYCH groups to evaluate generalizability. The WCST variables for the logistic regression were selected based off of the least correlated variables. A Spearman Rank Order (Rho) correlation was run to find what WCST variables were least correlated. The three least correlated WCST were used in order to minimize multicollineraity.

Next, to evaluate classification accuracy, Receiver Operator Characteristic (ROC) analyses were calculated for the WCST variables (Number of Trials, Total Errors, Perseverative Errors, Perseverative Responses, NonPerseverative Errors, Trials to 1st Category, Failure to Maintain Set, Number of Categories Completed, and Total Number Correct). In addition, the WCSTCOMB was also calculated and analyzed in each groups. The ROC analysis explored both individual variables and regression derived variables to assess what would result in the most acceptable classification accuracy across the three samples. Area under the curve (AUC) is reported in Table 10 for the mTBI sample, Table 11 for the PSYCH sample, and Table 12 for the STBI sample. Next, sensitivity (SN), specificity (SP), negative predictive power (NPP), and positive predictive power (PPP) were calculated for variables that had AUC of .70 or higher and are presented in Tables 13, 14, and 15 (Larrabee & Berry, 2007). Finally, recommended cut-off scores using 90% specificity is highlighted for each WCST variable (Larrabee & Berry, 2007). The specificity is set high to limit the occurrences of false-positive identifications. In other words, the specificity was set at 90% in order to avoid misclassifying a participant as having non-credible performance who truly had credible performance.

Procedures

All participants signed an informed consent at time of evaluation stating that they understand and consent their data to be used for future research purposes. The participants were informed that they could retract their data from participation in future research at any time. As a part of a comprehensive neuropsychological battery all participants were administered at least one freestanding PVT (TOMM or WMT), at least two embedded PVTs, and the WCST in a standardized fashion by a trained and competent examiner and supervised by a board certified neuropsychologist. Participants were informed that the neuropsychological evaluations included PVT's. Per Boone (2009) recommendations, all embedded and freestanding PVT's were administered throughout the entire evaluation.

Instruments. The Wisconsin Card Sorting Test (WCST) is used as a measure of executive functioning in neuropsychological assessment (Heaton et al., 1993). The WCST consists of 128 cards and four stimulus cards of various forms (crosses, circles, triangles, or stars), colors (blue, red, yellow, or green) and number of figures (one, two, three, or four). The client is instructed to match the deck of cards to the target cards. The examiner provides feedback of correct or incorrect. Without warning, the examiner changes the sorting principle and the client must learn from the examiners feedback. The test is discontinued when either the patient completes all 6 matching principles or finishes the deck of cards. This test is known to measure cognitive flexibility, problem solving, strategy, planning, and utilizing environmental feedback to shift cognitive sets (Heaton et al., 1993). Factor analysis found the WCST consists of the following three factors: cognitive flexibility, problem-solving, and response maintenance (Greve, Bianchini, et al., 2002). The normative sample for the WCST was created from 899 participants of six distinct samples (453 children and adolescents, 49 adults, 150 adults, 50 older adults, 124 commercial airline pilots, and 73 adults) (Heaton et al., 1993). The adult sample consisted of 384 participants that were of 20 years of age and older (M = 49.89, SD = 17.94). The mean education level was 14.95 (SD = 2.97). Overall, studies have confirmed that the WCST is a valid measure of executive functioning in the neuropsychological population (Heaton et al., 1993). Axelrod et al. (1992) found that interscorer reliability coefficients ranged from .895 to 1.000 for all 11 scores on the WCST. Whenever an individual is administered the WCST there is a 95% chance that his or her true score with be within 1.96 standard error of measurement of their obtained score (Heaton et. al., 1993).

Freestanding PVTs. The Test of Memory Malingering is a freestanding measure of effort (Tombaugh, 1996). The TOMM is a 50-item recognition test for adults, which includes two learning trials followed by recognition trial. The patients are shown pictures of common objects for three seconds each. An optional retention trial is available as well. When is it used Tombaugh (1996) indicated that this test is useful in aiding to identify poor effort. The large number of stimuli gives the impression that it is a difficult test, and additionally corrective feedback given allows for patients with good effort to increase their score and those with poor effort to negatively adjust their performance. The TOMM was normed with non-clinical participants in two phases. In the first phase, a preliminary version was used and the second phase was modified and used the two-choice format for the recognition format and feedback for correctness of each response. 405 participants ranged from 16 to 84 years (M = 54.8, SD = 20.2), and average education level was 13.1 years (SD = 3.2). The sample consisted of 47% males. On the first trial of the TOMM, 47.5/50 or 94% of the targets were correctly identified. On the other two trials the average number of targets correctly identified exceeded 99%. The second phase in the development and validation of the TOMM used a clinical sample of varying cognitive abilities, which included cognitive impairment, aphasia, traumatic brain injury, and dementia.

Results demonstrated similar highly accurate performance with non-impaired adults (Tombaugh, 1996). The dementia patients scored lower, but obtained a score of greater than 92% on trial 2. A criterion of 90% or 45/50 on the second recognition trial is suggested to assist in identifying patients with non-credible performance. This cutoff score had 100% specificity and sensitivity in a validation study (Tombaugh, 1997). Another study by Haber and Fichtenberg (2006) confirmed that using a cutoff score of 45 on Trial 2 had acceptable sensitivity and sensitivity to those exhibiting non-credible performance from those cognitively impaired from traumatic brain injury.

The Green's Word Memory Test (WMT) is another empirically validated freestanding PVT (Green et al., 1996). This computer administered assessment uses forced-choice paradigm where the patient chooses responses from a set of options. Semantically related word pairs are presented on the computer screen two times. Then there is an immediate recognition trial (IR) and a delayed recognition trial (DR) where the examinee is to choose the second word of the pair from the presented options. Feedback is provided for the correctness of responses. After the administration, the computer generates a report, which reports if the examinee "passed" or "failed" the subtests. A clear pass denotes scores that are above 90% correct. A clear fail denotes scores that are less than 82.5% correct, and caution indicates scores that are between 83% and 90% correct. The 82.5% cut-off for clear fail was created from 112 patients with moderate to severe TBI.

Norms for the WMT were created between 1995 and 2003 (Green, 2005). Sixty-one reference groups with over 3,000 participants available for comparison. For example, the normal adult controls (healthy volunteers) consisted of 40 participants with a mean age of 36.7 and mean education of 14 years. All participants in the adult normal group who passed the WMT effort subtests scored 90% or better on IR and DR. Research have examined its sensitivity and specificity as a PVT (Green, Iverson, & Allen, 1999; Green et al., 2001). Green et al. 1999 found that in 298 TBI patients (64 with moderate to severe TBI) that these individuals averaged over 90% correct on all of the WMT measures. Furthermore, Green, Flaro, & Courtney, (2009) demonstrated that false error rates are very rare.

Embedded PVTs. Recent research found acceptable classification accuracy in detecting non-credible performance in the Conner's Continuous Performance Test-II (CPT-II), Brief Testing Attention (BTA), and Trail Making Test (TMT) (Busse & Whiteside, 2012). Specifically, Busse and Whiteside's (2012) recommended cutoffs for suspected non-credible performance are as follows: BTA less than 16, TMT part B greater than 131 seconds, and CPT-II greater than 13 omission errors. Similarly, Whiteside, Wald, and Busse (2011) also found visuospatial measures such as the Judgment of Line Orientation (JLO), Benton Facial Recognition, and Rey Copy to have acceptable classification accuracy in detecting non-credible performance. The following cut-off scores were recommended: less than 16 on the JLO, less than 39 on the Benton Facial Recognition, and less than 23.5 on the Rey Copy.

In addition, the motor domain has also been explored as PVTs. Arnold et al. (2005) found the finger-tapping test to be acceptable in the detection of non-credible performance. The authors recommended a cut-off score of less than 33 for men and less than 28 for women.

Hypothesis Testing

If the tested WCST variables yield a AUC of .7 or greater and have acceptable sensitivity and specificity, then the hypothesis of this study will be supported that the WCST will have acceptable sensitivity and specificity in predicting effort performance among adults in a neuropsychological sample. That would provide evidence that the WCST can be used in conjunction with other indicators as a measure of non-credible performance among mild TBI patients.

Chapter IV: Results

The study consisted of 335 participants. The means and standard deviations for the demographic variables in each of the three samples were computed first. In the mTBI group, of 111 participants, 18 were placed in the FAIL group, for a FAIL base rate 16.2%. The mean age of the mTBI group was 45.4 (standard deviation (SD) = 12.3) with a mean of 13.4 years of education (SD = 2.1). In the STBI group, of 69 participants, 7 were placed in the FAIL group, for a FAIL base rate of 10.1%. The mean age of the STBI group was 44.2 years (SD = 12.1) with 13.2 years of education (SD = 2.5). In the PSYCH group, of 155 participants, 19 were placed in the FAIL group, for a FAIL base rate of 12.3%. The PSYCH group's mean age was 46.7 (SD = 13.6) with a mean of 13.5 (2.4) years of education. The means and standard deviations of age, education, and percentage of participants who were compensation seeking for all samples are provided in Table 3. Table 4 displays other demographic characteristics, including the percentages of female participants, ethnicity marital status, and handedness of all samples. Table 5 shows a list of psychiatric diagnoses for each of the three samples.

The demographic differences between the mTBI, STBI, and PSYCH groups can be found in Table 6. Within each sample, no significant demographic differences were found between the PASS and FAIL group. Group means and standard deviations for each of the three samples (mTBI, STBI, and PSYCH) were calculated for each of the WCST variables. No group differences were found between the three samples in regards to age, education, ethnicity, marital status, handedness, or PASS/FAIL base rate as presented in Table 7. However, significant group differences were found in gender and compensation status between the three samples and are also presented in Table 7. There were significantly more males in the STBI sample compared to the mTBI and PSYCH samples. Furthermore, there was significantly more compensation seeking referrals in the mTBI group compared to the PSYCH and STBI groups.

This chapter will first discuss the group differences between the PASS and FAIL groups using Mann-Whitney U, and then the classification accuracy of the WCST variables using ROC analyses for the primary sample of interest, the mTBI group. Using the MTBI sample, a logistic regression analysis was computed to derive a combined WCST variable, and this will be discussed in terms of the Mann-Whitney U and ROC analysis.

To evaluate the generalizability of the WCST individual and logistically derived combined measures, the Mann-Whitney U and ROC analyses will be discussed for the known neurological injury sample, the STBI group, and then the PSYCH sample. In the ROC analyses of all three groups, all variables with acceptable classification accuracy are defined as having an AUC of at least. 70. Recommended cutoff scores' sensitivity, specificity, positive predictive power, and negative predictive power will be discussed.

