

Theses and Dissertations

2012

A re-examination of the Meyers and Volbrecht motor equation for the identification of suspect effort

Ashley R. Curiel

Follow this and additional works at: <https://digitalcommons.pepperdine.edu/etd>

Recommended Citation

Curiel, Ashley R., "A re-examination of the Meyers and Volbrecht motor equation for the identification of suspect effort" (2012). *Theses and Dissertations*. 243.
<https://digitalcommons.pepperdine.edu/etd/243>

This Dissertation is brought to you for free and open access by Pepperdine Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Pepperdine Digital Commons. For more information, please contact Katrina.Gallardo@pepperdine.edu, anna.speth@pepperdine.edu, linhgavin.do@pepperdine.edu.

Pepperdine University
Graduate School of Education and Psychology

A RE-EXAMINATION OF THE MEYERS AND VOLBRECHT MOTOR
EQUATION FOR THE IDENTIFICATION OF SUSPECT EFFORT

A clinical dissertation submitted in partial satisfaction
of the requirements for the degree of
Doctor of Psychology

by

Ashley R. Curiel

March, 2012

Cary Mitchell, Ph.D. – Dissertation Chairperson

This clinical dissertation, written by

Ashley R. Curiel

under the guidance of a Faculty Committee and approved by its members, has been submitted to and accepted by the Graduate Faculty in partial fulfillment of the requirements for the degree of

DOCTOR OF PSYCHOLOGY

Doctoral Committee:

Cary Mitchell, Ph.D., Chairperson

Kyle Boone, Ph.D.

Carolyn Keatinge, Ph.D.

Po Lu, Psy.D.

© Copyright by Ashley R. Curiel (2012)

All Rights Reserved

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
VITA	viii
ABSTRACT	xiv
INTRODUCTION	1
METHOD	22
Subjects	22
Instruments/Procedures	25
Analyses	25
RESULTS	26
DISCUSSION	29
REFERENCES	33
TABLES	44

LIST OF TABLES

	Page
Table 1. Measures of Suspect Effort	44
Table 2. Descriptive Statistics and Group Differences for Individual Test and Formula Scores	45
Table 3. Comparison of Sensitivity Levels and Associated Cut-Scores of the Meyers and Volbrecht (2003) Formula and Individual Subtests at $\pm 90\%$ Specificity	46
Table 4. Group Differences for Individual Test Scores Between Credible Men and Women	46
Table 5. Group Differences for Individual Test Scores Between Noncredible Men and Women	47

DEDICATION

To my husband, Mark, who read every draft along the way. His patience and generosity are unmatched. To my parents, George and Martha Rudd, who sacrificed so much for my education. Finally, to my canine angel, Paco, who kept my lap warm for hours while I wrote at my computer. His companionship and love are constant sources of comfort and joy.

ACKNOWLEDGMENTS

I would like to thank my chairperson, Dr. Cary Mitchell, for his time, guidance, and support throughout my years at Pepperdine. I still remember my graduate school interview with Dr. Mitchell years ago. Today, as I near completion of my degree, I am grateful that he accompanied me on the journey.

I would also like to thank Dr. Kyle Boone, who inspired this project. She invested her time in me when I had no assessment or research experience; I will always be grateful to her for introducing me to the world of neuropsychology. Her expertise and achievement are inspiring and her generosity is truly exceptional. Without her outstanding guidance, this dissertation would not have been possible.

Thank you to Dr. Carolyn Keatinge and Dr. Po Lu, who kindly and generously shared their time and expertise to realize the completion of this work. I sincerely appreciate your mentorship.

I would also like to extend my deepest gratitude to my professors at Pepperdine and my clinical supervisors and research mentors who have so graciously provided me knowledge and opportunities beyond my imagination. Each of you has shaped my career immeasurably. Thank you to my peers and colleagues, who continue to offer incredible support, friendship, and motivation; I am certain that you will improve the world through your contributions to psychology.

Finally, I wish to thank my family and friends, who have patiently supported me throughout graduate school. Your loyalty and faith in me gives me courage and strength.

VITA

ASHLEY R. CURIEL

EDUCATION

- 2008 – Present Doctoral Candidate in Clinical Psychology, Psy.D. Program
Pepperdine University, Los Angeles, California
APA Accredited Program
GPA: 4.00
- 2005 – 2008 Master of Arts in Clinical Psychology: Family Psychology
Azusa Pacific University, Azusa, California
APA Accredited Program
GPA: 3.98
- 1997 – 2001 Bachelor of Arts: Spanish
Duke University, Durham, North Carolina
Cum Laude Honors
GPA: 3.69

DISSERTATION

A Re-examination of the Meyers and Volbrecht Motor
Equation for the Identification of Suspect Effort
Proposal Completed: June 2010
Defense Completed: October 2011

CLINICAL EXPERIENCE

- 08/11 – Present Doctoral Practicum Extern
Memory Rehabilitation Training and Brain Boot Camp
UCLA Longevity Center
Los Angeles, California
- 07/11 – Present Doctoral Practicum Extern
Geriatric Neuropsychology, UCLA
Los Angeles, California
- 04/11 – 05/11 Doctoral Practicum Extern
West Los Angeles Veterans Affairs Hospital (WLA VA)
Department of Geriatric Psychiatry
Los Angeles, California

- 01/11 – 03/11 Doctoral Practicum Extern
West Los Angeles Veterans Affairs Hospital (WLA VA)
Department of Geriatric Evaluation & Management
Geriatric Research Education & Clinical Center
Los Angeles, California
- 09/10 – 01/11 Doctoral Practicum Extern
West Los Angeles Veterans Affairs Hospital (WLA VA)
Department of Neuropsychological Assessment
Los Angeles, California
- 09/09 – 07/10 Doctoral Practicum Extern
Medical Psychology Assessment Center (MPAC), UCLA
Los Angeles, California
- 09/09 – 07/10 Doctoral Practicum Extern
The Help Group, UCLA
Sherman Oaks, California
- 09/09 – Present Doctoral Practicum Extern
Pepperdine Counseling and Educational Center
West Los Angeles Campus Clinic
Pepperdine University, Los Angeles, California
- 03/09 – 07/09 Doctoral Practicum Extern
Aging and Memory Research Center, UCLA
Los Angeles, California
- 08/08 – 02/09 Doctoral Practicum Extern
Alzheimer's Disease Research Center, UCLA
Los Angeles, California
- 09/07 – 08/08 Doctoral Practicum Extern
Harbor-UCLA Medical Center, Department of Neuropsychology
Torrance, California
- 09/06 – 06/07 Doctoral Practicum Extern
Child and Family Development Center
Azusa Pacific University, Azusa, California
- 09/06 – 06/07 Doctoral Practicum Extern
Azusa Unified School District
Longfellow Elementary School, Azusa, California

01/06 – 05/06 Doctoral Practicum Extern
 Child and Family Development Center
 Azusa Pacific University, Azusa, California

RESEARCH EXPERIENCE

06/11 – Present Research Associate
 Geriatric Neuropsychology, UCLA
 Los Angeles, California

09/09 – Present Research Associate
 Medical Psychology Assessment Center (MPAC), UCLA
 Los Angeles, California

02/09 – Present Research Associate
 Aging and Memory Research Center, UCLA
 Los Angeles, California

12/08 – 12/09 Research Associate
 Geriatric Psychiatry, UCLA
 Los Angeles, California

05/08 – 03/09 Research Associate
 Brain Mapping Center/Center for Cognitive Neuroscience, UCLA
 Los Angeles, California

04/08 – Present Research Associate
 Harbor-UCLA Medical Center, Division of Vascular Surgery
 Torrance, California

01/07 – 08/08 Research Associate
 Harbor-UCLA Medical Center, Department of Neuropsychology
 Torrance, California

PUBLICATIONS

Curiel, A. R., Miller, K. J., Pollard, K., Kim, J., Kravitz, J. & Small, G. W. (2012). Anxiety and verbal memory performance in APOE-4 carriers and noncarriers aged 50 years and above. *Aging Health, 8*, 99-104.

Neugebauer, Q., Mack, S., Roubin, A., & Curiel, A. R. (2011). Primary prevention of eating disorders in children and a proposed parent education program. *Graduate Student Journal of Psychology, 13*, 53-60.

Harris, S., Kaplan, J. T., Curiel, A. R., Bookheimer, S. Y., Iacoboni, M., & Cohen, M. S. (2009). The neural correlates of religious and nonreligious belief. *PLoS ONE*, 4(10), Article e7272. Retrieved January 4, 2010, from <http://www.plosone.org/article/citationList.action;jsessionid=FA0A587710C1EB5AA88BE4EFA48DEAEE?articleURI=info%3Adoi%2F10.1371%2Fjournal.pone.0007272>

Dean, A. C., Boone, K. B., Kim, M. S., Curiel, A. R., Martin, D. J., Victor, T. L., Zeller, M. A., & Lang, Y. K. (2008). Examination of the impact of ethnicity on the MMPI-2 Fake Bad Scale. *The Clinical Neuropsychologist*, 22, 1054-1060.

PRESENTATIONS

Curiel, A. R., Pollard, K., Miller, K. J., Kim, J., & Small, G. W. (2010). Using executive functioning measures to predict cognitive status. Presented at the 15th annual UCLA Research Conference on Aging, Los Angeles, CA, June, 2010.

Curiel, A. R., Pollard, K., Miller, K. J., Kim, J., & Small, G. W. (2010). Predicting cognitive status using executive functioning measures. Presented at the 38th annual International Neuropsychological Society meeting in Acapulco, Mexico, February, 2010.

Curiel, A. R. (2009). Semantic dementia: A case study. Seminar presented to UCLA neuropsychology trainees, Los Angeles, CA, October, 2009.

Curiel, A. R., Pollard, K., Miller, K. J., Kim, J., & Small, G. W. (2009). The effects of chronic distress in combination with the presence of the APOE-4 allele on cognitive status in older adults. Presented at the 7th annual American Academy of Clinical Neuropsychology Conference in San Diego, CA, June, 2009.

Curiel, A. R., Pollard, K., Miller, K. J., Kim, J., & Small, G. W. (2009). The effects of genetic risk and chronic distress on cognition in older adults. Presented at the 14th annual UCLA Research Conference on Aging, Los Angeles, CA, June, 2009.

Curiel, A. R., Pollard, K., Miller, K. J., Kim, J., & Small, G. W. (2009). Cortical versus subcortical dementia: An analysis through examination of memory scores. Presented at the 14th annual Research Conference on Aging, UCLA, Los Angeles, CA, June, 2009.

Curiel, A. R. (2009). Cognitive and behavioral changes associated with Alzheimer's disease. Seminar presented at Circle of Care Leeza's Place, Sherman Oaks, CA, June, 2009.

