9-2019

The Formulation of a RBANS Effort Supplement

Joshua Seth Goldberg

LOMA LINDA UNIVERSITY
School of Psychology
in conjunction with the
Faculty of Graduate Studies

_____

The Formulation of a RBANS Effort Supplement

by

Joshua Seth Goldberg

_____

A Dissertation submitted in partial satisfaction of
the requirements for the degree
Doctor of Philosophy in Psychology

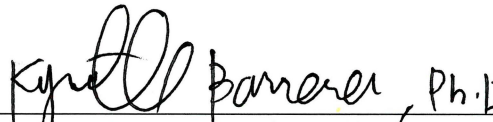
_____

September 2019

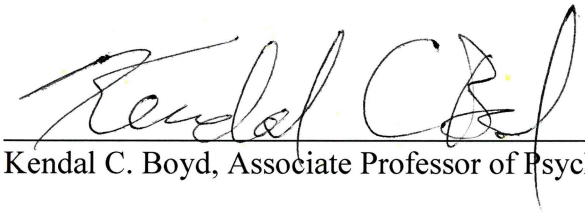
Each person whose signature appears below certifies that this dissertation in his/her opinion is adequate, in scope and quality, as a dissertation for the degree Doctor of Philosophy.

Chairperson

Grace J. Lee, Assistant Professor of Psychology



Kyrstie D. Barrera, Assistant Professor, Physical Medicine and Rehabilitation



Kendal C. Boyd, Associate Professor of Psychology

Travis G. Fogel, Assistant Professor, Physical Medicine and Rehabilitation

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Dr. Grace Lee for continuously providing me with unconditional support to pursue excellence in my graduate studies and beyond. Your mentorship is truly unparalleled and it has been one of my life's great joys to have learned from you and to have worked with you so closely throughout my graduate school career.

I would also like to thank Dr. Travis Fogel for introducing me to the world of clinical neuropsychology. Your enthusiasm and passion for clinical neuropsychology cultivated my interests in cognitive rehabilitation and effort detection. I cannot articulate how thankful I am for your guidance and your willingness to share this project's vision with me.

This project also would not have been possible without Dr. Kyrstle Barrera and Dr. Kenny Boyd. Thank you both for the time and effort you dedicated to this study. I most assuredly could not have completed this project without your advice and direction.

This project also would not have been possible without my family and friends. Your support propelled me along my graduate career. I extend a special thank you to my parents, Maureen and Harry, for their unconditional love, generosity, and encouragement. And of course, I am eternally grateful for my fiancé, Sarah Barney; thank you for always believing in me.

CONTENT

FIGURES

Figure Page

Chapter 4

TABLES

Tables Page

ABBREVIATIONS

MMPI-2	Minnesota Multiphasic Personality Inventory-2
L	Lie
F	Infrequency
Fb	F Back
K	Correction
VRIN	Variable Response Inconsistency
TRIN	True Response Inconsistency
F-K	F Minus K
S	Superlative Self-Presentation
Fp	Infrequency/Psychopathology
SVT	Symptom Validity Test
PVT	Performance Validity Test
TOMM	Test of Memory Malingering
DCT	Dot Counting Test
RDS	Reliable Digit Span
WAIS-IV	Wechsler Adult Intelligence Scale- Fourth Edition
CVLT-II	California Verbal Learning Test- Second Edition
WMS-IV-LM	Wechsler Memory Scales- Fourth Edition- Logical Memory
WAIS-R	Wechsler Adult Intelligence Scale-Revised
RBANS	Repeatable Battery for the Assessment of Neuropsychological Status

MCI	Mild Cognitive Impairment
AD	Alzheimer's Disease
AUC	Area under the Curve
ROC	Receiver Operating Characteristics
EI	Effort Index
ES	Effort Scale
RES	RBANS Effort Supplement
CNO	Clinical Neuropsychology Outpatients
SEG	Suboptimal Effort Group
ADHD	Attention-Deficit/Hyperactivity Disorder
TBI	Traumatic Brain Injury
MND	Major Neurocognitive Disorder
PD	Parkinson's Disease
LD	Learning Disorder
MS	Multiple Sclerosis
KR-20	Kuder-Richardson 20
ANCOVA	Analysis of Covariance

ABSTRACT OF THE DISSERTATION

The Formulation of a RBANS Effort Supplement

by

Joshua Seth Goldberg

Doctor of Philosophy, Graduate Program in Clinical Psychology
Loma Linda University, September 2019
Dr. Grace J. Lee, Chairperson

Assessment of effort detection is an essential component of a neuropsychological evaluation to ensure results of testing are valid indicators of an individual's true level of cognitive functioning. Effort detection in the initial screening process provides neuropsychologists information regarding patients' test engagement prior to administering longer testing batteries. Two effort measures are embedded in the Repeatable Battery for Assessment of Neuropsychological Status (RBANS), a neuropsychological screening assessment, but both have demonstrated elevated false positive rates for classifying individuals with memory impairment as those putting forth poor effort. These embedded measures rely on cut-off scores on digit span and memory subtests. In contrast, this RBANS Effort Supplement (RES) utilizes several forced-choice subtests, reflective of current research emphasizing the importance of multiple methods of effort detection; subtests in this measure included list learning forced-choice, figure copy forced-choice, picture naming forced-choice, a coding task, and a story recognition component utilized for face validity of memory assessment. Fifty-nine participants were recruited from an outpatient neuropsychology facility in conjunction with 14 poor effort simulators; each participant was administered the RBANS, the RES, and the Dot Counting Test (DCT). Results supported the RES'

reliability at the individual decision-making level. Validity analyses demonstrated that the RES exhibited strong convergent validity with established effort detection measures and that individuals putting for poor effort scored significantly lower on the RES than individuals who put forth adequate effort, as delineated by the established DCT cutoff score of 17. In summary, the RES was shown to be a valid indicator of effort detection. Clinical implications of the RES include reduction of time and costs involved in neuropsychological assessment.

CHAPTER ONE

INTRODUCTION

Cognitive assessment within the realm of a neuropsychological framework is a useful tool in diagnosing prominent neurological disorders. Often, patients are referred for neuropsychological assessment from a neurologist or primary care physician as changes within cognitive functioning become more apparent to either the patient and/or their surrounding community. As part of the assessment, patients are asked to put forth their best effort throughout the administration of cognitive testing so that valid data may be compiled that is an accurate representation of their cognitive functioning. Occasionally, patients can consciously or unconsciously fail to provide adequate effort resulting in invalid testing data.

Effort's Relevance to Neuropsychology

There are three prominent psychological occurrences that may explain the manifestation of suboptimal effort in neuropsychological testing. The unconscious failure to provide adequate effort as a reflection of an unidentified need or conflict is labeled as a somatoform disorder. The conscious need for a patient to assume a sick role is defined as a factitious disorder. Finally, malingering is typically defined as intentionally poor effort in order to maximize an external incentive (Larrabee, 2007). Malingering is more typically suspected within the medical-legal context, when there is a significant discrepancy between the individual's claimed symptomatology and objective findings, the presence of antisocial personality disorder, and an individual's lack of overall cooperation in neuropsychological testing. Within the clinical setting, researchers suggest

utilizing alternative phrasing rather than malingering, as the rationale for improper effort during testing may not be definitively identifiable. Thus, researchers within the neuropsychological field suggest using phrasing such as the mobilization of effort and test investment when referring to possible cases of malingering (Carone, Iverson, & Bush, 2010). Regardless, the predominant focus of this study was effort detection within neuropsychological testing.

Glenn Larrabee (2007) explains that suboptimal effort is not necessarily uncommon in neuropsychological settings. It is estimated that cases of poor effort occur in 29% of personal injury cases, 30% of disability cases, 19% of criminal cases, 38.5% of personal injury cases, and 8% of general medical cases involving symptom exaggeration. Thus, suboptimal effort occurs at relatively high rates in typical neuropsychological settings and as such, there is a necessity for valid measures of poor effort to distinguish between individuals who have genuine impairments and those whose symptoms may be attributed to other factors. Neuropsychologists agree that effort measurement is an integral part of both forensic and clinical settings (Martin, Schroeder, & Odland, 2015) and it is estimated that approximately 79 percent of neuropsychologists utilize effort measures in forensic type assessments (Slick, Tan, Strauss, & Hultsch, 2004). Determination of suboptimal effort within clinical neuropsychology differs somewhat across settings; however, a large consensus of neuropsychologists agrees that more confidence in definitively diagnosing poor effort occurs through multiple effort measures with little methodological overlap to limit redundancy (Larrabee, 2008; Mittenberg, Patton, Canyock & Condit, 2002).

The assessment of suboptimal effort can be achieved through several different modalities. Effort may be assessed through self-report measures, most prominently through notable personality inventories such as the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). Personality assessments often utilize subscales that identify when subjects are exaggerating psychological symptomatology, the most common of which are somatic subscales (Heilbronner, Sweet, Morgan, Larrabee, Millis, & Participants, 2009), or reporting symptomatology that is rare even among those with confirmed psychiatric illnesses. Specifically, the MMPI-2 has developed particular scales for individuals attempting to fake good behavior (L scale), faking psychological impairment (F and Fb scales), answering defensively (K scale), answering questions inconsistently across similar questions (VRIN), answering all questions indiscriminately as true or false (TRIN), not answering questions honestly (F-K), attempting to present as excessively good (S), and overreporting of psychopathological symptoms (Fp). The validity scales for the MMPI-2 and other self-report measures that simultaneously measure feigned symptomatology are considered symptom validity tests (SVT's), whereas assessment of effort typically resembles tests of cognitive performance, known as performance validity tests (PVT's; Larrabee, 2012). A primary focus of this study was to formulate an efficacious and time-efficient PVT to be utilized within the initial neuropsychological screening process.

Despite the abundance of effort detection methods currently available to neuropsychologists, there is a lack of consensus regarding when to use specific PVT's. Often, neuropsychologists utilize clinical judgment when incorporating a PVT's into their

testing batteries (Bigler, 2012). However, the use of standalone PVT's in forensic evaluations is strongly encouraged where there is a high risk of invalid responding. Standalone measures are often lengthier than comparative PVT's, but such time is considered medically necessary given the risk of suboptimal effort in these clinical contexts (Heilbronner et al, 2009).

Symptom Validity testing is also encouraged in cases where an individual presents with subjective cognitive abilities usually associated with mood concerns. Disorder-specific inventories with incorporated validity scales are recommended for targeted analysis of an individual's mood concerns and their association with their current cognitive symptoms. General personality inventories with validity scales are also encouraged when time is available to more fully grasp an individual's response bias tendencies (Heilbronner et al, 2009).

Common Modalities of Effort Detection

The measurement of effort in neuropsychology occurs in many different modalities and formats. Primarily, effort is analyzed through either standalone measures or embedded measures. Standalone measures, such as the Test of Memory Malingering (TOMM; Tombaugh, 1997) and the Dot Counting Test (DCT; Boone, Lu, & Herzberg, 2002) are tests specifically designed to measure effort that can be utilized independently without incorporating information present in any other neuropsychological test within the test battery. Standalone measures may be designed using an encoding/memory recognition format (TOMM) or a visual perceptual format (DCT).

An embedded memory measure is an analysis of effort utilizing data collected within an existing neuropsychological test that may be originally designed to assess a different aspect of cognitive functioning (Institute of Medicine, 2015). Examples of embedded measures include Reliable Digit Span from the Wechsler Adult Intelligence Scale- Fourth edition (RDS; WAIS-IV; Wechsler, 2008) and the California Verbal Learning Test-Second Edition forced-choice condition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000). Effort measures of this nature typically utilize memory recognition (CVLT-II) or attention (RDS) to assess an individual's concerted effort.

Effort measures may also utilize either a forced-choice or non-forced-choice paradigm. Forced choice measures appear to be difficult but are in fact, easy tasks. Typically, in forced-choice, an individual is asked to encode a series of pictures or words and then later asked to select each of the target pictures or words from two choices. Participants who perform below chance levels (i.e., below 50%) are identified as individuals who may been putting forth suboptimal effort. However, commercially available neuropsychological effort measures typically do not rely on comparing the total correct responses to the number expected by guess. Rather prominent neuropsychological effort measures typically examine poor effort as falling in a range of scores that would be expected by guessing throughout the measure (Frederick & Speed, 2007).

Non-forced-choice paradigms utilize a variety of methods including memory recognition, such as Logical Memory II Recognition from the Wechsler Memory Scale-Fourth Edition, motor skills (The B Test; Boone, Lu, & Herzberg, 2002), and perceptual skills (DCT). Neuropsychological batteries often include several methods of effort detection to create a more comprehensive approach. The purpose of this study was to

incorporate various methods of effort methodology into one brief but comprehensive supplement to aid in effort analysis within the initial neuropsychological screening process. To fully examine the framework of the measure, it is necessary to discuss the prominent methodologies and existing assessments of effort currently being utilized within the field of neuropsychology.

Forced-Choice Recognition

Forced choice recognition is becoming an increasingly popular method of analyzing effort. As explained previously, in forced-choice measures, the target stimuli are presented, after which the original targets are presented together with a foil and the subject is asked to choose which of the two items was presented previously. Forced choice recognition of memory malingering typically assesses how the examinee performs according to chance level (Grote & Hook, 2007). If an examinee performs below chance levels (i.e., less than 50% accuracy), it is thought that the examinee must knowingly be choosing the wrong answer, as an individual with no previous exposure to the original stimuli would still be expected to perform at chance levels. Research has indicated there is no significant correlation between memory capacity and forced-choice performance (Root, Robbins, Chang, & Van Gorp, 2006). Clinically referred patients in neuropsychology clinics routinely performed at near perfect levels within the forced-choice paradigm. Thus, performance at below levels of chance is characterized as suboptimal effort.

One forced-choice recognition test that is commonly utilized within the clinical neuropsychological field is the Test of Memory Malingering (TOMM). When given the

TOMM, subjects are presented a series of 50 pictures of objects. Immediately following the initial presentation, the examinee is presented with each object along with a foil and asked to pick the picture they saw previously. Subjects are corrected on incorrect responses. Following the first trial, the subject is once again presented with the same pictures but in a different order. Immediately following the second presentation of pictures, the subject is once again asked to pick the correct pictures from foils. After a fifteen-minute delay, an optional retention trial can be administered where the subject is once again administered the forced-choice paradigm between original images and foils, but without being presented with the original stimuli. Results from TOMM research studies have found that the test is considered relatively easy for individuals with depression, chronic pain, and dementia. The TOMM is considered a good screener for overall effort but is often criticized for being too easy and too long (Strauss, Sherman, & Spreen, 2006). Like the majority of forced-choice measures, it is recommended that the TOMM be utilized in conjunction with other measures of effort**.** In a cognitively impaired setting, the TOMM achieved high sensitivity of 90% when diagnosis of dementia was ruled out. However, when accounting for dementia diagnoses, the TOMM misclassified patients with dementia as putting forth suboptimal effort by over 70%, suggesting that the measure may be overly sensitive for individuals with dementia (Teicher & Wagner, 2004). These contrasts findings suggesting that the TOMM, along with the CVLT-II forced-choice, is reliably sensitive to suboptimal effort in cases of feigned traumatic brain injury (Moore & Donders, 2004). Thus, this current study will aim to provide a globally valid measure of effort within a neuropsychological setting.

As alluded to previously, another forced-choice measure is an embedded measure in the second edition of the California Verbal Learning Test. The CVLT-II is a verbal memory test where subjects are presented with a list of 16 words and asked to recall them over several trials at several different time points, both spontaneously and with category cues (immediate free recall, immediate cued recall, short delay free recall, long delay free recall, long delay cued recall, and yes/no recognition). Following the yes/no recognition portion of the CVLT-II and a ten-minute delay thereafter, subjects can be given a forced-choice recognition trial.