Mild Traumatic Brain Injury Sample

Mean group differences. As reported in Table 8, the Independent-Samples Mann-Whitney U Test indicated no significant group differences between the PASS and FAIL groups on any of the individual WCST variables, including Number of Trials, Number Correct, Total Errors, Perseverative Errors, Perseverative Responses, Nonperseverative Errors, Trials to First Category, Number of Categories, and FMS.

Logistic regression analysis. In addition to examining the individual WCST variables, a combined WCST variable was calculated using a logistic regression and the results are presented in Table 9. To reduce the risk of multicolinearity, Spearman's rho correlations were calculated between the WCST variables and variables that had the lowest correlations. The analysis found the following WCST variables to be least correlated: Number Correct, Perseverative Responses, and Trials to First Category. Number Correct and Perseverative Responses had a correlation of .061, Number Correct and Trials to First Category had a correlation of -.003, and Perseverative Responses and Trials to First Category had a correlation of .326. The remaining WCST variables were excluded from the logistic regression analysis due to significant correlations of 0.70 or higher. Therefore, only Number Correct, Perseverative Responses and Trials to First Category were entered into the logistic regression to calculate the combined WCST variable, called WCSTCOMB. For the logistic regression, a test of the full model against a constant only model was statistically reliable, likelihood ration (LR) ($x^2(3) = 306.97$, $p < .05$), indicating that the predictors as a set reliably differentiated between credible and non-credible performance on the WCST. The variance accounted for by this model was small, Nagelkerke's $R^2 = .03$. Overall, the model correctly identified 86.1% of cases included. The Wald Criterion indicated that only Perseverative Responses ($p < .05$) made significant contributions to the predictive power of the model while Trials to First Category ($p < .92$ and Number Correct ($p < .66$) was not a significant predictor The Exp (B) value indicates that when Perseverative Responses is raised by one unit, the odds ratio is .99 times as large.

Thus, the following logistically derived variable was calculated: WCSTCOMB = NumberCorrect*.005+PerseverativeResponses*-.014=TrialsToFirstCategory*-.001+1.796. The same logistically derived combined variable was then used in the analyses with the STBI and PSYCH samples. A Mann-Whitney U was calculated for the WCSTCOMB, which was not significantly different between tbe PASS and FAIL groups for the mTBI sample. The results of this analysis are presented in Table 10.

Receiver Operating Curve (ROC) analysis. The ROC analysis was calculated to evaluate the overall classification accuracy of the individual WCST variables and WCSTCOMB. For this study, the standard ROC interpretation criterion was used, specifically that area under the curve (AUC) between .7-.79 is considered to have "acceptable accuracy" (Hosmer & Lemeshow, 2000; Larrabee & Berry, 2007). Furthermore, predictor variables with an AUC of .5 are considered to have "chance level" accuracy and measures with an AUC between .8-.89 are "excellent" and .9 and above are "outstanding". Therefore, only variables that have an AUC of at least .7 had sensitivity, specificity, negative predictive power, positive predictive power, and cutoff scores calculated.

The ROC analysis for the mTBI sample for both the individual WCST variables and the three combined variables are presented in Table 11. No "acceptable" classification accuracy levels were found for any of the individual or combined predictor variables.

Moderate to Severe Traumatic Brain Injury Sample

Mean group difference. Independent-Samples Mann Whitney U Tests found significant differences between the PASS and FAIL groups for WCST variables as presented in Table 12. Number of Trials, Total Errors, Perseverative Errors, Perseverative Responses, and Number of Categories were significantly different between the PASS and FAIL group. Independent t-tests and Chi Square analyses found no significant demographic differences between the PASS and FAIL groups in the STBI sample and are presented in Table 3. A Mann-Whitney U was calculated for the WCSTCOMB, which was significantly different between tbe PASS and FAIL groups. The results of this analysis are presented in Table 11.

Receiver Operating Curve (ROC) analysis. The ROC analysis was then calculated to evaluate the overall classification accuracy of the individual WCST variables and for the logistically derived WCSTCOMB variable. Unlike the mTBI sample, in the STBI sample there were several variables that reached "acceptable" to "excellent" classification accuracy levels and are presented in Table 13. The following variables were found to have "acceptable" classification: Number of Trials = .74, Total Errors = .76, Number of Categories = .71, and WCSTCOMB = .79. Perseverative Responses and Perseverative Errors were found to have "excellent" classification accuracy of .81 and .80 respectively. The recommended cutoff scores, sensitivity, specificity, positive predictive power (PPP), and negative predictive power (NPP) were calculated for these variables and are presented in Table 14. As noted in Table 14, a 10% PVT failure base rate was used for the calculation of PPP and NPP.

PSYCH Sample

Mean group differences. Independent-Samples Mann Whitney U Tests were used to analyze mean group differences between the PASS and FAIL groups among the individual and combined WCST variables and are presented in Table 15. Significant PASS versus FAIL group differences was found for Number of Trials, Total Errors, Perseverative Errors, Perseverative Responses, and Number of Categories. A Mann-Whitney U was calculated for the WCSTCOMB, which was not significantly different between tbe PASS and FAIL group. The results of this analysis are presented in Table 10.

Receiving Operating Curve (ROC) analysis. The ROC analysis calculated the overall classification accuracy of the WCST variables in the PSYCH sample. As shown in Table 16, Perseverative Responses reached "acceptable" classification accuracy levels of .71, while WCSTCOMB also had "acceptable" classification accuracy of .70. Therefore, the recommended cutoff score, sensitivity, specificity, PPP, and NPP were calculated and are provided in Table 17. As shown on Table 17, a 14% failure PVT base rate was used for this sample. The base rate of 14% reflects the estimated percentage of non-credible performance in psychiatric populations (Mittenberg et. al, 2002).

Chapter V: Discussion

Purpose of Study

The purpose of this study was to expand the literature on the WCST as an embedded PVT in neurological and non-neurological samples. There has been limited research on the WCST as a PVT and the current literature has found inconsistent results in regards to the accuracy of the WCST as a validated PVT. Most research on PVTs has been in the memory domain, and it is important to assess for multiple PVTs across a variety of cognitive domains as cognitive effort fluctuates over the course of a neuropsychological evaluation (Boone, 2009). Therefore, additional research is needed to determine if the WCST is a validated PVT in the executive functioning domain. In addition, the study aimed to add to the literature by using two comparison groups (PSYCH and STBI) in addition to the mTBI sample of interest. So the study sought to determine if the WCST as a PVT could be generalized to both neurological and non-neurological populations. Finally, the study sought to assess if the individual WCST variables and combined WCST variables reached "acceptable" classification accuracy.

Hypothesis Testing

The first hypothesis, that the FAIL and PASS groups would perform significantly worse on the WCST measures was only partially supported. Specifically, only the STBI and PSYCH samples, and not the mTBI sample had statistically significant differences on the individual and combined WCST measures between the various PASS And FAIL groups. Furthermore, the second hypothesis was not supported as none of the individual or combined WCST variables reached acceptable classification accuracy in the mTBI sample. However, both the third and fourth hypotheses were supported as "acceptable" to "excellent" classification accuracy was found in the STBI group and "acceptable" classification accuracy was found in the PSYCH group for at least some of the individual WCST measures and WCSTCOMB.

Summary of Results

The results of this study failed to support the use of the WCST as an embedded PVT in an mTBI sample. Furthermore, the first hypothesis was not fully supported, as there were only small to negligible group differences that were not statically significant between the PASS and FAIL groups in the mTBI sample. However, there were statistically significant differences in the WCST between the PASS and FAIL groups in the STBI and PSYCH samples. In addition, hypothesis two was not supported due to lack of significant findings on individual WCST predictor variables and a logistically derived combined predictor variable also failed to reach acceptable classification accuracy among the mTBI sample. Unlike the insignificant findings in the mTBI group, the STBI and PSYCH samples found support for the use of the WCST as a PVT.

It is interesting that the WCST was not accurate in detecting non-credible performance among those with mTBI. While it is not possible to identify the specific reason for these non-significant findings, true deficits in WCST performance in mTBI patients are unlikely. This is because individuals with mTBI typically do not present with genuine executive functioning deficits or significant impairments in the WCST (Rohling et al. 2011). Therefore, low scores on the WCST are not expected in the mTBI population. Individuals with uncomplicated mTBI typically recover to baseline levels of functioning by three months after their injury (Carone & Bush, 2013; McCrea, 2008; Rohling et al., 2011), and all of the participants in this study were evaluated after three months post-injury. Consequently, the insignificant findings in the mTBI sample are likely not due to confounding factors such as true neurocognitive deficits.

Larrabee and Berry (2007) speculated why the sensitivity is low among the WCST as a PVT. The authors reasoned that even without neurocognitive deficits that significant individuals differences are present in the population. He stated, "for the WCST, demographic factors accounted for as much as 20% of the observed variance" (Larrabee & Berry, 2007, p. 213). Heaton et al. (1993) also highlighted the importance of age and education was especially important. The groups were well matched for age and education and there was not a statistical significance difference between the PASS and FAIL groups. However, the use of raw scores may have influenced the results. It is possible that different result could occur if standard WCST scores were used.

A second possibility is that individuals who are exhibiting non-credible performance may not see the WCST as relevant to brain injury (Greve & Bianchini, 2002; Greve et al., 2009; Larrabee & Berry, 2007;). The WCST is less face valid than other neuropsychological measures and individuals may fail to perceive the WCST as a measure of neurocognitive dysfunction. Instead, individuals may choose to exhibit poor effort on face valid tasks of memory, which is a more common complaint amongst mTBI patients. As for individuals who are not malingering but rather just exhibiting poor effort, the WCST may be engaging compared to other traditional pen and paper neuropsychological tasks. The WCST may standout compared to other tasks due to the game like appearance with the use of playing cards.

A third possibility is that individuals with mTBI may lack homogeneity in non-credible performance on the WCST. For example, an individual may decide to avoid getting too many consecutive correct responses, or avoid matching to any of the categories, or not incorporate feedback from the examiner and continue to sort to only one principle. Furthermore, an individuals may decide to use all or a combination of these strategies. Therefore, the variety of ways to perform poorly on the WCST is varied and therefore, this may lower the sensitivity of the WCST as a PVT among the mTBI population.

