THE RELATIONSHIP BETWEEN SCORES ON THE REPEATABLE BATTERY FOR

THE ASSESSMENT OF NEUROPSYCHOLOGICAL STATUS AND THE TEST OF

MEMORY MALINGERING

By

NEILA RENAE SACKETT

Bachelor of Arts
University of Missouri-Kansas City
Kansas City, MO
2004

Master of Science
Oklahoma State University
Stillwater, OK
2007

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
In partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2011

THE RELATIONSHIP BETWEEN SCORES ON THE REPEATABLE BATTERY FOR

THE ASSESSMENT OF NEUROPSYCHOLOGICAL STATUS AND THE TEST OF

MEMORY MALINGERING

Dissertation Approved:

Dr. Alfred Carlozzi

Dissertation Adviser

Dr. John Romans

Dr. Valerie McGaha

Dr. Steve Harrist

Dr. Mark E. Payton

Outside Committee Member

Dr. Mark E. Payton
Dean of the Graduate College

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Neuropsychologists have long been engaged in determining the validity of patient

responses on the various instruments which constitute the neuropsychological test battery. As

the field of neuropsychology has become increasingly forensically oriented in recent years,

issues of malingering and suspect effort have received increased attention (Boone, 2007; Essig,

Mittenberg, Petersen, Strauman, & Cooper, 2001). Legal referrals related to worker's

compensation, disability compensation, and personal injury damages constitute a substantial

number of cases seen by neuropsychologists (Bianchini, Mathias, & Greve, 2001; Vickery,

Berry, Inman, Harris, & Orey, 2001). The frequency of forensic cases being seen by

neuropsychologists warrants significant attention and caution in assessing for malingering.

Failure to detect malingering can result in unwarranted monetary rewards, evasion of

prosecution, and inappropriate awards of worker's compensation or disability benefits.

Conversely, falsely labeling a person as malingering can have detrimental consequences such as

denial of benefits, misdiagnosis, refusal of services, wrongful prosecution, or malpractice suits

against the clinician.

In order to establish the prevalence of malingering, Mittenberg, Patton, Conyock, and

Condit (2002) surveyed members of the American Board of Clinical Neuropsychology (ABCN).

Thirty-seven percent (144) of those surveyed responded, providing estimates of symptom

exaggeration on a total of 33, 531 cases which involved personal injury, disability, criminal

cases, or medical matters. Clinicians supported their diagnosis of malingering with multiple sources of data including severity of cognitive impairment inconsistent with the condition (65%), pattern of impairment inconsistent with the condition (64%), scores below empirically derived cutoffs on forced-choice tests (57%), implausible self-report of symptoms in the interview (46%), implausible changes in test scores across repeated examinations (45%), and validity scales on objective personality tests (38%). Results indicated that using the aforementioned criteria, clinicians suspected patients of malingering or symptom exaggeration in 29% of personal injury cases, 30% of disability cases, 19% of criminal cases, and 8% of medical cases. These findings are consistent with previous research on base rates of malingering in differing contexts (Green, Rohling, Lees-Haley, & Allen., 2001). These results demonstrate that malingering is a widespread problem of significant import.

Malingering is defined in the DSM-IV-TR as "the intentional production of false or grossly exaggerated physical or psychological symptoms motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" (American Psychiatric Association, 2000, p. 739). The DSM-IV-TR notes that malingering should be strongly suspected if any combination of the following factors are noted: (a) medicolegal context, (b) marked discrepancy between claimed stress or disability and objective findings, (c) lack of cooperation during the diagnostic evaluation as well as in complying with the prescribed regimen for treatment, (d) presence of Antisocial Personality Disorder. Malingering differs from Factitious Disorder in the respect that malingering involves symptom production for purposes of obtaining some external incentive, whereas Factitious Disorder is not motivated by external incentive, but rather by the need to assume the sick role. Further, malingering differs from Conversion Disorder and other Somatoform Disorders in that

the production of symptoms is both intentional and motivated by external incentives. For the purposes of the current study, which is retrospective and utilizes archival data, it will not be possible to perform extensive differential diagnosis. However, whether false symptoms are produced unwittingly or intentionally, the methodology used herein will still serve to differentiate those with genuine neuropsychological impairment from those whose impairment stems from other sources. Further, findings of suspected or probable malingering are only the beginning of the diagnostic process. These findings should be examined in conjunction with neuroimaging, patient history, the clinical interview, and collateral interviews where possible in order to determine the underlying cause of inconsistent or improbable neuropsychological findings (Larabee, 2007).

In addition to these differential diagnoses, there are a number of terms which are used somewhat inconsistently in the literature concerning malingering and response bias. Sub-optimal or incomplete effort, for example, suggests that the individual did not perform to the best of his or her abilities (Rogers, 2008). However, suboptimal effort may or may not be intentional, but rather may be attributed to causes such as fatigue, emotional distress, or co-morbid pathology. Feigning and dissimulation are also often discussed in conjunction with malingering. Feigning is comparable to malingering in that the fabrication or exaggeration of symptoms is deliberate, but it differs in that feigning does not make assumptions concerning the underlying motivation to respond in this fashion (Rogers & Bender, 2003). Thus, this term has less application in forensic contexts where the motivation to obtain external incentives is of primary importance. Similarly, dissimulation is used to define the purposive misrepresentation of symptoms, but again without calling motivation to do so into question (Rogers, 2008). The current study will refer to the construct as symptom exaggeration or suspicion of malingering.

Clinicians typically assess for the potential that a client may be malingering or exaggerating symptoms through two primary methods: analyzing patterns of performance on standard neuropsychological tests and through the use of freestanding symptom validity tests (Larrabee, 2007). Pattern analysis involves examining an individual's performance across tests for patterns which may be indicative of malingering. Examples may include failing easy items while passing difficult ones, uniformly poor performance on all tests, or failing tests of abilities which typically remain intact even in brain-damaged groups (Larrabee, 2007). Analyzing patterns of performance on standard neuropsychological tests can be an efficient means of assessing effort because it does not necessitate the administration of additional tests. Pattern analysis also offers an advantage in the respect that it is far more difficult for an individual to consistently and convincingly feign impairment on an entire battery of tests than on a single measure such as a symptom validity test. This approach also offers the opportunity to examine test results for poor effort retroactively in the event that effort should come into question and the original battery lacked an effort measure (Larrabee, 2007). A disadvantage of this approach is that it is more complex and requires a greater level of skill and investment of time for analysis on the part of the clinician. Also, many neuropsychologists use flexible batteries which frequently change, making it difficult to establish a uniform approach to pattern analysis.

In addition to or in lieu of analyzing patterns of performance on standard measures, clinicians often utilize symptom validity tests (SVTs). SVTs are freestanding measures of response bias and effort which typically employ a two-item, forced-choice technique originally designed by Pankratz and his associates (Pankratz, 1979; Pankratz, 1983). By this method, a stimulus is shown to the test taker, then after a brief delay, two stimuli are presented and the test taker is asked to identify the stimuli seen previously. These tests were originally developed to

assess the veracity of claimed sensory impairment and have since been adapted for use in assessing memory complaints (Pankratz & Binder, 1997). The rationale for the forced-choice model is that, as with the toss of a coin, the test taker would be expected to answer approximately one half of the questions correctly by chance alone (Grote & Hook, 2007). Thus, a patient who scores significantly less than chance is likely intentionally avoiding the correct response. However, researchers have found the less than chance criteria to be of limited clinical utility in assessing effort due to the fact that even subjects who are asked to deliberately feign memory impairment perform at above chance levels on these tests. Despite the fact that the performance of these subjects was greater than chance, it still fell significantly below the performance of both brain-damaged subjects and neurologically-intact subjects. For this reason, researchers have suggested that the use of norm-based criteria may be more useful in the detection of malingering (Tombaugh, 1997). The most common of these criteria include (a) scores occurring 1.3 standard deviations below the average traumatic brain injury score, (b) scores falling below the lowest scores observed in a brain-damaged group, and (c) scores lower than 90% correct. However, none of these criteria have been universally adopted to date (Tombaugh, 1997).

Varying types of SVTs, such as the Portland Digit Recognition Test (PDRT), the Computerized Assessment of Response Bias (CARB), and the Test of Memory Malingering (TOMM) have been designed as supplemental means of detecting suboptimal effort and malingering. These tests employ various types of stimuli. Tests such as Warrington's Recognition Memory Test for Words, the 21-item test, and the Word Memory Test use words as stimuli. Other tests use digit recognition as the stimulus, as seen with the CARB and the PDRT. Recent research has demonstrated that individuals have a high capacity for storing and retrieving

visual information (Tombaugh, 1997). This led to the development of the Test of Memory Malingering (TOMM), which utilizes picture stimuli. The TOMM is currently one of the most widely used assessments for evaluating malingering (Lally, 2003).

Despite the fact that symptom validity tests are specifically designed to detect malingering, they too are imperfect and insubstantial as a stand-alone approach. Considering the risks inherent to falsely labeling an individual as malingering as well as the risks involved with failing to detect malingering on neuropsychological batteries, clinicians generally agree that it is necessary to use numerous sources of data, including multiple neuropsychological tests and/or effort tests in order to accurately and confidently assess client effort (Greffenstein, Baker & Gola, 1996; Larrabee, 2007). The purpose of the current study was to explore a two-pronged approach of this nature. Specifically, the current study evaluated client effort by examining patterns of performance on the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) in conjunction with scores on the Test of Memory Malingering (TOMM).

Statement of the Problem

There is a clear need for consistent use of actuarial indicators of effort in the majority of neuropsychological evaluations. Although numerous symptom validity tests exist for this purpose, a gold standard for the assessment of effort has not yet been established (Larrabee, 2007; Stulemeijer, Andriessen, Brauer, Vos, & Van Der Werf, 2007). There is also no clear agreement as to whether the use of symptom validity tests or examination of patterns of performance on standard neuropsychological measures is the preferable means of assessing effort. Rather, it is suggested that clinicians use multiple measures to establish whether or not symptom exaggeration or malingering may be affecting the neuropsychological evaluation at hand.