Curiel, A. R., Pollard, K., Miller, K. J., Kim, J., & Small, G. W. (2009). The effects of chronic distress in combination with the presence of the APOE-4 allele on cognitive status in older adults. Presented at the 3rd annual Alzheimer's Disease and Aging Symposium, hosted by the Universal Gerontology and Alzheimer's Disease Awareness (UGADA) Association at UCLA in Los Angeles, CA, May, 2009.

Curiel, A. R., Pollard, K., Miller, K. J., Kim, J., & Small, G. W. (2009). Subcortical dementia: A case presentation of Parkinson's disease. Presented at the 3rd annual Alzheimer's Disease and Aging Symposium, hosted by the Universal Gerontology and Alzheimer's Disease Awareness (UGADA) Association at UCLA in Los Angeles, CA, May, 2009.

Curiel, A. R., & Lin, E. (2009). Alzheimer's disease: A neuropsychological perspective. Seminar presented to UCLA neuropsychology trainees, Los Angeles, CA, April, 2009.

Curiel, A. R. (2009). Alien Hand Syndrome: A case study. Seminar presented to UCLA neuropsychology trainees, Los Angeles, CA, January, 2009.

Curiel, A. R. (2008). An introduction to the California Verbal Learning Test – II (CVLT-2). Seminar presented to UCLA neuropsychology trainees, Los Angeles, CA, October, 2008.

Curiel, A. R., Boone, K. B., Ziegler, E. A., Victor, T. L., & Zeller, M. A. (2008). An examination of the effectiveness of the Meyers and Volbrecht (2003) motor function formula in assessment of response bias. Presented at the 6th annual American Academy of Clinical Neuropsychology Conference in Boston, MA, June, 2008.

Dean, A. C., Victor, T. L., Boone, K. B., Curiel, A. R., & Zeller, M. A. (2007). Preliminary data on the use of Finger Tapping as an effort test in demented populations. Presented at the 27th annual National Academy of Neuropsychology Conference in Scottsdale, AZ, November, 2007.

PROFESSIONAL AFFILIATIONS

American Academy of Clinical Neuropsychology (AACN) – Student member
 American Psychological Association (APA) – Student member
 APA Division 40 – Clinical Neuropsychology – Student member
 APA Division 20 – Adult Development and Aging – Student member
 American Psychological Association of Graduate Students (APAGS) – Student member
 International Neuropsychological Society (INS) – Student member
 Los Angeles County Psychological Association (LACPA) – Student member

National Academy of Neuropsychology (NAN) – Student member

LEADERSHIP AND VOLUNTEER EXPERIENCE

2011 – Present	Duke University Los Angeles Women’s Forum – Steering Committee member
2011 – Present	Junior League of Los Angeles – Board of Directors member
2011 – Present	Junior League of Los Angeles – Assistant to President
2011 – Present	Junior League of Los Angeles – Community Leadership Forum member
2010 – 2011	Junior League of Los Angeles – Development Co-Chair for Provisional Group
2006 – Present	Alumni Admissions Interviewer for Duke University
2008 – 2010	Duke University Los Angeles Women’s Forum – Secretary
2005 – 2008	Contributor to Azusa Pacific University’s publication <i>PsychoLinguistics</i>
2005 – 2006	Student Coordinator of Integration Seminar dinners at Azusa Pacific University

HONORS

Psi Chi National Honor Society in Psychology – inducted in 2010, Pepperdine University
 Dean’s List with Distinction – 1999, Duke University
 Dean’s List – 1998, 1999, 2000, Duke University

ABSTRACT

Neuropsychologists are frequently called upon to evaluate cognitive functioning and to participate in determining disability status, particularly in the wake of traumatic brain injuries, strokes, and other health events that compromise central nervous system functioning. A critical component of each evaluation is effort assessment. Ideally, the methods for assessing the credibility of effort are neither obvious to test-takers nor vulnerable to coaching. One of the promising ways to evaluate effort is to use a combination of test scores that assess a common domain, such as motor functioning. The purpose of the present study was to cross validate a linear regression formula developed by Meyers and Volbrecht (2003) to evaluate the credibility of effort on selected tests of motor functioning. The formula utilized scores from the Rey-Osterrieth Complex Figure Test, the WAIS-III Digit Symbol and Block Design subtests, and the Finger Tapping Test. The advantages of such a formula for evaluating effort include that it relies upon embedded measures, resulting in heightened efficiency and greater subtlety of assessment.

The current archival study re-examined the Meyers and Volbrecht (2003) formula using 281 ethnically diverse patients who were referred for neuropsychological evaluation. The sample included 101 patients who met criteria for noncredible effort and 180 patients who met criteria for credible effort. Cut-off scores for the formula were selected to maintain specificity in the credible patients of at least 90%. The associated sensitivity rate when the original cut-offs were applied to the noncredible group was 30.7%. Closer examination of the individual tests that comprised the formula revealed that the Finger Tapping Test had unacceptably low sensitivity (29.7%). Therefore, the

Finger Tapping component of the equation was removed, which increased the formula's sensitivity to 70.3% while maintaining specificity of at least 90%. The revised formula provides neuropsychologists a novel way to assess effort that is neither vulnerable to coaching nor adds time to the testing battery. Other findings, limitations of the current study, and recommendations for future research are discussed.

Introduction

The purpose of the present study was to provide additional validation of a linear regression formula developed by Meyers and Volbrecht (2003) to detect noncredible effort in cognitive evaluations. Meyers and Volbrecht's method evaluates effort using a combination of scores from tests examining a common domain (i.e., motor functioning), and therefore it represents a promising new approach in malingering research. Prior to discussing the present study, an overview of essential malingering/effort concepts and measures will be provided.

Definition of Malingering

Malingering is defined by the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR)* as "the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives" (*DSM-IV-TR*, 2000, p. 739). In a survey of American Board of Clinical Neuropsychology members, Mittenberg, Patton, Canyock, and Condit (2002) estimated the base rates of malingering and symptom exaggeration of cognitive symptoms to be 29% in personal injury, 30% in disability, 19% in criminal, and 8% in medical cases. A clinician's lack of awareness regarding the credibility of a patient's effort threatens the validity of a psychological evaluation, as suboptimal effort significantly affects cognitive test scores (Green, 2007).

Types of Malingering

Individuals can demonstrate poor effort in a variety of ways. Resnick, West, and Payne (2008) emphasized that malingering is comprised of subtypes, including pure malingering, partial malingering, and false imputation. In pure malingering, a disorder

that does not exist is feigned. In partial malingering (the most common type), an individual has actual symptoms, but deliberately exaggerates them. In false imputation, an individual deliberately attributes actual symptoms to a cause that is unrelated to the symptoms. These subtypes can be found in cases of physical, psychiatric, and cognitive malingering.

Physical malingering. Individuals have malingered physical symptoms in a variety of ways, including feigned pain, blindness, deafness, and deafblindness. Each of these types of malingering is considered below.

Pain. The motivation to feign pain symptoms could be quite great for two reasons. First, detection of malingered pain is difficult because it is a subjective experience that is easily simulated, prone to exaggeration, and influenced by emotion (Cunnien, 1997). Second, pain complaints are frequently compensable, particularly if due to a work accident, motor vehicle accident, assault, or negligence of another party (McGuire & Shores, 2001). In a review of literature from 1961 to 1999, Fishbain, Cutler, Rosomoff, and Rosomoff (1999) highlighted that no reliable detection method of malingered or exaggerated pain existed. However, they estimated that 1.25% to 10.4% of pain claimants feigned or exaggerated pain complaints.

Blindness. Beatty (1999) estimated that visual complaints without known pathology accounted for approximately 1% of visual complaints seen by ophthalmologists. He used the term “non-organic visual loss” to describe visual disturbance with no evidence of dysfunction of the structures between the cornea and the occipital cortex. Beatty explained that non-organic visual loss, more common among females and younger patients, can be psychogenic (caused by higher cortical structures

responsible for visual awareness and outside the patient's control) or the result of malingering. He identified qualitative features that distinguish psychogenic patients from malingerers: psychogenic patients tend to be highly suggestible, while malingerers tend to exert too much effort to convince the physician of the dysfunction in question. Beatty described simple tests that can help to distinguish between malingered visual loss and organic visual loss. First, an organically blind person can easily sign his or her name, while a malingerer is often unable to do so. Second, tests of proprioception can also aid in a differential diagnosis. For example, a blind person can bring his index fingers together in front of his eyes, while the malingerer (falsely believing the test to be vision-dependent) cannot.

Mavrakanas and Schutz (2009) concluded that feigned visual loss was characterized by a number of features, including significant monocular vision loss that was not explained by ocular pathology; onset of visual loss days, weeks, or months after the trauma; progressive worsening of vision loss months to years after the trauma; fluctuating or inconsistent visual acuity or field test results; multiple inconsistent diagnoses by other ophthalmologists; normal optic nerves on CT or MRI neuroimaging; non-physiological or bizarre symptoms; significant symptoms reportedly occurring for years that were not documented in other medical records; and compensation-seeking for the injury or loss of vision.

Deafness. Rickards and De Vidi (1995) evaluated the rate of exaggerated hearing loss in 333 individuals seeking worker's compensation for noise-induced hearing loss. Researchers determined the rate of exaggerated hearing loss to be 17.7%; however, there was a large degree of difference between raters when using a subjective hearing measure.

Rickards and De Vidi emphasized the need for raters to follow appropriate test procedures as well as follow up with a second hearing test.

Deafblindness. Miner and Feldman (1998) presented a case study on two female patients who feigned both blindness and deafness. Both women presented with inconsistent, vague, and sometimes completely contradictory personal histories, demonstrated strange behavior inconsistent with deafblindness (e.g., both patients wore two pairs of dark sunglasses simultaneously), seemed unconcerned about the severity of their claimed symptoms, and performed activities that one with true deafblindness could not accomplish, especially when they did not know they were being observed.

Psychiatric malingering. Individuals have also malingered psychiatric symptoms in a number of ways, including feigned psychosis and posttraumatic stress disorder (PTSD). These two examples of psychiatric malingering are described below.

Psychosis. In a review of malingered psychosis, Resnick and Knoll (2008) cited that the unknown true prevalence of feigned psychosis is due to the fact that those who successfully feign are not captured in the statistics. They found that those who malingered psychosis often presented with symptoms of suicidal ideation, visual hallucinations, and memory problems, but were unlikely to present with negative symptoms such as flat affect, alogia, avolition, or impaired interactions, as these symptoms are more difficult to feign and less obvious. Additionally, malingerers typically presented with vague symptoms and symptoms that did not cluster into known diagnostic categories, often endorsed rare and atypical symptoms, and tended to draw attention to their symptoms.