STBI group results. Unlike the mTBI group, the STBI sample revealed statistically significant differences between the PASS and FAIL groups for the following WCST's variables: Number of Trials, Total Errors, and Perseverative Errors. Next, hypothesis number four was supported in that the ROC analysis found "acceptable" classification accuracy for the following individual WCST variables: Number of Trials, Number of Categories Completed, Total Errors, and Perseverative Errors. Additionally, the logistically derived combined variable, WCSTCOMB, that combined Number Correct, Perseverative Responses, and Trials to First category, also found "acceptable" classification accuracy. However, both Number of Trials and Number of Categories Completed failed to reach acceptable sensitivity of 90%, and therefore recommended cutoff scores were not determined. Total Errors had a specificity of 90%, sensitivity of 14%, positive predictive power (PPP) of .13, Negative Predictive Power (NPP) of .90, and a recommended cutoff score of 74. Perseverative Errors had a specificity of 94%, sensitivity of 29%, PPP of .35, NPP of .92, and a recommended cutoff score of 43. Even more significant, Perseverative Responses alone were found to have "excellent" classification accuracy. Perseverative Responses had a specificity of 92%, specificity of 43%, PPP of .37, NPP of .94, and a recommended cutoff score of 47. Finally, the WCSTCOMB had a specificity of 90%, sensitivity of 57%, PPP of .39, NPP of .95, and recommended a cutoff score of 1.53. The PSYCH sample found fewer significant findings.

PSYCH group results. The PSYCH sample was statistically significantly different from the PASS and FAIL group for the following WCST variables:

Number of trials, Total Errors, Perseverative Errors, Perseverative Responses, Trials to First Category, and number of Categories Completed. Hypothesis three was partially supported as the ROC analysis found perseverative responses and WCSTCOMB to have "acceptable" classification accuracy in the PSYCH sample. Perseverative Responses had a specificity of 91%, sensitivity of 11%, PPP of.17, NPP of .86, and a recommended cutoff score of 46. Finally, WCSTCOMB had a specificity of 90%, sensitivity of 16%, PPP of .21, NPP of .87, and a recommended cutoff score of 1.44.

Clinical Applications

This study provided support for the use of both individual and combined variables of WCST as an embedded PVT in a neurological population, specifically those with moderate to severe traumatic brain injury, and for a non-neurological population, specifically psychiatric outpatients. Unfortunately, the study did not provide support for the WCST in mTBI patients. No research has previously been completed on the WCST as an embedded PVT among an outpatient non-neurological psychiatric population. Furthermore, the significant findings for the STBI sample were not consistent with previous literature. For example, many studies found the false positive error rates to be unacceptably high for WCST variables among those with moderate to severe TBI (Donders, 1999; Greve & Bianchini, 2002; Heinly et al.; 2006). However, most of these studies used both the Bernard and Suhr formulas with individuals with moderate and severe TBI and this study also found unacceptable classification accuracy for the Bernard and Suhr formulas among those with moderate to severe TBI. Therefore, none of the studies explored the classification accuracy, sensitivity, and specificity of the current individual WCST variables (Number of Trials, Total Errors, Perseverative Errors, Perseverative Responses, and WCSTCOMB) that were found to be acceptable in this study.

The current literature on the WCST as an embedded PVT has found inconsistent results (Greve & Bianchini, 2002). Therefore, the results of this study were only partially consistent with the existing literature. This study was not consistent with the findings on the Bernard or Suhr formulas for those with mTBI (Bernard et al., 1996; Suhr & Boyer, 1999). Furthermore, the Suhr and Bernard et al. formulas were created using simulators who were asked to feign malingering. The current sample utilizes clinical participants, which may account for the insignificant findings in the mTBI group compared to the Bernard et al. (1996) and Suhr and Boyer (1999) studies. The current findings in this study were more consistent with the Heinly et al. (2006) findings that the Bernard formula was not effective in accurately classifying poor effort among those with mTBI. However, Heinly et al. (2006) found that the WCST as a PVT is more accurate among those with mTBI over those with moderate and severe TBI, which are the opposite findings of this current study. This study also conflicted with Greve et al. (2009), which found the WCST might be helpful in detecting malingering among those with mTBI. Another study found that the Suhr and Bernard formulas were not effective in identifying malingers with mTBI, which is consistent with this study (Sweet & Nelson, 2007).

Finally, this study examined the sensitivity, specificity, positive predictive power, and negative predictive power of WCST variables that demonstrated "acceptable" or "excellent" classification accuracy. The specificity for these measures with adequate classification accuracy has been set at the standard level or at least 90% specificity. For those with outpatient psychiatric conditions without neurological disorders, a raw cutoff score of 48 is recommended for perseverative responses, which had a specificity of 90% and sensitivity of 15%. For those post-acute outpatient individuals with moderate to severe TBI the following raw cutoff scores are recommended: 74 for Total Errors (sensitivity of 20% and specificity of 90%), 39 for Perseverative Errors (sensitivity of 30% and specificity of 90%), and 47 for Perseverative Responses (sensitivity of 40% and specificity of 90%). In addition to the individual WCST measures, the regression derived WCSTCOMB had a sensitivity of 20% and specificity of 90% among the STBI sample. However, it is strongly recommended with these embedded WCST measures that they are not used alone, but in combination with other empirically validated embedded and freestanding PVTs (Boone, 2009). This is because embedded PVTs are known to have lower sensitivity than most free standing PVTs (Larrabee & Berry, 2007). Therefore, only using one embedded PVT often has an unacceptably low sensitivity for detecting non-credible performance. However, the specificity of embedded PVTs is high. It is recommended to add multiple measures of effort to increase the sensitivity and accuracy of detecting individuals with non-credible performance (Boone, 2009; Larrabee & Berry, 2007). Also, it is not rare to fail one embedded PVT (Dean, Victor, Boone, Philpott, & Hess, 2008). Therefore, in order to reduce false negatives in identifying non-credible effort, using multiple embedded measures of effort is recommended.

Limitations and Future Directions

There are several limitations to this study. First, medication status was not available for analysis for this study. Therefore, medication effects are known to impact neuropsychological testing and could have perhaps lowered the sensitivity of the WCST in all three samples. It would be beneficial if future studies included medication status in their analysis.

The sample was not diverse in terms of ethnicity. The database from which this study was conducted contains a higher percentage of Caucasian participants. Therefore, there is question of the generalizability of this study to ethnic and racial backgrounds other than Caucasian. There were only 2 out of 98 individuals from a different ethic background other than Caucasian in the PASS group and 3 out of 21 in the FAIL group. It is recommended that future WCST PVTs studies contain ethnically and racially diverse samples in order to determine generalizability of these findings. This is especially significant, as ethnic differences in assessing credible performance have not been explored. For example, Lilienfel, Thames, & Watts (2013) have highlighted that the impact of ethnicity and culture on the interpretation of PVTs has not been explored and the authors strongly urged for future research in this area. Furthermore, there are assessments that use different cutoffs and scores based off of ethnicity. For example, the Heaton norms factors ethnicity into determining standard scores (Heaton, Miller, Taylor, & Grant, 2004). It would be interesting to see if using different PVT cutoffs are needed for different ethnicities.

It is also important to note that the participants were primarily a product of the U.S. educational system and therefore the generalizability to individuals educated in different cultures is unknown. In future studies, the generalizability of the WCST as a PVT in other cultures would be beneficial to examine. In addition, the sample was more highly educated compared to the average population. However, the generalizability to individuals without a high school degree would be worth exploring in future research. Future research could examine if lower education impacts credible versus non-credible performance in the WCST.

Additionally, the occurrence and length of loss of consciousness in the mTBI sample was not available for this study. Therefore, it may be helpful for future studies to compare mTBI individuals with and without loss of consciousness. Another limitation of this study is that neuroimaging was not available for this study. For example, in the moderate to severe traumatic brain injury group, the localization of the injury and specific damage to regions of the brain were unknown. It would be beneficial if future studies could compare the regions of the brain affected to see if that has an impact on the WCST as a PVT.

Another limitation of this study was the unavailability of some WCST measures. For example, this study did not have Unique Responses available in the archival database. This measures the number of times an individual sorts the cards to a category that is not possible. Additionally, this study did not have percentage of Conceptual Level Responses available. Finally, raw WCST scores were used in this study instead of standard scores. Standard scores control for age and education in the WCST. Furthermore, Heaton et al. (1993) indicated that age and education plays a significant role in WCST performance. Therefore, future research utilizing WCST standard scores would be beneficial.

For the first time, this study examined a non-neurological mixed outpatient psychiatric population in studying the WCST as a PVT. This is the first known study that has explored the impact of the WCST as a PVT in this population and therefore there is a need for future research. Future research could replicate this study to see if similar positive findings are found within this population. Furthermore, it would be clinically valuable to determine if there is a difference between inpatient and outpatient psychiatric populations in credible

WCST performance. Future studies should also control for medication and see if psychotropic medication status influences credible and non-performance on the WCST.

Finally, future research should further examine the heterogeneity in terms of non-credible performance styles in the WCST. Research has acknowledged and discussed that there are multiple ways that an individuals could display non-credible performance on the WCST (avoid getting too many correct in a row, perseverating, or having unique responses) (Greve, Bianchini, et al., 2002; Larrabee & Berry, 2007). Future research could examine if there are factors that can accurately predict what way an individual will perform non-credibly. If the heterogeneity of non-credible performance is better understood, then the sensitivity of the WCST as a PVT may increase.

References

Alvarez, J. A., & Emory, E. (2006). Executive function and the fontal lobes: A meta-analytic review. *Neuropsychology Review, 16,* 17-42. doi:10.1007/s11065-006-9002-x

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders: DSM-IV-TR. Washington, DC: Author.

Anderson, S. W., Damasio, H., Jones, R. D., & Tranel, D. (1991). Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. *Journal of Clinical and Experimental Neuropsychology, 13,* 909-922. doi:10.1080/01688639108405107

Archer, R. P., Buffington-Vollum, J. K., Strendy, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87,* 84-94. doi:10.1207/s15327752jpa8701_07

Arnold, G., Boone, K., Lu, P., Dean, A., Wen, J., Nitch, S., & McPhersons (2005). Sensitivity and specificity of finger tapping test scores for the detection of suspect effort. *The Clinical Neuropsychologist, 19,* 105-120. doi:10.1080/13854040490888567

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences, 8,* 170-177. doi:10.1016/j.tics.2004.02.010

Ashendorf, L., Constantinou, M., & McCaffery, R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives of Clinical Neuropsychology, 19,* 125-130. doi:10.1093/arclin/19.1.125

Ashendorf, L, O'Bryant, S. E., & McCaffrey, R. J. (2003). Specificity of malingering detection strategies in older adults using the CVLT and WCST. *The Clinical Neuropsychologist, 17,* 255-262. doi:10.1076/clin.17.2.255.16502

Axelrod, B. N., Goldman, B. S., & Woodard, J. L., (1992). Interrater reliability in scoring the Wisconsin Card Sorting Test. *The Clinical Neuropsychologist, 6,* 143-155. doi:10.1080/13854049208401851

Baddeley, A. (1992). Working memory. *Science, 255,* 556-559. doi:10.1126/science.1736359

Baddeley, A., Sala, S. D., & Robbins, T. W. (1996). Working memory and executive control. *Biological Sciences, 351,* 1378-1388. doi:10.1098/rstb.1996.0123

Barceló, F. (2001) Does the Wisconsin Card Sorting Test measure prefrontal function? *The Spanish Journal of Psychology, 4,* 79-100. doi:10.1017/S1138741600005680

Barcelo, F., & Knight, R. T. (2002). Both random and perspective errors underlie WCST deficits in prefrontal patients. *Neuropsychologia, 43*, 349-356. doi:10.1016/S0028-3932(01)00110-5

Batt, K., Shores, E. A., & Chekaluk, E. (2008). The effect of distraction on the Word Memory Test and Test of Memory Malingering performance in patients with a severe brain injury. *Journal of International Neuropsychological Society, 14,* 1074-1080. doi:10.1017/S135561770808137X

Benge, J. F., Caroselli, J. S., & Temple, R. O. (2007). Wisconsin Card Sorting Test: Factor structure and relationship to productivity and supervision needs following severe traumatic brain injury. *Brain Injury, 21,* 395-400. doi:10.1080/02699050701311091

Benton, A. L., Sivan, A. B., Hamsher, K. D., Varney, N. R., & Spreen, O. (1994). *Contributions to neuropsychological assessment: A clinical manual second edition.* New York, NY: Oxford University Press.