Adding multiple effort tests to already extensive neuropsychological test batteries is often not a viable option. It is not uncommon for test batteries to involve 1-2 full days of testing even without the use of multiple effort measures. Adding tests to this process may be taxing for clients as well as clinicians in terms of time, costs incurred, and effort expended. In addition, many settings such as correctional facilities, HIV clinics, substance abuse treatment programs, and hospitals may use brief general-purpose neuropsychological screening batteries as a preliminary step in assessing patient needs. Research has demonstrated the RBANS to be a reliable and valid measure for use in such situations, but it does not contain any empirically validated means of assessing effort.

Two of the RBANS subtests, Digit Span and List Recognition, are adaptations of other established neuropsychological measures. These particular measures have been shown to be relatively insensitive to neuropsychological insult while also being quite sensitive to poor effort or malingering. Thus, researchers have used these two subtests in an attempt to establish and Effort Index for the RBANS. However, the Effort Index has been shown by several researchers to be confounded with variables such as age, education, cognitive ability (Duff et al., in press; Hook, Marquine, & Hoelzle, 2009). The current study examined these subtests as well as all of the remaining subtest and index scores of the RBANS for potential patterns indicative of malingering. This study also used the TOMM, a brief, well-validated symptom validity test, as a preliminary step through which to identify individuals suspected of malingering. The TOMM produces scores for 2-3 trials, depending on whether the optional retention trial is administered. The majority of studies which utilize the TOMM use TOMM Trial 2 scores for the assessment of malingering status, and the current study adhered to that standard as well.

Significance of the Study

The rationale behind a multiple methods approach to malingering assessment is to generate multiple data points in order to allow for convergent validity and greater accuracy in the assessment of malingering or symptom exaggeration. The purpose of the current study was to establish a means of assessing and confirming malingering or symptom exaggeration in the context of a brief neuropsychological screening test, specifically, the RBANS. Results from this study provided information concerning potential patterns indicative of malingering or symptom exaggeration on the RBANS which may be used in conjunction with data from the TOMM. With the knowledge accrued in this study, clinicians have the empirical data necessary to conduct a battery for cognitive screening and assess effort from two vantage points in less than one hour.

Many studies on the construct of malingering have used simulation designs to produce their findings. The current study is of greater clinical utility as it utilized recent archival data from a clinical sample of neuropsychological batteries administered as part of a comprehensive neuropsychological evaluation. In addition, results from this study are generalizable to a wide audience as they are drawn from a heterogeneous sample of patients presenting for outpatient psychological and/or neuropsychological evaluation.

Hypotheses

The following hypotheses were examined:

Hypothesis 1

There would be significant correlations between scores on the RBANS and scores on the TOMM. It was hypothesized that there would be a particularly strong relationship between TOMM Trial 2 scores and scores on the RBANS Digit Span subtest and List Recognition subtest, as tests of this nature have been demonstrated to be sensitive to poor effort.

Hypothesis 2

Variance on TOMM Trial 2 scores would be accounted for by scores on subtests of the RBANS.

It was hypothesized that scores on the Digit Span subtest and List Recognition subtests would

have the greatest predictive utility, given their sensitivity to poor effort.

Hypothesis 3

Group membership (whether individuals were suspected of malingering/symptom exaggeration

versus not suspected of malingering/symptom exaggeration) would vary as a function of scores

on the RBANS (low RBANS scores predicting malingering group membership).

Hypothesis 4

There would be a significant correlation between TOMM Trial 2 scores and RBANS Effort

Index scores.

CHAPTER II

METHOD

Participants

One hundred eleven outpatients served as participants in this study. Subsequent to institutional review board approval, this study examined recent archival data from outpatients referred to a Midwestern neuropsychology clinic for the purpose of undergoing a complete neuropsychological and/or psychological evaluation. Patients seen in this clinic are referred for a variety of neurological disorders including head injury, intracranial neoplasm, cerebral vascular accidents, and dementia. The clinic also sees a large number of forensic cases, including personal injury and worker's compensation cases. Among the sample used for this study, 29% (n=32) acknowledged that they were pursuing litigation at the time of the evaluation. Patients are heterogeneous in age and education. The study of such subjects is of particular importance to the practice of clinical neuropsychology as clinicians are seldom asked to evaluate "normals" with no psychiatric or neuropsychological symptoms. Inclusion criteria were being 18 years of age or older and successful completion of the RBANS as well as the TOMM. Patients who were diagnosed with moderate to severe dementia were excluded in light of the fact that the TOMM has been demonstrated to have questionable validity with such individuals (Teichner & Wagner, 2004; Tombaugh, 1997). Patients with aphasia and patients whose primary language is not English were also excluded. Table I below contains basic demographic information for participants.

Table I

Participant Demographic Information

	N	Mean	Std. Deviation
GENDER			
Male	53		
Female	58		
AGE (years)		51.20	10.95
EDUCATION (years)		13.74	3.07
RACE			
Caucasian	94		
African American	7		
Native American	4		
Hispanic	4		
Asian American	1		
Other			
Unknown	1		

Procedure

Data were obtained from the aforementioned outpatient psychiatry clinic. Participants

presented for evaluation within the last three years and were administered the RBANS and the

TOMM, along with other supplemental neuropsychological measures, as part of a

comprehensive neuropsychological test battery. These assessments were administered and scored

by licensed psychologists and trained clinicians who work under the supervision of Board

certified neuropsychologists.

Measures

Test of Memory Malingering (TOMM) (Tombaugh, 1996). The TOMM is a forced-choice

symptom validity test used to test for suboptimal effort. The test consists of two learning trials,

each of which is followed by a test phase, as well as an optional retention trial. During each

learning trial, the patient is presented with 50 black-and-white line-drawn pictures (targets) for 3

seconds with a 1-second interval between pictures (Tombaugh, 2002). The same 50 pictures are

used for each learning trial, though the order of presentation is varied. During the test phase, each target is presented with another line drawing not previously presented (distracter) and the subject is asked to select the correct picture. Following each individual response, the examiner provides feedback concerning the correctness of the response. This is done for each trial as well as the retention trial, if administered. The rationale for providing feedback is that (a) if test takers are malingering, feedback allows them to track their performance and adjust accordingly and (b) feedback provides another opportunity for those who are genuinely motivated to learn the targets (Tombaugh, 1997). One point is given for each correct answer, with a maximum score of 50 per trial. The manual suggests that the optional retention trial be administered 15 minutes after Trial 2 if the examinee obtains a score below 45 on Trial 2. The current study will not utilize the retention trial, as it is seldom administered at the clinic which provided data for this study. Scores on the TOMM have been demonstrated to be relatively unaffected by age, education, and moderate cognitive impairment. Based on extensive empirical data, a criterion score of 45/50 on Trial 2 is suggested for identifying patients suspected of malingering. This criterion score should be used as a guideline in assessing effort and not as a stand-alone assessment of malingering (Tombaugh, 1997).

Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)
(Randolph, 1998). Although this measure was originally developed for the assessment of dementia in elderly patients, its utility as a screening for neurocognitive functioning in a much broader population was realized early on and standardization was modified to include normative data on 540 individuals ages 20-89 (Randolph, 1998). The test is available in two equivalent forms to allow for tracking abilities over time while controlling for practice effects (Randolph, 1998). The test yields an aggregate measure of overall performance (Total Scale) as well as five

12

index scores which are derived from scores on twelve subtests: Immediate Memory (subtests are List Learning and Story Memory), Visuospatial/Constructional index (Figure Copy, Line Orientation), Language index (Picture Naming, Semantic Fluency), Attention index (Digit Span, Coding), and a Delayed Memory index (List Recall, List Recognition, Story Recall, and Figure Recall) (Randolph, 1998). Each index is recorded as a standard score with a mean of 100 and a standard deviation of 15 (Lezak, 2004). Index scores are also summed and converted into a Total Scale Score. Scores are corrected for age but not for education (Randolph, 1998). However, the manual does provide means and standard deviations for the five indexes which are broken down according to age and education (Randolph, 1998). RBANS subtests, indices and total score served as independent variables in this study. The battery indices and subtests are as follows:

(1) Immediate Memory Index. This index consists of 2 subtests: (a) List Learning: the test administrator reads a list of 10 semantically unrelated, early age-of-acquisition, high imagery, phonemically unique words and asks the examinee to repeat back as many as he or she can remember, (b) Story Memory: the test administrator reads a series of two stories aloud. After each story, the examinee is asked to repeat back as much of the story as they can. Recall is scored using a verbatim criterion.

(2) Visuospatial/Constructional Index. This index consists of 2 subtests: (a) Figure Copy: the examinee is given a copy of a geometric figure comprised of ten parts and asked to copy it as exactly as possible. The reproduction is scored for accuracy and placement of figure elements, (b) Line Orientation: this task entails a radiating array of 13 lines spanning 180 degrees. Below the array are 2 target lines which are identical to two of the lines from the array. The examinee must identify the matching lines.

13

(3) Language Index. This index consists of 2 subtests: (a) Picture Naming: this task presents ten line drawings and asks the examinee to name them. Semantic clues may be given if the object is clearly misperceived (e.g., for "well" the clue would be "you get water from it", (b) Semantic Fluency: the examinee is asked to produce the names of as many items in a given category (e.g., fruits and vegetables) as possible within 60 seconds.

(4) Attention Index. This index consists of 2 subtests: (a) Digit Span: analogous to digits forward on the WAIS. There are two strings of digits in each item, with lengths increasing from 2 to 9 digits. The second string is only administered if the first was failed, (b) Coding: comparable to the Digit Symbol subtest of the WAIS-R. In this case, the numbers are copied instead of the symbols in order to avoid possible detrimental effects of constructional apraxia.

(5) Delayed Memory Index. This index consists of four subtests: (a) List Recall: after a delay, the examinee is asked to freely recall (without prompting) as many words as possible from the earlier list learning task, (b) List Recognition: the test administrator reads a list of words and the examinee indicates whether or not they were on the list learning task at the beginning of the test by indicating yes or no, (c) Story Recall: after a delay during which other tests are administered, the examinee is asked to recall as many details of the story they heard earlier as possible, (d) Figure Recall: after a delay during which other tests are administered, the examinee is asked to draw as much of the complex figure they drew earlier as they can recall (Randolph, 1998).