In a medicolegal setting, the CVLT-II forced-choice paradigm performed similarly to the TOMM, in that it was suggested to be very sensitive (ranging from 81-93%), and only moderately specific (32-60%). Furthermore, it was recommended that the forced-choice component of the CVLT-II not be used for individuals suffering from frank dementia (Root et al., 2006). Thus, caution should be taken when definitively diagnosing poor effort as reflected by the CVLT-II forced-choice, especially in settings assessing for cognitive impairment. This is inherently problematic considering that dementia cases are extremely common within neuropsychological practices. Additionally, it is also worth mentioning that forced-choice measures should not be used in isolation to identify faulty effort. Although researchers have stated that forced-choice measures are the most effective modality of assessing suboptimal effort, it is recommended that forced-choice measures be utilized in combination with other effort measures to provide a more comprehensive overview of an individual's effort during testing (Strauss, Sherman, & Spreen, 2006). A forced-choice measure alone would not adequately define an individual's effort as suboptimal. This study will attempt to create a globally specific and

sensitive effort measure utilizing both forced-choice and non-forced-choice paradigms to optimize the detection of suboptimal effort.

Non-Forced-Choice Measures

Despite the effectiveness of forced-choice measures in assessing for poor effort, there are some limitations that warrant utilization of additional measures. As discussed previously, researchers have highlighted the importance of examining multiple non-interrelated measures of effort in order to validly examine definitive inadequate effort (Boone & Lu, 2007). Additionally, some standalone forced-choice measures often require lengthy durations to properly administer, and many can also be overly sensitive to legitimate memory impairment. Neuropsychologists often bolster stand-alone forced-choice malingering measures with non-forced-choice embedded effort measures.

Individuals putting forth suboptimal effort on embedded measures of effort, specifically those involving recognition, tend to exaggerate poor performance on memory tasks after delayed recall (Axelrod, Fichtenberg, Millis, & Wertheimer, 2006). On the Wechsler Memory Scales, Fourth Edition Logical Memory (WMS-IV-LM; Wechsler, 2009) patients are presented with two stories that they are asked to recall immediately and after a 20-30-minute delay. After the delay, patients are asked to recall the story from memory and are then administered a series of yes/no questions designed to see if they can recognize story details in this format. Literature has indicated that yes/no questions are different from forced-choice recognition in that they present targets and foils one after another, as opposed to forced-choice recognition measures that present targets and foils simultaneously (Bayley, Wixted, Hopkins, & Squire, 2008).

Recognition is typically easier than spontaneous recall **(**McDougall, 1904; Postman, 1963), as researchers have identified that recalling an item from memory requires more memory storage than simply recognizing the item via prompt. Thus, many patients (apart from those with severe dementia) who have difficulty recalling story details during the delayed recall trials tend to perform better on the recognition trial, when questions are posed in a yes/no format. In a study examining simulators acting as malingerers in comparison to individuals of mixed etiology and healthy controls, simulated malingerers performed significantly worse on the WMS-IV-LM Recognition test than both patients with mixed etiology and the healthy controls (Bouman, Hendriks, Schmand, Kessels, & Aldenkamp, 2016). Such findings indicated that individuals who feign impairments commonly overestimate the extent of cognitive deficiencies of patients who have true disorders.

Another prominent embedded measure is the Reliable Digit Span (RDS). RDS was originally derived from the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) Digit Span subtest, a measure of attention and working memory (Greiffenstein, Baker, & Gola, 1994). Within this subtest, examinees are asked to repeat a series of digits, initially forwards and then in backwards order until they provide incorrect responses on both trials of any given length of digit sequence. RDS is calculated by taking the sum of the length of the longest consecutive strings successfully repeated forward and backward. RDS is utilized within neuropsychological effort testing because it is based on the assumption that digit span appears to be a test on which brain-injured patients may exhibit difficulty but in reality, it is relatively preserved among patients with brain dysfunction including amnesia (Etherton, Bianchini, Greve, & Heinly, 2005).

Research has demonstrated RDS is moderately sensitive and specific to poor effort in a forensic setting (Sensitivity = 63%, Specificity = 86%) and can distinguish individuals who provide suboptimal effort from individuals with appropriate effort by more than one pooled standard variation (Jasinski, Berry, Shandera, & Clark, 2011: Larrabee & Berry, 2007). Thus, RDS seems to be an adequate measure of detecting poor effort in conjunction with additional embedded recognition tasks such as WMS-IV-LM Recognition and forced-choice measures. Despite the documented utility of embedded non-forced-choice measures of effort such as those included in the Wechsler scales, a standalone non-forced-choice effort measure was an ideal choice for optimizing effort detection in the current study.

Another method of analyzing effort is through the usage of standalone non-forced-choice measures. A commonly used effort measure that is neither embedded nor of the forced-choice variety is the Dot Counting Test (DCT; Boone, Lu, & Herzberg, 2002). On the DCT, patients are presented a series of cards with dots and are asked to count the dots as quickly as possible without committing any errors. Cards one through six contain dots disseminated randomly across the page, whereas cards 7-12 contain dots that are organized in clusters. A composite score (E-score) is computed based on the patient's average time to complete cards 1-6, summed with the patient's average time on cards 7-12 and total number of errors. Patients who may attempt to feign impairments often overestimate the difficulty of the DCT, and consequently take an inordinate amount of time to complete each item and/or commit numerous counting errors (Strauss, Sherman, & Spreen, 2006). The DCT has been shown to have moderate sensitivity (70%) and high specificity (90%) within clinical settings and is highly correlated with simple

digit span (56% shared variance). However, like other effort measures, it is recommended that this assessment be used in conjunction with other effort measures.

Neuropsychological research has identified improved accuracy in malingering detection when multiple measures of heterogeneous methodology are utilized together in order to substantiate effort claims. Current recommendations for neuropsychological practice suggest utilizing several effort indicators throughout a testing battery to definitively confirm suspect effort (Heilbronner et al., 2009). Specifically, research has indicated that failure on two effort measures likely suggests the presence of feigned impairment (Larrabee, 2003). Chaining measures of independent methodology increases the likelihood of correctly identifying suspect effort, whereas chaining effort measures with methodological overlap may inflate such probability (Grimes & Schulz, 2005). This study similarly aimed to create an effort measure utilizing multiple methods of analyzing effort within the neuropsychological screening process.

The RBANS

Effort measures can often be integrated into initial consultations along with neuropsychological screening measures to help identify the cognitive capacities of new patients as well as determine whether interpretations and future testing may be needed after the initial consult. A commonly administered screening measure is the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998). The RBANS, originally developed to detect dementia, consists of five domains: immediate memory, visuospatial/constructional skills, language, attention, and delayed memory. One of the key utilities of the RBANS is that it is highly correlated with longer

neuropsychological assessments, such as the Wechsler Adult Intelligence Scale IV Full Scale Intelligence Quotient (r = .75), but it only requires thirty minutes to complete (Hartman, 2009). Research within the last decade has also revealed that among commonly used dementia screening measures, total RBANS performance is one of the better measures in predicting total brain volume (Paul et al., 2011). Despite the RBANS lack of sensitivity (Total Scale Sensitivity = 0.55) towards classification of individuals with and without Mild Cognitive Impairment (MCI; Duff, Hobson, Beglinger, & O'Bryant, 2010), the RBANS does seem to be a valid diagnostic indicator of more pronounced neurologic disease. In terms of its diagnostic accuracy for Alzheimer's disease (AD), the RBANS demonstrated high probability of correctly classifying individuals with and without AD across all index scores (Duff, Clark, O'Bryant, Mold, Schiffer, & Sutker, 2008). Specifically, Duff et al. analyzed areas under the curve (AUC) of Receiver Operating Characteristics (ROC) to examine the diagnostic utility of the RBANS to correctly classify individuals with and without AD according to their performance on all RBANS indices within a 0 to 1 scale. High diagnostic accuracy was reflected on all RBANS indices, including visuospatial/constructional (AUC = 0.74), language (AUC = 0.83), and attention (AUC = 0.81), and particularly high accuracy on immediate memory (AUC = 0.96), delayed memory (AUC = 0.98), and total index score (0.98).

The RBANS' relevance to dementia screening in neuropsychology has warranted the development of accompanying effort measures to detect feigned impairment. However, the effort measures associated with the RBANS currently do not completely capture feigned impairment when it occurs in testing. As the RBANS' utility as a

cognitive screener has become more established, neuropsychological researchers have attempted to develop embedded malingering measures within its framework. Two such measures include the RBANS Effort Index (EI; Silverberg, Wertheimer, Fichtenberg, 2007) and the RBANS Effort Scale (ES; Novitski, Steele, Karantzoulis, & Randolph, 2012). The EI is calculated by combining the digit span subtest and list recognition scores into weighted scores, based on the utility of digit span and recognition formats that have been previously validated for symptom validity measurement. The ES utilizes the same subtests as the EI but includes an additional adjustment based on free recall scores (ES = List Recognition – (List Recall + Story Recall + Figure Recall) + Digit Span). Novitski et al. (2012) formulated the ES in this manner in order to discriminate between memory impairment and feigned impairment, as patients with true memory impairment are likely to have extremely low free recall scores (close to zero) by the time recognition scores begin to drop.

Despite the empirically validated research from which these measures were constructed, research has demonstrated that their validity may be somewhat limited. Research has illustrated that although the EI exhibits good specificity for simulated malingerers with a false-positive rate of 19% or less at selected cutoffs, it has only moderate sensitivity (66%), which risks the possibility of misdiagnosing malingerers with memory-related conditions (Crighton, Wygant, Holt, & Granacher, 2015). Additionally, the EI has been shown to have an elevated false-positive rate within populations of individuals suffering from dementias (Novitski et al., 2012; Duff et al, 2011). Concurrently, the ES has misclassified participants as malingerers due to its heavy emphasis on subtracting free recall scores as an overall reflection of its focus on patients

with amnesia and has also reflected high false-positive rates as well (Crighton et al., 2015). Thus, despite the presence of current embedded effort measures within the RBANS, such measures have exhibited limitations in correctly categorizing good effort and poor effort in dementia populations. There would appear to be a need for a more valid measure of effort detection utilizing the RBANS.

In this study, a new measure, the RBANS Effort Supplement, was formulated and assessed for reliability and validity to detect suboptimal effort through the sole usage of the RBANS assessment. The formulation of the RES had several particular advantages. It was designed to be a quick measure to administer, with the opportunity for cost-efficiency in that a subsequent longer evaluation would not be needed if effort were found to be a significant issue. It also included different methods/formats of malingering detection: forced-choice and memory recognition, reflective of Glenn Larrabee's research concerning how the aggregation of varying measures of effort provide a more definitive finding of suboptimal effort (Larrabee, 2008). Thus, the primary aim of this study was to establish the reliability and validity of the RBANS Effort Supplement (RES). It was hypothesized that the RES would be specific and sensitive towards detecting suboptimal effort in a simulator group compared to a generalized clinical neuropsychology population.

CHAPTER TWO

AIMS AND HYPOTHESES

The primary aim of this study was to determine if the RBANS Effort Supplement (RES) was a reliable and valid measure of effort. To measure the RES' reliability, the RES was assessed for internal consistency utilizing the Kuder-Richardson 20 method (Kuder & Richardson, 1937). It was hypothesized that the RES would be internally consistent. Following reliability analysis, the construct validity of the RES was examined. Specifically, The RES was assessed for convergent validity utilizing partial correlations controlling for age and years of education. It was hypothesized that the RES would exhibit convergent validity with the RBANS Effort Index, RBANS Effort Scale, and the Dot Counting Test. Further, we hypothesized that participants within the experimental malingering sample would score significantly lower on the RES in comparison to clinical groups.

An exploratory aim of this study was to examine the specificity and sensitivity of the RES. Such analyses were conducted utilizing ROC curve analyses according to a RES cut-off score to be determined according to frequency characteristics of the RES itself. It was hypothesized that the RES will be specific and sensitive in correctly classifying individuals engaging in suboptimal performance clinical groups.

CHAPTER THREE

METHODS

Participants and Procedures

Our study included two independent samples, a clinical neuropsychology outpatient (CNO) and a comparative suboptimal effort group. The CNO group was comprised of 59 outpatients, who were referred for neuropsychological testing at the Loma Linda University Medical Center East Campus neuropsychology service. Our suboptimal effort group was recruited from Loma Linda University and included 15 students from the graduate student population. All subjects fell within the age range of 20-89 and all spoke English fluently. One participant was excluded utilizing the outlier labelling rule on the RES total score to help correct for the skewness of the data. As such our analyses included 59 individuals included in the clinical outpatient group in comparison to 14 individuals included in the suboptimal effort group.

Participants involved in the CNO group were individuals who had been referred for clinical neuropsychological services for various reasons, including mild cognitive impairment, traumatic brain injury, stroke, epilepsy, ADHD, and varying mood disorders. After participants had completed a structured clinical interview as part of their neuropsychological referral, they were asked to participate in the current study. Agreeing participants completed the informed consent process and gave permission to use the results of their clinical testing (i.e. RBANS, RES, Dot Counting Test) for the current study. Participants then completed a brief additional structured interview asking for basic demographic information (i.e. age, ethnicity, education, referral complaint, handedness, engagement in previous neuropsychological testing, and current legal involvement).

Participants were administered the RBANS as part of their routine neuropsychological assessment and were additionally administered the RES immediately thereafter.

Participants enrolled in the suboptimal effort group (SEG) were recruited from Loma Linda University's graduate population. Subjects were recruited from various departments in the university through department-wide email notifications and campus-wide postings. Participants completed the informed consent process and a brief structured interview of demographic information. Participants were given the following script (DenBoer & Hall, 2007), prompting them to approach the neuropsychological tests as if they were trying to appear brain damaged in order to receive financial compensation in an ongoing lawsuit:

> You are about to take some cognitive tests that examine mental abilities such as attention, memory, thinking and reasoning skills, and your ability to think quickly. While responding to the tests, please pretend that you have experienced brain damage from a car accident involving a head-on collision. You hit your head against the windshield and were knocked out for 15 minutes. Afterwards, you felt ''dazed'' so you were hospitalized overnight for observation. Because the driver of the other car is at fault, you have decided to go to court to get money from the person responsible. During the next few months following the accident, the negative effects from your head injury disappear. Your lawsuit has not been settled yet, and your lawyer has told you that you may get more money if you look like you are still suffering from brain damage. As you pretend to be this car accident victim, try to respond to each test as a patient who is trying to appear brain damaged in order to get money from the lawsuit. Thus, your performance on the tests should convince the examiner as well as the people involved in deciding the outcome of your lawsuit that you are still suffering from brain damage.

Approval for the study was obtained from the Loma Linda University Human Subjects Committee Institutional Review Board, and written informed consent was acquired from all participants upon enrollment. It should be noted that Loma Linda University associated legal counsel stated that the RBANS Effort Supplement was considered legally permissible as long as primary investigators did not attempt to earn a profit from the measure itself. The RES was only utilized for the purposes of this study.

Instruments

Prior to the examination of participants, examiners interviewed participants using a standardized questionnaire (See Appendix B) in order to gather relevant demographic information, including age, years of education, gender, ethnicity, referral, handedness, prior neuropsychological testing, and engagement in ongoing litigation. Of note, none of the participants were involved in previous litigation and three participants had engaged in previous neuropsychological testing (one participant in 2016, another in 1985, and the third at an unknown time) .