Berman, K. F., Ostrem, J. L., Randolph, C., Gold, J., Goldberg, T., E., Coppola, R.,... Weinberger, D. R. (1995). Physiological activation of a cortical network during performance on the Wisconsin Card Sorting Test: A positron emission tomography study. *Neuropsycholgia, 33,* 1027-1046. doi:10.1016/0028-3932(95)00035-2

Bernard, L. C., McGrath, M. J., & Houston, W. (1996). The differential effects of simulating malingering, closed head injury, and other CNS pathology on the Wisconsin Card Sorting Test: Support for the "patter of performance" hypothesis. *Archives of Clinical Neuropsychology, 11,* 231-245. doi:10.1093/arclin/11.3.231

Binder, L. M. (1997). A review of mild head trauma: II. Clinical implications. *Journal of Clinical and Experimental Neuropsychology, 19,* 432-457. doi:10.1080/01688639708403871

Binder, L. M., & Rohling, M. L. (1996). Money matters: A meta-analytic review of the effects of financial incentives on recovery after closed head injury. *American Journal of Psychiatry, 153,* 7-10. doi:10.1097/00001199199608000-00012

Binder, L. M., Rohling, M. L., & Larrabee, G. J. (1997). A review of mild head trauma. Part I: Meta-analytic review of neuropsychological studies. *Journal of Clinical and Experimental Neuropsychology, 19,* 421-431. doi:10.1080/01688639708403870

Bledowski, C. N., Kaiser, J., & Rahm, B. (2010). Basic operations in working memory: Contributions from functional imagining studies. *Behavioral Brain Research, 214,* 172-179. doi:10.1016/j.bbr.2010.05.041

Bodin, D., Yeates, K. O., & Klamar, K. (2012). Definition and classification of concussion. In J.N. Apps, & K.D. Walter (Eds.), *Pediatric and*

Adolescent Concussion: Diagnosis, Management and Outcomes. New York, NY: Springer.

Boone, K. B. (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological examinations. *The Clinical Neuropsychologist, 23,* 729-741. doi:10.1080/13854040802427803

Borkowska, S. C., & Rybakowski, J. (2001). Neuropsychological frontal lobe tests indicate that bipolar depressed patients are more impaired than unipolar. *Bipolar Disorders, 3,* 88-94. doi:10.1034/j.1399-5618.2001.030207.x

Busse, M., & Whiteside, D. M. (2012). Detecting suboptimal cognitive effort: Classification accuracy of the Conner's Continuous Performance Test-II, Brief Test of Attention, and Trail Making Test. *The Clinical Neuropsychologist, 26,*1-13. doi:dx.doi.org/10.1080/13854046.2012.679623

Carone, D. A., & Bush, S. S. (2013). *Mild traumatic brain injury: Symptom validity assessment and malingering.* New York, NY: Springer Publishing.

Carroll, L. J., Cassidy, J. D., Holm, L., Kraus, J., & Coronado, V. G. (2004). Methodological issues and research recommendations for mild traumatic brain injury: The WHO Collaborating Centre Task Force on Mild

Traumatic Brain Injury. *Journal of Rehabilitation Medicine, 43,* 113-125. doi:10.1080/16501960410023877

Centers for Disease Control and Prevention (CDC) (2003). *National Center for Injury Prevention and Control. Report to Congress on mild traumatic brain injury in the United States: Steps to prevent a serious public health problem.* Retrieved from: http://www.cdc.gov/ncipc/pubres/mtbi/mtbireport.pdf

Chung, H. J., Weyandt, L. L., & Swentosky, A. (2014). The physiology of executive functioning. In S. Goldstein, & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 13-27). New York, NY: Springer.

Conners, C. K., & MHS Staff (Eds.) (2000). *Conners' Continuous Performance Test II: Computer program for Windows technical guide and software manual.* North Tonwanda, NY: Multi-Health Systems.

Dean, A. C., Victor, T. L., Boone, K. B., Philpott, L. M., & Hess, R. A. (2008). Dementia and effort test performance. *The Clinical Neuropsychologist, 23,* 133-152. doi:10.1080/13854040701819050

Demakis, G. J. (2003). A meta-analytic review of the sensitivity of the Wisconsin Card Sorting Test to frontal and lateralized frontal brain damage. *Neuropsychology, 17,* 255-264. doi:10.1037/0894-4105.17.2.255

Donders, J. (1999). Brief report: Specificity of a malingering formula for the Wisconsin Card Sorting Test. *Journal of Forensic Neuropsychology, 1,* 35-42. doi:10.1300/J151v01n03_03

Eling, P., Derckx, K., & Maes, R. (2008). On the historical and conceptual background of the Wisconsin Card Sorting Test. *Brain and Cognition, 67,* 247-253. doi:10.1016/j.bandc.2008.01.006

Frencham, K. A., Fox, A. M., & Maybery, M. T. (2005). Neuropsychological studies of mild traumatic brain injury: A meta-analytic review of research since 1995. *Journal of Clinical and Experimental Neuropsychology, 27,* 334-351. doi:10.1080/1303390490520328

Goldman-Rakic, P. S., Thierry, A. M., Glowinski, J. & Christen, Y. (1994). *Motor and cognitive function of the prefrontal cortex.* Berlin, Germany: Springer.

Goldstein, S., Naglieri, J. A., Princiotta, D., & Otero, T. M. (2014). Introduction: A history of executive functioning as a theoretical and clinical construct. In S. Goldstein, & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 3-12). New York, NY: Springer.

Green, P. (2005). *Green's Word Memory Test for Windows user's manual.* Edmonton, Canada: Green's Publishing Inc.

Green, P., Allen, L. M., & Astner, K. (1996). *The Word Memory Test: A user's guide to the oral and computer-administered forms, US version 1.1* Durham, NC: Cognisyst.

Green, P., Flaro, L., & Courtney, J. (2009). Examining false positives on the Word Memory Test in adults with mild traumatic brain injury. *Brain Injury, 23,* 741-740. doi:10.1080/02699050903133962

Green, P., Iverson, G. L., & Allen, L. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury, 3,* 813-819. doi:10.1080/026990599121205

Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M., III. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury, 15,* 1045-1060. doi:10.1080/02699050110088254

Greiffenstein, M. F., & Baker, W. J. (2006). Miller was (mostly) right: Head injury severity inversely related to simulation. *Legal and Criminological Psychology, 11,* 131-145. doi:10.1348/135532505X49828

Greve, K. W., & Bianchini, K. J. (2002). Using the Wisconsin Card Sorting Test to detect malingering: An analysis of the specificity of two methods in nonmalingering normal and patient samples. *Journal of Clinical and Experimental Neuropsychology, 24,* 48-54. doi:10.1076/jcen.24.1.48.968

Greve, K. W., Bianchini, K. J., Mathias, C. W., Houston, R. J., & Crouch, J. A. (2002). Detecting malingered performance with the Wisconsin Card Sorting Test: A preliminary investigation in traumatic brain injury. *The Clinical Neuropsychologist, 16,* 179-191. doi:10.1076/clin.16.2.179.13241

Greve, K. W., Heinly, M. T., Bianchini, K. J., & Love, J. M. (2009). Malingering detection with the Wisconsin Card Sorting Test in mild traumatic brain injury. *The Clinical Neuropsychologist, 23,* 343-362. doi:10.1080/13854040802054169

Greve, K. W., Love, J. M., Sherwin, E., Mathias, C. W., Ramzinksi, R., & Levy, J. (2002). Wisconsin Card Sorting Test in chronic severe traumatic brain injury: factor structure and performance subgroups. *Brain Injury, 16,* 29-40. doi:10.1080/02699050110088803

Haber, A. H., & Fichtenberg, N. L. (2006). Replication of the Test of Memory Malingering (TOMM) in a traumatic brain injury and head trauma sample. *The Clinical Neuropsychologist, 20,* 524-532. doi:10.1080/13854040590967595

Haines, J., Hoffman, W., Hanada, J., Brown, P., Labs, S., Loberg, D., & Walter, C. T. (1994). Further evaluation of prefrontal lobe function data in schizophrenic patients during Wisconsin Card Sorting Test. *American Journal of Psychiatry, 151,* 1842-1843. Retrieved from: http://ajp.psychiatryonline.org/article.aspx?articleID=170724&resultClic =3

Halstead, W. C., & White, J. (1950). *Manual for the Halstead Battery of Neuropsychological Tests.* Chicago, IL: University of Chicago Press.

Hashimoto, K., Uruma, G., & Abo, M. (2008). Activation of the prefrontal cortex during the Wisconsin Card Sorting Test (Keio Version) as measured by two-channel near-infrared spectroscopy in patients with traumatic brain injury. *European Neurology, 59,* 24-30. doi:10.1159/000109257

Heaton, R. K. (1981). *A manual for the Wisconsin Card Sorting Test.* Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test Manual: Revised and expanded.* Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Miller, W., Taylor, M., & Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted norms for African American and Caucasian adults (HRB).* Lutz, FL: Psychological Assessment Resources, Inc.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. E., Millis, S. R., & Conference Participants (2009). American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23,* 1093-1129. doi:10.1080/13854040903155063

Heinly, M. T., Greve, K. W., Love, J. M., & Bianchini, K. J. (2006, February). *Sensitivity and specificity of Wisconsin Card Sorting Test indicators of* *potential malingered neurocognitive dysfunction in traumatic brain injury.* Poster presented at the 34^{th} annual meeting of the International Neuropsychological Society, Boston, MA.