Design and Statistical Analyses

The current study utilized correlation analyses, multiple regression, and binary logistic regression. Descriptive statistics were generated for all subjects to include mean scores and standard deviations on Trial 1 and Trial 2 of the TOMM as well as on Total Score, all five index scores, and all 12 subtests scores for the RBANS. The magnitude and direction of the relationships between performance on Trial 2 of the TOMM and RBANS Total Score, Indices, and subtests was examined through computation of a series of bivariate Pearson linear correlation coefficients. Significance was evaluated at the .05 level. The second analysis utilized multiple backwards stepwise regression to determine which RBANS scores accounted for the greatest variance in TOMM scores. This analysis was conducted in two parts, one examining the predictive utility of subtest scores and another examining that of index scores. The rationale for this two-part analysis was to maintain independence of variables, as index scores are directly derived from subtest scores.

Following this analysis, subjects were divided into two groups according to their scores on Trial 2 of the TOMM in order to allow for direct comparison between those subjects suspected of malingering and those not suspected. Those scoring below the established cutoff of 45 were placed in the "suspected of malingering group" (SM) and those scoring 45 or higher were placed in a "not suspected of malingering group" (NSM). Binary logistic regression analysis was utilized to determine how group membership varied as a function of scores on the RBANS. This analysis was chosen due to the binary, categorical nature of group membership. A backwards stepwise approach was utilized due to the exploratory nature of the research question being addressed and to produce a model which included only the most significant predictors of malingering status. Again, separate analyses were used to examine the predictive utility of

subtest scores and index scores. At each step, the backwards stepwise regression model removed an RBANS variable that did not show a significant relationship with the dependent variable of group membership. In order to explore the difference between the SM and NSM groups, Z-scores were produced for each group on each of the 12 subtests. In addition, percentile scores for the five index scores were generated for the two groups and this information was presented in table format. A final analysis examined the relationship between the RBANS Effort Index, a relatively newly proposed means of assessing effort on the RBANS, and the TOMM, a well established measure of effort on neuropsychological measures.

CHAPTER III

RESULTS

The purpose of the current study was to evaluate client effort by examining patterns of performance on the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) in conjunction with scores on the Test of Memory Malingering (TOMM). Table 2 below contains basic descriptive statistics for the dependent and independent variables.

Table 2.
Descriptive Statistics for Dependent and Independent Variables

Variable	N	Range	Mean	Std.
Tomm Trial 2	111	18-50	47.63	5.89
RBANS Total Scale Score	111	45-120	84.39	17.84
Immediate Memory Index	111	40-129	83.16	20.36
Visuospat/Construc Index	111	50-126	95.87	17.81
Language Index	111	11-120	89.21	15.39
Attention Index	111	40-128	86.59	20.77
Delayed Memory Index	111	40-119	83.78	19.38
List Learning	111	6-38	23.55	6.20
Story Memory	111	3-24	14.31	5.15
Figure Copy	111	2-20	17.90	2.74
Line Orientation	111	5-20	15.50	3.51
Picture Naming	111	6-10	9.52	.88
Semantic Naming	111	4-31	17.41	5.41
Digit Span	111	1-16	9.67	2.86
Coding	111	0-78	39.29	13.23
List Recall	111	0-10	4.27	2.86
List Recognition	111	9-20	18.33	2.15
Story Recall	111	0-12	6.69	3.46
Figure Recall	111	0-20	10.97	5.21

The first hypothesis for this study stated that there would be significant correlations between TOMM Trial 2 scores and scores on the RBANS, particularly the Digit Span and List Recognition subtests of the RBANS. In order to test this hypothesis, correlation analyses were conducted to determine the relationship between scores on the RBANS and suspicion of malingering as measured by TOMM Trial 2 scores. Table 2 depicts results from the analysis, which revealed significant relationships between TOMM Trial 2 scores and several variables on the RBANS. Digit Span demonstrated a significant but relatively small correlation with TOMM scores (r=.265, p=.0050), while List Recognition demonstrated a significant and moderate

correlation (r=.305, p=.0011). Several other variables were found to be significantly correlated

with TOMM scores as well and included Total Scale Score (r=.414, p<.0001), Immediate

Memory Index (r=.331, p=.0004), Visuospatial/Constructional Index (r=.267, p=.0046),

Language Index (r=.228, p=.0159), Attention Index (r=.355, p=.0001), Delayed Memory Index

(r=.419, p<.0001), List Learning (r=.373, p=<.0001), Story Memory (r=.280, p=.0029), Line

Orientation (r=.274, p=.0036), Semantic Fluency (r=.310, p=.0009), Coding (r=.381, p=<.0001),

List Recall (.305, p=.0011), Story Recall (r=.339, p=.0003), and Figure Recall (r=.315, p=.0007).

There was not a significant relationship between TOMM Trial 2 scores and Figure Copy or

Picture Naming. Data for the correlation analysis may be found in Table 3 below.

Table 3.
Correlations of RBANS variables with TOMM Trial 2 scores

RBANS Variable	TOMM	Significance
Total Scale Score	.414	<.0001*
Immediate Memory Index	.331	.0004*
Visuospatial/Constructional Index	.267	.0046*
Language Index	.228	.0159*
Attention Index	.355	.0001*
Delayed Memory	.419	<.0001*
List Learning	.373	.0001*
Story Memory	.280	.0029*
Figure Copy	.184	.0528
Line Orientation	.274	.0036*
Picture Naming	.129	.1789
Semantic Fluency	.310	.0009*
Digit Span	.265	.0050*
Coding	.381	<.0001*
List Recall	.305	.0011*
List Recognition	.339	.0003*
Story Recall	.275	.0035*
Figure Recall	.315	.0007*

*Correlation is significant at the .05 level

The next analysis consisted of backwards stepwise multiple regression analysis in order to determine which RBANS scores accounted for the greatest variance in TOMM Trial 2 scores. Specifically, it was hypothesized that RBANS Digit Span and List Recognition subtest scores would account for the greatest variance in TOMM scores, given that tests of this nature are known to be sensitive to poor effort. An initial regression analysis was conducted to examine the ability of scores on the twelve subtests of the RBANS to predict variance in TOMM Trial 2 scores. Significance was evaluated at the .10 level in order to avoid extracting variables which should remain part of the model. Results of the backwards stepwise regression indicated that three predictor variables accounted for 21% of the variance in TOMM scores (R^2=.207, F (3,110)=9.28, p<.0001). It was found that the Digit Span (β=.331, p=.08), List Recognition (β=.563, p=.03), and Coding (β=.110, p=.01) subtests significantly predicted TOMM scores.

Table 4.
Backwards Stepwise Regression Model of Variance in TOMM Predicted by RBANS Subtest Scores

RBANS Variable	Step	R^2	F	P
All variables		.239		
Story Memory	-1	.239	.08	Ns
List Recall	-2	.238	.10	Ns
Picture Naming	-3	.236	.17	Ns
Line Orientation	-4	.233	.43	Ns
Figure Copy	-5	.231	.27	Ns
List Learning	-6	.227	.52	Ns
Story Recall	-7	.225	.37	Ns
Semantic Fluency	-8	.217	.99	Ns
Figure Recall	-9	.207	1.44	Ns
Digit Span	Final		3.22	.0758*
Coding	---		6.28	.0137*
List Recognition	---		4.65	.0332*

*Note: Variables in bold type were retained. Variables with (-) steps were ejected from the model. Significance was evaluated at the .10 level.

Given that the index scores are directly derived from the subtest scores, a separate regression analysis was conducted in order to determine the extent to which variance in TOMM scores could be attributed to the various index scores. Results of the backwards stepwise regression indicated that two predictor variables accounted for 20% of the variance in TOMM scores (R^2=.204, F (2,110)=13.84, p<.0001). It was found that the Attention Index (β=..055, p=.0524), and the Delayed Memory Index (β=.098, p=.0015), significantly predicted TOMM scores. These results were somewhat expected, given that the Coding and Digit Span subtests are the two subtests comprising the Attention Index and that the List Recognition subtest is one of four subtests which are used to generate the Delayed Memory Index.

Table 5.
Backwards Stepwise Regression Model of Variance in TOMM Predicted by RBANS Index Scores

RBANS Variable	Step	R^2	F	P
All variables		.206		
Language Index	-1	.206	.00	ns
Visuospatial/Constructional Index	-2	.206	.02	ns
Immediate Memory Index	-3	.204	.27	ns
Attention Index	Final		3.85	.0524*
Delayed Memory Index	---		10.59	.0015*

*Note: Variables in bold type were retained. Variables with (-) steps were ejected from the model. Significance was evaluated at the .10 level.

The next analysis utilized binary backward stepwise logistic regression to establish a model which best predicted group membership (whether individuals were suspected of malingering or not suspected of malingering on the basis of their TOMM scores) from among the RBANS subtest scores (independent variables). It was hypothesized that scores on the RBANS, particularly the subtest scores for Digit Span and List Recognition, would account for a statistically significant amount of variation in group membership. The best resulting model

included a combination of two predictor variables and the equation demonstrated statistical significance (χ^2= 14.27; df=2; p=.0008). The best two predictor variables were found to be scores on the RBANS List Learning subtest (*B*= -.12, Wald χ^2= 4.81, df=1, p=.03, d=.38; odds ratio [OR]=.89, 95% confidence interval [CI]=.80-.99) and Coding (*B*= -.06, Wald χ^2= 4.19, df=1, p=.04, d=.63; OR=.95, CI=.90-1.00). This model correctly classified 91 of 93 subjects not suspected of malingering (98% specificity), but correctly classified only 4 of 18 subjects who were malingering, yielding a sensitivity of 22%. Ideally, this model should have been tested with an independent data set. However, no such data set was available. As such, it should be noted that this assessment of the model may be somewhat flawed given that the same data were used to predict as were used to create the model.

Table 6.
Final logistic regression model of variance in group membership predicted by RBANS subtest scores

RBANS Variable	*B*	SE	Wald	P
List Learning	-.12	.05	4.81	.0283*
Coding	-.06	.03	4.19	.0407*

*p<.05

Again, a separate regression analysis was conducted in order to determine the extent to which variance in group membership could be attributed to the various index scores. The best resulting model included a combination of two predictor variables and the equation demonstrated statistical significance (χ^2= 15.44; df=2; p=.0004). The best two predictor variables were found to be scores on the RBANS Attention Index (*B*= -.03, Wald χ^2= 3.10, df=1, p=.08, d=.34; odds ratio [OR]= .97, 95% confidence interval [CI]=.94-1.00) and the Delayed Memory Index (*B*= -.04, Wald χ^2= 6.56, df=1, p=.01, d=.49; OR=.96, CI=.93-1.00). Comparable to the results found with subtests scores, this model yielded a specificity of 97% and 22% sensitivity. While this

model provides some utility for accurately classifying subjects who are not malingering, it fails

to identify many individuals who malinger poor performance.