The Dot *Counting* Test

Boone, Lu, and Herzberg's Dot Counting Test (2002) is a measure of symptom validity and malingering. Participants are presented with a series of twelve dotted cards and are asked to count the number dots as quickly as possible and relay to the examiner the number of dots that they counted. On cards one through six, the dots on the cards are disseminated in no organizational fashion. In cards seven through twelve, the dots on the cards are grouped in such a way that it is easier to count the number of dots quickly. An E-score is tabulated according to the participant's response times and number of errors on the test itself (lower E-scores reflect fewer errors and faster response times). Research has identified that the DCT is an adequate measure of suspect effort, with moderate sensitivity and high specificity of identifying possible malingerers. It has been encouraged that the DCT be utilized in conjunction with other measures when assessing for symptom validity (Strauss, Sherman, & Spreen, 2008). Previous research has

suggested a general cut-off score of ≥17 for classification of suboptimal effort (Boone et al., 2002).

The *RBANS*

Randolph's RBANS (1998) is a neuropsychological assessment used to test the cognitive status of individuals suffering from neurological diseases or head trauma. One of the core advantages to using the RBANS is its brevity. The RBANS takes approximately 30 minutes to administer, as opposed to other cognitive assessments that require a much longer duration to fully administer.

The RBANS is comprised of five indices (immediate memory, delayed memory, visuospatial ability, language, and attention) and twelve subtests (list learning, story memory, figure copy, line orientation, digit span, symbol digit coding, picture naming, semantic fluency, list recall, list recognition, story recall, and figure recall). All index scores are comprised of two subtests except for the delayed memory domain, which consists of four subtests. The RBANS total score provides an overall outcome statistic for an individual's overall neuropsychological functioning. In addition to the total score, individual subscale scores for immediate memory, visuospatial ability, language, attention, and delayed memory can be calculated. All subtests are given a subtest raw score. Raw scores of subtests within each domain are added and converted to an age-corrected index score. Index scores can also be converted to percentile scores, according to the age-based normative conversions from the RBANS manual.

Immediate *Memory*

The Immediate Memory domain assesses an individual's ability to remember and recall a small amount of information directly after it has been presented. The immediate memory domain is assessed using two subtests:

List Learning

List Learning consists of a list of 10 unrelated words, read for immediate recall over four trials, with a maximum score of 40. Words used in the List Learning task are considered moderate-high imagery words with relatively low age of acquisition. The high imagery levels and low age of acquisition of these words is considered helpful in reducing education effects on neuropsychological performance and allows for easing language translation difficulties.

Story Memory

This subtest is comprised of a story with 12 itemized details; the story is read for immediate recall over two trials, for a total maximum score of 24.

Visuospatial Ability

The Visuospatial domain prompts participants to examine, comprehend, and recreate spatial relations. Notably, this domain assesses participants' ability to estimate distance and depth and navigate the surrounding environment. The subtests used to analyze visuospatial/constructional ability are as follows:

Figure Copy.

The Figure Copy subtest prompts participants to draw an exact copy of a complex figure comprised of geometric shapes. The subtest itself is considered very similar yet less complex to the Rey-Osterrieth Complex Figure Test (Meyers & Meyers, 1995). The RBANS figure is comprised of 10 components, and a structured simplified scoring guide, which provides for a maximum score of 20.

Line Orientation

On this subtest, participants are presented with an arrangement of 13 lines, beginning at a common point of origin and fanning out across 180 degrees, which serves as the reference figure. Each item consists of two target lines that are shown beneath the reference figure. Subjects must correctly identify which two lines in the reference match the two target lines. Line orientation consists of 10 items, each comprised of two target lines, for a total maximum score of 20.

Delayed Memory

The Delayed Memory domain of the RBANS requires participants to recall information for an extended length of time. These subtests are presented to the participants approximately 20 minutes after initial presentation.

List Learning Free Recall

Free recall of the words from the List Learning subtest (max = 10).

L**ist *Learning Recognition***

Yes/No recognition of the words from the List Learning subtest, with 10 foils (max = 20).

Story Memory Free Recall

Free recall of the story from the Story Memory subtest (max=12).

Figure Free Recall

Free recall of the figure from the Figure Copy subtest (max = 20).

Language

The language domain prompts participants to execute communication skills to verbally name and retrieve previously learned semantic information. Two subtests are included in this domain:

Picture Naming

Picture Naming is considered a confrontation-naming task, with 10-line drawings of objects that the participant must name.

Semantic Fluency

Participants are allotted one minute to provide as many examples from a semantic category as possible (i.e., animals, fruits).

Attention

The RBANS attention domain assesses an individual's ability to select a component of information to focus on in subsequent processing and integration tasks. The attention domain prompts the participant to manipulate previously presented material (visual and oral) that has been stored within the individual's short-term memory. This domain includes the following subtests:

Digit Span

Subjects are asked to repeat a series of numbers, with stimulus items increasing in length from 2 digits to 9 digits. The items are presented in order of length (shortest to longest), and the test itself is discontinued when the participant fails all trials within a given string length. It should be noted that there is no digit span backwards on the RBANS.

Coding

Coding is an assessment of an examinee's processing speed that is very similar to the Coding subtest of the Wechsler Adult Intelligence Scale. Subjects are asked to fill in digits matching with corresponding shapes on a coding key as fast as they can. After practice items are completed, participants have 90 seconds to complete as many items as possible.

Total Scale

The Total Scale is the overall outcome statistic for an individual's overall neuropsychological functioning as comprised by the sum of all the index scores of the RBANS (Attention, Immediate Memory, Delayed Memory, Visuospatial/Constructional, and Language).

RBANS Effort Supplement

The RES is comprised of one Yes-No Recognition component (Story Memory) and four components in Forced-Choice Recognition format: List Learning, Picture Naming, Figure Copy, and Coding. It should be noted that the RES has never been utilized in previous research. The RES was constructed utilizing the stimuli in RBANS form A, with all non-target stimuli for verbal and nonverbal information derived from alternative forms of the RBANS.

Story Memory Recognition

Participants were administered 12 questions in a yes/no format regarding details from the story that was read to them twice previously in the RBANS Story Memory subtest (max = 12). This subtest was not included in the final RES Total score and was meant to serve as face valid indicator of memory performance.

List Learning Forced Choice

Participants were administered a forced-choice task involving the 10 words from the List Learning subtest. For each item, participants were prompted with two words, one

word from the original list and one novel word, and subsequently asked to select the word that appeared on the original list (max =10).

Picture Naming Forced Choice

Participants were administered a forced-choice task involving the 10 objects from the Picture Naming subtest. For each item, participants were prompted with two pictures, one that was presented during the Picture Naming task and one that was not and asked to select the picture they had seen previously. It should be noted that the non-target pictures were pictures from alternate forms of the RBANS. (max =10).

Figure Copy Forced Choice

Participants were administered a forced-choice task involving the Figure Copy subtest. On each item, participants were prompted with two figures, one that was a component of the original figure presented during the Figure Copy task and one that was not and asked to select the component they had seen previously. It should be noted that figures that were presented that were not components of the original complete figure were figure components from alternate forms of the RBANS (max = 12).

Coding Task

Participants were administered a task involving the 9 symbols from the Coding subtest. Participants were asked to select 9 coding symbols from a larger set, which they thought matched those they had seen during the previous administration of the RBANS Coding subtest. Participants were also asked to recall where each symbol was located in

the original key; this location task was not included in the final RES Total Score and was meant to serve as a ruse that the measure appeared to be more difficult than it actually was. It should be noted that symbols that were presented that were not components of the original complete figure were symbols used in alternate forms of the RBANS (max = 9).

RES Total Score

The Total RES score was computed by adding all total scores except for RES Recognition (Max = 41).

The RBANS Effort Scale

The RBANS Effort Scale (Novitski et al., 2012) is an existing embedded measure in the RBANS, which is calculated by subtracting delayed free recall scores from recognition and then adding the score from the RBANS digit span subtest. The measure was validated on a population of individuals with amnestic disorders and compared against a mild traumatic brain injury group who had failed a second measure of effort. ES scores less than 12 are considered suspicious for poor effort. However, a limitation of the ES is that it yields significantly negative scores when individuals perform at a high level on measures of delayed free recall and has been cautioned to only be utilized in circumstances where effort during testing is in question.

The RBANS Effort Index

The RBANS Effort Index (Silverberg, Wertheimer, & Fichtenberg, 2007) is another embedded effort measure in the RBANS. Primary investigators for the EI

converted raw scores into a common metric based on their relative infrequency in a derivation sample with true cognitive impairment and then summed these weighted scores to arrive at an index score. More infrequent scores on digit span and list recognition were assigned higher weighted values. The EI is then calculated by using weighted scores on RBANS raw scores of digit span and list recognition and computed by adding the sum of these weighted scores. Thus, a higher EI score indicates worse effort. The measure was validated on a clinical neurological disorders population and compared against a mild traumatic brain injury group in conjunction with three "suboptimal" groups. EI scores greater than 3 are considered suspicious for suboptimal effort.

CHAPTER FOUR

RESULTS

Participant Demographic Information

The demographic characteristics of participants in the CNO and SEG are shown in Table 1. In sum, 73 participants were included in analyses for this study. The CNO was comprised of 59 participants (50.9% male) with an average age of approximately 54 years (M = 53.54, SD = 20.23). The majority of participants were Caucasian (66.1%) with an average of approximately 15 years of education (M = 14.89 years, SD = 2.49). In contrast, the SEG included 14 participants (36.7% male) with an average age of approximately 30 years (M = 30.29, SD = 12.02). The majority of participants were Caucasian (28.6%) with an average of approximately 16 years of education (M = 16.42, SD = 1.16)). Of note, the SEG was significantly younger and had more years of education than the CNO group, $p < .05$.

The distributions of outcome measures (e.g. RES, Dot Counting Test, RBANS Effort Scale and RBANS Effort Index) were examined. The RES was found to be negatively skewed. To correct for skewness, logarithmic transformations of RES were used; the RES was then normally distributed. We found that the Dot Counting Test (DCT) and RBANS Effort Index (EI) were positively skewed. We then performed logarithmic transformations of these outcome measures as well, resulting in normal distributions for both outcome measures. The RBANS Effort Scale (ES) was normally distributed and did not require transformations.

Table 1. Demographic Statistics for Experimental Groups

Total N = 73	Clinical Group N = 59	Actor Group N = 14	Statistic	*p Value*
Gender (%)				
Male	30 (50.9)	5 (36.7)	$\chi^2 = 1.04$	.31
Female	29 (49.1)	9 (63.3)		
Age (SD)	53.54 (20.23)	30.29 (12.02)	$t = -4.11$	.00**
Ethnicity (%)			$\chi^2 = 18.39$	.00**
Caucasian	39 (66.1)	4 (28.6)		
African American	7 (11.9)	2 (14.3)		
Latino	5 (8.5)	3 (21.4)		
Asian	2 (3.4)	4 (28.6)		
Indian	1 (1.6)	1 (0.4)		
Other	5 (8.5)	0 (0)		
Education Years (SD)	14.93 (2.49)	16.42 (1.16)	$t = 2.32$	.03*
Diagnosis (%)				
Suboptimal Effort	-	14 (100)		
MCI	20 (33.9)	-		
Somatoform	7 (11.9)	-		
Normal	6 (10.2)	-		
ADHD	6 (10.2)	-		
TBI	5 (8.1)	-		
MND	4 (6.8)	-		
PD	3 (5.1)	-		
Mood	3 (5.1)	-		
LD	2 (3.4)	-		
MS	1 (1.7)	-		
Epilepsy	1 (1.7)	-		
Executive dysfunction	1 (1.7)	-		

Note. *denotes significance at $p < .05$. ** denotes significance at $p < .01$

Independent Variables of Interest

Descriptive statistics calculated for all experimental groups on RBANS indices are shown in Table 2. Additionally, descriptive statistics on relevant outcome measures are shown in Table 3.

Table 2. Descriptive Statistics for Experimental Groups on RBANS Indices

	Immediate	Visuospatial	Language	Attention	Delayed	Total Scale
Clinical Groups (Total)	77.46 (15.34)	88.92 (16.59)	92.81 (13.29)	89.44 (16.83)	82.63 (20.87)	82.25 (14.94)
MCI	75.00 (15.04)	91.35 (18.39)	91.55 (12.55)	97.20 (14.70)	81.70 (20.79)	83.55 (14.20)
Somatoform	78.14 (13.40)	89.14 (21.24)	94.29 (9.36)	75.57 (17.03)	80.00 (20.73)	78.86 (16.64)
Normal	88.33 (20.39)	90.17 (13.57)	101.50 (9.48)	97.33 (22.12)	91.83 (19.29)	91.33 (24.11)
ADHD	80.83 (6.31)	91.50 (13.53)	94.83 (16.33)	84.17 (8.84)	89.67 (10.69)	84.33 (8.94)
TBI	82.40 (14.54)	90.80 (14.69)	89.80 (11.67)	88.00 (19.90)	82.00 (34.76)	83.20 (21.42)
MND	59.00 (6.93)	67.50 (5.80)	77.75 (16.46)	79.00 (10.68)	58.75 (14.64)	61.00 (8.60)
PD	68.67 (6.35)	88.67 (13.50)	93.33 (7.09)	95.00 (6.25)	93.33 (8.08)	83.67 (5.51)
Mood	87.67 (28.10)	97.67 (12.50)	107.67 (10.97)	88.33 (22.19)	84.67 (21.36)	91.67 (24.11)
LD	77.00 (5.66)	97.00 (7.07)	100.00 (11.31)	73.00 (12.73)	77.00 (5.66)	78.00 (16.97)
MS	65.00 (-)	64.00 (-)	99.00 (-)	91.00 (-)	65.00 (-)	75.00 (-)
Epilepsy	78.00 (-)	72.00 (-)	74.00 (-)	72.00 (-)	78.00 (-)	72.00 (-)
Executive dysfunction	94.00 (-)	92.00 (-)	71.00 (-)	100.00 (-)	94.00 (-)	87.00 (-)
Actor Group	65.93 (15.32)	65.07 (11.17)	70.00 (27.47)	62.00 (20.36)	61.43 (20.20)	59.50 (15.47)

Notes. Scores are standard scores (M = 100, SD = 15). Abbreviations: MCI (Mild Cognitive Impairment), ADHD (Attention-Deficit Hyperactivity Disorder), TBI (Traumatic Brain Injury), MND (Major Neurocognitive Disorder), PD (Parkinson's Disease), LD (Learning Disorder), MS (Multiple Sclerosis).

Table 3. Descriptive Statistics for Experimental Groups on Effort Outcome Measures

	RES	ES	EI	DCT
Clinical Groups (Total)	39.59 (2.29)	7.69 (10.15)	0.95 (1.46)	11.76 (4.10)
MCI	39.20 (2.82)	13.10 (8.78)	0.70 (1.34)	11.60 (4.02)
Somatoform	39.29 (2.63)	2.29 (5.74)	2.86 (1.95)	14.29 (5.96)
Normal	40.50 (0.84)	2.67 (6.06)	0.33 (0.82)	8.17 (2.04)
ADHD	40.83 (0.41)	-3.00 (4.10)	1.17 (1.32)	10.17 (2.14)
TBI	39.20 (2.68)	6.40 (13.10)	1.00 (1.41)	11.40 (3.05)
MND	37.00 (2.16)	21.00 (6.88)	0.75 (0.96)	15.00 (2.45)
PD	41.00 (0.00)	12.33 (5.86)	0.00 (0.00)	13.67 (5.51)
Mood	40.00 (1.73)	2.67 (9.29)	1.00 (1.73)	11.00 (5.57)
LD	40.50 (0.71)	4.00 (9.43)	1.00 (1.41)	11.50 (2.12)
MS	40.00 (-)	17.00 (-)	0.00 (-)	9.00 (-)
Epilepsy	40.00 (-)	-2.00 (-)	0.00 (-)	17.00 (-)
Executive dysfunction	41.00 (-)	-6.00 (-)	0.00 (-)	12.00 (-)
Actor Group	33.14 (8.05)	5.43 (5.40)	4.79 (4.84)	19.79 (7.57)

Note. Abbreviations: MCI (Mild Cognitive Impairment), ADHD (Attention-deficit hyperactivity disorder), TBI (Traumatic Brain Injury), MND (major neurocognitive disorder), PD (Parkinson's Disease), LD (Learning Disorder), MS (Multiple Sclerosis)

RES Reliability Analyses

To analyze the primary aim of assessing the internal consistency of the RES, the Kuder-Richardson Formula 20 (KR-20; Kuder & Richardson, 1937) was utilized. The KR-20 is recommended over the split half method of internal consistency reliability because the split-half method artificially reduces a test's reliability by its division of the analysis into two parts. Additionally, the KR-20 is recommended for a test that is dichotomously scored such as the RES (Cortina, 1993). Our internal consistency analysis revealed that the 41-item RES with picture naming, figure copy, coding, and word list subtests had a reliability coefficient of $\alpha = 0.91$, which is in accordance with acceptable standards for individual decision-making (Nunnally, 1978). Individual reliability analyses for individual subtests were as follows: RES picture naming $\alpha = 0.81$, RES figure copy $\alpha = 0.72$, RES coding $\alpha = 0.65$, RES word list $\alpha = 0.81$. As such, no individual subtest alone demonstrated an acceptable reliability for individual decision-making. Considering the low reliability level of the RES coding, the RES' reliability was assessed once again after extracting the coding subtest, which revealed similar reliability, $\alpha = 0.91$.