Heubrock, D., & Petermann, F. (1998). Neuropsychological assessment of suspected malingering: Research results, evaluations techniques, and further directions of research and application. *European Journal of Psychological Assessment, 14,* 211-225. doi:10.1027/1015-5759.14.3.211

Holtz, J. L. (2011). *Applied Clinical Neuropsychology.* New York, NY: Springer.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression.* New York, NY: Wiley.

Ilonen, T., Taiminen, T., Karlsson, H., Lauerma, H., Tuimala, P., Leinonen, K M.,...Salokangas, R. (2000). Impaired Wisconsin Card Sorting Test performance in first-episode severe depression. *Nordic Journal of Psychiatry, 54,* 275-280. doi:10.1080/080394800448156

Iverson, G. L., Page, J. L., Koehler, B. E., Shojania, K., & Badii, M. (2007). Test of Memory Malingering (TOMM) scores are not affected by chronic pain or depression in patients with fibromyalgia. *The Clinical Neuropsychologist, 21,* 532-546. doi:10.1080/13854040600611392

Kawasaki, Y., Maeda, Y., Suzuki, M., Urata, K., Higashima, M., Kiba, K.,...Hisada, K. (1993). SPECT analysis of-regional cerebral blood flow changes in patients with schizophrenia during the Wisconsin Card Sorting

Test. *Schizophrenia Research, 10,* 109-116. doi:10.1016/0920-9964(93)90045-K

King, N. S., & Kirwilliam, S. (2011). Permanent post-concussion symptoms after mild head injury. *Brain Injury, 25,* 462-470. doi:10.3109/02699052.2011.558042

King, J. H., Sweet, J. J., Sherer, M., Curtiss, G., & Vanderploeg, R. D. (2002). Validity indicators within the Wisconsin Card Sorting Test: Application of new and previously researched multivariate procedures in multiple traumatic brain injury samples. *The Clinical Neuropsychologist, 16,* 506-523. doi:10.1076/clin.16.4.506.13912

Konishi, S., Nakajima, K., Uchida, I., Kameyama, M., Nakahara, K., Sekihara, K., & Miyashita, Y. (1998). Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nature Neuroscience, 1,* 80-84. doi:10.1038/283

Larrabee G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist, 17,* 410-425. doi:10.1076/clin.17.3.410.18089

Larrabee G. J. (Ed.) (2012). *Forensic Neuropsychology: A scientific approach.* Oxford, England: University Press.

Larrabee G. J., & Berry, D. T. (2007). Diagnostic classification statistics and diagnostic validity of malingering assessment. In G. J., Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 14-25). Oxford, England: University Press.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological Assessment.* Oxford, England: Oxford University Press.

Lilienfeld, S. O., Thames, A. D., & Watts, A. L. (2013). Symptom validity testing: Unresolved questions, future directions. *Journal of Experimental Psychopathology, 4,* 78-87. doi:http://dx.doi.org/10.5127/jep.028312

Marenco, S., Coppola, R., Daniel, D. G., Zigun, J. R., & Weinberger, D. R. (1993). Regional cerebral blood flow during the Wisconsin Card Sorting Test in normal subjects studied by xenon-133 dynamic SPECT: Comparison of absolute values, percent distribution values, and covariance analysis. *Psychiatry Research: Neuroimaging, 50,* 177-192. doi:10.1016/0925-4927(93)90029-H

Martin, D. J., Oren, Z., & Boone, K. (1991). Major depressives' and dysthymics' performance on the Wisconsin Card Sorting Test. *Journal of Clinical Psychology, 47,* 684-690. doi:0.1002/1097-4679(199109)47:5<684::AID JCLP2270470509>3.0.CO;2-G

McCrea, M. A. (2008). *Mild traumatic brain injury and postconcussion syndrome: The new evidence base for diagnosis and treatment* (AACN workshop series). New York, NY: Oxford University Press.

Mentzel, H. J., Gaser, C., Volz, H. P., Rzanny, R., Hager, F., Sauer, H., & Kaiser, W. A. (1998). Cognitive stimulation with the Wisconsin Card Sorting Tests: Functional MR imaging at 1.5 T. *Radiology, 207*, 399-404. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/9577487?dopt=Citation

Meyers, J., & Meyers, K. (1995). *The Meyers scoring system for the Rey Complex Figure and the Recognition Trial: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Mild Traumatic Brain Injury Committee, American Congress of Rehabilitation Medicine (ACRM) Head Injury Interdisciplinary Special Interest Group. (1993). Definition of mild traumatic brain injury. *Journal of Head Trauma Rehabilitation, 8,* 86-87. doi:10.1097/00001199-199309000-00010

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology, 9,* 100-110. doi:10.1001/archneur.1963.00460070100010

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24,* 1094-1102. doi:10.1076/jcen.24.8.1094.8379

Mittenberg, W., Rotholc, A., Russell, E., & Heilbronner, R. (1996). Identification of malingered head injury on the Halstead-Reitan Battery. *Archives of Clinical Neuropsychology, 11,* 271-281. doi:10.1093/arclin/11.4.271

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A., H., Gowerter, A., & Wagner, T. D., (2000). The unity and diversity of executive functions and their contributions to complex and frontal lobe tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49-100. doi:10.1006/cogp.1999.0734

Moritz, S., Birkner, C., Kloss, M., Jahn, H., Hand, I., Haasen, C., & Krausz, M. (2002). Executive functioning in obsessive-compulsive disorder, unipolar depression, and schizophrenia. *Archives of Clinical Neuropsychology, 17,* 477-483. doi:10.1093/arclin/17.5.477

Nagahama, Y., Fukuyama, H., Yamaguchi, H., Katsumi, Y., Magata, Y., Shibasaki, H., & Kimura, J. (1997). Age-related changes in cerebral blood flow activation during performance of a card sorting test. *Experimental Brain Research, 114,* 571-577. doi:10.1007/PL00005665

Nagahama, Y. Fukuyama, H., Yamaguchi, H., Matsuzaki, S., Konishi, J., Shibasaki, H., & Kimura, J. (1996). Cerebral activation during the performance of a card sorting test. *Brain, 119,* 1667-1675. doi:10.1093/brain/119.5.1667

Nagahama, Y., Sadato, N., Yamauchi, H., Katsumi, Y., Hayashi, T., Fukuyama, H.,...Yonekura, Y. (1998). Neural activity during attention shifts between object features. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 9,* 2633-2638. doi:10.1097/00001756-199808030-00038

Newman, S. D., Carpenter, P. A., Varma, S., & Just, M. A. (2003). Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologia, 41,* 1668-1682. doi:10.1016/S0028-3932(03)00091-5

Nyhus, E., & Barceló (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition, 71,* 437-451. doi:10.1016/j.bandc.2009.03.005

O'Bryant, S. E., Finlay, C. G., & O'Jile, J. R. (2007). TOMM performances and self-reported symptoms of depression and anxiety. *Journal of Psychopathology and Behavioral Assessment, 29,* 111-114. doi:10.1007/s10862-006-9034-9

Ord, J. S., Greve, K. W., Bianchini, K. J., & Aguerrevere, L. E. (2010). Executive dysfunction in traumatic brain injury: The effects of injury severity and effort on the Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology, 32,* 132-140. doi:10.1080/13803390902858874

Otero, T. M., & Barker, L. A. (2014). The frontal lobes and executive functioning. In S. Goldstein, & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 29-44). New York, NY: Springer.

Parellada, E., Catafau, A. M., Bernardo, M., Lomenta, F., Catarineu, S., & Gonzalez-Monclus, E. (1998). The resting and activation issue of hypofrontality: A single photon emission computed tomography study in neuroleptic-naïve and neuroeptic-free schizophrenic female patients. *Biological Psychiatry, 44,* 787-790. doi:10.1016/S0006-3223(98)00057-2

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55-85). Hillsdale, NJ: Lawrence Erlbaum.

Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology, 20,* 33-65. doi:10.1016/j.acn.2004.02.005

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Malingering Test. *Archives of Clinical Neuropsychology, 16,* 501-606. doi:10.1093/arclin/16.5.501

Rohling, M. L., Binder, L. M., Demakis, G. J., Larrabee, G. J., Ploetz, D. M., & Langhinrichsen-Rohling, J. (2011). A meta-analysis of neuropsychological outcome after traumatic mild injury: Re-analysis and consideration of

Binder et al. (1997), Frenchman et al. (2005), and Pertab et al. (2009). *The Clinical Neuropsychologist, 25,* 608-623. doi:10.1080/13854046.2011.565076

Rubenstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive process in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 763-797. doi:10.1037//00961523.27.4.763

Ruiz, M. A., Drake, E. B., Glass, A., Marcotte, D., & van Gorp, W. G. (2002). Trying to beat the system: Misuse of the internet to assist in avoiding the detection of psychological symptom dissimulation. *Professional Psychology: Research and Practice, 33,* 294-299. doi:10.1037//07357028.33.3.294

Schretlen, D. (1997). *The Brief Test of Attention professional manual.* Odessa, FL: Psychological Assessment Resources, Inc.

Schutte, C. & Axelrod, B. N. (2013). Use of embedded cognitive symptom validity measures in mild traumatic brain injury cases. In D. A., Carone, & S. S., Bush (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering.* New York, NY: Springer Publishing.

Shute, G. E., & Huertas, V. (1990). Developmental variability in frontal lobe function. *Developmental Neuropsychology, 6,* 1-11. doi:10.1080/87565649009540445

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingering neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13,* 545-561. doi:10.1076/1385-4046(199911)13:04;1-Y;FT545

Smitherman, T. A., Huerkamp, J. K., Miller, B. I., Houle, T. T., & O'Jile, J. R. (2007). The relation of depression and anxiety to measure of executive functioning in a mixed psychiatric sample. *Archives of Clinical Neuropsychology, 22,* 647-654. doi:10.1016/j.acn.2007.04.007

Stemme, A., Deco, G., & Busch, A. (2007). The neuronal dynamics underlying cognitive flexibility in set shifting tasks. *Journal of Computational Neuroscience, 23,* 313-331. doi: 10.1007/s10827-007-0034-x

Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary, third edition.* New York, NY: Oxford University Press.