Table 7.
Final logistic regression model of variance in group membership predicted by RBANS index scores

RBANS Variable	B	SE	Wald	p
Attention Index	-.03	.02	3.10	.0784*
Delayed Memory Index	-.04	.02	6.56	.0104*

*p<.05

Finally, correlation analyses were conducted to determine whether a significant

relationship existed between TOMM Trial 2 scores and Effort Index scores. It was hypothesized

that there would be a significant correlation between the two variables, given that each is

intended to detect poor effort on neuropsychological examinations.   Results indicated that these

two variables are indeed significantly negatively correlated (r= -.415, p=<.0001). This supports

the hypothesized relationship between the variables, as low TOMM scores (<45) are indicative

of potential malingering, whereas high scores on the Effort Index (≥1) are considered indicative

of the same. Table 8 below provides group comparisons on effort measures and tests related to

effort for the two respective groups.

Table 8.
Effort scores for the two groups

	Suspected of malingering	Not suspected of malingering
TOMM Trial 1	31 (7.28)	47 (3.49)
TOMM Trial 2	36 (7.70)	50 (.60)
RBANS Digit Span (raw)	8 (3.50)	10 (2.64)
RBANS List Recognition (raw)	17 (2.35)	19 (1.98)
RBANS Effort Index	1	0

*Figures represent Mean (SD)

Variables accounting for group membership

In the interest of further exploring the differences between scores of those falling in the suspected of malingering group versus those not suspected, mean scores for all of the RBANS variables were calculated for each respective group and analyzed in order to produce Z scores. It should be noted that given the small sample size of the malingering group (N=18), these comparisons have poor statistical power and these results should be interpreted with caution. The table below illustrates differences between the two groups on the twelve subtests of the RBANS. Among the malingering group, scores on eleven of the twelve subtests fell in the low average, borderline, and impaired ranges. By contrast, those in the non-malingering group scored in the average to high average range on eleven of the twelve subtests, with scores falling into the low average range on only one subtest (story recall). On average, the non-malingering group scored approximately .38 standard deviations below the mean on subscales of the RBANS, whereas those in the malingering group scored an average of 1.6 standard deviations below the mean. These results are consistent with the findings of Constantinou et al. (2005), which found that subjects putting forth optimal effort tended to score approximately one-third of a standard deviation below the mean, while those demonstrating suboptimal effort typically scored more than one standard deviation below the mean.

Table 9.
RBANS subtest Z-scores for the malingering versus non-malingering groups

RBANS Subtest score	Malingering group (n=18)	Non-malingering group (n=93)
List Learning	-1.94 (borderline)	-0.53 (average)
Story Memory	-1.76 (borderline)	-0.68 (average)
Figure Copy	-1.21 (low average)	-0.14 (average)
Line Orientation	-1.03 (low average)	-0.14 (average)
Picture Naming	-0.44 (average)	.67 (high average)
Semantic Fluency	-1.40 (borderline)	-0.60 (average)
Digit Span	-1.04 (low average)	-0.21 (average)
Coding	-1.94 (borderline)	-0.60 (average)
List Recall	-1.67 (borderline)	-0.48 (average)
List Recognition	-2.80 (impaired)	-0.50 (average)
Story Recall	-1.86 (borderline)	-0.95 (low average)
Figure Recall	-1.97 (impaired)	-0.45 (average)

Similar group differences were observed on the five indices of the RBANS. For index scores, the scoring program generates percentile scores in lieu of Z-scores. Whereas those in the non-malingering group typically fell in the 21st—37th percentile (low aveage—average range), subjects in the malingering group scores fell in the $<1^{st}$ percentile—16th percentile (impaired—low average range). These results demonstrate markedly different profiles between the two groups. This analysis should be replicated with a larger sample size in order to determine whether the same distinct differences are observed, which may allow for the establishment of cutoff scores to assess for malingering on the various subscales of the RBANS.

Table 10.

RBANS Index Score Percentiles for the Malingering Versus Non-malingering Groups

RBANS Index score	Malingering group (n=18)	Non-malingering group (n=93)
Immediate Memory Index	1.00 (impaired)	20.68 (low average)
Visuospatial/Constructional Index	8.23 (borderline)	39.60 (average)
Language Index	15.73 (low average)	36.54 (average)
Attention Index	1.81 (borderline)	36.54 (average)
Delayed Memory Index	.40 (impaired)	36.54 (average)

CHAPTER IV

DISCUSSION

The purpose of this study was to establish a means of assessing and confirming

malingering or symptom exaggeration in the context of a brief neuropsychological screening test,

specifically, the RBANS. Correlation analyses were conducted in order to explore whether

scores on the RBANS rose or fell in accordance with subject effort as defined by the TOMM.

The majority of scores on the RBANS were found to be significantly correlated with effort as

determined by the TOMM. Additionally, subtest and index scores on the RBANS were

examined to determine their ability to account for variation in TOMM scores. Consistent with

the hypothesized relationship, Digit Span and List Recognition subtest scores were found to be

predictive of TOMM scores. These two scores, in conjunction with Coding subtest scores, were

found to account for 21% of variance in TOMM Scores. The predictive utility of RBANS index

scores were examined in a separate regression equation and results indicated that the Attention

Index and Delayed Memory Index accounted for 20% of variance in TOMM scores. A third

analysis explored the extent to which RBANS subtest and index scores were predictive of group

status (i.e., suspected of malingering versus not suspected of malingering.) Two of the RBANS

index scores, the Attention Index and Delayed Memory Index, were found to be predictive of

group membership. This study also examined the relationship between the Effort Index, a

relatively recently proposed means for assessing effort on the RBANS as compared with the

TOMM, a well established measure of malingering on neuropsychological examinations. Results

27

showed the RBANS Effort Index and TOMM Trial 2 scores to be significantly correlated. Overall, while results from this study did not fully support all of the initial hypotheses, several significant relationships between variables on the RBANS and the TOMM were established.

The relation between neuropsychological scores and malingering

This study generated several significant findings. Exploratory correlation analyses indicated that sixteen of the 18 RBANS variables were significantly related to scores on the TOMM. Specifically, the Total Score, Immediate Memory Index, Visuospatial/Constructional Index, Language Index, Attention Index, Delayed Memory Index, as well as subtest scores for List Learning, Story Memory, Line Orientation, Semantic Fluency, Digit Span, Coding, List Recall, List Recognition, Story Recall and Figure Recall were all shown to be related to malingering scores. That is, as the likelihood that subjects are malingering poor performance increases, their scores on the aforementioned subtests and indices decrease. These results suggest that individuals who attempt to malinger poor performance on neuropsychological exams may score lower than their non-malingering counterparts on the majority of domains. This type of nearly uniformly poor performance, versus scoring poorly in specific neurocognitive domains, is a well established pattern among those attempting to malinger poor performance (Constantinou, 2005; Larabee, 2007; O'Bryant, Duff, Fisher & McCaffrey, 2004; Rogers & Bender, 2003). Many individuals who seek to appear impaired on such measures are unaware of which cognitive abilities are unlikely to be compromised in patients with genuine neurocognitive impairment. Thus, subjects seeking to malinger poor performance will often inadvertently perform poorly on tests of abilities which typically remain intact among brain damaged individuals (Rogers & Bender, 2003).

Utility of RBANS index and subtest scores to predict malingering scores

The primary aim of this study was to identify patterns of scoring on the RBANS which might be indicative of malingering. RBANS subtests Digit Span and List Recognition are the RBANS' version of two tests which have are well known to be relatively insensitive to cognitive impairment, while also being quite sensitive to poor effort (Axelrod, Fichtenberg, Millis, & Wertheimer, 2006; Griffenstein, Baker, & Gola, 1994; Iverson & Franzen, 1996; Mathias, Greve, Bianchini & Crouch, 2002). For this reason, it was hypothesized that these scores would account for the greatest amount of variance among malingering scores. Consistent with the hypothesized relationship, these subtests were among the RBANS scores found to be most predictive of malingering scores. These two scores, in conjunction with Coding subtest scores, accounted for 21% of the variance in TOMM scores. Although the predictive utility of Coding was not specifically hypothesized, this finding is not surprising given that Coding, along with Digit Span are the two measures which comprise the RBANS Attention Index. As seen with Digit Span, basic attention measures such as this one are relatively insensitive to even severe neurocognitive insult, and of particular utility with regards to assessing effort and/or malingering (Larrabee, 2007).

This study also examined the predictive utility of index scores in predicting TOMM scores. Certain RBANS subtest scores have a limited range of scores and skewed distributions in normal subjects, warranting caution in interpreting these scores. Conversely, index scores are based on two or more subtest scores and may represent a more thorough sampling of the neurocognitive domains they address. In keeping with the aforementioned findings related to subtest scores, the RBANS Attention Index and Delayed Memory Index were found to predict 20% of the variance in TOMM scores. Again, attention is a domain known to be sensitive to

29

malingering and relatively insensitive to neurological insult. Similarly, the predictive utility of

the Delayed Memory Index is consistent with previous studies which have found that despite the

tendency for recognition memory (which is one means of assessing delayed memory) to remain

intact in the face of neurological insult, subjects who attempt to malinger poor performance

frequently score below individuals with traumatic brain injury. Although the TOMM has been

found to be indicative of malingering in other neurocognitive domains in addition to memory, it

was expected that individuals failing the Test of Memory Malingering would also score poorly

on measures of delayed memory. In fact, the majority of the indices which measure malingering

in the field of neuropsychology have been established via memory tasks, in large part due to the

fact that memory complaints are commonly observed among individuals seeking secondary gain

(Binder & Rohling, 1996; Paniak et al., 2002).