RES Validity

To determine convergent validity, partial correlations were used between the RES total score to assess for associations with existing effort measures such as the DCT, ES, and EI controlling for age and years of education. Analyses revealed that the RES was negatively associated with the EI ($r = - 0.83$, $p < .01$) and the DCT ($r = -0.52$, $p < .01$). As

such, higher scores on the RES were associated with lower scores on the EI and DCT. It was not significantly associated with the ES, $p > .05$.

Additionally, partial correlations were utilized for all individual RES subtests to examine their associations with the DCT, ES, and EI, again controlling for age and years of education. RES picture naming was negatively associated with the EI ($r = -0.86$, $p < .01$) and the DCT ($r = -0.53$, $p < .01$) but was not significant associated with the ES, $p > .05$. RES figure copying was negatively associated with the ES ($r = -0.28$, $p < .01$), the EI ($r = -0.73$, $p < .01$), and the DCT ($r = -0.56$, $p < .01$). The RES word list was negatively associated with the EI ($r = -0.85$, $p < .01$) and the DCT ($r = -0.52$, $p < .01$) but was not significantly associated with the ES, $p > .05$. RES coding was significantly associated with the ES ($r = -0.35$, $p < .01$) and the EI ($r = -0.42$, $p < .01$) but was not significantly associated with the DCT, $p > .05$, see Table 4.

Table 4. Partial Correlations among RES and Effort Indices

	RES Total	Picture Naming	Coding	Figure	List	ES	EI	DCT
RES Total	-	0.93**	0.71**	0.93**	0.90**	-0.22	-0.83**	-0.52**
Picture Naming	-	-	0.48**	0.83**	0.88**	-0.07	-0.86**	-0.53**
Coding	-	-	-	0.59**	0.44**	-0.35**	-0.42**	-.18
Figure	-	-	-	-	0.76**	-0.28**	-0.73**	-0.56**
List	-	-	-	-	-	-0.06	-0.85**	-0.52**
ES	-	-	-	-	-	-	-0.15	0.13
EI	-	-	-	-	-	-	-	.51**
DCT	-	-	-	-	-	-	-	-

Notes.* denotes significance at $p < .05$. ** denotes significance at $p < .01$ level.

Because the RES coding subtest demonstrated the weakest reliability (α = 0.65) and weakest associations with existing effort detection measures in this study, an additional exploratory analysis was included. After eliminating coding from the RES, the RES was more significantly associated with the EI (r = -.86, p <.01) and the DCT (r = -0.57, p <. 01).

To assess the RES's criterion validity, an Analysis of Covariance (ANCOVA) was utilized to examine how the RES could accurately differentiate between participants groups espousing adequate and suboptimal effort, see Table 5. Because of the possibility that members of the CNO would also provide suboptimal effort on neuropsychological testing, it was decided to recategorize the groups according to the more established DCT E-score. Previous research has suggested a general cut-off score of ≥17 for classification of suboptimal effort (Boone et al., 2002), which was used for our reclassification of variables. As such, we re-classified our data into two groups (good and poor effort according to DCT E score) and compared the two groups on their RES performance. Following this reclassification, 17 participants were left in the suboptimal effort group and 56 participants in the adequate effort group. Using the log-based transformation for the RES to conform with the univariate assumption of normality, the ANCOVA was significant [F (1,69) = 14.87, p < .01, r^2 = .19]. As such, individuals engaging in adequate effort (M = 39.41, SD = 3.01) scored significantly higher on the RES than individuals who engaged in suboptimal effort (M = 34.88, SD = 7.30), p <.01 which suggests that the full RES was a valid indicator of effort detection on neuropsychological testing, see Table 6. Similarly, when RES Coding was extracted from the full RES analyses, the

adequate effort group continued to perform significantly better than the suboptimal effort group [F (1,69) = 16.48, $p < .01$] with an equivalent effect size ($r^2 = .19$).

Similar analyses were examined on log-based transformations of individual RES subtests. The adequate effort group performed significantly better than the suboptimal effort group on RES Picture Naming [F (1,69) = 38.99, $p < .01$, $r^2 = .39$], RES Figure Copy , [F (1,69) = 23.15, $p < .01$, $r^2 = .25$], RES List Learning, [F (1,69) = 21.81, $p < .01$, $r^2 = .26$], and RES Coding , [F (1,69) = 8.30, $p < .01$, $r^2 = .17$], see Table 6. Analyses indicated that RES Picture Naming demonstrated the largest effect among individual subtests ($r^2 = .39$), whereas coding demonstrated the smallest effect ($r^2 = .17$).

Additional analyses indicated that individuals diagnosed with mild cognitive impairment or dementia did not perform significantly differently on the RES than other clinical populations, $p > .05$.

Table 5. Descriptive Statistics of RES Performance by Raw Subtest and Total Score

	RES Picture Naming	RES Coding	RES Figure	RES List	RES Total Score
Clinical Groups	9.92 (0.28)	8.44 (1.21)	11.48 (0.86)	9.76 (0.73)	39.59 (2.29)
MCI	9.90 (0.31)	8.35 (1.39)	11.25 (1.07)	9.70 (0.57)	39.20 (2.82)
Somatoform	9.71 (0.49)	8.86 (0.38)	11.71 (0.49)	9.00 (1.73)	39.29 (2.63)
Normal	10.00 (0.00)	8.67 (0.82)	11.83 (0.41)	10.00 (0.00)	40.50 (0.84)
ADHD	10.00(0.00)	9.00 (0.00)	11.83 (0.41)	10.00 (0.00)	40.83 (0.41)
TBI	10.00 (0.00)	8.00 (1.73)	11.20 (1.10)	10.00 (0.00)	39.20 (2.68)
MND	9.75 (0.50)	7.00 (2.16)	10.50 (1.00)	9.75 (0.50)	37.00 (2.16)
PD	10.00 (0.00)	9.00 (0.00)	12.00 (0.00)	10.00 (0.00)	41.00 (0.00)
Mood	10.00 (0.00)	8.33 (1.15)	11.67 (0.58)	10.00 (0.00)	40.00 (1.73)
LD	10.00 (0.00)	8.50 (0.71)	12.00 (0.00)	10.00 (0.00)	40.50 (0.71)
MS	10.00 (-)	9.00 (-)	11.00 (-)	10.00 (-)	40.00 (-)
Epilepsy	10.00 (-)	8.00 (-)	12.00 (-)	10.00 (-)	40.00 (-)
Executive dysfunction	10.00 (-)	9.00 (-)	12.00 (-)	10.00 (-)	41.00 (-)
Actor Group	7.96 (2.34)	7.93 (1.27)	9.42 (2.41)	7.86 (2.57)	33.14 (8.05)

Note. Abbreviations: MCI (Mild Cognitive Impairment), ADHD (Attention-deficit hyperactivity disorder), TBI (Traumatic Brain Injury), MND (major neurocognitive disorder), PD (Parkinson's Disease), LD (Learning Disorder), MS (Multiple Sclerosis)

Table 6. RES Descriptive Statistics for Effort Groups

	Picture Naming	Coding	Figure	List	Total Score
Adequate Effort	9.84 (0.63)	8.41 (1.30)	11.45 (1.03)	9.71 (0.76)	39.41 (3.01)
Suboptimal Effort	8.53 (2.18)	8.12 (0.93)	9.88 (11.45)	8.35 (2.52)	34.88 (7.30)

Exploratory Analyses

Exploratory analyses included ROC curves examining the sensitivity and specificity of the RES with and without the coding subtest. When examining the full RES, our analyses revealed a cutoff score of 39.50 was associated with moderate sensitivity (sensitivity = 0.73) with moderate specificity (specificity = 0.59), see Figure 1. When excluding the coding subtest, a cut-off of 30.50 (out of a total of 32 points) was associated with moderate sensitivity (sensitivity = .80) and moderate specificity (specificity = .53), see Figure 2.

In comparison to the RES, the EI also had moderate sensitivity (sensitivity = 0.65) and moderate specificity (specificity = 0.68) at a cut-off at 0.5, see Figure 3. The ES had moderate sensitivity (sensitivity = 0.71) and moderate specificity (specificity = 0.54)at a cutoff at 3.50, see Figure 4.

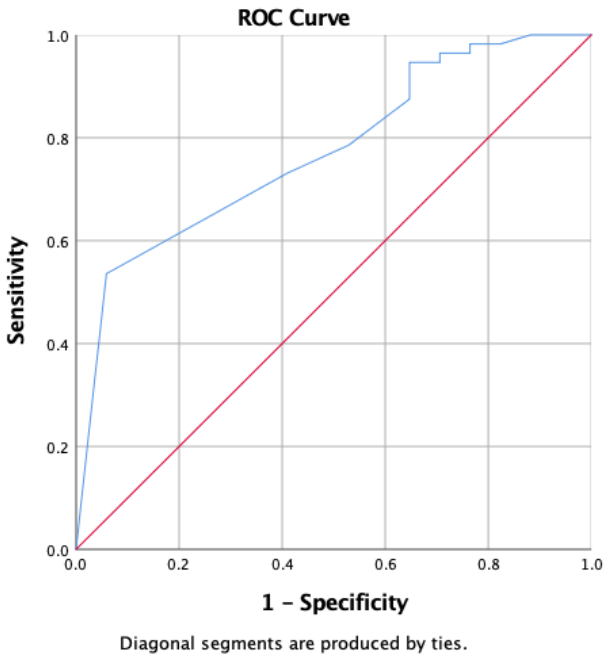


Figure 1. *ROC curve analyzing RES sensitivity and specificity*.

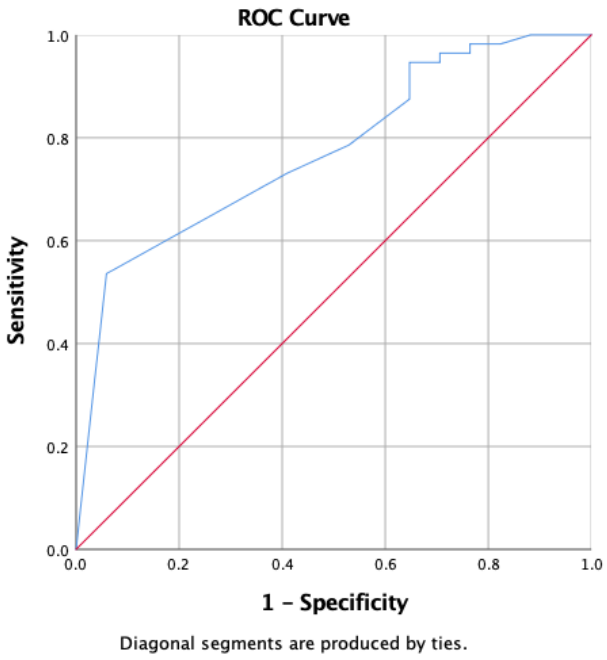


Figure 2. ROC curve analyzing RES sensitivity and specificity.

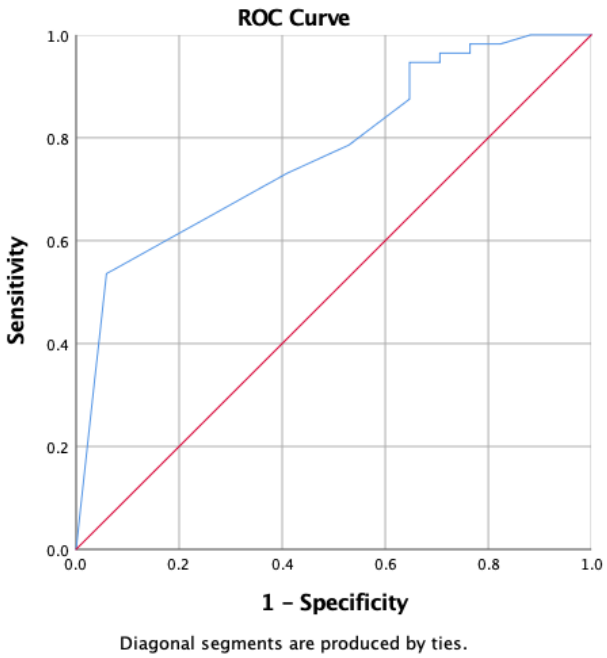


Figure 3. ROC Curve analyzing EI sensitivity and specificity.

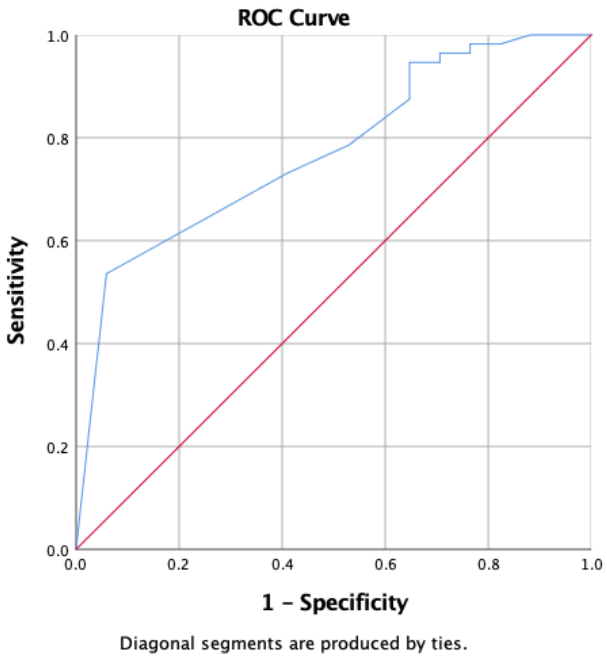


Figure 4. ROC curve analyzing ES sensitivity and specificity.

CHAPTER FIVE

DISCUSSION

This study analyzed the reliability and validity of a performance validity supplement to the RBANS. Data was collected for this study from September 2018 until April of 2019. This study analyzed data from 59 clinical neuropsychology outpatients from Loma Linda University Medical Center's Clinical Neuropsychology Clinic and 14 experimental suboptimal effort actors from Loma Linda University's graduate student population.

The purpose of this study was to build upon existing measures of effort detection within the initial screening process. Researchers have developed embedded effort detection measures in the RBANS, namely the RBANS Effort Scale (2012) and the RBANS Effort Index (2007), which estimate effort through analysis of recall and digit span scores. Both measures have been found to be sensitive but limited in specificity when classifying clinical patients from individuals exhibiting suboptimal effort. As such, this study centered around the validation of a new supplement, which incorporated multiple forced-choice paradigms to create a more well-rounded effort-detection measure.