Stuss, D. T., & Alexander, M. P. (2007). Is there a dysexecutive system? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362,* 901-915. doi:10.1098/rstb.2007.2096

Suhr, J. A., & Boyer D. (1999). Use of the Wisconsin Card Sorting Test in the detection of malingering in student simulator and patient samples. *Journal of Clinical and Experimental Neuropsychology, 21,* 701-708. doi:10.1076/jcen.21.5.701.868

Suhr, J. A., & Gunstad, J. (2000). The effects of coaching on the sensitivity and specificity of malingering measures. *Archives of Neuropsychology, 15,* 415-424. doi:10.1093/arclin/15.5.415

Sumitani, S., Tanaka, T., Tayoshi, S. Y., Ota, K., Kameoka, N., Ueno, S., & Ohmori, T. (2006). Activation of the Prefrontal Cortex during the Wisconsin Card Sorting Test as Measured by Multichannel Near-Infrared Spectroscopy: *Neuropsychobiology, 53*, 70-76. doi:10.1159/000091722

Sweet, J. J., & Nelson, N. W. (2007). Validity Indicators within Executive Function Measures: Use and Limits in Detection of Malingering. In Boone, K. B. (Ed.), *Assessment of Feigned Cognitive Impairment: A Neuropsychological Perspective.* New York, NY: The Guildford Press.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. *Lancet, 2,* 81-84. doi:10.1016/S01406736(74)91639-0

Thomas, M. L., & Youngjohn, J. R. (2009). Let's not get hysterical: Comparing the MMPI-2 validity, clinical and RC scales in TBI litigants tested for effort. *The Clinical Neuropsychologist, 23,* 1067-1084. doi:10.1080/13854040902795000

Tien, A. Y., Schlaepfer, T. E., Orr, W., & Pearlson, G. D. (1998). SPECT brain flow changes with continuous ligand infusion during previously learned

ECST performance. *Psychiatric Research and Neuroimaging, 82,* 47-52. doi:10.1016/S0925-4927(98)00003-1

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM).* New York, NY: Multi Health Systems.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9,* 260-268. doi:10.1037//1040-3590.9.3.260

Volz, H., Gaser, C., Häger, F., Rzanny, R., Mentzel, H., Kreitschmann Andermahr, I.,...Sauer, H. (1997). Brain activation during cognitive stimulation with the Wisconsin Card Sorting Test-A functional MRI study on healthy volunteers and schizophrenics. *Psychiatry Research: Neuroimaging, 75,* 145-157. doi:10.1016/S0925-4927(97)00053-X

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale fourth edition.* San Antonio, TX: Pearson.

Weinberger, D. R., Berman, K. F., & Zec, R. F. (1986). Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia: I. Regional cerebral blood flow (rCBF) evidence. *Archives of General Psychiatry, 43,* 114-124. doi:10.1001/archpsyc.1986.01800020020004

Wetter, M., & Corrigan, S. (1995). Providing information to clients about psychological tests: A survey of attorneys and law students' attitudes.

Professional Psychology: Research and Practice, 26, 474-477. doi:10.1037//0735-7028.26.5.474

Whiteside, D. M., Kogan, J., Wardin, L., Philips, D., Franzwa, M. G., Rice, L...Roper, B. (2014). *Boston Naming Test and Controlled Oral Word Association Test as embedded performance validity measures in traumatic brain injury and psychiatric samples.* Manuscript submitted for publication.

Whiteside, D. M., Wald, D., & Busse, M. (2011). Classification accuracy of multiple visualspatial measures in the detection of suspect effort. *The Clinical Neuropsychologist, 25,* 287-301. doi:10.1080/13854046.2010.538436

Yanez, Y. T., Fremouw, W., Tennant, J., Strunk, J. & Coker, K. (2006). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology, 21,* 161-165. doi:10.1016/j.acn.2005.07.0009

Youngjohn, J. R., Davis, D., & Wolf, I. (1997). Health injury and the MMPI-2: Paradoxical severity effects and the influence of litigation. *Psychological Assessment, 9,* 177-184. doi:http://dx.doi.org/10.1037//1040-3590.9.3.177

Zakzanis, K. K., Mraz, R., & Graham, S. J. (2005). An fMRI study of the Trail Making Test. *Neuropsychologia, 43,* 1878-1886. doi:10.1016/j.neuropsychologia.2005.03.013

Table 1
Known Studies Utilizing the Wisconsin Card Sorting Test as a Performance Validity Test

Study	Participants	Predictor Variable	Cutoff Score	SN	SP
Bernard et al. (1996)	Simulated MND (*n*=24)	Bernard DFA[a]			
	Controls (*n*= 21)			100%	92%
	Mixed neurological diagnosis (*n*=89)			58%	100%
	Closed head TBI (*n*= 70)			86%	94%
Suhr et al. (1999)	Simulated MND (*n*= 41)	Suhr Formula[b]			
	Controls (*n*=31)			70.7%	87.1%
	Probable MND (*n*=17)			82.4%	
	mTBI and MTBI (*n*=16)				93.3%
Donders et al. (1999)	Mixed TBI non-MND (*n*=130)	Bernard DFA[a]			95%
Sweet & Nelson (2007)	STBI no incentive (*n*=30)	Bernard DFA[a]		0%	95%
		Suhr formula[b]		0%	95%
	mTBI no incentive (*n*=30)				
	mTBI with incentive (*n*=30)				
Greve, Bianchini, et al. (2002)	89 patients with mixed TBI	Bernard DFA[a]	>-3	38%	39%
			>0	16%	94%
		Suhr formula[b]	>1.90	47%	89%
	Controls (no incentive)		>3.86	34%	94%
	(*n*=17)	Unique Responses	≤1	35%	94%
			>1	22%	100%

	Probable MND (*n*=32)	Perfect matches missed	>0 >1	0% 9%	100% 100%
	Suspected MND (*n*=30)				
	Incentive only (*n*=10)				
Greve & Bianchini (2002)	College controls (*n*=133)	Bernard DFA[a] Suhr formula[b]			97.7% 100%
	College controls 2 (*n*=76)				84.2% 98.7%
	Inpatient substance abuse (*n*=44)				79.5% 81.8%
	Residential STBI treatment (*n*=69)				58.3% 73.9%
	Inpatient stroke rehab (*n*=83)				88% 73.5%
	Mixed neurological (*n*=128)				92.2% 85.9%
	Mixed neurological 2 (*n*=360)				86.7% 79.7%
King et al. (2002) Study 1	Probable MND (*n*=27)	Bernard DFA[a]		63%	94%
	Credible MTBI and STBI (*n*=33)	Suhr formula[b]	≥3.16	59%	88%
		King formula[c]	≥2.13	70%	82%
King et al. (2002) Study 2	MTBI and STBI (*n*=75)	Bernard DFA[a] Suhr formula[b] King formula[c]			73% 75% 97%
King et al (2002) Study 3	Mild to severe TBI (*n*=130)	Bernard DFA[a] Suhr formula[b] King Formula[c]			95% 85% 99%
Ashendorf (2003)	Non-MND healthy older adults (*n*=197)	Bernard DFA[a]			55-91.4%

		Suhr formula[b]			47.6-62.7%
		Unique Responses	>5		91.4%
		Perfect matches missed	>1		95.1%
Larrabee (2003)	MND (n=26)	Suhr formula[b]	>0	64%	52%
			>2.41	40%	87%
	MTBI to STBI (n=31)				
		FMS	>1	48%	87%
	Mixed neurological and psychiatric (n=27)		>2		96%
Heinly et al. (2006)[d]	mTBI (n=137)	Bernard DFA[a]	>3	32%	87%
			>0	24%	91%
	MTBI to STBI (n=139)	Suhr formula[b]	>0	57%	60%
			>1.90	37%	88%
			>2.41	33%	88%
	Chronic STBI (n=101)		>3.16	29%	90%
			>3.68	22%	93%
	General clinical patients (n>1,000)		≥4.00	22%	95%
			≥4.50	9%	98%
		FMS	>1	35%	85%
			>2	24%	92%
			>3	16%	100%
Greve et al. (2009)[d]	mTBI non-MND (n=55)	Bernard DFA[a]	≥ -1	26%	89%
		Suhr formula[b]	≥3	32%	89%
	mTBI MND (n=38)				
		King formula[c]	>0	32%	94%
	MTBI and STBI non-MND (n=92)	FMS	≥3	29%	89%
	STBI non-MND (n=91)	Perseverative Responses	≥45	16%	89%
	General non-TBI clinical patients (n=766)	Categories Completed	≤2	26%	87%
		Unique Responses	≤2	34%	87%

Note. SN = sensitivity; SP = specificity; DFA = Discriminant function formula; MND = malingered neurocognitive dysfunction; TBI = traumatic brain injury; mTBI = mild traumatic brain injury; MTBI = moderate traumatic brain injury; STBI = severe traumatic brain injury; *n* = number in sample.

[a]Bernard DFA consists of categories completed and perseverative errors.

[b]Shur formula consists of a logistic regression number of categories and FMS.

[c]King formula consists of a logistic regression of number of categories completed, failure to maintain set, and percent conceptual level responses.

[d]Sensitivity and Specificity are listed for the mTBI group.

Table 2

Performance Validity Test Cutoff Scores and References

Test	Cutoff Scores	Reference
Benton Facial Recognition	< 39	Whiteside, Wald, & Busse (2011)
BTA	< 16	Busse & Whiteside (2012)
CPT-II Omission Errors	>13	Busse & Whiteside (2012)
Digit Span Adjusted Scaled Score	<5	Greiffenstein, Baker, & Gola (1994) and Babikian et al. (2006)
Finger Tapping Test[a]	Males: <33 Females: <28	Arnold et al. (2005)
JLO	<18	Whiteside, Wald, & Busse (2011)
Rey Copy	< 23.5	Whiteside, Wald, & Busse (2011)
TOMM	<45 on Trial 2 or Delayed	Tombaugh (1996)
WMT	Failure on any PVT section	Green, Allen, & Astner (1996)

Note. BTA = Brief Test of Attention; CPT-II = Continuous Performance Test-2nd Edition; JLO = Judgment of Line Orientation; TOMM = Test of Memory Malingering; WMT = Word Memory Test; PVT = performance validity test.

[a]Dominant hand.

Table 3

Demographics Test Statistics of mTBI, STBI, and PSYCH Samples

Sample	*n*	Pass PVT/Fail PVT (Base Rate[a])	Compensation Seeking[b]	Age (SD)	Education (SD)
mTBI	111	93/18 (16.2%)	18.9%	45.4 (12.3)	13.4 (2.1)
STBI	69	62/7 (10.1%)	7.2%	44.2 (12.1)	13.2 (2.5)
PSYCH	155	136/19 (12.3%)	7.1%	46.7 (13.6)	13.5 (2.4)

Note. mTBI = mild traumatic brain injury; STBI = severe to moderate traumatic brain injury; PSYCH = psychiatric diagnosis; *n* = sample number; PVT = performance validity test; SD = standard deviation.

[a]PVT fail base rate.

[b]Percent yes for compensation seeking status.

Table 4

Demographic Characteristics of mTBI, STBI, and PSYCH Samples

Sample	Gender (Percent Female)	Ethnicity (Percent Caucasian)	Martial Status (Percent Married)	Handedness (Percent right)
mTBI	64.0%	96.4%	57.7%	79.3%
STBI	43.5%	94.2%	60.9%	79.7%
PSYCH	60.0%	94.8%	55.5%	85.8%

Note. mTBI = mild traumatic brain injury; STBI = severe to moderate traumatic brain injury; PSYCH = psychiatric diagnosis.