PTSD. The concept of psychological symptoms following trauma is not new; past terms used to describe the experience include “nervous shock” and “post-traumatic neurosis” in the 1880s, “shell shock” during World War I, and “battle fatigue” in World War II (Resnick & Knoll, 2008). PTSD is the only psychiatric disorder that requires a causal link to an external event—an attribution that has been widely utilized by plaintiffs’ attorneys (Resnick & Knoll, 2008). Detecting feigned PTSD is quite difficult as the symptoms are subjective and widely known by the public. Resnick and Knoll (2008) highlighted a number of indicators of malingered PTSD, including overacting the part, providing excessively dramatic reports, being too eager to share a story or being excessively vague, hesitating to discuss a return to work or possible monetary compensation, indicating that a story is “too long” or “too complicated” to report, attempting to intimidate the interviewer or becoming hostile, possessing antisocial or psychopathic traits, and presenting with poor premorbid social and occupational functioning.

Resnick and Knoll (2008) explained that while some memory impairment is expected in PTSD, some characteristics indicate feigned amnesia, including overacting memory deficits, inability to remember over-learned information, claimed impairment of procedural memory, markedly poor performance on tests identified as memory measures, worse than chance performance on memory measures, and an ability to clearly recall the trauma despite claiming memory loss. Resnick and Knoll also advised psychologists to look for internal inconsistencies (e.g., an individual provides conflicting information to the same examiner), and external inconsistencies (e.g., an individual alleges social detachment, yet is seen happily participating in social and recreational activities). Finally,

Resnick and Knoll advised examiners to be wary if a patient's report conflicts or is inconsistent with hospital or police records of the traumatic event in question.

Cognitive malingering. Individuals have malingered cognitive symptoms in multiple ways, including feigned mental retardation (MR), amnesia, and cognitive loss due to traumatic brain injury (TBI). Each of these malingering domains is considered below.

MR. According to the DSM-IV-TR (APA, 2000):

The essential feature of Mental Retardation [MR] is significantly subaverage general intellectual functioning [IQ of 70 or below] that is accompanied by significant limitations in adaptive functioning in at least two of the following skill areas: communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health, and safety. The onset must occur before age 18 years. (p. 41)

Feigned MR was estimated to be between 13% and 17% in a competency to stand trial sample (Victor & Boone, 2007). Further, in 2002, the United States Supreme Court ruled that it was "cruel and unusual" punishment to subject a mentally retarded individual to the death penalty; thus, there may be great incentive to feign MR among criminal defendants (Victor & Boone, 2007).

The use of many frequently utilized measures in an MR evaluation is problematic. Most effort test measures are normed on individuals with "normal" intelligence; very little research has examined the use of effort measures with an MR population (Victor & Boone, 2007). Dean, Victor, Boone, and Arnold (2008) showed that individuals with an IQ of 60-69 failed approximately 44% of effort indicators, as compared to a 17% failure

rate in individuals with borderline intelligence (70-79), and a less than 10% failure rate in individuals with low average intelligence or higher (i.e., ≥ 80). Victor and Boone (2007) noted that simply lowering the cut-off scores has proven problematic, as an attempt to increase specificity in the detection of noncredible effort often results in lowered and unacceptable sensitivity. They concluded that unique effort tests are needed to evaluate those with MR, rather than adapting existing measures to an MR population.

Amnesia. Sweet, Condit, and Nelson (2008) emphasized the importance of understanding the characteristics associated with true memory impairment as such knowledge can help to identify feigned impairment. Genuine memory disorders are generally divided into two categories: those caused by medical conditions and those caused by psychological conditions. Amnesia usually refers to a loss of anterograde memory (i.e., ability to learn new information after a certain point in time); this memory loss is severe and significantly more severe than any other cognitive impairment that might be present (Sweet et al., 2008). Sweet et al. used the term “neurogenic” to refer to all amnesia and memory loss that is brain-based and occurring as a result of primary neurological disorder. Those with neurogenic amnesia often demonstrate a significant difference between free-recall and recognition performance, while other nonmemory functions are often preserved, including pre-illness memory in terms of intellectual, social, and language functioning, and previously acquired motor and cognitive abilities (Sweet et al., 2008). There are few psychological conditions that are associated with credible amnesia and memory loss; these few include dissociative/psychogenic amnesia, dissociative fugue, dissociative identity disorder, and factitious disorder (Sweet et al., 2008).

Feigned memory loss of a neurogenic etiology is more common in civil contexts, while feigned memory loss of a psychogenic or dissociative etiology is more common in criminal contexts (Sweet et al., 2008). For example, criminal defendants attempting to use an insanity defense are more likely to over-endorse psychopathology on measures like the MMPI-2, while civil litigants claiming memory impairment are more likely to underreport psychopathology on the same measures (Sweet et al., 2008). Amnesia for criminal activity is frequently claimed; up to 65% of individuals who commit murder are estimated to claim amnesia for the event, and the base rate for claims of at least partial amnesia for criminal activity is estimated to be at least 25% (Sweet et al., 2008).

Feigned cognitive loss associated with TBI. TBI is typically defined by trauma to the brain (through blunt or acceleration-deceleration forces) and subsequent signs and symptoms (Bender, 2008). There are approximately 1.5 million new cases of TBI every year in the United States; nearly 80% of those cases are classified as “mild” as defined by scores of 13-15 on the Glasgow Coma Scale (GCS), loss of consciousness less than 30 minutes, post traumatic amnesia (PTA) <24 hours, and normal brain imaging. Research (including five meta-analyses) shows that patients who experience mild brain trauma have returned to baseline by weeks to months post-injury (133 studies, $N = 1463$, Belanger, Curtiss, Demery, Lebowitz, & Vanderploeg, 2005; 21 studies, $N = 790$, Belanger & Vanderploeg, 2005; 120 studies, Carroll, Cassidy, Holm, Kraus, & Coronado, 2004; 17 studies, $N = 634$, Frencham, Fox, & Maybery, 2005; 39 studies, $N = 1716$, Schretlen & Shapiro, 2003), and if mild brain injury is associated with any long-term sequelae, it would be on the order of 1/8 standard deviation, or the equivalent of 2 IQ points (Millis & Volinsky, 2001), an essentially inconsequential change.

Moderate to severe TBI, GCS scores <13, brain imaging abnormalities, and PTA >24 hours, can result in highly variable scores on neuropsychological measures (Bender, 2008). TBI generally follows a dose-response curve, which means that the more severe the injury, the longer the expected recovery and the poorer the expected outcome (Bender, 2008). Bender (2008) reported that 20% to 40% of those with severe TBI are expected to make a “good” recovery; although some symptoms can persist indefinitely, symptoms are not expected to worsen over time unless fueled by a secondary psychological problem such as depression.

Bender (2008) noted that mild TBI is the most common diagnosis seen in forensic cases. Mittenberg et al. (2002), in survey data obtained from experienced neuropsychologists, found estimated base rates of malingering of cognitive symptoms in mild TBI of 40%, with estimates of feigning in the context of moderate to severe TBI of 9%. Signs that suggest possible neurocognitive malingering include reported impairment beyond what should be expected considering the injury severity, a degree of functional disability that is inconsistent with impairment, reported symptoms and/or resulting cognitive profiles that do not make sense neurologically, test scores inconsistent with known cognitive profiles, discrepant scores on tests that measure similar abilities, frequent near misses, passing more difficult items while failing easy ones, and quick “I don’t know” responses from the patient (Bender, 2008).

Research Approaches

The research design utilized in effort assessment studies is of critical importance. Rogers (2008) discussed three different research approaches, including simulation

studies, known-group comparison studies, and a combined approach that utilizes both simulation and known-group comparison studies.

Simulation studies. In simulation studies, a subset of analogue research, nonclinical participants randomly assigned to control groups are compared with convenient clinical samples and experimental groups instructed to simulate/feign (Rogers, 2008). Rogers (2008) cautions that simulation designs are of limited relevance as they “do not differentiate feigned from genuine disorders” (p. 413). Rogers is particularly critical of studies that utilize convenient student samples. Moreover he states, “Analogue feigning research that does not include the relevant comparison groups is fatally flawed and should not be published” (p. 413).

Known groups. Known-group comparison studies involve two distinct phases. First, criterion groups must be established, which includes identifying actual patients and malingerers. Second, the two criterion groups are systematically analyzed for similarities and dissimilarities (Rogers, 2008).

Combined. Rogers (2008) recommends the use of both the simulation and known-group designs in combination when validating dissimulation assessment methods. He explains, “The respective strengths of both designs are complimentary. Well-designed simulation studies address satisfactorily the need for experimental rigor (internal validity)...In contrast, known-groups comparisons address sufficiently the need for clinical relevance (external validity)” (p. 427).

Assessment of Noncredible Cognitive Symptoms

Definition of terms. Sensitivity, specificity, positive predictive power, and negative predictive power are key terms in evaluating effort. Sensitivity refers to the

proportion of actual positives correctly classified as such, while specificity is the proportion of actual negatives correctly classified as such. Positive predictive power is the probability that a positive test result is reflective of an actual positive result, while negative predictive power represents the probability that a negative test result is reflective of an actual negative result.

Dedicated/free-standing effort tests. There are two formats of dedicated/free-standing effort tests: forced-choice and non-forced-choice. Both methods are described below.

Forced-choice. Forced-choice tests involve the initial presentation of a series of stimuli, and a second presentation of the original stimulus alongside a wrong answer, or foil (Grote & Hook, 2007). The patient must identify the original stimulus from the distractor item. As a patient has a 50% chance of responding correctly on each item, interpretations regarding effort can be made as a patient's score deviates from chance (Grote & Hook, 2007). For example, a score significantly below chance is indicative of intentional poor performance, as one would be expected to answer approximately 50% of the items correctly even without viewing the initial stimulus presentation (Grote & Hook, 2007).

Test of Memory Malingering (TOMM). During the TOMM (Tombaugh, 1996) patients are shown 50 line drawings, one item at a time. Then, in Trial 1, the patient is presented 50 more pages with two drawings on each page: one of the original drawings and a foil. The patient must identify the original drawing, and is provided feedback from the examiner on whether the response is correct. Next, the original 50 drawings are presented again in a different order, followed by Trial 2, which includes 50 additional

pages, each page with an original drawing and a new foil. The TOMM also includes an optional Retention Trial, which is presented 15 minutes after Trial 2, but the original 50 drawings are not re-administered prior to the trial. Greve, Ord, Curtis, Bianchini, and Brennan (2008) found sensitivity rates on Trial 2 to be between 40% and 48% (98% specificity) when the cutoff score was <45 and between 50% (98% specificity) and 70% (93% specificity) when the cutoff score was ≤ 48 .