The primary hypothesis of this study was that the RES would be a reliable and valid measure of effort detection. KR-20 analyses revealed that our hypothesis was confirmed from a reliability standpoint. However, none of the individual subtests alone reached acceptable alpha levels for individual decision-making. RES Coding demonstrated the lowest alpha level and after extracting it from the total RES, the RES had an equivalent alpha level. Validity analyses confirmed our hypothesis that the RES

would demonstrate convergent validity with existing measures of effort detection including the EI and DCT. It should be noted that the RES was not significantly correlated with ES; this may be emblematic of the primary caveat of the ES in that individuals who excel on free recall on the RBANS have significantly negative scores on their ES composite score. Individual subtests demonstrated similarly significant associations with the EI and the DCT, with RES Picture Naming having the strongest correlation among subtests with existing effort measures. RES Coding had the weakest correlation with existing effort measures and after extracting it from the total RES score, the RES' associations with the EI and the DCT slightly improved. The RES also demonstrated construct validity; participants who had been classified into a suboptimal effort group according to DCT E-score performed significantly worse than their counterparts in the similarly classified in the adequate effort group. All individual subtests reflected similar group differences, with RES Picture Naming again demonstrating the strongest effect and RES Coding demonstrating the weakest effect. When extracting RES Coding from the RES Total score, the effect size was equivalent.

Notably, the no significant differences were detected in the Total RES and the RES without Coding scores were between individuals with a memory disorder and other clinical participants. Participants presenting with memory impairment are not expected to perform significantly worse than individuals without memory impairment on the RES, as the RES is not a memory measure. These results demonstrate the RES' strength as an effort detection measure, despite its face validity as memory measure. Given these results, the RES appears to be a true measure of effort, and not a measure of memory function.

Exploratory analyses indicated that the Total RES was moderately sensitive and specific at a cutoff of 39.50; when coding was extracted, the measure was slightly more sensitive and slightly less specific. It should be noted that the RES demonstrated greater sensitivity and specificity than the ES and the EI in this study.

Clinical Implications

There are many exciting clinical implications from this study. Given the RES' observed reliability and validity, our study demonstrates its utility in the initial neuropsychological screening process. The RES' compilation of several effort measures in one supplement may provide clinicians with a more well-rounded analysis and characterization of their patient's effort. The RES' multifactorial detection of effort is also a measure that can be completed in approximately 10-15 minutes with the opportunity to give providers valuable information prior to committing to a full neuropsychological evaluation. This notion may be associated with significant cost reduction while also saving significant time. Additionally, the RES may provide clinicians the opportunity to immediately discuss effort from a multidimensional standpoint when it is in question. Such discussions may conjunctively be useful in determining whether to pursue further neuropsychological testing as well.

Broader implications include the importance of assessing effort in most if not at all clinical contexts. Effort detection options are widely available for neuropsychologists to utilize with most referral questions. Effort detection also validates the nature of neuropsychological services in that clinicians can validate an individual's diagnosis and subsequently provide appropriate recommendations for on their behalf. Effort detection

rules out the possibility of feigned impairment for personal gain and essentially provides credence to the field of neuropsychology.

Limitations

This study is not without limitations. Primarily, a control group would have provided a baseline comparison to both the clinical and suboptimal effort groups. Additionally, the study would have benefitted from a larger sample size in general with larger representation from common neuropsychological referrals. Understanding performance from various neuropsychological perspectives would be helpful in analyzing RES performance trends from a diversity of neuropsychological presentations. Our experimental groups differed significantly in terms of sample size, which may have contributed to the skewness of the original raw data. Additionally, sampling in itself may have been a confounding issue. Specifically, participants in the experimental suboptimal effort group were highly educated, averaging over 16 years of education, and were actively participating in graduate education. Most graduate programs in Loma Linda University emphasize a broad academic curriculum and it is possible that participants may have had prior knowledge of suboptimal effort presentations on neurocognitive testing.

Research Implications and Future Directions

This study leads to several questions regarding future research. It may be useful to consider including a digits backward component to the RES; this may allow for the computation of reliable digits similar to the WAIS-IV and would add yet another

component of effort detection to the supplement. Additionally, it is recommended that the RES be analyzed for reliability and validity in other clinical settings as well. The RES would certainly benefit significantly from replication in other settings and among a wide variety of clinical and demographic populations.

Conclusion

In summary, this study analyzed the reliability and the validity of a novel measure of effort and motivation, the RBANS effort supplement. This study found that the RES was a reliable measure of effort detection. Additionally, the RES exhibited convergent validity with an established embedded effort detection measure from the RBANS (the RBANS Effort Index) and the DCT, which is another well-established independent effort detection measure. The RES demonstrated construct validity in that participants who were classified in the suboptimal effort group according to their performance on the DCT performed significantly worse on the RES than did individuals who had been classified into the adequate effort group. A ROC curve analysis was performed and demonstrated that the RES exhibited moderate sensitivity and specificity at a cut-off score of 39.50. Clinical implications of this study include the potential for screening for effort from a multifactorial approach during the initial neuropsychological screening process, which may significantly reduce costs and save a significant amount of time. Key limitations include a lack of a control group, small sample size, and lack of greater representation from common outpatient referral sources. Future research directions include replication of reliability and validity analyses in a different neuropsychological setting. This study identified the RES as a useful measure in detecting effort, but further research is

undoubtedly necessary to fully understand the extent of its utility in a clinical neuropsychological setting.

REFERENCES

Axelrod, B. N., Fichtenberg, N. L., Millis, S. R., & Wertheimer, J. C. (2006). Detecting Incomplete Effort with Digit Span from the Wechsler Adult Intelligence Scale—Third Edition. *The Clinical Neuropsychologist*, *20*(3), 513-523.

Bayley, P. J., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2008). Yes/No Recognition, Forced-choice Recognition, and the Human Hippocampus. *Journal of Cognitive Neuroscience*, *20*(3), 505–512. http://doi.org/10.1162/jocn.2008.20038

Beatty, W. W., Mold, J. W., & Gontkovsky, S. T. (2003). RBANS performance: influences of sex and education. *Journal of Clinical and Experimental Neuropsychology*, *25*(8), 1065-1069.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom Validity Testing: A Critical Review. *The Clinical Neuropsychologist*,*15*(1), 19-45. doi:10.1076/clin.15.1.19.1907

Bigler, E. D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society*, *18*(4), 632-640.

Boone, K. B. (Ed.). (2007). *Assessment of feigned cognitive impairment: a neuropsychological perspective*. New York, NY: The Guilford Press.

Boone, K. B., Lu, P. and Herzberg, D. S. (2002). *The b Test Manual*, Los Angeles, CA: Western Psychological Services.

Boone, K., Lu, P., & Herzberg, D. S. (2002). *The Dot Counting Test.* Los Angeles: Western Psychological Services.

Boone, K. B., & Lu, P. H. (2007). Non-Forced-Effort Measures. In G. J. Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 27-44). New York, NY: Oxford University Press.

Boone, K. B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., ... & Warner-Chacon, K. (2002). Sensitivity and specificity of the Rey Dot Counting Test in patients with suspect effort and various clinical samples. *Archives of Clinical Neuropsychology*, *17*(7), 625-642.

Bouman, Z., Hendriks, M. P., Schmand, B. A., Kessels, R. P., & Aldenkamp, A. P. (2016). Indicators of suboptimal performance embedded in the Wechsler Memory Scale–Fourth Edition (WMS–IV). *Journal of Clinical and Experimental Neuropsychology*,*38*(4), 455-466. doi:10.1080/13803395.2015.1123226

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A, & Kaemmer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Carone, D. A., Iverson, G. L., & Bush, S. S. (2010). A Model to Approaching and Providing Feedback to Patients Regarding Invalid Test Performance in Clinical Neuropsychological Evaluations. *The Clinical Neuropsychologist,24*(5), 759-778. doi:10.1080/13854041003712951

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology,78*(1), 98-104. doi:10.1037//0021-9010.78.1.98

Crighton, A. H., Wygant, D. B., Holt, K. R., & Granacher, R. P. (2015). Embedded Effort Scales in the Repeatable Battery for the Assessment of Neuropsychological Status: Do They Detect Neurocognitive Malingering? *Archives of Clinical Neuropsychology,30*(3), 181-185. doi:10.1093/arclin/acv002

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). California Verbal Learning Test—2nd edition: Manual. San Antonio: The Psychological Corporation.

DenBoer, J.W. & Hall, S. (2007). Neuropsychological test performance of successful brain injury simulators. *The Clinical Neuropsychologist, 21*(6). 943-955.

Duff, K., Hobson, V.L, Beglinger, L.J, & O'Bryant, S. E. (2010). Diagnostic Accuracy of the RBANS in mild cognitive impairment: limitations on assessing milder impairments. *Archives of Clinical Neuropsychology, 25*(5), 429-441.

Duff, K., Humphreys Clark, J. D., O'Bryant, S. E., Mold, J. W., Schiffer, R. B., Sutker, P.B. (2008). Utility of the RBANS in detecting cognitive impairment associated with Alzheimer's disease: Sensitivity, specificity, and positive and negative predictive powers. *Archives of Classical Neuropsychology 23*, 603-612.

Duff, K., Spering, C. C., O'Bryant, S. E., Beglinger, L. J., Moser, D. J., Bayless, J. D., . . . Scott, J. G. (2011). The RBANS Effort Index: Base Rates in Geriatric Samples. *Applied Neuropsychology,18*(1), 11-17. doi:10.1080/09084282.2010.523354

Etheron, J.L., Bianchini, K.J., Greve, K.W., & Heinly, M.T. (2005). Sensitivity and specificity of reliable digit span in malingered pain-related disability. *Assessment, 12* (2), 130-136.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:10.3758/bf03193146

Frederick, R. I., & Speed, F. M. (2007). On the interpretation of below-chance responding in forced-choice tests. *Assessment*, *14*, 3-11.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*,*6*(3), 218-224. doi:10.1037//1040-3590.6.3.218

Grimes, D., A., & Schulz, K. F. (2005). Refining clinical diagnosis with likelihood ratios. *The Lancet, 365,* 1500-1505.

Grote, C. L., & Hook, J. N. (2007). Forced-Choice Recognition Tests of Malingering. In G. J. Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 44-80). New York, NY: Oxford University Press.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology*, *143*, 29–36.

Hartman, D. E. (2009). *Applied Neuropsychology*, *16*(WAIS-IV), 85–87. doi:10.1080/09084280802644466

Heilbronner, R., Sweet, J., Morgan, J., Larrabee, G., Millis, S., & Participants, C. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the Neuropsychological Assessment of Effort, Response Bias, and Malingering. *The Clinical Neuropsychologist*,*23*(7), 1093-1129. doi:10.1080/13854040903155063

Institute of Medicine. (2015). *Psychological Testing in the Service of Disability Discrimination.* The National Academies Press. doi: 10.17226/21704.

Jasinski, L. J., Berry, D. T., Shandera, A. L., & Clark, J. A. (2011). Use of the Wechsler Adult Intelligence Scale Digit Span subtest for malingering detection: A meta-analytic review. *Journal of Clinical and Experimental Neuropsychology*,*33*(3), 300-314. doi:10.1080/13803395.2010.516743

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*(3), 151-160.

Larrabee, G. J. (2014). Assessment of Performance and Symptom Validity and the Diagnosis of Malingering. In M. W. Parsons & T. A. Hammeke (Eds.), *Clinical Neuropsychology: A Pocket Handbook for Assessment* (3rd ed., pp. 90-113). American Psychological Association.

Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society, 18* (4), 625–630.

Larrabee, G. J. (2008). Aggregation across multiple indicators improves the detection of malingering: relationship to likelihood ratios. *The Clinical Neuropsychologist*, *22* (4), 666-679. DOI: 10.1080/13854040701494987

Larrabee, G. J. (2007). Introduction: Malingering, Research Design, and Base Rates. In G. J. Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 3-14). New York, New York: Oxford University Press.

Larrabee, G. J., & Berry, D. T. (2007). Diagnostic Classification Statistics and Diagnostic Validity of Malingered Assessment. In G. J. Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 14-27). New York, New York: Oxford University Press**.**

Larrabee, G.J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist, 17* (3), 410-425.

Martin, P.K, Schroeder, R. W., & Odland, A.P. (2015). Neuropsychologists validity testing beliefs and practices: a survey of North American professionals. *The Clinical Neuropsychologist*, (*29*) 6, 741-776. doi: 10.1080/13854046.2015.1087597

McDougall, R. (1904). Recognition and recall. *Journal of Philosophy, Psychology, and Scientific Methods, 1*, 229-233.

Meyers, J. E., & Meyers, K.R. (1995). Rey Complex Figure Test and Recognition Trial. Odessa, Florida: Psychological Assessment Resources.

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base Rates of Malingering and Symptom Exaggeration. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A),24*(8), 1094-1102. doi:10.1076/jcen.24.8.1094.8379

Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury,18*(10), 975-984. doi:10.1080/02699050410001672350

Morgan, J. E., & Sweet, J. J. (2015). *Neuropsychology of malingering casebook*. New York, NY: Psychology Press.

Nelson, N.W., Boone, K., Dueck, A., Wagener, L., Lu, P., & Grills, C. (2003). Relationships between eight measures of suspect effort. *The Clinical Neuropsychologist, 17* (2), 263-272.

Novitski, J., Steele, S., Karantzoulis, S., & Randolph, C. (2012). The Repeatable Battery for the Assessment of Neuropsychological Status Effort Scale. *Archives of Clinical Neuropsychology,27,* 190-195.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Paul, R., Lane, E. M., Tate, D. F., Heaps, J., Romo, D. M., Akbudak, E., … Conturo, T. E. (2011). Neuroimaging signatures and cognitive correlates of the montreal cognitive assessment screen in a nonclinical elderly sample. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists, 26*(5), 454–60. doi:10.1093/arclin/acr017

Postman, L. (1963). One trial learning. In C. F. Cofer & B. S. Musgrave (Eds.), Verbal behavior and Learning (pp. 295-332). New York: McGraw-Hill.

Randolph, C. (1998). *Repeatable Battery for the Assessment of Neuropsychological Status*. San Antonio, TX: The Psychological Corporation.

Root, J. C., Robbins, R. N., Chang, L., & Gorp, W. G. (2006). Detection of inadequate effort on the California Verbal Learning Test-Second edition: Forced choice recognition and critical item analysis. *Journal of the International Neuropsychological Society,12*(05). doi:10.1017/s1355617706060838

Silverberg, N. D., Wertheimer, J. C., & Fichtenberg, N. L. (2007). An effort index for the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). *The Clinical Neuropsychologist, 21*(5), 841-854.

Slick, D.J., Tan, J.E., Strauss, E.H., & Hultsch, D.F. (2004). Detecting malingering: a survey of experts' practices. *Archives of Clinical Neuropsychology, 19*, 465-473.

Strauss, E., Sherman, E. M., Spreen, O. (2006). *A compendium of neuropsychological tests: administration, norms, and commentary*. Oxford: Oxford University Press.

Teicher, G., & Wagner, M. T. (2004). The Test of Memory Malingering (TOMM): Normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology, 19*, 455–464.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment,9*(3), 260-268. doi:10.1037//1040-3590.9.3.260

Wechsler, D. (2009). Advanced clinical solutions for the WAIS-IV and WMS-IV. San Antonio, TX: Pearson.

Wechsler, D. (2008). Wechsler Adult Intelligence Scale—Fourth Edition. San Antonio, TX: Pearson.

Wechsler, D. (2009). Wechsler Memory Scale–Fourth Edition. Pearson; San Antonio, TX: Pearson.

Wechsler, D. (1981). Wechsler Adult Intelligence Scale—Revised. San Antonio, TX: The Psychological Corporation.