Table 5

Psychiatric Diagnoses of mTBI, STBI, and PSYCH Samples

Diagnosis	mTBI	STBI	PSYCH
Mood Disorders	1	20	144
Anxiety	17	11	8
Somatization/Pain Disorders	62	6	0
PTSD	6	6	2
Adjustment Disorder	7	6	1
Personality Disorder	2	0	0
Cognitive Disorder NOS	8	14	0
No Diagnosis	8	6	0
Total	111	69	155

Notes. mTBI= mild traumatic brain injury; STBI = moderate to severe traumatic brain injury; PSYCH = mixed psychiatric diagnosis; PTSD = Post-Traumatic Stress Disorder; ADHD = Attention Deficit Hyperactivity Disorder.

Table 6

Demographic Test Statistics Between PASS and FAIL Groups for mTBI, STBI, and PSYCH Samples

Demographic	mTBI	STBI	PSYCH
	t-test/ Chi Square	t-test/Chi Square	t-test/Chi Square
Age	t(109)= -.21, p = .84	t(67) = - .37, p = .72	t(153)= -.81, p = .42
Education	t(109)= -.81, p = .42	t(67)= -1.30, p = .20	t(153) = .10, p = .92
Gender	$x^2(1) = .76, p = .38$	$x^2(1) = .70, p = .40$	$x^2(1) = .04, p = .84$
Ethnicity	$x^2(1) = 3.49, p = .06$	$x^2(1) = 1.03, p = .31$	$x^2(1) = 1.79, p = .18$
Compensation Status[a]	$x^2(1) = 2.91, p = .09$	$x^2(1) = .61, p = .44$	$x^2(1) = 2.48, p = .12$

Note. mTBI = mild traumatic brain injury; STBI = severe and moderate traumatic brain injury; PSYCH = mixed mood/anxiety disorder psychiatric sample.

None were significant at the p<0.05 level.

[a] Refers to either clinical or forensic referrals.

Table 7

Demographic Test Statistics of mTBI, STBI, and PSYCH Samples

Demographic	t-test/ Chi Square
Age	t(2)= 3.33
Education	t(2)= 1.08
Gender	$x^2(2) = .02$*
Ethnicity	$x^2(2) = .48$
PASS/FAIL rate	$x^2(2) = .46$
Compensation Status[a]	$x^2(2) = 10.48$*
Marital Status	$x^2(12) = .80$
Handedness	$x^2(2) = .48$

Note. mTBI = mild traumatic brain injury; STBI = severe and moderate traumatic brain injury; PSYCH = mixed mood/anxiety disorder psychiatric sample.

* Refers to significance at the $p<0.05$ level.

[a] Refers to either clinical or forensic referrals.

Table 8

mTBI Sample Means and Effect Sizes for Wisconsin Card Sorting Test Variables

	Group	*n*	Mean (SD)	Mann-Whitney U	Effect Size Cohen's d (Descriptor[a])
Number of Trials	Pass	93	100.96 (23.12)	640.5	.44 (Medium)
	Fail	18	111.00 (22.71)		
Number Correct	Pass	93	71.02 (11.83)	766.5	.24 (Small)
	Fail	18	67.33 (18.40)		
Total Errors	Pass	93	29.86 (23.11)	569.5	.52 (Medium)
	Fail	18	43.67 (29.67)		
Perseverative Errors	Pass	93	16.67 (17.52)	695.5	.20 (Small)
	Fail	18	20.00		

			(15.32)		
Perseverative	Pass	93	18.50	731.5	.11 (Small)
Responses			(21.18)		
	Fail	18	20.72		
			(18.50)		
Nonperseverative	Pass	93	15.10	592.5	.48 (Medium)
Errors			(14.92)		
	Fail	18	23.44		
			(19.77)		
Trials to 1st	Pass	93	16.63	684.0	.06 (Small)
Category			(19.88)		
	Fail	18	18.22		
			(28.85)		
FMS	Pass	93	1.05	778.0	.22 (Small)
			(1.45)		
	Fail	18	.78 (1.00)		
Number of	Pass	93	5.08	625.0	.49 (Medium)
Categories			(2.05)		
	Fail	18	4.00		
			(2.38)		

Note. mTBI = mild traumatic brain injury; n = sample; SD = standard deviation;

FMS = failure to maintain set; * = $p < 0.05$.

[a]Effect size descriptor.

Table 9

Logistic Regression Using Wisconsin Card Sorting Test Predictor Variable

Variable	B (S.E.)	Wald	df	*P*<	Exp (B)
WCSTCOMB					
Perseverative Responses	-.005 (.012)	3.93	1	.047	.99
Trials to 1st Category	-.001 (.007)	.009	1	.923	.999
Number Correct	.005 (.007)	.19	1	.66	1.01
Constant	1.796 (.942)	3.693	1	.056	6.027

Notes. B = beta weights coefficient; S.E. = standard error of coefficient; df = degrees of freedom; Exp (B) = exponentiation of the B coefficient; FMS = failure to maintain set; WCSTCOMB = logistic regression consisting of perseverative responses, trials to 1st category and number correct.

Table 10

Sample Means and Effect Sizes for WCSTCOMB Variable

Sample	Group	*n*	Mean (SD)	Mann-Whitney U	Cohen's d (Descriptor[a])
WCSTCOMB in mTBI	Pass	93	1.88 (.33)	667.0	.18 (Small)
	Fail	18	1.82 (.34)		
WCSTCOMB in PSYCH	Pass	136	1.87 (.27)	781.5	.54 (Medium)
	Fail	19	1.71 (.32)		
WCSTCOMB in STBI	Pass	62	1.90 (.27)	90.5*	1.09 (Large)
	Fail	7	1.61 (.26)		

Note. PSYCH = mTBI = mild traumatic brain injury; mixed psychiatric diagnoses; STBI = moderate to severe traumatic brain injury; n = sample; SD = standard deviation; FMS = failure to maintain set;

* $p<0.05$

[a]Effect size descriptor.

Table 11

Area Under Curve for Wisconsin Card Sorting Test Variables in the mTBI Sample

Predictor Variable	AUC	Asymptotic Significance
Number of Trials	.62	.12
Total Errors	.64	.05
Perseverative Errors	.61	.16
Perseverative Responses	.56	.40
Nonperseverative Errors	.65	.05
Number of Categories	.63	.09
Number Correct	.54	.57
Trials to 1st Category	.41	.22
FMS	.47	.64
WCSTCOMB	.60	.17

Notes. mTBI = mild traumatic brain injury; AUC = area under curve; FMS = failure to maintain set; WCSTCOMB = logistic regression of number correct, perseverative responses, and trails to first category.

Table 12

STBI Sample Means and Effect Sizes for Wisconsin Card Sorting Test Variables

	Group	*n*	Mean (SD)	Mann-Whitney U	Effect Size Cohen's d (Descriptor[a])
Number of Trials	Pass	62	97.55 (23.00)	114.5*	.86 (Large)
	Fail	7	116.29 (20.41)		
Number Correct	Pass	62	67.61 (14.23)	199.5	.12 (Small)
	Fail	7	66.27 (6.32)		
Total Errors	Pass	62	29.53 (25.62)	106.0*	.39 (Medium)
	Fail	7	50.00 (22.69)		
Perseverative Errors	Pass	62	14.79 (14.89)	86.5*	1.06 (Large)
	Fail	7	30.43		

			(14.58)		
Perseverative	Pass	62	15.90	82.5*	1.17 (Large)
Responses			(15.79)		
	Fail	7	35.86		
			(18.20)		
Nonperseverative	Pass	62	15.03	163.0	.26 (Small)
Errors			(14.30)		
	Fail	7	18.43		
			(11.96)		
Trials to 1st	Pass	62	16.02	208.5	.06 (Small)
Category			(21.11)		
	Fail	7	15.00		
			(9.06)		
FMS	Pass	62	.86 (1.20)	156.0	.44
	Fail	7	1.43		(Moderate)
			(1.40)		
Number of	Pass	62	4.79	126.0*	.86 (Large)
Categories			(1.99)		
	Fail	7	3.00		
			(2.16)		

Note. STBI = moderate to severe traumatic brain injury; n = sample; SD = standard deviation; FMS = failure to maintain set; * = $p < 0.05$.

[a]Effect size descriptor.

Table 13

Area Under Curve for STBI Sample

Predictor Variable	Raw Cutoff Score	AUC	*P*<	95% Confidence Interval	
				LB	UB
Number of Trials		.74	.04	.57	.90
Total Errors	74	.76	.03	.60	.92
Perseverative Errors	39	.80	.01	.64	.97
Perseverative Responses	47	.81	.01	.64	.98
NonPerseverative Errors	n/a	.62	.28	.42	.83
Trials to 1st Category	n/a	.48	.87	.21	.75
FMS	n/a	.64	.23	.43	.85
Number Correct	n/a	.54	.73	.38	.70
Number of Categories	n/a	.71	.07	.50	.92

WCSTCOMB	1.37	.79	.01	.62	.96

Notes. STBI = moderate to severe traumatic brain injury; AUC = area under curve; *p*< = asymptotic significance; LB = lower bound; UB = upper bound; n/a = not applicable due to area under the curve >.7; FMS = failure to maintain set; WCSTCOMB = logistic regression of number correct, perseverative responses; and trails to first category.

Table 14

Classification Accuracy of Wisconsin Card Sorting Test Variables in the STBI Sample

Raw Cutoff Score (AUC)	Sensitivity	Specificity	10% Base Rate	
			PPP	NPP
Total Errors (.76)				
75	.00	.90	.00	.89
74*	.14	.90	.13	.90
73	.14	.89	.13	.90
68	.14	.86	.10	.90
62	.29	.86	.19	.92
58	.43	.86	.25	.93
56	.71	.84	.33	.96
55	.71	.81	.29	.96
52	.71	.79	.27	.96
47	.71	.76	.25	.96
Perseverative Errors (.80)				
55	.00	.97	.00	.89
53	.00	.95	.00	.90
49	.00	.94	.00	.90

46	.14	.94	.21	.91
43*	.29	.94	.35	.92
40	.29	.92	.29	.92
38	.29	.90	.24	.92
34	.43	.90	.32	.93
31	.43	.89	.30	.93
30	.57	.87	.32	.95
Perseverative Responses (.81)				
52	.14	.95	.24	.91
50	.29	.95	.39	.92
49	.43	.95	.49	.94
48	.43	.94	.44	.94
47*	.43	.92	.37	.94
45	.43	.90	.32	.93
43	.43	.89	.30	.93
38	.57	.89	.37	.95
33	.57	.87	.33	.95
32	.57	.86	.31	.95
Number of Trials (.74)				
123	.71	.69	.20	.96

117	.71	.68	.20	.95
111	.71	.66	.18	.95
105	.71	.65	.18	.95
102	.71	.63	.18	.95
Number of Categories (.71)				
0	.00	.95	.00	.90
1	.29	.86	.19	.92
2	.57	.82	.20	.94
3	.71	.79	.27	.96
4	.71	.69	.20	.96
5	.71	.68	.20	.95
6	1.00	.00	.10	N/A
WCSTCOMB (.79)				
1.33	.14	.94	.21	.91
1.37	.14	.92	.16	.91
1.40	.14	.90	.13	.90
1.43	.29	.90	.24	.92
1.47	.43	.90	.32	.90
1.53*	.57	.90	.39	.95
1.56	.57	.89	.37	.95

1.62	.57	.87	.33	.95
1.68	.57	.86	.31	.95
1.69	.57	.84	.28	.95

Notes. STBI = moderate to severe traumatic brain injury; PPP = positive predictive power; NPP = negative predictive power; AUC = area under curve; WCSTCOMB = logistic regression of number correct, perseverative responses and trials to 1st category, * = recommended cutoff.