Word Memory Test (WMT). The WMT (Green, Allen, & Astner, 1996) is designed to assess both effort and verbal memory and can be presented orally by the examiner or administered on the computer. The patient is presented two learning trials of 20 semantically-related word pairs. Then the patient is presented with an immediate recognition trial in which the target word must be identified from foils. Next, following a 30-minute delay, the patient must again identify the target words from distractors in a delayed recognition trial. Following the delayed recognition trial, the patient is presented with three memory trials, including two cued memory trials and a free recall measure. Then, 20 minutes after the last free recall measure, the patient is administered a final free recall measure. Greve et al. (2008) found 78% sensitivity with 70% specificity at the original published cutoff score of ≤ 82.5 (WMT Consistency). An adjusted cutoff score of ≤ 72.5 (WMT Consistency) increased specificity to 93%, yet lowered sensitivity to 63% (Greve et al., 2008).

The Warrington Recognition Memory Test. The Warrington Recognition Memory Test, which is comprised of two subtests (Faces and Words; Warrington, 1984), was originally designed to assess verbal and nonverbal memory (Lu, Rogers, & Boone, 2007). The Words subtest has shown the most promise as an effort indicator. On this task, the

subject is shown 50 words, one at a time, and asked to rate each word as to its pleasantness. The subject is then presented with a page with 50 pairs of words; the person is asked to report which word in each pair was previously shown in the booklet. Kim et al. (2010) documented sensitivity of 89% (at $\geq 90\%$ specificity) using an accuracy cut-score of ≤ 42 , and 65% sensitivity using a time cut-off of ≥ 207 " in a large "real world" sample of noncredible and credible neuropsychological clinic patients.

Additional types of forced-choice effort tests include the Victoria Symptom Validity Test (VSVT; Slick, Hopp, Strauss, & Spellacy, 1996; Slick, Hopp, Strauss, & Thompson, 1997), the Portland Digit Recognition Test (PDRT; Binder, 1993a; Binder, 1993b; Binder & Willis, 1991), and the Computerized Assessment of Response Bias (CARB; Allen, Condor, Green, & Cox, 1997; Condor, Allen, & Cox, 1992).

Non-forced-choice. Nitch and Glassmire (2007) cited a number of potential advantages of non-forced-choice measures. First, they are less identifiable as effort tests and can more easily blend into a neuropsychological test battery. Second, a "typical" performance by someone who is truly impaired is not obvious to a test taker who intends to feign impairment; thus, an attempt to feign will likely result in rare or unbelievable response patterns. Third, the aforementioned measures are brief and easy to insert throughout the battery to assess effort continuously in the testing session, which is the current recommended practice (Bush et al., 2005).

Rey 15-Item Memory Test. Developed to evaluate the validity of visual memory complaints, the Rey 15-Item Memory Test (Rey, 1964) is one of the most frequently used effort tests (Nitch & Glassmire, 2007). The Rey 15-Item Memory Test consists of one page that contains 15 items (five rows of three items each); subjects are instructed to

study the page (presented for ten seconds), and then to reproduce as many of the items as possible once the page is removed. Using the original cut-off of <9 , Boone, Salazar, Lu, Warner-Chacon, and Razani (2002) found that only approximately 46% of noncredible subjects were correctly identified, but incorporating a recognition trial following the recall trial boosted sensitivity to 71% (cut-off of <20 for recall plus recognition total minus false positives on recognition). However, recent cross-validation data show a drop in sensitivity with the recognition trial to 58%, possibly related to test overexposure in the past decade (Boone & Lu, 2007).

Dot Counting Test. The Dot Counting Test (Boone, Lu, & Herzberg, 2002b; Rey, 1941), originally developed to identify malingered cognitive performance, consists of 12 cards with varying amounts of dots (half grouped and half ungrouped) that the patient is instructed to count as quickly as possible. Use of a cut-off of ≥ 17 (mean ungrouped dot counting time plus mean grouped dot counting time plus number of errors) was found to identify 78% of noncredible subjects (at $\geq 90\%$ specificity; Boone et al., 2002b), with recent cross validation data showing sensitivity of 73% (Boone & Lu, 2007).

b Test. The b Test, developed to assess feigned impairment in processing speed and recognizing over-learned information, requires an examinee to circle all the lowercase b letters that are intermingled between other letters that look similar to the letter b in a several-page booklet (Boone, Lu, & Herzberg, 2002a). Distortions of the letter b, combined with stimulus pages in which letters become progressively smaller, make the test appear difficult, although it is actually quite easy. Noncredible performance is characterized by slower completion time, failing to circle b letters, and incorrectly circling non-b letters. Use of a cut-off of ≥ 120 (mean time per page plus number of

omission errors plus [number of commission errors plus number of “d” commission errors x 10]) identified 74% of noncredible subjects (at $\geq 90\%$ specificity across most clinical comparison groups; Boone et al., 2002a).

Rey Word Recognition Test. In the 1940s, Andre Rey also developed the Rey Word Recognition Test to detect suspect effort on cognitive tests (Nitch, Boone, Wen, Arnold, & Alfano, 2006). In this test, the patient is read a list of 15 words and then asked to identify the words presented from a list of 15 targets and 15 foils. Nitch and colleagues (2006) found a significant gender effect, requiring separate gender cut-offs. Using a cut-off of ≤ 7 correct in women, 81% of noncredible women were detected (at 90% specificity), although the cut-off had to be reduced to ≤ 5 in men to achieve the same specificity, resulting in sensitivity of 63% in identifying noncredible men. Use of a cut-off of ≤ 9 for a combination equation (in which recognition of the first eight words in the list was double-weighted) resulted in 82% sensitivity in a mild TBI subset of the noncredible sample (Nitch et al., 2006).

Embedded effort indicators. The aforementioned effort measures are free-standing and ultimately add length to the testing battery. Researchers have begun to develop embedded effort indicators derived from tests that are already part of the standard neuropsychological battery.

The increasing availability of embedded effort indices derived from standard cognitive tests, some with sensitivity values equal or higher than those of some free-standing effort measures ... provides the opportunity to increase the number of effort indicators without adding to test battery time. (Boone, 2009, pp. 737-738)

Attention. Many patients believe the Digit Span subtest of the Wechsler Adult Intelligence Scale (WAIS; Revised, III, and IV; Wechsler 1955, 1981, 1997a, 2008), in which they are asked to repeat a string of digits recited by the examiner, to be a measure of memory (although it is a measure of attention) and attempt to demonstrate their “memory impairments” on this subtest (Babikian & Boone, 2007). Babikian and Boone (2007) reviewed a number of studies on the use of the Digit Span subtest to detect malingering and documented several detection strategies, including the use of: Digit Span Age-Corrected Scaled Score (ACSS), Reliable Digit Span (RDS; the sum of the longest number of digits correctly recited over two trials, both forward and backward), and time to recite digits forward. For Digit Span ACSS, sensitivity rates ranged from 36% to 47% (with specificity $\geq 90\%$) when the cutoff was set to ≤ 5 ACSS; for an RDS cutoff of ≤ 7 , sensitivity rates ranged from 50% to 87% (but demonstrated compromised specificity for moderate to severe traumatic brain injury); for timed forward digit span, data suggested sensitivity rates between 37% and 50%.

Executive. The Wisconsin Card Sorting Test (WCST; Berg, 1948; Grant & Berg, 1948) is a measure of executive functioning that requires patients to match cards according to category with limited instruction and feedback from the examiner. The categories shift and the patient must shift set based on examiner feedback. In one of the earliest studies investigating the use of the WCST as an effort measure, Bernard, McGrath, and Houston (1996) developed a series of discriminant functions to detect malingering. *Categories completed* was the only WCST index to successfully detect malingered performance (100% sensitivity and 92% specificity). *Categories completed* and *perseverative errors* were entered into a discriminant analysis and yielded 86%

sensitivity and 94% specificity when discriminating between malingerers and closed head injury patients. More recent studies have yielded mixed results, with variables such as age and severity of head injury complicating the ability to detect invalid performance (Sweet & Nelson, 2007). Sweet and Nelson (2007) argue that the most effective use of the WCST in the detection of insufficient effort involves the use of multivariate formulae, but they warn the examiner to use caution when interpreting WCST performance in terms of effort, as much of the research has included head-injured patient samples and might not be appropriate for use with other patient populations.

Other tests of executive functioning that have been investigated for use as effort measures include the Category Test, Booklet Category Test, verbal and figural fluency tasks, Stroop Color-Word Test, and the Trailmaking Test (see Sweet & Nelson, 2007).

Sensory/motor. The Finger Tapping Test (Halstead, 1947; Reitan & Wolfson, 1993) is a motor functioning measure that utilizes a lever (finger tapper) mounted on a board. Patients must use the index finger of each hand (with the remaining fingers, hand, and wrist in a flat and still position) to tap the lever as many times as possible in a 10-second period (Arnold et al., 2005). Trials are completed for both the dominant and nondominant hand and scores are averaged across the trials. Arnold et al. (2005) found that men tapped faster than women (which required the groups be divided by gender), and the dominant hand score was more sensitive to noncredible performance, particularly in women. Arnold et al. found that a dominant hand cutoff score of ≤ 35 yielded 50% sensitivity for men, while a cutoff score of ≤ 28 yielded 61% sensitivity for women (when specificity was set at 90%).

In a study of head-injured patients, Binder, Kelly, Villanueva, and Winslow (2003) found that tactile sensation was more affected by motivation than by the severity of the head injury. Tactile Finger Recognition (Finger Agnosia; Reitan & Wolfson, 1993) requires that, with eyes closed, patients identify which finger the examiner touches; a total of 4 trials are obtained for each finger in a randomized order. Binder et al. (2003) found that a score of more than 4 errors was 93% specific in the moderate-severe head-injured group with good motivation, 82% specific in the mild head-injured group with good motivation, and 56% sensitive to poor effort in the mild head-injured group with poor motivation.

Verbal memory. The Rey Auditory-Verbal Learning Test (RAVLT; Rey, 1964; Schmidt, 1996), a word list-learning task, consists of a five-time list presentation, an interference trial, a delayed-recall trial, and a recognition trial. Research shows that while number correct for the recognition trial is effective in identifying suboptimal effort (e.g., 67% sensitivity with a cut-off of <10), the most sensitive measure involves double-weighting recognition for the first five words on the list (recognition minus false positives plus the number of words recognized from the first five words on the list; Boone, Lu, & Wen, 2005). Using a cut-off of ≤ 12 , 74% of noncredible subjects were identified (at 90+% specificity).