APPENDIX A

RES STUDY DEMOGRAPHICS QUESTIONNAIRE

RES Study Demographics Questionnaire

Date of birth: ___

Gender: _____

Ethnicity: _____

Primary Language: _____

If English is not primary language: ____ Completely fluent in English

____ No difficulty with conversational English

____ Some difficulty with self-expression in English

Highest degree completed: _____

Years of education: _____

Handedness: _____

Previous Evaluation? Yes No If yes, date of previous evaluation: _____

The following questions for patient groups only:

Reason for your evaluation today: _____

Date of Injury (if applicable): _____

Involved in legal case associated with injury? Yes No

If yes, are you seeking any of the following? ____ Financial disability settlement

____ Social Security

____ Worker's Compensation

____ Academic Accommodation

____ Other

Loma Linda University Health
Institutional Review Board
Approved 10/5/2018
IRB# 5180167

APPENDIX B

IRB APPROVAL DOCUMENTS

LOMA LINDA UNIVERSITY
HEALTH

INSTITUTIONAL REVIEW BOARD
RESEARCH PROTECTION PROGRAMS
24887 Taylor Street • Suite 202 • Loma Linda, CA 92350
(909) 558-4531 (voice) • (909) 558-0131 (fax)

Initial Approval Notice - Expedited ***IRB# 5180167***

To: **Lee, Grace**
Department: **Psychology**
Protocol: **The formulation of a RBANS effort supplement**

This study was reviewed and approved administratively on behalf of the IRB. This decision includes the following determinations:

Risk to research subjects: **Minimal**
Approval period begins: **04-Sep-2018** and ends **03-Sep-2019**
Stipulations of approval:
See attached list of items (if applicable).
See Appendix A for Conditions of Approval.

Adverse events and unanticipated problems must be reported in accord with the attached Adverse Event Reporting Matrix **A.**

All investigators are responsible for assuring that studies are conducted according to the approved protocol. Principal investigators are responsible for the actions of sub-investigators and staff with regard to this approval.

Please note the PI's name and the assigned IRB number, as indicated above, on any future communications with the IRB.

Direct all communications to the IRB c/o Research Protection Programs.

Thank you for your cooperation in LLU's shared responsibility for the ethical use of human subject in research.

IRB Chair/Designee

9/5/18
Date

Loma Linda University Health holds Federalwide Assurance (FWA) No. 00006447 with the U.S. Office for Human Research Protections and the IRB registration no. is IOR00000226. This Assurance applies to the following: Loma Linda University, Loma Linda University Medical Center (including Loma Linda University Children's Hospital, LLUMC East Campus Hospital), Loma Linda University Behavioral Medicine, and affiliated medical practices groups.

IRB Chair:
Travis Losey, MD
Department of Neurology
(909) 558-4531 • tlosey@llu.edu • Pager #4290 for emergencies

IRB Administrator:
Amy L. Casey, MBA
Responsible Research
Ext 14658 • Fax 80131 • acasey@llu.edu

LOMA LINDA
UNIVERSITY
HEALTH

CONDITIONS OF IRB APPROVAL

IRB REQUIREMENTS FOR IMPLEMENTING FEDERAL AND INSTITUTIONAL IRB POLICIES

General

1. No subjects may be involved in any study procedure prior to the date indicated in the IRB approval notice or after the expiration date. This includes any testing solely to determine eligibility.
2. Unless IRB has stipulated a waiver of informed consent, the current IRB-approved consent form must be used to enroll subjects. The official consent accompanying this letter may be used as the master for making copies to provide to prospective study participants.
3. All recruitment materials and methods must be approved by the IRB prior to being used.
4. All modifications (e.g. changes to the protocol, the informed consent document(s)/process, or number of subjects) must be IRB approved prior to implementation (see http://research.llu.edu/changerequest.asp)
5. No study activities may be conducted after the IRB approval end date unless determined to be medically necessary by the PI.
6. To avoid disruption of a study, the PI should request renewal of IRB approval by promptly submitting a research report form (see http://research.llu.edu/researchreport.asp)

Records retention

Minimum requirements

- All records relating to this project, including signed consent forms, must be kept on file for at least 3 years following completion of the study, unless special conditions apply (see below).
- Study records MUST remain at the institution under the jurisdiction of a responsible party to be able to locate records after study closure. This includes approved IRB documents, as well as case-report forms, tapes, or transcripts, and all other data collection instruments and source documents.

Other special conditions

- Records involving the generation, disclosure, and/or use of Protected Health Information (PHI) should be retained for six years.
- Minors in research: records must be retained for seven years after all minors enrolled in the study reach the age of majority [age 18 in California]
- Records pertaining to in vitro fertilization studies or research involving pregnant women must be retained for a total of 25 years after study closure.
- Records generated from sponsored clinical trials may not be disposed until written permission is obtained from the sponsor.

FDA-regulated studies - In addition to minimum requirements:

- For drugs and devices with approved marketing applications, the retention period is two years after FDA approval.
- For drugs and devices where no application is filed or the application is not approved, the retention period is two years after the investigation is discontinued and FDA is notified. 21 CFR 312.57 and 21 CFR 812.140.

ver. 8/1/2014

APPENDIX C

MANUSCRIPT FOR ARCHIVES OF CLINICAL NEUROPSYCHOLOGY JOURNAL

The Formulation of a RBANS Effort Supplement: A Validation Study

Joshua S. Goldberg, Travis G. Fogel, Kyrstle D. Barrera, Kendal C. Boyd,

and Grace J. Lee

Loma Linda University

Author Note

Joshua Goldberg, Department of Psychology, Loma Linda University

Correspondence concerning this article should be addressed to Joshua Goldberg, Department of Psychology, Loma Linda University, School of Psychology, Loma Linda, CA, 92354

Email: jgoldberg@llu.edu

Abstract

Assessment of effort detection is an essential component of a neuropsychological evaluation to ensure results of testing are valid indicators of an individual's true level of cognitive functioning. Effort detection in the initial screening process provides neuropsychologists information regarding patients' test engagement prior to administering longer testing batteries. Two effort measures are embedded in the Repeatable Battery for Assessment of Neuropsychological Status (RBANS), a neuropsychological screening assessment, but both have demonstrated elevated false positive rates for classifying individuals with memory impairment as those putting forth poor effort. These embedded measures rely on cut-off scores on digit span and memory subtests. In contrast, this RBANS Effort Supplement (RES) utilizes several forced-choice subtests, reflective of current research emphasizing the importance of multiple methods of effort detection; subtests in this measure included list learning forced-choice, figure copy forced-choice, picture naming forced-choice, a coding task, and a story recognition component utilized for face validity of memory assessment. Fifty-nine participants were recruited from an outpatient neuropsychology facility in conjunction with 14 poor effort simulators; each participant was administered the RBANS, the RES, and the Dot Counting Test (DCT). Results supported the RES' reliability at the individual decision-making level. Validity analyses demonstrated that the RES exhibited strong convergent validity with established effort detection measures and that individuals putting for poor effort scored significantly lower on the RES than individuals who put forth adequate effort, as delineated by the established DCT cutoff score of 17. In summary, the RES was shown to be a valid indicator of effort detection.

Clinical implications of the RES include reduction of time and costs involved in neuropsychological assessment

.

The Formulation of a RBANS Effort Supplement

As part of the neuropsychological assessment, patients are asked to put forth their best effort throughout the administration of cognitive testing so that valid data may be compiled that is an accurate representation of their cognitive functioning. Occasionally, patients can consciously or unconsciously fail to provide adequate effort resulting in invalid testing data.

Glenn Larrabee (2007) explains that suboptimal effort is not necessarily uncommon in neuropsychological settings. It is estimated that cases of poor effort occur in 29% of personal injury cases, 30% of disability cases, 19% of criminal cases, 38.5% of personal injury cases, and 8% of general medical cases involving symptom exaggeration. Determination of suboptimal effort within clinical neuropsychology differs somewhat across settings; however, a large consensus of neuropsychologists agrees that more confidence in definitively diagnosing poor effort occurs through multiple effort measures with little methodological overlap to limit redundancy (Larrabee, 2008; Mittenberg, Patton, Canyock & Condit, 2002).

Effort measures may utilize either a forced-choice or non-forced-choice paradigm. Forced choice measures appear to be difficult but are in fact, easy tasks. Typically, in forced-choice, an individual is asked to encode a series of pictures or words and then later asked to select each of the target pictures or words from two choices. Participants who perform below chance levels (i.e., below 50%) are identified as individuals who may been putting forth suboptimal effort.

Effort measures can often be integrated into initial consultations along with neuropsychological screening measures to help identify the cognitive capacities of new

patients as well as determine whether interpretations and future testing may be needed after the initial consult. A commonly administered screening measure is the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998). Research within the last decade has also revealed that among commonly used dementia screening measures, total RBANS performance is one of the better measures in predicting total brain volume (Paul et al., 2011). The RBANS' relevance to dementia screening in neuropsychology has warranted the development of accompanying effort measures to detect feigned impairment. Two such embedded measures include the RBANS Effort Index (EI; Silverberg, Wertheimer, Fichtenberg, 2007) and the RBANS Effort Scale (ES; Novitski, Steele, Karantzoulis, & Randolph, 2012).

Despite the empirically validated research from which these measures were constructed, research has demonstrated that their validity may be somewhat limited. Research has illustrated that although the EI exhibits good specificity for simulated malingerers with a false-positive rate of 19% or less at selected cutoffs, it has only moderate sensitivity (66%), which risks the possibility of misdiagnosing malingerers with memory-related conditions (Crighton, Wygant, Holt, & Granacher, 2015). Concurrently, the ES has misclassified participants as malingerers due to its heavy emphasis on subtracting free recall scores as an overall reflection of its focus on patients with amnesia and has also reflected high false-positive rates as well (Crighton et al., 2015).

In this study, the RBANS Effort Supplement was formulated and assessed for reliability and validity to detect suboptimal effort through the sole usage of the RBANS assessment. It was designed to be a quick measure to administer, with the opportunity for cost-efficiency in that a subsequent longer evaluation would not be needed if effort were

found to be a significant issue. It also included different methods/formats of malingering detection, reflective of Glenn Larrabee's research concerning how the aggregation of varying measures of effort provide a more definitive finding of suboptimal effort (Larrabee, 2008). Thus, the primary aim of this study was to establish the reliability and validity of the RBANS Effort Supplement (RES)

.

Aims and Hypotheses

The primary aim of this study was to determine if the RBANS Effort Supplement (RES) was a reliable and valid measure of effort. To measure the RES' reliability, the RES was assessed for internal consistency utilizing the Kuder-Richardson 20 method (Kuder & Richardson, 1937). It was hypothesized that the RES would be internally consistent. Following reliability analysis, the construct validity of the RES was examined. Specifically, The RES was assessed for convergent validity utilizing partial correlations controlling for age and years of education. It was hypothesized that the RES would exhibit convergent validity with the RBANS Effort Index, RBANS Effort Scale, and the Dot Counting Test. Further, we hypothesized that participants within the experimental malingering sample would score significantly lower on the RES in comparison to clinical groups.

An exploratory aim of this study was to examine the specificity and sensitivity of the RES. Such analyses were conducted utilizing ROC curve analyses according to a RES cut-off score to be determined according to frequency characteristics of the RES itself. It was hypothesized that the RES will be specific and sensitive in correctly classifying individuals engaging in suboptimal performance clinical groups.

Methods

Participants and Procedures

Our study included two independent samples, a clinical neuropsychology outpatient (CNO) and a comparative suboptimal effort group. The CNO group was comprised of 59 outpatients, who were referred for neuropsychological testing at the Loma Linda University Medical Center East Campus neuropsychology service. Our suboptimal effort group was recruited from Loma Linda University and included 15 students from the graduate student population. All subjects fell within the age range of 20-89 and all spoke English fluently. One participant was excluded utilizing the outlier labelling rule on the RES total score to help correct for the skewness of the data. As such our analyses included 59 individuals included in the clinical outpatient group in comparison to 14 individuals included in the suboptimal effort group.

Participants involved in the CNO group were individuals who had been referred for clinical neuropsychological services for various reasons, including mild cognitive impairment, traumatic brain injury, stroke, epilepsy, ADHD, and varying mood disorders. After participants had completed a structured clinical interview as part of their neuropsychological referral, they were asked to participate in the current study. Agreeing participants completed the informed consent process and gave permission to use the results of their clinical testing (i.e. RBANS, RES, Dot Counting Test) for the current study. Participants then completed a brief additional structured interview asking for basic demographic information (i.e. age, ethnicity, education, referral complaint, handedness, engagement in previous neuropsychological testing, and current legal involvement).

Participants were administered the RBANS as part of their routine neuropsychological assessment and were additionally administered the RES immediately thereafter.

Participants enrolled in the suboptimal effort group (SEG) were recruited from Loma Linda University's graduate population. Subjects were recruited from various departments in the university through department-wide email notifications and campus-wide postings. Participants completed the informed consent process and a brief structured interview of demographic information. Participants were given the following script (DenBoer & Hall, 2007), prompting them to approach the neuropsychological tests as if they were trying to appear brain damaged in order to receive financial compensation in an ongoing lawsuit:

> You are about to take some cognitive tests that examine mental abilities such as attention, memory, thinking and reasoning skills, and your ability to think quickly. While responding to the tests, please pretend that you have experienced brain damage from a car accident involving a head-on collision. You hit your head against the windshield and were knocked out for 15 minutes. Afterwards, you felt ''dazed'' so you were hospitalized overnight for observation. Because the driver of the other car is at fault, you have decided to go to court to get money from the person responsible. During the next few months following the accident, the negative effects from your head injury disappear. Your lawsuit has not been settled yet, and your lawyer has told you that you may get more money if you look like you are still suffering from brain damage. As you pretend to be this car accident victim, try to respond to each test as a patient who is trying to appear brain damaged in order to get money from the lawsuit. Thus, your performance on the tests should convince the examiner as well as the people involved in deciding the outcome of your lawsuit that you are still suffering from brain damage.

Approval for the study was obtained from the Loma Linda University Human Subjects Committee Institutional Review Board, and written informed consent was acquired from all participants upon enrollment. It should be noted that Loma Linda University associated legal counsel stated that the RBANS Effort Supplement was considered legally permissible as long as primary investigators did not attempt to earn a profit from the measure itself. The RES was only utilized for the purposes of this study.

Instruments

Prior to the examination of participants, examiners interviewed participants using a standardized questionnaire (See Appendix B) in order to gather relevant demographic information, including age, years of education, gender, ethnicity, referral, handedness, prior neuropsychological testing, and engagement in ongoing litigation. Of note, none of the participants were involved in previous litigation and three participants had engaged in previous neuropsychological testing (one participant in 2016, another in 1985, and the third at an unknown time) .

The Dot Counting Test.

Boone, Lu, and Herzberg's Dot Counting Test (2002) is a measure of symptom validity and malingering. Participants are presented with a series of twelve dotted cards and are asked to count the number dots as quickly as possible and relay to the examiner the number of dots that they counted. On cards one through six, the dots on the cards are disseminated in no organizational fashion. In cards seven through twelve, the dots on the cards are grouped in such a way that it is easier to count the number of dots quickly. An E-score is tabulated according to the participant's response times and number of errors on the test itself (lower E-scores reflect fewer errors and faster response times). Research has identified that the DCT is an adequate measure of suspect effort, with moderate sensitivity and high specificity of identifying possible malingerers. It has been encouraged that the DCT be utilized in conjunction with other measures when assessing for symptom validity (Strauss, Sherman, & Spreen, 2008). Previous research has suggested a general cut-off score of $\geq$17 for classification of suboptimal effort (Boone et al., 2002).

Repeatable Battery for the Assessment of Neuropsychological Status.