Table 15

PSYCH Sample Means and Effect Sizes for Wisconsin Card Sorting Test Variables

	Group	*n*	Mean (SD)	Mann-Whitney U	Cohen's d (Descriptor[a])
Number of Trials	Pass	136	102.06 (23.53)	813.5*	-.73 (Large)
	Fail	19	117.32 (18.11)		
Number Correct	Pass	136	70.77 (11.74)	1206.0	-.06 (Small)
	Fail	19	71.53 (14.01)		
Total Errors	Pass	136	33.33 (37.51)	802.5*	-.39 (Small)
	Fail	19	45.68 (24.15)		
Perseverative Errors	Pass	136	15.93 (13.34)	797.0*	-.61 (Medium)
	Fail	19	24.74		

			(15.38)		
Perseverative Responses	Pass	136	18.71 (18.53)	744.5*	-.55 (Medium)
	Fail	19	29.34 (19.76)		
Nonperseverative Errors	Pass	136	14.55 (11.54)	1041.5	-.36 (Small)
	Fail	19	19.58 (15.83)		
Trials to 1st Category	Pass	136	14.94 (10.46)	953.0	-.53 (Medium)
	Fail	19	24.47 (23.28)		
FMS	Pass	136	.83 (1.23)	1153.0	-.18 (Small)
	Fail	19	1.05 (1.22)		
Number of Categories	Pass	136	5.15 (1.85)	841.5*	.64 (Medium)
	Fail	19	4.00 (1.76)		

Note. PSYCH = mixed psychiatric diagnoses; n = sample; SD = standard deviation; FMS = failure to maintain set;

* $p<0.05$

[a]Effect size descriptor.

Table 16

Area Under Curve for PSYCH Sample

Predictor Variable	Raw Cutoff Score	AUC	*P*<	95% Confidence Interval	
				LB	UB
Number of Trials	n/a	.69	.009	.574	.797
Total Errors	n/a	.69	.008	.569	.810
Perseverative Errors	n/a	.69	.007	.570	.813
Perseverative Responses	48	.71	.003	.602	.822
NonPerseverative Errors	n/a	.60	.172	.467	.727
Trials to 1st Category	n/a	.63	.064	.483	.779
FMS	n/a	.55	.448	.410	.697
Number Correct	n/a	.47	.639	.320	.614
Number of Categories	n/a	.67	.014	.544	.805

WCSTCOMB	n/a	.70	.005	.576	.819

Notes. PSYCH = psychiatric diagnosis sample; AUC = area under curve; *p*< = asymptotic significance; LB = lower bound; UB = upper bound; n/a = not applicable due to area under the curve >.7; FMS = failure to maintain set; WCSTCOMB = logistic regression of number correct, perseverative responses; and trails to first category.

Table 17

Classification Accuracy of Wisconsin Card Sorting Test Variables in the PSYCH Sample

Raw Cutoff Score (AUC)	Sensitivity	Specificity	14% Base Rate	
			PPP	NPP
Perseverative Responses (.71)				
36	.37	.85	.29	.89
37	.37	.88	.33	.90
38	.26	.88	.26	.88
40	.26	.88	.26	.88
42	.26	.89	.28	.88
43	.21	.90	.25	.87
45	.11	.90	.15	.86
46	.11	.91	.17	.86
49	.11	.92	.18	.86
51	.11	.93	.20	.87
53	.11	.95	.26	.87
WCSTCOMB (.70)				
1.31	.05	.94	.12	.86
1.33	.05	.93	.10	.86

1.35	.05	.93	.10	.86
1.36	.11	.93	.20	.87
1.39	.11	.92	.18	.86
1.42	.16	.91	.22	.87
1.44	.16	.90	.21	.87
1.45	.16	.89	.19	.87
1.47	.16	.88	.18	.87

Notes. PSYCH = psychiatric diagnosis; AUC = area under curve; PPP = positive predictive power; NPP = negative predictive power.

Appendix A

March 25, 2013

Dear Lydia Wardin,

The Institutional Review Board evaluated the changes to your application, proposal **#13-037**, *Classification Accuracy of the WCST in Determining Suboptimal Cognitive Effort Among Neuropsychological Patients.* Your application has now received **Full Approval**. This decision means that you may proceed with your plan of research as it is proposed in your application.

Please note that if you wish to make changes to your procedures or materials, you must provide written notification to the IRB in advance of the changes, co-signed by your Dissertation Chair, Dr. Whiteside. **You may not implement those changes until you have received a Full Approval letter from the IRB**. Please feel free to contact myself or other IRB committee members should you have any questions.

Sincerely,

Peter Ji, Ph.D
Core Faculty, Psy.D. Program in Clinical Psychology
Co-Chair, Institutional Review Board
Adler School of Professional Psychology

Appendix B

ADLER SCHOOL of Professional Psychology

DATA USE AGREEMENT BETWEEN

Adler Collaborative Database Project
And
Lydia Wardin
Adler School of Professional Psychology

This Data Use Agreement is made and entered into on **2/13/2013,** between Douglas Whiteside, PhD, ABPP, hereafter "Holder" and **Lydia Wardin**, hereafter "Recipient."

1. This agreement sets forth the terms and conditions pursuant to which Holder will allow recipient access to the requested data.

   1.1 The data set is an existing database housed at **Adler School of Professional Psychology**. The data set is part of Adler School of Professional Psychology. The use of the database for outcomes related research was approved Adler Institutional Review Board (IRB). The dataset includes the following information: demographic data, neuropsychological data and psychological assessment data. All patient identifiers are removed from the dataset.

2. Permitted Uses and Disclosures

   2.1 Except as otherwise specified herein, Recipient may make all uses and disclosures of the data set necessary to conduct the research described herein:

   The purpose of this research project is to **explore the classification accuracy of the Wisconsin Card Sorting Test as a measure of embedded effort among an adult neuropsychological population.**

   The data set will not be used as part of a publication or oral/poster presentation without consulting with the Holder.

3. The data set is allowed to be taken off the original primary holder's site.
   X☐Yes, data set is allowed off site.
   3.1 ☐No, all data analyses should be contacted at the original primary holder's site. Only aggregate output can be taken off site for the purpose of report writing.

   3.2 In addition to the Recipient, the individuals, or classes or individuals, who are permitted to use or receive the aggregate outputs from statistical analysis and related reports for purposes of the dissertation include: **Lydia Wardin**. The IRB at the Adler School of Professional Psychology is also permitted to review the data.

4. Recipient Responsibilities

4.1 Recipient will not use or disclose the data set for any purpose other than permitted by this Agreement pertaining to the dissertation project or as required by law;

4.2 Recipient will use appropriate administrative, physical and technical safeguards including password protected files to prevent use or disclosure of the data set other than as provided for by this Agreement including keeping with guidelines consistent with those of Adler School of Professional Psychology, and with APA guidelines for data retention of seven (7) years (notwithstanding any already existing data use agreements);

4.3 Recipient will report to the Holder any use or disclosure of the data set not provided for by this Agreement of which the Recipient becomes aware within 15 days of becoming aware of such use or disclosure;

4.4 A summary report of the data analysis findings will be provided to the Holder.

5. Term and Termination

5.1 The terms of this Agreement shall be effective as of **2/13/2013**, and shall remain in effect until the data set is destroyed or returned to the Holder.

5.2 Upon the Holder's knowledge of a material breach of this Agreement by the Recipient, the Holder shall notify the Recipient of the Holder's knowledge of such material breach in writing within 45 days and provide an opportunity for Recipient to cure the breach or end the violation. If efforts to cure the breach or end the violation are not successful within 60 days of Recipients receipt of Holder's notice of breach, the Holder shall report the problem to the dissertation chair, Douglas Whiteside, PhD, ABPP and the Adler School of Professional Psychology's Institutional Review Board (IRB) representative Peter Ji, PhD or David Castro-Blanco, PhD, ABPP.

5.3 Holder agrees to submit a continuing approval notice to the Adler School of Professional Psychology 1 year after the project approval date and every year thereafter until the project is closed.

6. General Provisions

6.1 Each party agrees that it will be responsible for its own acts and the results thereof to the extent authorized by law and shall not be responsible for the acts of the other party or the results thereof;

6.2 This Agreement (a) is the complete agreement of the Parties concerning the subject matter hereof and supersedes any prior agreements, understanding or discussions with respect to the subject matter hereof; and (b) may not be amended or in any manner modified except by a non-electronic written instrument signed by authorized representatives of both Parties.

6.3 If any provision of this Agreement is found unenforceable, the remainder shall be enforced as fully as possible and the unenforceable provision shall be deemed modified to the limited extent required to permit its enforcement in a manner most closely representing the intention of the Parties as expressed herein.

6.4 A signed copy of this agreement will be given to both the Holder and Recipient

7. Contact Information of Recipient: Lydia Wardin (lwardin@my.adler.edu), 847-217-2671

8. Adler IRB Contact Information

Dr. Castro-Blanco (dcastroblanco@adler.edu, 312-662-4333) and Dr. Ji. (pji@adler.edu, 312-662-4354), Adler IRB Co-chairs

IN WITNESS WHEREOF, the parties hereto execute this agreement as follows:

Signature: ____________________

Date: 2-13-13

By: Douglas Whiteside

Signature: ____________________

Date: 2/13/13

By: Lydia Wardin

Primary Investigator, Student, Adler School of Professional psychology

Signature: ____________________

Date: 2-13-13

By: Douglas Whiteside, Ph.D., ABPP

Dissertation Chair, Core Faculty, Adler School of Professional Psychology