The California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987) is a list-learning task similar to the RAVLT, except that the CVLT has 16 words that belong to one of four semantic categories and there are two additional recall trials; the California Verbal Learning Test – Second Edition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000) includes a new word list and an additional forced-choice recognition trial.

Use of a cutoff score of 14 on forced-choice recognition yielded a sensitivity of 44% with 93% specificity (Root, Robbins, Chang, & van Gorp, 2006).

Visual memory. The Rey-Osterrieth Complex Figure Test (ROCFT; Rey, 1941) is a frequently used measure of visual memory and visuoconstructive skills (Lu, Boone, Cozolino, & Mitchell, 2003). The recent addition of a recognition memory test (with 12 design portions from the original ROCFT stimulus mixed with 12 foils; Meyers & Meyers, 1995) has increased the test's utility as an effort measure, as patients often mistakenly believe that recognition memory is as impaired as free recall in brain injury (Lu et al., 2003). Lu et al. (2003) found that utilizing a combination score that included the copy, true positive recognition, and atypical recognition error scores yielded 74% sensitivity, resulting in only 4% of verbal memory impaired patients, 12% of visual memory impaired patients, and 3% of nonmemory impaired patients being misclassified.

The Continuous Visual Memory Test (CVMT; Trahan & Larrabee, 1988) is a test of visual recognition memory in which a patient is presented a series of abstract visual designs in different categories at the rate of one every two seconds over six trial blocks (Larrabee, 2009). Nine stimuli appear only once, while seven designs recur; the patient must identify the stimuli as "new" (appearing for the first time) or "old" (previously presented). The initial presentation is followed by a 30-minute delay. With a total of 103 scoreable items, Larrabee (2009) identified 20 items on the CVMT that discriminated between litigants identified as definite malingerers and those with traumatic brain injury. Larrabee found that a score of <14 correctly classified 83.3% of the malingerers, with a false positive rate of 11.1% in the traumatic brain injury group (88.9% specificity).

The Faces subtest of the Wechsler Memory Scale – Third Edition (WMS-III; Wechsler, 1997b) consists of 24 color pictures of human faces. Patients are shown the 24 pictures and then immediately are shown another set of 48 faces, half of which are the original pictures and the other half are new pictures. The patient must identify with a “yes” or “no” response if the face was one of the original pictures shown (Glassmire, 2003). The subtest includes a 25- to 35-minute delay with 48 additional trials. Glassmire (2003) sought to develop a malingering index for the Faces subtest and found that the total raw score yielded the strongest classification accuracy. A raw score cutoff of 31 achieved 93.3% sensitivity and 80.0% specificity in nonlitigating traumatic brain injury patients and controls in a simulation design.

Visual perceptual spatial. The Benton Visual Form Discrimination (VFD; Benton, Sivan, Hamsher, Varney, & Spreen, 1994) test is a measure of visual perception requiring that the patient examine a series of line drawings and find their identical matches in a four-choice array. Larrabee (2003) found that a raw score of < 26 on the VFD test accurately classified 12 of 25 known malingerers (48%) and 27 of 29 (93.1%) of those with moderate to severe closed head injury. Further, Larrabee found that no closed head injury patient scored less than 24 on VFD.

Combinations of scores. Recent studies have found that the use of a combination of scores is considerably more effective in the detection of suspect effort than the use of individual scores alone by offering increased sensitivity without lowering specificity rates (Boone et al., 2002; Boone et al., 2005; Lu et al., 2003). Lu et al. (2003) found that the use of a combination of ROCFT scores increased sensitivity by 50% while maintaining a high specificity. Other advantages of using embedded effort indicators in the form of a

combination of scores include the ability to assess effort without adding time or testing measures to the battery (Boone, 2009; Lu et al., 2003), as well as the approach's resistance to coaching. For example, it would be impossible for someone to know how to perform when multiple scores are compared and weighted in a formula or discriminant function.

Meyers and Volbrecht (2003) developed a linear regression formula from a database of 650 neuropsychological patients with varied diagnoses to examine effort using a combination of scores from tests that evaluate a common function (i.e., motor). Specifically, they examined the Rey-Osterrieth Complex Figure Test (copy) and the WAIS-III Digit Symbol and Block Design subtests in predicting Finger Tapping scores. An actual Finger Tapping score more than 10 points below the estimated score served as an indication of suspect effort. The exact formula was as follows: $[(\text{ROCFT raw score} \times .185) + (\text{Digit Symbol scale score} \times .491) + (\text{Block Design scale score} \times .361) + 31.34]$ to calculate an estimated Finger Tapping score. Meyers and Volbrecht did not discuss how the formula was developed or the rationale for the specific tests used, other than their evaluation of a common function (i.e., motor). Meyers and Volbrecht developed a total of nine individual methods to detect malingering incorporating commonly administered neuropsychological tests, thus evaluating effort without adding test measures or time to the battery.

Meyers and Volbrecht (2003) investigated the use of the aforementioned formula along with the additional eight methods using a clinical sample of 796 participants. Participants included mild, moderate, and severe TBI patients, chronic pain and depressed patients, community controls, malingering simulators, and institutionalized and

non-institutionalized patient groups; participants also included both litigating and non-litigating groups. They found that failure on any two of the nine malingering tests suggested suspect effort/malingering with 83% sensitivity and 100% specificity (resulting in 0 false positive identifications). Meyers and Volbrecht emphasized that these measures are particularly appropriate for assessing effort in those with brain injury, chronic pain, and depression. They cautioned that the measures might be inappropriate for those who are neuropsychologically unable to be tested, reside in 24-hour institutional care, present with cerebrovascular accident (CVA) that affects one's ability to understand simple directions, or present with advanced dementia or mental retardation.

The only known documented use of the formula was in the Meyers and Volbrecht (2003) study. More research is needed to determine the usefulness and applicability of the formula in contemporary samples. The purpose of the current study was to cross-validate the Meyers and Volbrecht formula utilizing scores from a large group of patients documented as showing noncredible cognitive performances and a comparison group of credible patients with heterogeneous neurologic and psychiatric diagnoses. Empirical evidence of the viability of the formula may prove useful to neuropsychologists and researchers seeking effective methods for the detection of noncredible effort.

Method

Subjects

Subjects were referred for neuropsychological assessment to the Harbor-UCLA Medical Center Department of Psychiatry Outpatient Neuropsychology Service in Torrance, CA. Patients were primarily referred by treating psychiatrists or neurologists for diagnostic clarification, case management, and/or determination of appropriateness

for disability compensation. IRB approval to examine archival data was obtained from the hospital-affiliated research institute (Los Angeles Biomedical Institute) and from Pepperdine University's Graduate and Professional Schools IRB. Of the 281 cases identified, 101 were assigned to the noncredible group and 180 to the credible group. Criteria for inclusion and exclusion within noncredible and credible groups are described below. All participants were fluent in English and most were native English-speakers. On average, credible and noncredible groups were in their early 40s with 12-13 years of education. Representation of men and women was roughly equivalent in the sample as a whole. The majority of patients were Caucasian, African American, or Latino, although other ethnicities were also represented.

Patients with suspect effort. The 101 noncredible patients met Slick, Sherman, and Iverson (1999) criteria for probable malingered neurocognitive dysfunction. Specifically, all were seeking to obtain disability benefits for cognitive symptoms associated with alleged medical or psychiatric disorders; all failed \geq two independent effort indicators (tests and cut-offs listed in Table 1) not due to other psychiatric, neurologic, or developmental disorders; and low standard cognitive scores were at variance with evidence of normal functioning in activities of daily living. For example, noncredible patients scoring below 70 on the FSIQ were not excluded from the sample if their test performance appeared inconsistent with their demonstrated capabilities in other areas (e.g., job performance, ability to live independently), as their diminished performance was likely a result of effort rather than true intellectual deficiency. Data on age, education, gender, and ethnic composition of the sample are reported in Table 2.

Frequency of presenting diagnoses were psychosis/rule-out psychotic disorder/major depression with psychotic features ($n = 18$), mild TBI ($n = 14$), depression ($n = 13$), learning disability ($n = 10$), severe TBI ($n = 8$), stroke/aneurysm ($n = 6$), epilepsy/seizure disorder ($n = 5$), mental retardation ($n = 5$), anxiety/panic disorder ($n = 3$), substance abuse ($n = 3$), chronic pain ($n = 2$), vascular dementia ($n = 2$), electrocution ($n = 2$), somatoform disorder ($n = 2$), meningitis ($n = 1$), moderate TBI ($n = 1$), syncope ($n = 1$), dementia ($n = 1$), HIV ($n = 1$), mold exposure ($n = 1$), cognitive disorder NOS ($n = 1$), and anoxia ($n = 1$).

Credible patients. The 180 credible subjects were not in litigation or seeking to obtain disability benefits for cognitive symptoms and failed ≤ 1 effort indicator (tests and cut-offs listed in Table 1). Patients who failed one effort test were retained in the sample because research shows that failure on a single effort indicator among several is not unusual in credible populations (Victor, Boone, Serpa, Beuhler, & Ziegler, 2009). Patients with a FSIQ lower than 70 or a dementia or amnesic disorder diagnosis were excluded. Data on age, education, gender, and ethnic composition of the sample are reported in Table 2.

Final diagnoses (i.e., determined by history and cognitive test results) included depression/rule-out depression ($n = 33$), learning disability/rule-out learning disability ($n = 22$), somatoform disorder/rule-out somatoform disorder ($n = 16$), psychosis/major depression with psychosis/rule-out major depression with psychotic features ($n = 15$), seizure disorder/epilepsy ($n = 14$), bipolar disorder/bipolar disorder with psychosis ($n = 11$), severe TBI ($n = 11$), stroke/aneurysm ($n = 9$), substance abuse ($n = 9$), HIV ($n = 5$), anxiety disorder/panic disorder/obsessive-compulsive disorder ($n = 5$), anoxia/rule-out

anoxia ($n = 5$), ADHD ($n = 5$), cognitive disorder NOS/rule-out cognitive disorder NOS ($n = 4$), multiple sclerosis ($n = 3$), brain tumor ($n = 3$), mild TBI ($n = 1$), moderate TBI ($n = 1$), hydrocephalus ($n = 1$), PTSD ($n = 1$), encephalitis ($n = 1$), rule-out frontotemporal dementia ($n = 1$), prenatal substance exposure ($n = 1$), rule-out Asperger's disorder ($n = 1$), Klinefelter syndrome ($n = 1$), and end stage liver disease ($n = 1$).