Randolph's RBANS (1998) is a neuropsychological assessment used to test the cognitive status of individuals suffering from neurological diseases or head trauma. One of the core advantages to using the RBANS is its brevity. The RBANS takes approximately 30 minutes to administer, as opposed to other cognitive assessments that require a much longer duration to fully administer.

The RBANS is comprised of five indices (immediate memory, delayed memory, visuospatial ability, language, and attention) and twelve subtests (list learning, story memory, figure copy, line orientation, digit span, symbol digit coding, picture naming, semantic fluency, list recall, list recognition, story recall, and figure recall). All index scores are comprised of two subtests except for the delayed memory domain, which consists of four subtests. The RBANS total score provides an overall outcome statistic for an individual's overall neuropsychological functioning. In addition to the total score, individual subscale scores for immediate memory, visuospatial ability, language, attention, and delayed memory can be calculated. All subtests are given a subtest raw score. Raw scores of subtests within each domain are added and converted to an age-corrected index score. Index scores can also be converted to percentile scores, according to the age-based normative conversions from the RBANS manual.

Immediate Memory. The Immediate Memory domain assesses an individual's ability to remember and recall a small amount of information directly after it has been presented. The immediate memory domain is assessed using two subtests:

List Learning: List Learning consists of a list of 10 unrelated words, read for immediate recall over four trials, with a maximum score of 40. Words used in the List Learning task are considered moderate-high imagery words with relatively low age of acquisition. The

high imagery levels and low age of acquisition of these words is considered helpful in reducing education effects on neuropsychological performance and allows for easing language translation difficulties.

Story Memory: This subtest is comprised of a story with 12 itemized details; the story is read for immediate recall over two trials, for a total maximum score of 24.

Visuospatial Ability. The Visuospatial domain prompts participants to examine, comprehend, and recreate spatial relations. Notably, this domain assesses participants' ability to estimate distance and depth and navigate the surrounding environment. The subtests used to analyze visuospatial/constructional ability are as follows:

Figure Copy: The Figure Copy subtest prompts participants to draw an exact copy of a complex figure comprised of geometric shapes. The subtest itself is considered very similar yet less complex to the Rey-Osterrieth Complex Figure Test (Meyers & Meyers, 1995). The RBANS figure is comprised of 10 components, and a structured simplified scoring guide, which provides for a maximum score of 20.

Line Orientation: On this subtest, participants are presented with an arrangement of 13 lines, beginning at a common point of origin and fanning out across 180 degrees, which serves as the reference figure. Each item consists of two target lines that are shown beneath the reference figure. Subjects must correctly identify which two lines in the reference match the two target lines. Line orientation consists of 10 items, each comprised of two target lines, for a total maximum score of 20.

Delayed Memory. The Delayed Memory domain of the RBANS requires participants to recall information for an extended length of time. These subtests are presented to the participants approximately 20 minutes after initial presentation.

List Learning free recall: Free recall of the words from the List Learning subtest (max = 10).

List Learning Recognition: Yes/No recognition of the words from the List Learning subtest, with 10 foils (max = 20).

Story Memory Free Recall: Free recall of the story from the Story Memory subtest (max=12).

Figure Free Recall: Free recall of the figure from the Figure Copy subtest (max = 20).

Language. The language domain prompts participants to execute communication skills to verbally name and retrieve previously learned semantic information. Two subtests are included in this domain:

Picture Naming: Picture Naming is considered a confrontation-naming task, with 10-line drawings of objects that the participant must name.

Semantic Fluency: Participants are allotted one minute to provide as many examples from a semantic category as possible (i.e., animals, fruits).

Attention. The RBANS attention domain assesses an individual's ability to select a component of information to focus on in subsequent processing and integration tasks. The attention domain prompts the participant to manipulate previously presented material (visual and oral) that has been stored within the individual's short-term memory. This domain includes the following subtests:

Digit Span: Subjects are asked to repeat a series of numbers, with stimulus items increasing in length from 2 digits to 9 digits. The items are presented in order of length (shortest to longest), and the test itself is discontinued when the participant fails all trials

within a given string length. It should be noted that there is no digit span backwards on the RBANS.

Coding: Coding is an assessment of an examinee's processing speed that is very similar to the Coding subtest of the Wechsler Adult Intelligence Scale. Subjects are asked to fill in digits matching with corresponding shapes on a coding key as fast as they can. After practice items are completed, participants have 90 seconds to complete as many items as possible.

Total Scale. The Total Scale is the overall outcome statistic for an individual's overall neuropsychological functioning as comprised by the sum of all the index scores of the RBANS (Attention, Immediate Memory, Delayed Memory, Visuospatial/Constructional, and Language).

RBANS Effort Supplement.

The RES is comprised of one Yes-No Recognition component (Story Memory) and four components in Forced-Choice Recognition format: List Learning, Picture Naming, Figure Copy, and Coding. It should be noted that the RES has never been utilized in previous research. The RES was constructed utilizing the stimuli in RBANS form A, with all non-target stimuli for verbal and nonverbal information derived from alternative forms of the RBANS.

Story Memory Recognition: Participants were administered 12 questions in a yes/no format regarding details from the story that was read to them twice previously in the RBANS Story Memory subtest (max = 12). This subtest was not included in the final RES Total score and was meant to serve as face valid indicator of memory performance.

List Learning Forced Choice: Participants were administered a forced-choice task involving the 10 words from the List Learning subtest. For each item, participants were prompted with two words, one word from the original list and one novel word, and subsequently asked to select the word that appeared on the original list (max =10).

Picture Naming Forced Choice: Participants were administered a forced-choice task involving the 10 objects from the Picture Naming subtest. For each item, participants were prompted with two pictures, one that was presented during the Picture Naming task and one that was not and asked to select the picture they had seen previously. It should be noted that the non-target pictures were pictures from alternate forms of the RBANS. (max =10)

Figure Copy Forced Choice: Participants were administered a forced-choice task involving the Figure Copy subtest. On each item, participants were prompted with two figures, one that was a component of the original figure presented during the Figure Copy task and one that was not and asked to select the component they had seen previously. It should be noted that figures that were presented that were not components of the original complete figure were figure components from alternate forms of the RBANS (max = 12).

Coding Task: Participants were administered a task involving the 9 symbols from the Coding subtest. Participants were asked to select 9 coding symbols from a larger set, which they thought matched those they had seen during the previous administration of the RBANS Coding subtest. Participants were also asked to recall where each symbol was located in the original key; this location task was not included in the final RES Total Score and was meant to serve as a ruse that the measure appeared to be more difficult than it actually was. It should be noted that symbols that were presented that were not

components of the original complete figure were symbols used in alternate forms of the RBANS (max = 9).

The Total RES score was computed by adding all total scores except for RES recognition (Max = 41).

The RBANS Effort Scale (ES).

The RBANS Effort Scale (Novitski et al., 2012) is an existing embedded measure in the RBANS, which is calculated by subtracting delayed free recall scores from recognition and then adding the score from the RBANS digit span subtest. The measure was validated on a population of individuals with amnestic disorders and compared against a mild traumatic brain injury group who had failed a second measure of effort. ES scores less than 12 are considered suspicious for poor effort. However, a limitation of the ES is that it yields significantly negative scores when individuals perform at a high level on measures of delayed free recall and has been cautioned to only be utilized in circumstances where effort during testing is in question.

The RBANS Effort Index (EI).

The RBANS Effort Index (Silverberg, Wertheimer, & Fichtenberg, 2007) is another embedded effort measure in the RBANS. Primary investigators for the EI converted raw scores into a common metric based on their relative infrequency in a derivation sample with true cognitive impairment and then summed these weighted scores to arrive at an index score. More infrequent scores on digit span and list recognition were assigned higher weighted values. The EI is then calculated by using weighted scores on RBANS raw scores of digit span and list recognition and computed by adding the sum of these weighted scores. Thus, a higher EI score indicates worse

effort. The measure was validated on a clinical neurological disorders population and compared against a mild traumatic brain injury group in conjunction with three "suboptimal" groups. EI scores greater than 3 are considered suspicious for suboptimal effort.

Results

Participant Demographic Information

The demographic characteristics of participants in the CNO and SEG are shown in Table 1. In sum, 73 participants were included in analyses for this study. The CNO was comprised of 59 participants (50.9% male) with an average age of approximately 54 years (M = 53.54, SD = 20.23). The majority of participants were Caucasian (66.1%) with an average of approximately 15 years of education (M = 14.89 years, SD = 2.49). In contrast, the SEG included 14 participants (36.7% male) with an average age of approximately 30 years (M = 30.29, SD = 12.02). The majority of participants were Caucasian (28.6%) with an average of approximately 16 years of education (M = 16.42, SD = 1.16)). Of note, the SEG was significantly younger and had more years of education than the CNO group, $p < .05$.

The distributions of outcome measures (e.g. RES, Dot Counting Test, RBANS Effort Scale and RBANS Effort Index) were examined. The RES was found to be negatively skewed. To correct for skewness, logarithmic transformations of RES were used; the RES was then normally distributed. We found that the Dot Counting Test (DCT) and RBANS Effort Index (EI) were positively skewed. We then performed logarithmic transformations of these outcome measures as well, resulting in normal distributions for both outcome measures. The RBANS Effort Scale (ES) was normally distributed and did not require transformations.

Independent Variables of Interest

Descriptive statistics calculated for all experimental groups on RBANS indices are shown in Table 2 (See Appendix A). Additionally, descriptive statistics on relevant outcome measures are shown in Table 3.

RES Reliability Analyses

To analyze the primary aim of assessing the internal consistency of the RES, the Kuder-Richardson Formula 20 (KR-20; Kuder & Richardson, 1937) was utilized. The KR-20 is recommended over the split half method of internal consistency reliability because the split-half method artificially reduces a test's reliability by its division of the analysis into two parts. Additionally, the KR-20 is recommended for a test that is dichotomously scored such as the RES (Cortina, 1993). Our internal consistency analysis revealed that the 41-item RES with picture naming, figure copy, coding, and word list subtests had a reliability coefficient of $\alpha = 0.91$, which is in accordance with acceptable standards for individual decision-making (Nunnally, 1978). Individual reliability analyses for individual subtests were as follows: RES picture naming $\alpha = 0.81$, RES figure copy $\alpha = 0.72$, RES coding $\alpha = 0.65$, RES word list $\alpha = 0.81$. As such, no individual subtest alone demonstrated an acceptable reliability for individual decision-making. Considering the low reliability level of the RES coding, the RES' reliability was assessed once again after extracting the coding subtest, which revealed similar reliability, $\alpha = 0.91$.

RES Validity

To determine convergent validity, partial correlations were used between the RES total score to assess for associations with existing effort measures such as the DCT, ES, and EI controlling for age and years of education. Analyses revealed that the RES was

negatively associated with the EI (r = - 0.83, p <.01) and the DCT (r = -0.52, p <.01). As such, higher scores on the RES were associated with lower scores on the EI and DCT. It was not significantly associated with the ES, p > .05.

Additionally, partial correlations were utilized for all individual RES subtests to examine their associations with the DCT, ES, and EI, again controlling for age and years of education. RES picture naming was negatively associated with the EI (r = -0.86, p < .01) and the DCT (r = -0.53, p < .01) but was not significant associated with the ES, p >.05. RES figure copying was negatively associated with the ES (r = -0.28, p < .01), the EI (r = -0.73, p < .01), and the DCT (r = -0.56, p < .01). The RES word list was negatively associated with the EI (r = -0.85, p < .01) and the DCT (r = -0.52, p < .01) but was not significantly associated with the ES, p > .05. RES coding was significantly associated with the ES (r = -0.35, p < .01) and the EI (r = -0.42, p < .01) but was not significantly associated with the DCT, p > .05, see Table 4.

Because the RES coding subtest demonstrated the weakest reliability ($\alpha = 0.65$) and weakest associations with existing effort detection measures in this study, an additional exploratory analysis was included. After eliminating coding from the RES, the RES was more significantly associated with the EI (r = -.86, p <.01) and the DCT (r = -0.57, p <. 01).

To assess the RES's criterion validity, an Analysis of Covariance (ANCOVA) was utilized to examine how the RES could accurately differentiate between participants groups espousing adequate and suboptimal effort, see Table 5. Because of the possibility that members of the CNO would also provide suboptimal effort on neuropsychological testing, it was decided to recategorize the groups according to the more established DCT

E-score. Previous research has suggested a general cut-off score of $\geq$17 for classification of suboptimal effort (Boone et al., 2002), which was used for our reclassification of variables. As such, we re-classified our data into two groups (good and poor effort according to DCT E score) and compared the two groups on their RES performance. Following this reclassification, 17 participants were left in the suboptimal effort group and 56 participants in the adequate effort group. Using the log-based transformation for the RES to conform with the univariate assumption of normality, the ANCOVA was significant [F (1,69) = 14.87, p < .01, r^2 = .19]. As such, individuals engaging in adequate effort (M = 39.41, SD = 3.01) scored significantly higher on the RES than individuals who engaged in suboptimal effort (M = 34.88, SD = 7.30), p <.01 which suggests that the full RES was a valid indicator of effort detection on neuropsychological testing, see Table 6. Similarly, when RES Coding was extracted from the full RES analyses, the adequate effort group continued to perform significantly better than the suboptimal effort group [F (1,69) = 16.48, p < .01] with an equivalent effect size (r^2 = .19).

Similar analyses were examined on log-based transformations of individual RES subtests. The adequate effort group performed significantly better than the suboptimal effort group on RES Picture Naming [F (1,69) = 38.99, p < .01, r^2 = .39], RES Figure Copy , [F (1,69) = 23.15, p < .01, r^2 = .25], RES List Learning, [F (1,69) = 21.81, p < .01, r^2 = .26], and RES Coding , [F (1,69) = 8.30, p < .01, r^2 = .17], see Table 6. Analyses indicated that RES Picture Naming demonstrated the largest effect among individual subtests (r^2 = .39), whereas coding demonstrated the smallest effect (r^2 = .17).

Additional analyses indicated that individuals diagnosed with mild cognitive impairment or dementia did not perform significantly differently on the RES than other clinical populations, $p > .05$.

Exploratory Analyses

Exploratory analyses included ROC curves examining the sensitivity and specificity of the RES with and without the coding subtest. When examining the full RES, our analyses revealed a cutoff score of 39.50 was associated with moderate sensitivity (sensitivity = 0.73) with moderate specificity (specificity = 0.59), see Figure 1. When excluding the coding subtest, a cut-off of 30.50 (out of a total of 32 points) was associated with moderate sensitivity (sensitivity = .80) and moderate specificity (specificity = .53), see Figure 2.

In comparison to the RES, the EI also had moderate sensitivity (sensitivity = 0.65) and moderate specificity (specificity = 0.68) at a cut-off at 0.5, see Figure 3. The ES had moderate sensitivity (sensitivity = 0.71) and moderate specificity (specificity = 0.54)at a cutoff at 3.50, see Figure 4.

Discussion

This study analyzed the reliability and validity of a performance validity supplement to the RBANS. Data was collected for this study from September 2018 until April of 2019. This study analyzed data from 59 clinical neuropsychology outpatients from Loma Linda University Medical Center's Clinical Neuropsychology Clinic and 14 experimental suboptimal effort actors from Loma Linda University's graduate student population.

The purpose of this study was to build upon existing measures of effort detection within the initial screening process. Researchers have developed embedded effort detection measures in the RBANS, namely the RBANS Effort Scale (2012) and the RBANS Effort Index (2007), which estimate effort through analysis of recall and digit span scores. Both measures have been found to be sensitive but limited in specificity when classifying clinical patients from individuals exhibiting suboptimal effort. As such, this study centered around the validation of a new supplement, which incorporated multiple forced-choice paradigms to create a more well-rounded effort-detection measure.