Utility of RBANS index and subtest scores to predict group membership (suspected of

malingering or not suspected of malingering)

For this analysis, established cutoffs on the TOMM were used to classify subjects as

suspected of malingering or not suspected of malingering. Individuals scoring below 45 were

grouped as suspected of malingering and those scoring 45 or greater were grouped as not

suspected of malingering. Results from this analysis should be interpreted with caution, as only

18 of the 111 subjects were classified as suspected of malingering, giving the analysis poor

statistical power. That said, among the 12 subtests, results indicated that the List Learning

subtest and the Coding subtests best predicted group membership. It was expected that these

results would coincide more closely with the analysis examining subtest scores as predictive of

TOMM scores as a continuous variable, but only minimal overlap was found (the significance of

Coding). The poor statistical power of the analysis may have contributed in failing to find

statistical significance among the expected predictors. The two significant predictors, Coding and List Learning, are representative of the neurocognitive domains of attention (Coding) and immediate memory (List Learning). As mentioned previously, tests of attention are typically good indicators of poor effort. Similarly, individuals feigning neurocognitive injury frequently claim memory impairment.

When the binary logistic regression equation was run with the RBANS index scores, results indicated that the Attention Index and Delayed Memory Index were the best predictors of group membership. These results were consistent with the results of the analysis examining RBANS scores as predictive of TOMM scores as a continuous variable. As mentioned previously, RBANS index scores have greater interpretive value as they are a compilation of multiple subtests assessing the same domains.

The relationship between scores on the TOMM and the Effort Index

The present study found a significant relationship between malingering classification as determined by the TOMM and the Effort Index, a proposed means of evaluating effort on the RBANS which has demonstrated mixed results to date. In previous studies, the Effort Index has shown utility in distinguishing between malingerers and non-malingerers (Silverberg et al., 2007), but has raised concerns that the index may inaccurately classify individuals as malingering due to effects of education, age, and level of cognitive functioning (Duff et al., in press). The EI uses uncorrected raw scores, thus increasing the likelihood that older adults, who often scores lower than their younger counterparts, will be inaccurately labeled as putting forth suspect effort. The EI has also been shown to be significantly correlated with cognitive functioning, therefore confounding the issue of effort with cognitive abilities (Duff et al., in

press). An additional concern with regards to the EI is that of education effects. Duff et al. (in press) found significant education effects in 3 of the 4 clinical samples used in their study.

Limitations

The results of this study should be examined in the context of several identified limitations. First, this study did not utilize experimental design and random selection of participants was not possible. The sample for this study was derived from archival data maintained from outpatient neuropsychological assessments administered at the neuropsychology clinic of a major Midwestern medical center. This sample is representative of individuals who presented for neuropsychological examination for purposes ranging from diagnostic clarification to litigation. There may be fundamental differences between individuals who present for such examination as opposed to those who are not motivated to do so.

In addition, this study is limited due to the inability to definitely identify subjects as malingering or not malingering. In this study, participants were classified as suspected of malingering on the basis of their scores on trial 2 of the TOMM. Although the TOMM has been widely researched and found to be of sound reliability and validity, the possibility that subjects were erroneously classified as suspected or not suspected of malingering may exist. This study also did not seek to ascertain what sort of secondary gain may have influenced the subjects' effort scores.

In addition, the results of this study were limited in the respect that only 18 subjects were classified as suspected of malingering versus 93 subjects who were classified as not suspected of malingering. As previously mentioned, although this study did make comparisons between the two groups, such comparisons suffer from poor statistical power. In addition, there may have

been results which failed to reach significance due to the low number of subjects in the smaller of the two groups.

Future research

The results of this study suggest that examining patterns of performance on the RBANS in conjunction with a measure of effort such as the TOMM can provide multiple data points from which to draw conclusions concerning the potential for malingering. Although this study did not produce a predictive model with high sensitivity and specificity, several items emerged as being indicative of poor effort. The Attention Index and Delayed Memory Index were consistently found to be predictive of both TOMM scores and group membership. Although subtest scores were somewhat less consistent predictors, future research may provide further insight concerning the predictive utility of subtests. This study also yielded results which suggest that establishing cutoff scores for the subtests and index scales of the RBANS may be of considerable clinical use in the assessment of malingering. Consistent with findings in comparable studies with other measures, results from this study suggest that subjects suspected of malingering consistently score in excess of one standard deviation below the mean, whereas their non-malingering counterparts more often score approximately one-third of a standard deviation below the mean. In summary, findings from this study suggest that the establishment of cutoff scores for the various subtests and index scores of the RBANS as well as close examination of scores on the Attention Index and Delayed Memory Index may prove useful in the assessment of malingering. Future studies should seek to reproduce the results of this study with a larger sample size.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th Ed. Text Revision)*. Washington, DC: Author.

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*, 513-523.

Barker, M.D., Horner, M.D., & Bachman, D.L. (2010). Embedded indices of effort in the repeatable battery for the assessment of neuropsychological status (RBANS) in a geriatric sample. *The Clinical Neuropsychologist, 24*, 1064-1077.

Bauer, L., & McCaffery, R.J. (2006). Coverage of the Test of Memory Malingering, Victoria Symptom Validity Test, and Word Memory Test on the Internet: Is test security threatened? *Archives of Clinical Neuropsychology, 21*, 121-126.

Bender, S.D. & Rogers, R. (2004). Detection of neurocognitive feigning: development of a multi-strategy assessment. *Archives of Clinical Neuropsychology, 19,* 49-60.

Bianchini, K.J., Mathias, C.W., & Greve, K.W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist, 15,* 19-45.

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

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

Constantinou, M., Bauer, L., Ashendorf, L, Fisher, J.M., & McCaffrey, R.J. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests? *Archives of Clinical Neuropsychology, 20,* 191-198.

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

Duff, K., Beglinger, L., Schoenberg, M., Patton, D., Mold, J., Scott, J., et al. (2005). Test-retest stability and practice effects of the RBANS in a community dwelling elderly sample. *Journal of Clinical and Experimental Neuropsychology, 27*(5), 565-575.

Duff, K., Patton, D., Schoenberg, M., Mold, J., Scott, J., & Adams, R. (2003). Age- and education-corrected independent normative data for the RBANS in a community dwelling elderly sample. *The Clinical Neuropsychologist, 17*(3), 351-366.

Duff, K., Spering, C., O'Bryant, S., Beglinger, L., Moser, D., Bayless, J., Culp, K., Mold, J., Adams, R. & Scott, J. (in press). The RBANS effort index: Base rates in geriatric samples. *Applied Neuropsychology.*

Essig, S.M., Mittenberg, W., Petersen, R.S, Strauman, S., & Cooper, J.T. (2001). Practices in forensic neuropsychology: Perspectives of neuropsychologists and trial attorneys. *Archives of Clinical Neuropsychology, 16,* 271-279.

Gansler, D.A. & Moczynski, N.P. (1998). Five validation experiments of the Test of Memory Malingering (TOMM). *Psychological Assessment, 10*, 10-20.

Gervais, R., Rohling, M., Green, P., & Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. *Archives of Clinical Neuropsychology, 19*, 475-487.

Gold, J.M., Queern, C., Iannone, V.N., & Buchanan, R.B. (1999). Repeatable Battery for the Assessment of Neuropsychological Status as a screening test in schizophrenia, I: Sensitivity, reliability, and validity. *American Journal of Psychiatry, 158*, 1955-1950.

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

Griffenstein, M.F., Baker, W.J. & Gola, T. (1996). Comparison of multiple scoring methods for Rey's malingered amnesia measures. *Archives of Clinical Neuropsychology, 11*, 283-293.

Griffenstein, M.F., Baker, W.J. & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 6*, 218-224.

Grote, C.L, & Hook, J.N. (2007). Forced-choice recognition tests of malingering. In Larrabee, G.J. (Ed.) *Assessment of malingered neuropsychological deficits* (pp. 44-79). New York: Oxford.

Haines, M.E. & Norris, M.P. (2001). Comparing student and patient simulated malingerers' performance on standard neuropsychological measures to detect feigned cognitive deficits. *The Clinical Neuropsychologist, 15*, 171-182.

Heaton, R.K., Smith, H.H., Lehman, R.A.W. & Vogt, A.T. (1978). Prospects for faking believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 46*, 802-900.

Hobart, M.P., Goldberg, R., Bartko, J.J. & Gold, J.M. (1999). Repeatable battery for the assessment of neuropsychological status as a screening test in schizophrenia, II:

Convergent/discriminant validity and diagnostic group comparisons. *American Journal of Psychiatry, 156*, 1951-1957.

Hook, J.N., Marquine, M.J., & Hoelzle, J.B. (2009). Repeatable Battery for the Assessment of Neuropsychological Status effort index performance in a medically ill geriatric sample. *Archives of Clinical Neuropsychology, 24*, 231-235.

Iverson, G.L. & Franzen, M.D. (1996). Using multiple objective measures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology, 18,* 38-51.

Lally, S. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology: Research and Practice, 34*, 491-498.

Larrabee, G.J. (2007). *Assessment of malingered neuropsychological deficits.* 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.

Lezak, M. (2004). *Neuropsychological assessment (4th Ed.)* New York: Oxford University Press.

Lindem, K., White R.F., Heeren, T., Proctor, S.P., Krengel, M., Vasterling, J., Wolfe, J., Sutker, P.B., Kirkley, S., & Keane, T.M. (2003). Neuropsychological performance in gulf war era veterans: Motivational factors and effort. *Journal of Psychopathology and Behavioral Assessment, 25*, 129-138.

Mathias, C.W., Greve, K.W., Bianchini, K.J., Housotn, R.J., & Crouch, J.A. (2002). Detecting malingered neurocognitive dysfunction using the reliable digit span in traumatic brain injury. *Assessment, 9*(3), 301-308.

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

Nies, K., & Sweet, J., (1994). Neuropsychological assessment and malingering: A critical review of past and present strategies. *Archives of Clinical Neuropsychology, 9,* 501-552.

O'Bryant, S.E., Duff, K., Fisher, J., & McCaffrey, R.J. (2003). Performance profiles and cut-off scores on the Memory Assessment Scales. *Archives of Clinical Neuropsychology, 19*, 489-496.

O'Bryant, S.E. & Lucas, JA. (2006). Estimating the predictive value of the Test of Memory Malingering: An illustrative example for clinicians. *The Clinical Neuropsychologist, 20*, 533-540.

Paniak, C., Reynolds, S., Toller-Lobe, G., Melnyk, A., Nagy, J., & Schmidt, D. (2002). A longitudinal study of the relationship between financial compensation and symptoms after treated mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 24*, 187-193.