Instruments/Procedures

The Finger Tapping Test, WAIS-III Block Design and Digit Symbol subtests, and the ROCFT copy trial were administered in standard format as part of a comprehensive neuropsychological battery. The scores used for analysis were the ROCFT copy trial raw score, the Block Design scale score, the Digit Symbol scale score, the average Finger Tapping score for the dominant hand, and the full Meyers and Volbrecht (2003) formula $[(\text{ROCFT raw score} \times .185) + (\text{Digit Symbol scale score} \times .491) + (\text{Block Design scale score} \times .361) + 31.34]$ to calculate an estimated Finger Tapping score. A partial version of the formula excluding Finger Tapping scores was also evaluated.

Analyses

Groups were compared on age and education, and on all test scores, via independent t tests. To examine potential impact of gender on test scores and the Meyers and Volbrecht (2003) equation, performance of men and women was compared in each group separately. Correlational analyses were conducted separately within each group to examine whether test scores were significantly related to age or education. Correlations were also computed between all individual test scores within each group separately to examine extent of redundancy.

Score frequency counts were computed for the individual test scores and the Meyers and Volbrecht (2003) equation in each group separately for the purposes of assessing sensitivity of cut-scores selected for $\geq 90\%$ specificity. The 10% of credible subjects failing the equation cut-score were examined to determine if commonalities could be found that could be used to flag those credible patient groups who might be at risk for false positive identification despite applying credible effort.

Results

As can be seen in Table 2, groups did not differ in age but did differ in educational level (credible subjects averaged one more year of education than noncredible subjects); gender distribution was roughly equivalent across groups. Correlations between test scores and age and education for each group separately revealed in credible subjects significant correlations between educational level and Digit Symbol scale score, Block Design scale score, Rey-O copy trial raw score, and the partial Meyers formula (r^2 's = 0.198 to 0.300). No significant relationships were found between age or education and test scores in the noncredible group, and given that education accounted for less than 10% of test score variance in the credible group, it was not further considered in data analyses.

The means and standard deviations of scores obtained by credible and noncredible participants on the individual subtest scores that comprise the Meyers and Volbrecht (2003) formula as well as the whole and partial formula scores are shown in Table 2. As can be seen from the table, highly significant group differences, in favor of better performance in credible patients, were observed in independent t -test analyses.

Cut-offs for whole and partial Meyers and Volbrecht (2003) formula scores, Rey-O copy trial raw score, Digit Symbol scale score, Block Design scale score, and average Finger Tapping score for the dominant hand were selected to maintain specificity in the credible patients of at least 90%. The associated sensitivity rates when cut-offs were applied to the noncredible group are shown in Table 3. As can be seen, all of the individual scores with the exception of Finger Tapping outperformed the full Meyers and Volbrecht formula (30.7% sensitivity). When the cut-off recommended by Meyers and Volbrecht was employed (actual minus estimated Finger Tapping score <-10), sensitivity was slightly higher (38.6%), but specificity was unacceptable (85.6%).

Given concerns that the poor sensitivity rate for Finger Tapping was suppressing the effectiveness of the Meyers and Volbrecht (2003) formula, the Finger Tapping component was removed and the sensitivity rate for the remaining portion of the equation utilizing ROCFT copy, Block Design, and Digit Symbol was calculated. As can be seen from Table 3, this shortened equation yielded the highest sensitivity (70.3% at $\pm 90\%$ specificity), outperforming all individual subtests and the whole formula by a large margin. Application of the formula resulted in some apparent differences in ethnic group composition between the credible and noncredible groups.

To check for possible gender effects, independent *t*-tests were calculated comparing credible men and women (Table 4), and noncredible men and women (Table 5) on all scores. As can be seen from Table 4, credible men outperformed credible women on Block Design scaled score, dominant Finger Tapping, and the whole Meyers and Volbrecht (2003) formula. Noncredible women obtained higher Digit Symbol scaled

scores than noncredible men, while the latter scored higher than noncredible women on the whole Meyers and Volbrecht formula.

No significant differences were observed in gender comparisons in credible and noncredible groups on the partial Meyers and Volbrecht (2003) formula. However, when cut-offs for the first portion of the Meyers and Volbrecht formula were chosen to maintain specificity of $\geq 90\%$ in each gender separately, a cut-score of ≤ 10.47 could be used with men (77% sensitivity), while a cut-score of ≤ 9.94 was necessary with women (55% sensitivity).

The diagnostic and demographic characteristics of the nine (out of 94) credible female participants who fell below the cut-off of ≤ 9.94 on the partial Meyers and Volbrecht (2003) formula were examined for potential risk factors for false positive identification. On average, these participants were 49.22 years old ($SD = 7.29$), had a mean educational level of 11.22 years ($SD = 3.77$), and were ethnically diverse: Caucasian ($n = 2$), African American ($n = 2$), Latina ($n = 4$), and Biracial (Latina/Native American; $n = 1$). Final diagnoses for this group included bipolar disorder ($n = 1$), learning disability ($n = 2$), multiple sclerosis ($n = 1$), depression ($n = 1$), substance abuse ($n = 1$), psychosis ($n = 1$), brain tumor ($n = 1$), and rule-out somatoform disorder ($n = 1$).

The diagnostic and demographic characteristics of the eight (out of 86) credible male participants who fell below the cut-off of ≤ 10.47 on the partial Meyers and Volbrecht (2003) formula were also examined for potential risk factors for false positive identification. On average, these participants were 39.88 years old ($SD = 18.88$), had a mean educational level of 12.13 years ($SD = 1.46$), and again were ethnically diverse: Caucasian ($n = 3$), African American ($n = 2$), and Latino ($n = 3$). Final diagnoses for this

group included bipolar disorder ($n = 1$), severe TBI ($n = 1$), rule-out somatoform disorder ($n = 1$), anoxia ($n = 1$), learning disability ($n = 1$), encephalitis ($n = 1$), seizures ($n = 1$), and stroke ($n = 1$).

The only discernable pattern in those credible subjects failing the partial Meyers and Volbrecht (2003) formula was that all 17 subjects had a WAIS-III Performance IQ < 84. PIQ scores in this subgroup ranged from 64 to 83, with a mean of 76.24 as compared to a mean PIQ of 94.6 (SD = 13.7) in the credible sample as a whole. Of the credible subjects with PIQ <84, 43.6% fell below cut-offs on the partial formula.

We had been concerned that stroke patients might be at particular risk for failure given the possibility of weakness of the contralateral upper extremity that could interfere with psychomotor test execution. However, only one of nine stroke patients in the credible group fell below cut-offs (89% specificity).

Discussion

The present study re-examined the Meyers and Volbrecht (2003) motor formula for the detection of suspect effort using scores in a large “known groups” sample of credible and noncredible patients from an outpatient neuropsychology clinic. Findings from the current study revealed that the complete formula, developed to identify nonplausibly poor Finger Tapping scores, was ineffective at capturing poor effort (30.7% sensitivity at $\geq 90\%$ specificity). Scrutiny of the individual scores comprising the formula showed that Finger Tapping in isolation had a very low sensitivity rate comparable to that of the complete formula (29.7% versus 30.7%). It was reasoned that the ineffectiveness of Finger Tapping in identifying poor effort was suppressing the sensitivity rate for the entire formula. The decision was made to delete the estimated and actual Finger Tapping

scores and to retain only the first part of the formula: (ROCFT raw score x .185) + (Digit Symbol scale score x .491) + (Block Design scale score x .361). Application of a cut-off to this partial formula significantly increased sensitivity (70.3%) while still maintaining specificity of $\geq 90\%$.

While group comparisons on the partial formula yielded no significant gender differences, when cut-offs were chosen to maintain specificity of $\geq 90\%$ in each gender separately, a cut-score of ≤ 10.47 could be used with men (77% sensitivity in noncredible men), while a cut-score of ≤ 9.94 was necessary with women (55% sensitivity in noncredible women). These findings show that the partial Meyers and Volbrecht (2003) formula is a much more effective measure of response bias in men, and raise the very intriguing likelihood that men are more likely than women to target constructional/psychomotor tasks on which to display suboptimal effort. Some research shows male superiority on tasks of mental rotation (Mann, Sasanuma, Sakuma, & Masaki, 1990), motor learning (Dorfberger, Adi-Japha, & Karni, 2009), and speed of motor performance (Jiménez-Jiménez et al., 2011), and men more than women may perceive that poor performance in these areas will better demonstrate “disability.” The results indicate that gender-based cut-scores do not change classification accuracy in men over that found with non-gendered cut-offs, but are necessary to avoid slightly elevated misclassification of credible females as noncredible (application of the non-gendered cut-score of ≤ 10.3 was associated with only 87.2% specificity in women).

The diagnostic and demographic characteristics of the 17 (out of 180) credible female and male participants who fell below the gender-specific cut-offs on the partial Meyers and Volbrecht (2003) formula were examined for potential risk factors for false

positive identification. The only commonality was WAIS-III Performance IQ < 84; in fact, 43.6% of credible subjects with PIQ in this range fell below cut-offs (56.4% specificity). This is not unexpected given that two of the three tests comprising the partial formula are WAIS-III subtests. However, these findings indicate that use of the partial formula is problematic in populations with lowered PIQ, and should either not be used, or cut-offs need to be adjusted, although this will sacrifice sensitivity (e.g., for males, a cut-off of <6.98 is associated with 100% specificity and 25% sensitivity, while a cut-off of <7.65 yields 100% specificity for women at 28% sensitivity). The practical problem with avoiding use of the partial formula with individuals with low PIQ is that individuals applying suboptimal effort on neuropsychological exams also typically obtain spuriously low IQ scores. Therefore, the partial Meyers and Volbrecht formula is judged appropriate for use among test takers with evidence of intact premorbid intellectual functioning who subsequently claim conditions not shown to markedly lower IQ scores (e.g., mild TBI, depression, learning disability, ADHD, most substance abuse, chronic pain, anxiety, PTSD, etc.), regardless of obtained PIQ score.

The current study is not without limitations. First, credible and noncredible groups were not evenly distributed in terms of ethnicity. However, in credible samples, those who failed the equation represented each of the major ethnic groups that comprised the study's sample (e.g., Caucasian, African American, and Latino) and PIQ scores, not ethnicity, appeared to be the determining factor that underpinned failure on the equation. Thus, as previously discussed, use of the formula in populations with credible PIQ scores < 84 is not advised without caution. Second, as the study utilized an archival database, there was no opportunity to vary or alter the procedures used. Also, data was collected in

one treatment setting, which could limit the applicability and generalizability of the current findings to other communities and locales. Third, the formula was based on WAIS-III subtests (e.g., Digit Symbol and Block Design). Since the development of the equation, the WAIS-IV was published; it is uncertain how the formula might translate for use with WAIS-IV subtests.