The primary hypothesis of this study was that the RES would be a reliable and valid measure of effort detection. KR-20 analyses revealed that our hypothesis was confirmed from a reliability standpoint. However, none of the individual subtests alone reached acceptable alpha levels for individual decision-making. RES Coding demonstrated the lowest alpha level and after extracting it from the total RES, the RES had an equivalent alpha level. Validity analyses confirmed our hypothesis that the RES would demonstrate convergent validity with existing measures of effort detection

including the EI and DCT. It should be noted that the RES was not significantly correlated with ES; this may be emblematic of the primary caveat of the ES in that individuals who excel on free recall on the RBANS have significantly negative scores on their ES composite score. Individual subtests demonstrated similarly significant associations with the EI and the DCT, with RES Picture Naming having the strongest correlation among subtests with existing effort measures. RES Coding had the weakest correlation with existing effort measures and after extracting it from the total RES score, the RES' associations with the EI and the DCT slightly improved. The RES also demonstrated construct validity; participants who had been classified into a suboptimal effort group according to DCT E-score performed significantly worse than their counterparts in the similarly classified in the adequate effort group. All individual subtests reflected similar group differences, with RES Picture Naming again demonstrating the strongest effect and RES Coding demonstrating the weakest effect. When extracting RES Coding from the RES Total score, the effect size was equivalent.

Notably, the no significant differences were detected in the Total RES and the RES without Coding scores were between individuals with a memory disorder and other clinical participants. Participants presenting with memory impairment are not expected to perform significantly worse than individuals without memory impairment on the RES, as the RES is not a memory measure. These results demonstrate the RES' strength as an effort detection measure, despite its face validity as memory measure. Given these results, the RES appears to be a true measure of effort, and not a measure of memory function.

Exploratory analyses indicated that the Total RES was moderately sensitive and specific at a cutoff of 39.50; when coding was extracted, the measure was slightly more sensitive and slightly less specific. It should be noted that the RES demonstrated greater sensitivity and specificity than the ES and the EI in this study.

Clinical Implications

There are many exciting clinical implications from this study. Given the RES' observed reliability and validity, our study demonstrates its utility in the initial neuropsychological screening process. The RES' compilation of several effort measures in one supplement may provide clinicians with a more well-rounded analysis and characterization of their patient's effort. The RES' multifactorial detection of effort is also a measure that can be completed in approximately 10-15 minutes with the opportunity to give providers valuable information prior to committing to a full neuropsychological evaluation. This notion may be associated with significant cost reduction while also saving significant time. Additionally, the RES may provide clinicians the opportunity to immediately discuss effort from a multidimensional standpoint when it is in question. Such discussions may conjunctively be useful in determining whether to pursue further neuropsychological testing as well.

Broader implications include the importance of assessing effort in most if not at all clinical contexts. Effort detection options are widely available for neuropsychologists to utilize with most referral questions. Effort detection also validates the nature of neuropsychological services in that clinicians can validate an individual's diagnosis and subsequently provide appropriate recommendations for on their behalf. Effort detection

rules out the possibility of feigned impairment for personal gain and essentially provides credence to the field of neuropsychology.

Limitations

This study is not without limitations. Primarily, a control group would have provided a baseline comparison to both the clinical and suboptimal effort groups. Additionally, the study would have benefitted from a larger sample size in general with larger representation from common neuropsychological referrals. Understanding performance from various neuropsychological perspectives would be helpful in analyzing RES performance trends from a diversity of neuropsychological presentations. Our experimental groups differed significantly in terms of sample size, which may have contributed to the skewness of the original raw data. Additionally, sampling in itself may have been a confounding issue. Specifically, participants in the experimental suboptimal effort group were highly educated, averaging over 16 years of education, and were actively participating in graduate education. Most graduate programs in Loma Linda University emphasize a broad academic curriculum and it is possible that participants may have had prior knowledge of suboptimal effort presentations on neurocognitive testing.

Research Implications and Future Directions

This study leads to several questions regarding future research. It may be useful to consider including a digits backward component to the RES; this may allow for the computation of reliable digits similar to the WAIS-IV and would add yet another component of effort detection to the supplement. Additionally, it is recommended that the RES be analyzed for reliability and validity in other clinical settings as well. The RES

would certainly benefit significantly from replication in other settings and among a wide variety of clinical and demographic populations.

Conclusion

In summary, this study analyzed the reliability and the validity of a novel measure of effort and motivation, the RBANS effort supplement. This study found that the RES was a reliable measure of effort detection. Additionally, the RES exhibited convergent validity with an established embedded effort detection measure from the RBANS (the RBANS Effort Index) and the DCT, which is another well-established independent effort detection measure. The RES demonstrated construct validity in that participants who were classified in the suboptimal effort group according to their performance on the DCT performed significantly worse on the RES than did individuals who had been classified into the adequate effort group. A ROC curve analysis was performed and demonstrated that the RES exhibited moderate sensitivity and specificity at a cut-off score of 39.50. Clinical implications of this study include the potential for screening for effort from a multifactorial approach during the initial neuropsychological screening process, which may significantly reduce costs and save a significant amount of time. Key limitations include a lack of a control group, small sample size, and lack of greater representation from common outpatient referral sources. Future research directions include replication of reliability and validity analyses in a different neuropsychological setting. This study identified the RES as a useful measure in detecting effort, but further research is undoubtedly necessary to fully understand the extent of its utility in a clinical neuropsychological setting.

Additional Information

Funding

Funding for this study was provided by Loma Linda University.

References

Axelrod, B. N., Fichtenberg, N. L., Millis, S. R., & Wertheimer, J. C. (2006). Detecting Incomplete Effort with Digit Span from the Wechsler Adult Intelligence Scale—Third Edition. *The Clinical Neuropsychologist*, *20*(3), 513-523.

Bayley, P. J., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2008). Yes/No Recognition, Forced-choice Recognition, and the Human Hippocampus. *Journal of Cognitive Neuroscience*, *20*(3), 505–512. http://doi.org/10.1162/jocn.2008.20038

Beatty, W. W., Mold, J. W., & Gontkovsky, S. T. (2003). RBANS performance: influences of sex and education. *Journal of Clinical and Experimental Neuropsychology*, *25*(8), 1065-1069.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom Validity Testing: A Critical Review. *The Clinical Neuropsychologist*,*15*(1), 19-45. doi:10.1076/clin.15.1.19.1907

Bigler, E. D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society*, *18*(4), 632-640.

Boone, K. B. (Ed.). (2007). *Assessment of feigned cognitive impairment: a neuropsychological perspective*. New York, NY: The Guilford Press.

Boone, K. B., Lu, P. and Herzberg, D. S. (2002). *The b Test Manual*, Los Angeles, CA: Western Psychological Services.

Boone, K., Lu, P., & Herzberg, D. S. (2002). *The Dot Counting Test.* Los Angeles: Western Psychological Services.

Boone, K. B., & Lu, P. H. (2007). Non-Forced-Effort Measures. In G. J. Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 27-44). New York, NY: Oxford University Press.

Boone, K. B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., ... & Warner-Chacon, K. (2002). Sensitivity and specificity of the Rey Dot Counting Test in patients with suspect effort and various clinical samples. *Archives of Clinical Neuropsychology*, *17*(7), 625-642.

Bouman, Z., Hendriks, M. P., Schmand, B. A., Kessels, R. P., & Aldenkamp, A. P. (2016). Indicators of suboptimal performance embedded in the Wechsler Memory Scale–Fourth Edition (WMS–IV). *Journal of Clinical and Experimental Neuropsychology*,*38*(4), 455-466. doi:10.1080/13803395.2015.1123226

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A, & Kaemmer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Carone, D. A., Iverson, G. L., & Bush, S. S. (2010). A Model to Approaching and Providing Feedback to Patients Regarding Invalid Test Performance in Clinical Neuropsychological Evaluations. *The Clinical Neuropsychologist*,*24*(5), 759-778. doi:10.1080/13854041003712951

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*,*78*(1), 98-104. doi:10.1037//0021-9010.78.1.98

Crighton, A. H., Wygant, D. B., Holt, K. R., & Granacher, R. P. (2015). Embedded Effort Scales in the Repeatable Battery for the Assessment of Neuropsychological

Status: Do They Detect Neurocognitive Malingering? *Archives of Clinical Neuropsychology*,*30*(3), 181-185. doi:10.1093/arclin/acv002

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). California Verbal Learning Test—2nd edition: Manual. San Antonio: The Psychological Corporation.

DenBoer, J.W. & Hall, S. (2007). Neuropsychological test performance of successful brain injury simulators. *The Clinical Neuropsychologist*, *21*(6). 943-955.

Duff, K., Hobson, V.L, Beglinger, L.J, & O'Bryant, S. E. (2010). Diagnostic Accuracy of the RBANS in mild cognitive impairment: limitations on assessing milder impairments. *Archives of Clinical Neuropsychology*, *25*(5), 429-441.

Duff, K., Humphreys Clark, J. D., O'Bryant, S. E., Mold, J. W., Schiffer, R. B., Sutker, P.B. (2008). Utility of the RBANS in detecting cognitive impairment associated with Alzheimer's disease: Sensitivity, specificity, and positive and negative predictive powers. *Archives of Classical Neuropsychology 23*, 603-612.

Duff, K., Spering, C. C., O'Bryant, S. E., Beglinger, L. J., Moser, D. J., Bayless, J. D., . . . Scott, J. G. (2011). The RBANS Effort Index: Base Rates in Geriatric Samples. *Applied Neuropsychology*,*18*(1), 11-17. doi:10.1080/09084282.2010.523354

Etheron, J.L., Bianchini, K.J., Greve, K.W., & Heinly, M.T. (2005). Sensitivity and specificity of reliable digit span in malingered pain-related disability. *Assessment, 12* (2), 130-136.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:10.3758/bf03193146

Frederick, R. I., & Speed, F. M. (2007). On the interpretation of below-chance responding in forced-choice tests. *Assessment, 14*, 3-11.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment,6*(3), 218-224. doi:10.1037//1040-3590.6.3.218

Grimes, D., A., & Schulz, K. F. (2005). Refining clinical diagnosis with likelihood ratios. *The Lancet, 365,* 1500-1505.

Grote, C. L., & Hook, J. N. (2007). Forced-Choice Recognition Tests of Malingering. In G. J. Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 44-80). New York, NY: Oxford University Press.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology*, *143*, 29–36.

Hartman, D. E. (2009). *Applied Neuropsychology*, *16*(WAIS-IV), 85–87. doi:10.1080/09084280802644466

Heilbronner, R., Sweet, J., Morgan, J., Larrabee, G., Millis, S., & Participants, C. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the Neuropsychological Assessment of Effort, Response Bias, and Malingering. *The Clinical Neuropsychologist,23*(7), 1093-1129. doi:10.1080/13854040903155063

Institute of Medicine. (2015). *Psychological Testing in the Service of Disability Discrimination.* The National Academies Press. doi: 10.17226/21704.

Jasinski, L. J., Berry, D. T., Shandera, A. L., & Clark, J. A. (2011). Use of the Wechsler Adult Intelligence Scale Digit Span subtest for malingering detection: A meta-

analytic review. *Journal of Clinical and Experimental Neuropsychology*,*33*(3), 300-314. doi:10.1080/13803395.2010.516743

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*(3), 151-160.

Larrabee, G. J. (2014). Assessment of Performance and Symptom Validity and the Diagnosis of Malingering. In M. W. Parsons & T. A. Hammeke (Eds.), *Clinical Neuropsychology: A Pocket Handbook for Assessment* (3rd ed., pp. 90-113). American Psychological Association.

Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society, 18* (4), 625–630.

Larrabee, G. J. (2008). Aggregation across multiple indicators improves the detection of malingering: relationship to likelihood ratios. *The Clinical Neuropsychologist*, *22* (4), 666-679. DOI: 10.1080/13854040701494987

Larrabee, G. J. (2007). Introduction: Malingering, Research Design, and Base Rates. In G. J. Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 3-14). New York, New York: Oxford University Press.

Larrabee, G. J., & Berry, D. T. (2007). Diagnostic Classification Statistics and Diagnostic Validity of Malingered Assessment. In G. J. Larrabee (Ed.), *Assessment of Malingered Neuropsychological Deficits* (pp. 14-27). New York, New York: Oxford University Press**.**

Larrabee, G.J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist, 17* (3), 410-425.

Martin, P.K, Schroeder, R. W., & Odland, A.P. (2015). Neuropsychologists validity testing beliefs and practices: a survey of North American professionals. *The Clinical Neuropsychologist*, (*29*) 6, 741-776. doi: 10.1080/13854046.2015.1087597

McDougall, R. (1904). Recognition and recall. *Journal of Philosophy, Psychology, and Scientific Methods, 1*, 229-233.

Meyers, J. E., & Meyers, K.R. (1995). Rey Complex Figure Test and Recognition Trial. Odessa, Florida: Psychological Assessment Resources.

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base Rates of Malingering and Symptom Exaggeration. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A),24*(8), 1094-1102. doi:10.1076/jcen.24.8.1094.8379

Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury,18*(10), 975-984. doi:10.1080/02699050410001672350

Morgan, J. E., & Sweet, J. J. (2015). *Neuropsychology of malingering casebook*. New York, NY: Psychology Press.

Nelson, N.W., Boone, K., Dueck, A., Wagener, L., Lu, P., & Grills, C. (2003). Relationships between eight measures of suspect effort. *The Clinical Neuropsychologist, 17* (2), 263-272.

Novitski, J., Steele, S., Karantzoulis, S., & Randolph, C. (2012). The Repeatable Battery for the Assessment of Neuropsychological Status Effort Scale. *Archives of Clinical Neuropsychology,27*, 190-195.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Paul, R., Lane, E. M., Tate, D. F., Heaps, J., Romo, D. M., Akbudak, E., … Conturo, T. E. (2011). Neuroimaging signatures and cognitive correlates of the montreal cognitive assessment screen in a nonclinical elderly sample. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists*, *26*(5), 454–60. doi:10.1093/arclin/acr017

Postman, L. (1963). One trial learning. In C. F. Cofer & B. S. Musgrave (Eds.), Verbal behavior and Learning (pp. 295-332). New York: McGraw-Hill.

Randolph, C. (1998). *Repeatable Battery for the Assessment of Neuropsychological Status*. San Antonio, TX: The Psychological Corporation.

Root, J. C., Robbins, R. N., Chang, L., & Gorp, W. G. (2006). Detection of inadequate effort on the California Verbal Learning Test-Second edition: Forced choice recognition and critical item analysis. *Journal of the International Neuropsychological Society,12*(05). doi:10.1017/s1355617706060838

Silverberg, N. D., Wertheimer, J. C., & Fichtenberg, N. L. (2007). An effort index for the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). *The Clinical Neuropsychologist*, *21*(5), 841-854.

Slick, D.J., Tan, J.E., Strauss, E.H., & Hultsch, D.F. (2004). Detecting malingering: a survey of experts' practices. *Archives of Clinical Neuropsychology*, *19*, 465-473.

Strauss, E., Sherman, E. M., Spreen, O. (2006). *A compendium of neuropsychological tests: administration, norms, and commentary*. Oxford: Oxford University Press.

Teicher, G., & Wagner, M. T. (2004). The Test of Memory Malingering (TOMM): Normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology, 19*, 455–464.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment,9*(3), 260-268. doi:10.1037//1040-3590.9.3.260

Wechsler, D. (2009). Advanced clinical solutions for the WAIS-IV and WMS-IV. San Antonio, TX: Pearson.

Wechsler, D. (2008). Wechsler Adult Intelligence Scale—Fourth Edition. San Antonio, TX: Pearson.

Wechsler, D. (2009). Wechsler Memory Scale–Fourth Edition. Pearson; San Antonio, TX: Pearson.

Wechsler, D. (1981). Wechsler Adult Intelligence Scale—Revised. San Antonio, TX: The Psychological Corporation.