Pankratz, L. (1979). Symptom Validity Testing and symptom retraining: Procedures for the assessment and treatment of functional sensory deficits. *Journal of Consulting and Clinical Psychology* 47**,** 409–410

Pankratz, L. (1983). A new technique for the assessment and modification of feigned memory deficit. *Perceptual and Motor Skills* 57, 367–372.

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

Randolph, C., Tierney, M.C., Mohr, E., & Chase, T.N. (1998). The Repeatable Battery for the

Assessment of Neuropsychological Status (RBANS): Preliminary clinical validity.

Journal of Clinical and Experimental Neuropsychology, 20, 310-319.

Rees, L.M., Tombaugh, T.N., Gansler, D., & Moczynski, N. (1998). Five validation experiments

of the Test of Memory Malingering (TOMM). *Psychological Assessment, 10*, 10-20.

Rose, F. E., Hall, S., & Szalda-Petree, A. D. (1995). Computerized Portland Digit Recognition

Test: The detection of malingering is improved by measuring response latency. *The*

Clinical Neuropsychologist, *9*, 124-134.

Rogers, R. (2008). *Clinical assessment of malingering and deception* (3rd Ed.) New York:

Guilford Press.

Rogers, R., & Bender, S.D. (2003). Evaluation of malingering and deception. In A.M. Goldstein

& I.B. Weiner (Eds.), *Handbook of psychology: Volume II: Forensic psychology*. (pp.

109-129). Hoboken, NJ: John Wiley & Sons, Inc.

Ruocco, A. C., Swirsky-Sacchetti, T., Chute, D. L., Mandel, S., Platek, S. M., & Zillmer, E. A.

(2008). Distinguishing between neuropsychological malingering and exaggerated

psychiatric symptoms in a neuropsychological setting. *Clinical Neuropsychologist, 22*(3),

547-564.

Silverberg, N.D., Wertheimer, J.C., & Fichtenberg, N.L. (2007). An effort index for the

repeatable battery for the assessment of neuropsychological status. *The Clinical*

Neuropsychologist, 21, 841-854.

Slick, D.J., Hopp, G., Strauss, E., & Spellacy, F.J. (1996). Victoria Symptom Validity Test:

Efficiency for detecting feigned memory impairment and relationship to

neuropsychological tests and MMPI-2 validity scales. *Journal of Clinical and Experimental Neuropsychology, 18*(6), 911-922.

Slick, D.J., Sherman, E.M.S., & Iverson, G.L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13,* 545-561.

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

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

Stulemeijer, M., Teuntje, Andriessen, T.M.J.C., Brauer, J.M.P., Vos, P.E., & Van Der Werf, S. (2007). Cognitive performance after mild traumatic brain injury: the impact of poor effort on test results and its relation to distress, personality, and litigation. *Brain Injury, 21*, 309-318.

Teichner, 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. (1996). TOMM. *Test of memory malingering.* North Tonawanda, NY: Multi-Health Systems, Inc.

Tombaugh, T (1997). The test of memory malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9* (3), 260-268.

Tombaugh, T. (2002). Test of memory malingering (TOMM) in forensic psychology. *Journal of Forensic Neuropsychology*, 2, 69-96.

Vickery, C.D., Berry, D.T.R., Inman, T.H., Harris, M.J., & Orey, S.A. (2001). Detection of

inadequate effort on neuropsychological testing: A meta-analytic review of selected

procedures. *Archives of Clinical Neuropsychology, 16*, 45-73.

Victor, T., & Abeles, N. (2004). Coaching clients to take psychological and neuropsychological

tests: A clash of ethical obligations. *Professional Psychology: Research and Practice,*

35, 373-379.

APPENDIX A

REVIEW OF THE LITERATURE

Neuropsychology and Neuropsychological Assessment

Clinical neuropsychology is an applied science which is concerned with the behavioral expression of brain dysfunction (Lezak, 2004). While the study of brain-behavior relationships dates back much further, the field of neuropsychology really began to gain momentum with the advent of World War I (Lezak, 2004). Many servicemen were returning from war with brain injuries and associated behavioral disturbances. This spurred an increase in neuropsychology programs across the nation and the field was off and running. Educational psychologists such as Binet and Simon (1908) and Spearman (1904) also made significant contributions as they pioneered the assessment of intelligence. Two early neuropsychologists, Ward Halstead and Ralph Reitan, played a pivotal role in shaping neuropsychological assessment as we know it today. Ward Halstead developed the core tests of what is now the Halstead-Reitan Battery (HRB) in the 1940s for the assessment of brain damage. Ralph Reitan later compiled several of those tests into a fixed battery which continued to be improved upon throughout the years and is one of the most noteworthy fixed batteries in the field (Lezak, 2004). Another widely respected and utilized fixed battery is the Luria-Nebraska Neuropsychological Battery (LNNB), based on the work of Russian neuropsychologist Alexander Luria (Lezak, 2004). These batteries are referred to as fixed because they are designed to be given as a set in a structured, uniform manner.

In addition to fixed batteries, neuropsychologists may also elect to use "flexible batteries," so called because in this paradigm neuropsychologists assemble their own batteries and typically adjust which tests are administered according to the patient's presenting issue or to the circumstances of the evaluation (Lezak, 1994). The clinician typically uses a core set of tests according to diagnostic category, which they then modify or add to as needed. This approach is favored by approximately 70% of clinicians in North America (Sweet, Moberg, and Suchy, 2000 as cited in Lezak, 2004). When assembling such batteries, clinicians typically make their selection from various instruments which assess domains such as general cognitive functioning, academic achievement, attention and concentration, learning and memory, speech and language, sensory-perceptual, motor, constructional/visuospatial, executive functioning, emotion/personality, and effort or motivation. The assessment of effort on these batteries is of critical importance to determine whether test results are an accurate reflection of the client's abilities.

Malingering and Related Constructs

There are countless reasons why an individual might perform poorly on measures of neuropsychological functioning. Chief among these reasons are underlying brain disorders or injuries, but poor performance may also be accounted for by insufficient or suboptimal effort, which suggests that the test taker did not perform to the full extent of his or her capabilities. In many cases, this may be due to factors such as fatigue, frustration, or co-morbid psychological conditions such as schizophrenia, depression, or antisocial personality disorder (Rogers, 2008). That said, issues of suboptimal effort often are intentional. In such cases, the individual may deliberately produce or exaggerate symptoms. This deliberate fabrication or exaggeration of

43

symptoms is referred to in the literature as feigning, malingering, symptom exaggeration, or simply as suboptimal effort (Rogers & Bender, 2003; Rogers, 2008).

Frequently, individuals undergoing neuropsychological evaluation intentionally produce false or exaggerated symptoms in hopes of attaining some external incentive. Malingering is unique in that it assumes some external motivation on the part of the test taker. It is defined in the DSM-IV-TR as "the intentional production of false or grossly exaggerated physical or psychological symptoms motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" (American Psychiatric Association, 2000, p. 739). The DSM-IV-TR also notes that malingering should be strongly suspected if any combination of the following factors are noted: (a) medicolegal context, (b) marked discrepancy between claimed stress or disability and objective findings, (c) lack of cooperation during the diagnostic evaluation as well as in complying with the prescribed regimen for treatment, (d) presence of Antisocial Personality Disorder. Malingering also differs from Factitious Disorder, in which case motivation is not external, rather the individual is internally motivated by a need to assume the sick role. Further, external motivation and intentionality distinguish malingering from Conversion Disorder and other Somatoform Disorders, in which case symptom production is not volitional (American Psychiatric Association, 2000).

Rogers (2008) suggests the use of the term feigning when underlying goals cannot be ascertained. Feigning also involves symptom exaggeration or fabrication, but does not make assumptions as to motive (Rogers & Bender, 2003). Many of the aforementioned terms are used interchangeably in the literature, in large part due to the fact that it is difficult to establish motivation and definitively classify test takers as malingering. Standardized measures of effort

and symptom validity have not been validated for assessment of an individual's underlying motivations (Rogers, 2008). Thus, the current study will typically refer to the construct as symptom exaggeration or suspicion of malingering.

It is important to note that there is a distinction between neuropsychological malingering or symptom exaggeration and the malingering or exaggeration of psychiatric symptoms. Feigned cognitive impairment is fundamentally different with respect to the tasks required of the malingerer and the detection strategies employed by the clinician (Rogers & Bender, 2003). The principle task of those feigning cognitive impairment is the effortful failure of tests (Rogers & Bender, 2003). The examinee must try to convince the clinician that he or she is putting forth his or her best effort and that the inability to perform is genuine. Conversely, malingering psychiatric impairment requires the test taker to endorse more items indicative of dysfunction. Ruocco et al. (2008) studied neuropsychological malingering and psychiatric malingering in a sample of 105 patients referred for neuropsychological evaluation using the TOMM, Reliable Digit Span (RDS), and the Millon Clinical Multiaxial Inventory-III (MCMI-III). Factor analysis revealed two distinct factors for neurocognitive malingering and psychiatric malingering. Thus, clients may be malingering in one domain while putting forth good effort in the other. As such, clinicians should never assume that malingering of cognitive impairments also renders psychiatric measures invalid.

Malingering Research Design

Malingering is a difficult construct to study. Those who feign or exaggerate symptoms are unlikely to admit to doing so and thus it is difficult to study malingering directly. The two designs most commonly applied for the investigation of malingering are simulation designs and known-group designs, each of which have significant drawbacks. Perhaps the most often

criticized methodology is the use of simulated malingerers. This approach typically recruits

university students or other volunteers from the general population. One advantage of this

approach is that it allows for a quasi-experimental design in which subjects can be randomly

assigned to dissimulating versus non-dissimulating conditions (Larrabee, 2007). It also offers the

advantage that the experimenter can be certain that the subjects in the dissimulating group are in

fact malingering. The obvious disadvantage is the lack of a clinical comparison group. Another

disadvantage is that there is no guarantee that the simulators will perform comparably to actual

patients who are feigning cognitive impairment, which of course limits external validity to some

extent.