Recommendations for Future Research

Given the current findings, it is recommended that the formula's use be investigated in other heterogeneous real-world samples. For example, the present study utilized patient groups from an urban, largely underserved community. It would be beneficial to learn how the formula's use might translate to other settings. Such replication studies may also shed light on factors such as the impact of IQ, gender, and ethnicity on application of the formula. Additionally, future studies might investigate the formula's use with the Digit Symbol and Block Design subtests from the WAIS-IV to further increase the equation's potential utility.

References

- Allen, L. M., Condor, R. L., Green, P., & Cox, D. R. (1997). *CARB '97: Computerized Assessment of Response Bias. Manual*. Durham, NC: CogniSyst.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision; pp. 739-740). Washington, DC: Author.
- Arnold, G., Boone, K. B., Lu, P., Dean, A., Wen, J., Nitch, S., & McPherson, S. (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
- Babikian, T., & Boone, K. B. (2007). Intelligence tests as measures of effort. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment* (pp. 103-127). New York, NY: Guilford Press.
- Babikian, T., Boone, K., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspect effort. *The Clinical Neuropsychologist, 20*, 145-159. doi: 10.1080/13854040590947362
- Beatty, S. (1999). Non-organic visual loss. *Postgraduate Medical Journal, 75*, 201-207.
- Belanger, H. G., Curtiss, G., Demery, J. A., Lebowitz, B. K., & Vanderploeg, R. D. (2005). Factors moderating neuropsychological outcomes following mild traumatic brain injury: A meta-analysis. *Journal of the International Neuropsychological Society, 11*, 215-227. doi: 10.1017/S1355617705050277

- Belanger, H. G., & Vanderploeg, R. D. (2005). The neuropsychological impact of sports-related concussion: A meta-analysis. *Journal of the International Neuropsychological Society, 11*, 345-357. doi: 10.1017/S1355617705050411
- Bender, S. D. (2008). Malingered traumatic brain injury. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 69-86). New York, NY: Guilford Press.
- Benton, A. L., Sivan, A. B., Hamsher, K., Varney, N. R., & Spreen, O. (1994). *Contributions to neuropsychological assessment. A clinical manual* (2nd ed.). New York, NY: Oxford University Press.
- Berg, E. A. (1948). A simple objective treatment for measuring flexibility in thinking. *Journal of General Psychology, 39*, 15-22.
- 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 "pattern of performance hypothesis." *Archives of Clinical Neuropsychology, 11*(3), 231-245. doi: 10.1016/0887-6177(95)00025-9
- Binder, L. M. (1993a). An abbreviated form of the Portland Digit Recognition Test. *The Clinical Neuropsychologist, 7*, 104-107. doi: 10.1080/13854049308401892
- Binder, L. M. (1993b). Assessment of malingering after mild head trauma with the Portland Digit Recognition Test. *Journal of Clinical and Experimental Neuropsychology, 15*(2), 170-182. doi: 10.1080/01688639308402555

- Binder, L. M., Kelly, M. P., Villanueva, M. R., & Winslow, M. M. (2003). Motivation and neuropsychological test performance following mild head injury. *Journal of Clinical and Experimental Neuropsychology*, 25(3), 420-430. doi: 10.1076/jcen.25.3.420.13806
- Binder, L. M., & Willis, S. C. (1991). Assessment of motivation after financially compensable minor head trauma. *Psychological Assessment*, 3, 175-181. doi: 10.1037/1040-3590.3.2.175
- Boone, K. B. (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological examinations. *The Clinical Neuropsychologist*, 23(4), 729-741. doi: 10.1080/13854040802427803
- Boone, K. B., Lu, P., & Herzberg, D. (2002a). The b Test. Los Angeles, CA: Western Psychological Services.
- Boone, K. B., Lu, P., & Herzberg, D. (2002b). Rey Dot Counting Test. Los Angeles, CA: Western Psychological Services.
- Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002) The Rey 15-Item Recognition Trial: A technique to enhance sensitivity of the Rey 15-Item Memorization Test. *Journal of Clinical and Experimental Neuropsychology*, 24(5), 561-573. doi: 10.1076/jcen.24.5.561.1004
- Boone, K. B., Lu, P., & Wen, J. (2005). Comparison of various RAVLT scores in the detection of noncredible memory performance. *Archives of Clinical Neuropsychology*, 20, 301-319. doi: 10.1016/j.acn.2004.08.001

- Boone, K. B., & Lu, P. H. (2007). Non-forced-choice effort measures. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 27-43). New York, NY: Oxford University Press.
- Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., . . . Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity (NAN Policy and Planning Committees). *Archives of Clinical Neuropsychology, 20*, 419-426. doi: 10.1016/j.acn.2005.02.002
- 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 for Neurotrauma Task Force on Mild Traumatic Brain Injury. *Journal of Rehabilitation Medicine, 43*, 113-125. doi: 10.1080/16501960410023877
- Condor, R., Allen, L., & Cox, D. (1992). *Computerized Assessment of Response Bias test manual*. Durham, NC: CogniSyst.
- Cunnie, A. J. (1997). Psychiatric and medical syndromes associated with deception. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 23-46). New York, NY: Guilford Press.
- Dean, A. C., Victor, T. L., Boone, K. B., & Arnold, G. (2008). The relationship of IQ to effort test performance. *The Clinical Neuropsychologist, 22*, 705-722. doi: 10.1080/13854040701440493
- Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test (CVLT). Adult Version (Research ed.)*. San Antonio, TX: Psychological Corporation.

- Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test – Second Edition (CVLT-II)*. San Antonio, TX: Psychological Corporation.
- Dorfberger, S., Adi-Japha, E., & Karni, A. (2009). Sex differences in motor performance and motor learning in children and adolescents: An increasing male advantage in motor learning and consolidation phase gains. *Behavioural Brain Research, 198*, 165-171. doi: 10.1016/j.bbr.2008.10.033
- Fishbain, D. A., Cutler, R., Rosomoff, H. L., & Rosomoff, R. S. (1999). Chronic pain disability exaggeration/malingering and suboptimal effort research. *The Clinical Journal of Pain, 15*(4), 244-274. doi: 10.1097/00002508-199912000-00002
- Frencham, K. A. R., 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*(3), 334-351. doi: 10.1080/13803390490520328
- Glassmire, D. M. (2003). Using the WMS-III Faces subtest to detect malingered memory impairment. *Journal of Clinical and Experimental Neuropsychology, 25*(4), 465-481. doi: 10.1076/jcen.25.4.465.13875
- Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of the degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology, 38*, 404-411. doi: 10.1037/h0059831

- Green, P. (2007). Spoiled for choice: Making comparisons between forced-choice effort tests. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment* (pp. 50-77). New York, NY: Guilford Press.
- Green, P., Allen, L. M., & Astner, K. (1996). *The Word Memory Test: A user's guide to the oral and computer-administered forms*. Durham, NC: CogniSyst.
- Greve, K. W., Ord, J., Curtis, K. L., Bianchini, K. J., & Brennan, A. (2009). Detecting malingering in traumatic brain injury and chronic pain: A comparison of three forced-choice symptom validity tests. *The Clinical Neuropsychologist*, 22(5), 896-918. doi: 10.1080/13854040701565208
- Grote, C. L., & Hook, J. N. (2007). Forced-choice recognition tests of malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 44-79). New York, NY: Oxford University Press.
- Halstead, W. C. (1947). *Brain and intelligence*. Chicago, IL: University of Chicago Press.
- Iverson, G. L., & Franzen, M. D. (1998). Detecting malingered memory deficits with the Recognition Memory Test. *Brain Injury*, 12, 275-282. doi: 10.1080/026990598122575
- Jiménez-Jiménez, F. J., Calleja, M., Alonso-Navarro, H., Rubio, L., Navacerrada, F., Pilo-de-la-Fuente, B., ... Agúndez, J. A. (2011). Influence of age and gender in motor performance in healthy subjects. *Journal of the Neurological Sciences*, 302(1-2), 72-80. doi: 10.1016/j.jns.2010.11.021
- Kim, M. S., Boone, K., Victor, T., Marion, S., Amano, S., Cottingham, M., . . . Zeller, M. (2010). The Warrington Recognition Memory Test for Words as a measure

- of response bias: Total score and response time cutoffs developed on “real world” credible and noncredible subjects. *Archives of Clinical Neuropsychology*, 25, 60-70. doi: 10.1093/arclin/acp088
- Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, 17(3), 410-425. doi: 10.1076/clin.17.3.410.18089
- Larrabee, G. J. (2009). Malingering scales for the Continuous Recognition Memory Test and The Continuous Visual Memory Test. *The Clinical Neuropsychologist*, 23, 167-180. doi: 10.1080/13854040801968443
- Lu, P. H., Boone, K. B., Cozolino, L., & Mitchell, C. (2003). Effectiveness of the Rey-Osterrieth Complex Figure Test and the Meyers and Meyers Recognition Trial in the detection of suspect effort. *The Clinical Neuropsychologist*, 17(3), 426-440. doi: 10.1076/clin.17.3.426.18083
- Lu, P. H., Rogers, S. A., & Boone, K. B. (2007). Use of standard memory tests to detect suspect effort. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment* (pp. 128-151). New York, NY: Guilford Press.
- Mann, V. A., Sasanuma, S., Sakuma, N., & Masaki, S. (1990). Sex differences in cognitive abilities: A cross-cultural perspective. *Neuropsychologia*, 28(10), 1063-1077. doi: 10.1016/0028-3932(90)90141-A
- Mavrakanas, N. A., & Schutz, J. S. (2009). Feigned visual loss misdiagnosed as occult traumatic optic neuropathy: Diagnostic guidelines and medical-legal issues. *Survey of Ophthalmology*, 54(3), 412-416. doi: 10.1016/j.survophthal.2009.02.006

- McGuire, B. E., & Shores, E. A. (2001). Simulated pain on the Symptom Checklist 90-Revised. *Journal of Clinical Psychology, 57*(12), 1589-1596. doi: 10.1002/jclp.1121
- Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial*. Odessa, FL: Psychological Assessment Resources.
- Meyers, J. E., & Volbrecht, M. E. (2003). A validation of multiple malingering detection methods in a large clinical sample. *Archives of Clinical Neuropsychology, 18*, 261-276. doi: 10.1016/S0887-6177(02)00136-1
- Millis, S. R., & Volinsky, C. T. (2001). Assessment of response bias in mild head injury: Beyond malingering tests. *Journal of Clinical and Experimental Neuropsychology, 23*(6), 809-828. doi: 10.1076/jcen.23.6.809.1017
- Miner, I. D., & Feldman, M. D. (1998). Factitious deafblindness: An imperceptible variant of factitious disorder. *General Hospital Psychiatry, 20*, 48-51. doi: 10.1016/S0163-8343(97)00100-X
- 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*(8), 1094-1102. doi: 10.1076/jcen.24.8.1094.8379
- Nitch, S., Boone, K. B., Wen, J., Arnold, G., & Alfano, K. (2006). The utility of the Rey Word Recognition Test in the detection of suspect effort. *The Clinical Neuropsychologist, 20*(4), 873-887. doi: 10.1080/13854040590967603
- Nitch, S. R., & Glassmire, D. M. (2007). Non-forced-choice measures to detect noncredible cognitive performance. In K. B. Boone (Ed.), *Assessment of*