Another significant disadvantage to the simulation design is the question of external

motivation. Typically, simulators are offered minimal compensation if they are offered

compensation at all. Conversely, those who feign or exaggerate symptoms in the context of an

actual neuropsychological evaluation frequently have significant external motivation such as

financial incentive, evasion of prosecution, or excusal from military or other work duties. This

high level of motivation makes actual malingerers more likely to invest time educating

themselves about the nature of the symptoms associated with the injury they are feigning (Nies

& Sweet, 1994). Researchers have endeavored to work around these differences by providing

study participants with written scenarios providing details of a scenario in which they would be

motivated to malinger, and some researchers may also offer monetary incentive up to $50 for the

most convincing performance (Rees et al, 1998).

Known-groups comparisons are another research design commonly applied in research

concerning malingering. Known-groups comparisons consist of two primary phases: (a)

establishing criterion groups (bona-fide patients versus malingerers) and (b) analyzing the

similarities and differences between those groups (Rogers, 2008). The problem inherent to this design is the difficulty of accurately discerning bona-fide patients from malingerers. Group designation is typically determined by highly trained experts on the basis of full evaluations or by measures with stable cutoff scores and high rates of sensitivity and specificity (Rogers, 2008). This design sacrifices the capacity for random assignment, but offers the benefit of using a clinical sample. Such research also offers a high degree of clinical relevance because it is conducted in settings where malingering is expected to occur with individuals who have real-world reasons to dissimulate (Rogers, 2008). Known-group comparisons do not offer the luxury of knowing with absolute certainty that subjects are malingering, thus researchers often group subjects as suspected of malingering or not suspected of malingering (Rogers, 2008).

Patterns of Performance on Neuropsychological Tests as Indicators of Malingering

One means by which to assess for symptom exaggeration or malingering is through identifying atypical patterns of performance on standard neuropsychological tests (Larrabee, 2003; Larrabee, 2007). Analysis of atypical performance patterns provides the opportunity to examine an individual's performance across a battery of tests. There are several benefits to this approach. Perhaps one of the most obvious is the benefit of saving time (Larrabee, 2007). If effort can be assessed through the neurological test battery itself versus having to add freestanding measures of effort then both the clinician and the examinee benefit. Another advantage to this approach is that it is far more difficult for test takers to consistently and convincingly feign impairment across an entire battery of tests than it is to feign impairment on a single symptom validity test. Thus, the clinician stands a greater chance of identifying examinees feigning or exaggerating symptoms. In addition, there may also be cases where a neuropsychological battery has already been administered without incorporation of an effort

measure. Should circumstances later indicate questionable effort on the part of the test taker the data can be examined for patterns indicative of malingering (Larrabee, 2007). When used in conjunction with symptom validity tests, pattern analysis provides increased convergent validity in assessing effort.

With the explosion of forensic cases involving neuropsychological assessment, the issue of coaching has become a primary concern. When clinicians administer symptom validity tests for the assessment of malingering, they may be working under the assumption that the client has no knowledge of malingering tests or how they work (Larrabee, 2007). Coaching may take the form of clients coaching themselves as to how to "pass" malingering instruments or it may take the form of attorneys coaching their clients. Self-coaching may entail Internet searches, attending support groups for the condition they wish to feign, or using feedback from physicians and psychologists to refine symptom presentation in subsequent evaluations (Larrabee, 2007).

A number of recent studies have explored the threat the Internet may pose to test security (Larrabee, 2007). Bauer and McCaffrey (2006) performed a study using Google to see how easily clients could obtain information which might threaten the security of three widely used effort tests: the Test of Memory Malingering (TOMM), Victoria Symptom Validity Test (VSVT), and the Word Memory Test (WMT). Threat level categories were classified according to none (no information is provided), low (minimal or basic test information only), moderate (provides description of the test format or studies which have examined the test), and high (explains scoring, or provides specific cutoff scores). Among the sites they found containing information concerning these tests, they determined that 26% posed a moderate to high threat to test security. The TOMM results produced the most "high" threat level websites, followed by the VSVT search results, and WMT produced the fewest. The most potentially damaging

information located in the searches pertained to the TOMM. Unfortunately, the simplistic design of the TOMM may make it particularly conducive to understanding and making use of the available information (Bauer et al, 2006).

Attorneys are another significant source of coaching for those attempting to malinger neurological impairment (Larrabee, 2007; Victor & Abeles, 2004). Victor and Abeles (2004) surveyed members of the Association of Trial Lawyers and the National Academy of Neuropsychology concerning coaching practices. Researchers found that 75% of attorney admitted to spending an average of 25-60 minutes preparing their clients for neuropsychological evaluations by providing information concerning the tests and making suggestions as to how the client should respond. In addition, 44% of attorneys indicated that they inquired as to what tests would be given during the course of their client's evaluation and most of them reported receiving this information from the psychologist or neuropsychologist (Victor & Abeles, 2004). This research clearly illustrates how impairment can easily be feigned on a single measure of effort whether the individual is coached by someone with familiarity with the test or does his or her own research on how to fool the examiner. These threats to test security necessitate the use of more complex methods for the assessment of malingering. It would be extremely difficult for an individual to consistently feign impairment across many tests, even if coached. Thus, examining overall patterns of performance within and across neuropsychological measures is of considerable utility in the assessment of symptom exaggeration and malingering.

Although no studies to date have examined atypical patterns of performance on the RBANS in conjunction with the TOMM, several studies have evaluated effort through use of the TOMM or other symptom validity tests and analysis of performance patterns on neuropsychological measures. In one such study, O'Bryant, Duff, Fisher & McCaffrey (2004)

used the TOMM in conjunction with a pattern analysis approach in order to evaluate effort in traumatic brain injury (TBI) litigants. Researchers used archival data from 97 patients referred to private neuropsychological practices for neuropsychological evaluation due to traumatic brain injury. All subjects had been administered the TOMM, the Rey-15, or both, as well as the Memory Assessment Scales (MAS). Based on empirically validated cutoff scores (<45 on the TOMM, or <9 on Rey-15), subjects were grouped as suspected of symptom exaggeration/malingering or as not suspected of symptom exaggeration/malingering. Two separate MANCOVAs were conducted using education as a covariate and found significant main effects for group (malingering vs. non-malingering) on both summary scores and subtest scores. Follow-up analyses indicated that the group suspected of malingering yielded lower scores than their non-suspect counterparts on all four summary scores as well as all 12 subtests. Of particular interest was the finding that both groups had very similar profile patterns, although the group suspected of poor effort performed at a consistently lower level across indices. Thus, those suspected of malingering in this study could be differentiated from non-malingerers by consistently lower performance across subtests and summary scores. The researchers also generated cutoff scores, which typically fell two standard deviations below the mean, for each of the scales. However, at 70-80% accuracy, these cutoffs did not achieve optimal classification rates. These results are promising, however, and warrant further research. Overall, results indicate that patterns of poor performance across nearly all MAS indices may aid in the detection of potential symptom exaggeration when used in conjunction with significant findings on a symptom validity test such as the TOMM (O'Bryant et al, 2004).

Lindem et al. (2003) examined the relationship between suboptimal effort as indicated by the TOMM and performance on a battery of neuropsychological tests. Participants included 77

Gulf-War era veterans who had been deployed to either Germany or the Gulf War. Subjects

completed a comprehensive neuropsychological test battery which included the TOMM.

Researchers elected to use Trial 1 scores in order to save time and because they felt Trial 1

scores are often more strongly related to other indicators of suboptimal performance or effort.

Researchers also used a cutoff score of 47 as opposed to the typical cutoff scores of 45. A cutoff

score of 45 is considered by some to be a very conservative cutoff, which may result in false

negatives for poor effort. Subsequent to testing, results were grouped into optimal (48-50) and

suboptimal effort ($\leq$ 47) groups according to Trial 1 scores. Researchers then used univariate

regression analyses to compare neuropsychological test performance between groups. Results

indicated that the majority of subjects appeared to be well motivated, with approximately 79% of

subjects scoring in the 48-50 range. However, the remaining 21% fell in the suboptimal effort

group. Scores among subjects in this group were significantly associated with lower scores on

neuropsychological measures. Specifically, subjects displaying suboptimal effort obtained lower

scores on measures of attention, executive abilities, and memory. Overall, results indicate that

subjects obtaining lower scores on the TOMM performed more poorly than their optimal effort

counterparts on a number of neuropsychological measures. These findings illustrate the

importance of examining potential exaggeration of symptoms through both the use of a symptom

validity test such as the TOMM as well as patterns of performance across a battery assessing

cognitive abilities.

Some researchers have questioned the ability of the TOMM, which uses recognition

memory, to predict performance on other measures of other neuropsychological domains such as

attention and concentration or executive functioning (Nies & Sweet, 1994; Slick et al., 1996).

Constantinou, Bauer, Ashendorf, Fisher, & McCaffrey (2005) investigated this assertion by

evaluating effort using the TOMM as well as pattern of performance on the WAIS-R and the

Halstead-Reitan Neuropsychological Battery-A. Participants included archival data from 69

mild TBI litigants referred to two different private practices for neuropsychological evaluation.

Ages ranged from 18 to 72 (M=42.41, SD=12.45) and education ranged from seven to 22 years

(M=12.96, SD=2.61). From the neuropsychological instruments, researchers elected to extract

Verbal IQ, Performance IQ, Full Scale IQ, and subtest scaled scores from the WAIS-R and to

extract scores for the General Neuropsychological Deficit Scale (GNDS) and the Halstead

Impairment Index (HII) from the HRNB-A. Magnitude and direction of relationships among

variables were calculated through a series of bivariate Pearson r coefficients. Results indicated

that scores on Trial 2 of the TOMM were significantly associated with the overall pattern of

performance on both the HRNB-A and the WAIS-R. Specifically, poor performance on Trial 2

was associated with decreases in VIQ, PIQ, FSIQ and WAIS-R subtests. Poor TOMM Trial 2

scores were also associated with higher levels of impairment as measured by the HRNB-A

indices.

For further analysis, subjects were divided into an optimal effort (OE) group (TOMM

Trial 2 score below 45) or a suboptimal effort (SE) group (TOMM Trial 2 scores of 45 or above).

A series of t-test analyses indicated that those in the optimal effort group exhibited higher

cognitive functioning and were less impaired than the suboptimal group on both WAIS-R and

HRNB-A. Of particular interest, FSIQ, PIQ, and VIQ scores of the OE group deviate from

normative means by only a third of a standard deviation. Conversely, the same indices for the

the SE group deviated more than one standard deviation from the normative mean. In addition,

although subjects in both groups were classified as mildly brain-injured, those in the SE group

had GNDS and HII scores indicative of significantly higher levels of impairment than the OE

group. Overall, results indicate that individuals who perform poorly on recognition memory effort tests will perform poorly across a variety of neuropsychological domains (Constantinou et al, 2005).