- feigned cognitive impairment* (pp. 78-102). New York, NY: Guilford Press.
- Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical applications* (2nd ed.). Tucson, AZ: Neuropsychology Press.
- Resnick, P. J., & Knoll, J. L., IV. (2008). Malingered psychosis. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 51-68). New York, NY: Guilford Press.
- Resnick, P. J., West, S., & Payne, J. W. (2008). Malingering of posttraumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 109-127). New York, NY: Guilford Press.
- Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique [The psychological examination in cases of traumatic encephalopathy]. *Archives de Psychologie* [Archives of Psychology], 28, 286-340.
- Rey, A. (1964). *L'examen clinique en psychologie* [The clinical examination in psychology]. Paris, France: Presses Universitaires de France.
- Rickards, F. W., & De Vidi, S. (1995). Exaggerated hearing loss in noise-induced hearing loss compensation claims in Victoria. *The Medical Journal of Australia*, 163(7), 360-363.
- Rogers, R. (2008). Researching response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 411-434). New York, NY: Guilford Press.

- Root, J. C., Robbins, R. N., Chang, L., & van Gorp, W. G. (2006). Detection of inadequate effort on the California Verbal Learning Test (2nd ed.): Forced-choice recognition and critical item analysis. *Journal of the International Neuropsychological Society, 12*(5), 688-696. doi: 10.1017/S1355617706060838
- Schmidt, M. (1996). *Rey Auditory and Verbal Learning Test. A handbook*. Los Angeles, CA: Western Psychological Services.
- Schretlen, D. J. & Shapiro, A. M. (2003). A quantitative review of the effects of traumatic brain injury on cognitive functioning. *International Review of Psychiatry, 15*, 341-349. doi: 10.1080/09540260310001606728
- Sherman, D. S., Boone, K. B., Lu, P., & Razani, J. (2002). Re-examination of a Rey Auditory Verbal Learning Test/Rey Complex Figure discriminant function to detect suspect effort. *The Clinical Neuropsychologist, 16*(3), 242-250. doi: 10.1076/clin.16.3.242.13860
- Slick, D. J., Hopp, G., Strauss, E., & Spellacy, F. J. (1996). Victoria Symptom Validity Test: Efficiency for detecting feigned memory impairment and relationship to neuropsychological tests and MMPI-2 validity scales. *Journal of Clinical and Experimental Neuropsychology, 18*, 911-922. doi: 10.1080/01688639608408313
- Slick, D., Hopp, G., Strauss, E., & Thompson, G. B. (1997). *Victoria Symptom Validity Test*. Odessa, FL: Psychological Assessment Resources.
- Slick, D. J., Sherman, E. M., & Iverson, G. L. (1999). Diagnostic criteria for malingered 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
- Sweet, J. J., Condit, D. C., & Nelson, N. W. (2008). Feigned amnesia and memory loss. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 218-236). New York, NY: Guilford Press.
- Sweet, J. J., & Nelson, N. W. (2007). Validity indicators within executive function measures. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment* (pp. 152-177). New York, NY: Guilford Press.
- Tombaugh, T. N. (1996). *Test of Memory Malingering*. Los Angeles, CA: Western Psychological Services.
- Trahan, D. E., & Larrabee, G. J. (1988). *Continuous Visual Memory Test*. Odessa, FL: Psychological Assessment Resources.
- Victor, T. L., & Boone, K. B. (2007). Identification of feigned mental retardation. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment* (pp. 310-345). New York, NY: Guilford Press.
- Victor, T. L., Boone, K. B., Serpa, J. G., Buehler, J., & Ziegler, E. A. (2009). Interpreting the meaning of multiple symptom validity test failure. *The Clinical Neuropsychologist*, 23(2), 297-313. doi: 10.1080/13854040802232682
- Warrington, E. K. (1984). *Recognition Memory Test: Manual*. Berkshire, UK: NFER-Nelson.
- Wechsler, D. (1955). *Wechsler Adult Intelligence Scale manual*. New York, NY: The Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised manual*. New York, NY:

The Psychological Corporation.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale-III manual*. San Antonio, TX:

The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale. Third edition manual*. San Antonio, TX:

The Psychological Corporation.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale-IV manual*. San Antonio, TX:

The Psychological Corporation.

Table 1 Measures of suspect effort

Tests	Cut-off criteria	Study
1. Rey 15-Item Test	Combination score of <20	Boone et al. (2002)
2. Dot Counting Test	E-score of ≥ 17	Boone et al. (2002b)
3. Harbor-UCLA b Test	E-score of ≥ 155	Boone et al. (2002a)
4. Rey Word Recognition Test	Total recognized (without subtracting false positives) for men ≤ 5 , or total recognized (without subtracting false positives) for women ≤ 7 , or Nitch combination equation ≤ 9	Nitch et al. (2006)
5. Digit Span	Age-corrected scaled score ≤ 5 , Reliable Digit Span ≤ 6 , or Average time to repeat 3 digits forward ≥ 3 seconds, or average time to repeat 4 digits forward ≥ 5 seconds	Babikian, Boone, Lu, & Arnold (2006)
6. Rey Auditory Verbal Learning Test	Effort equation score ≤ 12 , or Rey-O/RAVLT discriminant function ≤ -0.40	Boone et al. (2005); Sherman et al. (2002)
7. The Finger Tapping Test	Dominant hand ≤ 35 for males, or ≤ 28 for females	Arnold et al. (2005)
8. Test of Memory Malingering	Trial 2 <45	Tombaugh (1996)
9. Rey-Osterrieth Complex Figure Test	Combination score <47	Lu et al. (2003)
10. Warrington Recognition Memory Test - Words	Total score ≤ 38	Iverson and Franzen (1998)

Table 2 Descriptive statistics and group differences for individual test and formula scores

Test/Formula	Credible (<i>n</i> = 180) <i>M</i> (<i>SD</i>)	Noncredible (<i>n</i> = 101) <i>M</i> (<i>SD</i>)	<i>t</i>	<i>df</i>	<i>p</i>
Age (years)	42.93(13.98)	43.41(11.14)	-0.295	279	0.768
Education (years)	13.38(2.77)	12.20(4.72)	2.654	279	0.008
Male/Female (%)	47.8/52.2	60.4/39.6			
Ethnicity (%)					
Caucasian	47.8	22.8			
African American	12.8	45.5			
Latino/a	22.8	18.8			
Asian	3.9	4.0			
Middle Eastern	3.3	2.0			
Native American	1.1	3.0			
Other	8.3	2.0			
Unknown	0	2.0			
Meyers Formula Partial	12.97(2.21)	9.08(3.22)	11.94	279	.0001
Meyers Formula Whole	-0.06(9.69)	-6.73(11.94)	5.08	279	.0001
ROCFT raw score	31.36 (3.71)	24.34 (8.59)	9.51	279	.0001
Digit Symbol Scale Score	7.60 (2.49)	4.55 (2.23)	10.21	279	.0001
Block Design Scale Score	9.51 (3.02)	6.23 (2.76)	9.00	279	.0001
Finger Tapping Dominant Hand Score	44.24 (10.18)	33.69 (11.84)	7.86	279	.0001

Table 3 Comparison of sensitivity levels and associated cut-scores of the Meyers and Volbrecht (2003) formula and individual subtests at $\pm 90\%$ specificity

Formula/subtest	Cut-score	Sensitivity
MV Formula Partial	≤ 10.3	70.3%
MV Formula Whole	≤ -14.09	30.7%
ROCFT raw score	≤ 25.5	48.5%
Digit Symbol scale score	≤ 4	62.4%
Block Design scale score	≤ 5	47.5%
Finger Tapping score	≤ 27.3	29.7%

MV Formula Partial = Meyers and Volbrecht (2003) formula, partial: (ROCFT raw score x .185) + (Digit Symbol scale score x .491) + (Block Design scale score x .361)

MV Formula Whole = Meyers and Volbrecht (2003) formula, whole: [(ROCFT raw score x .185) + (Digit Symbol scale score x .491) + (Block Design scale score x .361) + 31.34] to calculate an estimated Finger Tapping score. The estimated Finger Tapping score was subtracted from the actual Finger Tapping score; a score less than -10 was indicative of suspect effort

Finger Tapping score = average Finger Tapping score for the dominant hand

Table 4 Group differences for individual test scores between credible men and women

Test/Formula	Men (n = 86) M(SD)	Women (n = 94) M(SD)	t	df	p
ROCFT raw score	31.55 (3.67)	31.20 (3.76)	0.63	178	0.529
Digit Symbol Scale Score	7.50 (2.58)	7.69 (2.41)	-0.51	178	0.608
Block Design Scale Score	10.12 (3.22)	8.95 (2.72)	2.64	178	0.009
Finger Tapping Dominant Hand Score	48.08 (9.41)	40.73 (9.62)	5.17	178	0.000
Meyers Formula Partial	13.17 (2.27)	12.78 (2.15)	1.19	178	0.235
Meyers Formula Whole	3.57 (8.96)	-3.39 (9.16)	5.14	178	0.000

Table 5 Group differences for individual test scores between noncredible men and women

Test/Formula	Men (<i>n</i> = 61) <i>M</i> (<i>SD</i>)	Women (<i>n</i> = 40) <i>M</i> (<i>SD</i>)	<i>t</i>	<i>df</i>	<i>p</i>
ROCFT raw score	23.84 (9.49)	25.10 (7.03)	-0.72	99	0.475
Digit Symbol Scale Score	4.16 (2.00)	5.15 (2.45)	-2.22	99	0.029
Block Design Scale Score	6.26 (2.94)	6.18 (2.51)	0.15	99	0.878
Finger Tapping Dominant Hand Score	35.47 (11.74)	30.97 (11.62)	1.89	99	0.061
Meyers Formula Partial	8.87 (3.46)	9.40 (2.81)	-0.81	99	0.420
Meyers Formula Whole	-4.74 (12.30)	-9.77 (10.82)	2.11	99	0.038