DenBoer and Hall (2007) took an interesting twist on the pattern analysis approach. In a simulation study using 237 Introduction to Psychology students, participants were initially assigned to either a control group (n=146), uncoached brain injury simulator group (n=35), or coached brain injury simulator group (n=56). Participants from simulation groups were then classified as successful brain injury simulators (SBIS) (n=29) or detected brain injury simulators (n=62) according to their scores on the TOMM. All participants completed the TOMM, the Wisconsin Card Sorting Test (WCST), the Trail Making Test-A and B, as well as the Digit-Symbol Coding, Incidental Learning, and Free Recall subtests from the WAIS-III.

One-way ANOVAs found significant group differences for TOMM Trial 1 and Trial 2. A Tukey test revealed no significant differences between scores of control and SBIS. SBIS scored significantly higher than detected coached brain injury simulators and detected uncoached brain injury simulators, whose scores did not differ from each other. The TOMM identified 80% of uncoached, but only 60% of coached brain injury simulators. Although SBIS "passed" the TOMM, they did exhibit a pattern of performance different from controls. Specifically, they exhibited a somewhat lower Trial 1 performance before improving their scores to the level of controls on Trial 2. SBIS were also distinct from controls in the respect that they obtained significantly lower scores on Trails A and B as well as on WCST Total Errors and Failure to Maintain Set. However, SBIS typically performed better than their detected coached and detected uncoached counterparts. Overall, results suggest that coaching has a significant negative impact on the sensitivity of the TOMM. Results also suggest that the sensitivity of the TOMM

may be improved by taking into account the pattern of scores across trials. In addition, patterns of performance on Trails A and B as well as Failure to Maintain Set and Total Errors on the WCST should be examined for suspiciously low performance in the event that other neuropsychological data suggests an examinee may be exaggerating symptoms or malingering (DenBoer & Hall, 2007).

Test of Memory Malingering

The Test of Memory Malingering (TOMM) is a forced-choice symptom validity test developed by Tom Tombaugh and published in 1996. The test is designed to determine whether or not an individual is feigning or malingering memory impairment (Tombaugh, 2002). Several researchers have demonstrated that individuals have a remarkably high capacity for the storage and retrieval of visual information, even among older adults and neurologically impaired populations (Tombaugh, 1997). Thus, the TOMM was designed using 50 pictures of common items. The test consists of three trials of forced-choice recognition, two of which are learning trials and the third of which is an optional retention trial. The learning trials are administered by showing the examinee the pictures for approximately three seconds each with a one second interval between pictures. The same pictures are used for each trial, though in varying order. Immediately following each of the 2 learning trials, the examinee is presented with 50 pages consisting of two images each (one target image and one distracter image) and is asked to identify the image previously seen. The test administrator provides the examinee with feedback as to whether the response was correct or incorrect after each answer. The feedback is part of standardized administration and is intended to help well-motivated test takers learn the targets while also allowing malingerers to track their performance and adjust accordingly. The retention trial, which is optional, is typically administered 15 minutes after completion of Trial 2. It is

54

recommended that the retention trial be administered if an individual obtains a score below 45 on Trial 2. During the retention trial, the clinician does not administer the learning trial, rather, the examinee goes directly to the forced-choice recognition trial. Scores on the TOMM are computed by identifying the number of items correctly answered for each trial with a total of 50 points possible per trial (Tombaugh, 1996).

As with all forced-choice symptom validity tests, the statistical probability is that the test taker should be able to answer 50% of the items correctly by chance alone. Thus, a patient who scores significantly less than chance is likely intentionally avoiding the correct response and should be suspected of malingering (Tombaugh, 1997). However, researchers have found simulators, subjects asked to feign memory impairment, perform at above chance levels on these tests. Despite the fact that performance was greater than chance, it still fell significantly below the performance of both brain-damaged subjects and neurologically-intact subjects. For this reason, researchers have suggested that the use of norm-based criteria may be more useful in detection of malingering (Tombaugh, 1997). In standardization samples, Tombaugh found that subjects generally scored 45 or above regardless of age, education, neurological dysfunction, or co-morbid psychological conditions. Thus, he has suggested the use of a normative rather than a statistical cutoff score. Individuals scoring below 45 on Trial 2 should be suspected of malingering (Tombaugh, 1997).

Tombaugh conducted normative experiments with cognitively intact and cognitively impaired individuals. The first of these was conducted with a non-clinical sample of community-dwelling residents. Subjects performed with high accuracy on all trials, generating 99% response accuracy for Trial 2 with less than 2% of variance accounted for by age and education (Tombaugh, 1997). To establish face validity, participants were asked at the beginning

of each test phase how many of the pictures they thought they would be able to accurately

identify. Subjects' estimated scores were significantly lower than obtained scores, indicating

high face validity as a memory test.

The TOMM was also validated using inpatient and outpatient data from clinical samples

with various neurological impairments including traumatic brain injury, aphasia, dementia,

general cognitive impairment, and controls. Documentation of impairment was documented for

95% of patients through medical and radiological records. The patients ranged in age from 19 to

90 years (M=56.2, SD=18.8) and education ranged from 4 to 21 years (M=12.7, SD=2.8).

Scores among those with cognitive impairment, no cognitive impairment, aphasia, and traumatic

brain injury were comparable, with the majority of respondents obtaining 97% accurate

responses on Trial 2. Dementia patients, however, responded with an average score of 92%

correct, demonstrating that the TOMM does exhibit some sensitivity to dementia (Tombaugh,

1997). Four participants in the dementia group, all of whom suffered from moderate to severe

dementia, scored below 40 on the second trial, indicating that those with moderate to severe

dementia could be inaccurately classified by the TOMM. Overall, this study demonstrated that

the TOMM was relatively insensitive to genuine memory impairment, but somewhat sensitive to

moderate to severe dementia (Tombaugh, 1997).

Tombaugh also used university students to study the efficacy of the TOMM in a

simulation design. Participants included 41 volunteers from an introductory psychology class.

The mean age and education levels were 22.2 years (SD=3.9) and 13.3 years (SD=1.0)

respectively. All participants received course credit for study participation. Exclusion criteria

were medical history suggestive of central nervous system impairment. Subjects were randomly

assigned to a malingering group (n=20) or a control group (n=21). Those in the malingering

group were provided with a scenario one week prior to testing and were told to perform as though they had been in a head-on collision where they had hit their head against the windshield and were now involved in litigation. They were advised to convince the examiner they were brain-damaged but were cautioned against over-exaggeration. Results demonstrated that control participants scored higher than malingerers for each trial and that malingerers consistently scored below 42 on Trial 2, yielding sensitivity and specificity of 100% (Tombaugh, 1997).

Since these original validation studies, Tombaugh and others have conducted additional studies validating the TOMM with various populations and under various conditions. In the first of a series of validity studies, Rees, Tombaugh, Gansler & Moczynski (1998) used a within-subjects simulation design to determine whether prior exposure to the TOMM would influence results. All subjects participated in each of 2 conditions: one in which they were advised to "try their best" and another in which they were told to malinger. The order of the conditions was varied. Those in the malingering condition were advised that the individual who provided the most convincing simulation would receive a $50 prize so as to create genuine external incentive. Results demonstrated that prior exposure to the test did not impact performance on the TOMM . Rather, the measure maintained an excellent 95% sensitivity and 100% specificity (Rees et al, 1998).

Rees et al. (1998) also validated the TOMM when administered as part of a neuropsychological battery, given that this is the fashion in which it in typically administered in clinical practice. A sample of 44 introductory psychology students were randomly assigned to either a malingering group or a control group. The TOMM was administered approximately 30 minutes into a 2 hour battery consisting of 9 tests. Results of this study produced 100% specificity and 84% sensitivity, accurately detecting 21 of the 25 subjects instructed to malinger.

The Retention trial demonstrated equal sensitivity and an increased specificity of 88%. These findings are consistent with several studies of the TOMM which have reported high specificity and moderate to high sensitivity (Gervais, Rohling, Green & Ford, 2004).

In light of the fact that an undergraduate sample may not have working knowledge of traumatic brain injuries, Rees et al. also performed a simulation study using individuals who had previously suffered a traumatic brain injury. Participants were divided into a TBI malingering group (n=8) and a TBI control group (n=10). The study also used 10 cognitively intact participants from the previous study for a point of comparison. None of the subjects were involved in litigation at the time of the study. Utilizing the typical cutoff score of 45, results on Trial 2 demonstrated 96% specificity for the TBI control group and 100% specificity for both the malingering group and the cognitively intact group. Sensitivity and specificity were improved to 100% for all groups upon administration of the Retention Trial. Overall, results indicated that first-hand knowledge of the symptoms associated with TBI did not impact performance on the TOMM (Rees et al, 1998).

In a fourth experiment in this series, Rees et al. evaluated the validity of the TOMM in a sample consisting of litigating TBI patients (n=13), nonlitigating TBI patients (n=13), and cognitively intact hospital controls (n=13). The TBI subjects were selected from an outpatient neuropsychology assessment clinic, a neurology unit, or a neuropsychiatry unit. All participants were administered the TOMM as part of a comprehensive neuropsychological evaluation. Results indicated that on Trial 2, only litigating TBI patients scored below the established cutoff of 45. The nonlitigating TBI group scored at or above 47 and the control group all obtained perfect scores of 50. Thus, results with litigating and nonlitigating samples parallel results found when using simulated malingering and control groups (Rees et al, 1998).

Several researchers have analyzed the utility of response latencies in identifying

individuals suspected of malingering (Bender & Rogers, 2004; Rogers, 2008). Rose, Hall, and

Szalda-Petree (1995) found that response latencies measured by a computerized administration

of the Portland Digit Recognition Test (PDRT) increased the sensitivity of the test as compared

with detecting malingering on the basis of number of correct responses alone (as cited in Rees et

al, 1998). In order to incorporate response latencies times in relation to the TOMM, Rees et al.

(1998) conducted a fifth experiment utilizing a computerized version of the TOMM which

captured response latency. Forty participants were assigned to either a computer malingering

(n=20) or a computer control group (n=20). The computerized version was comparable to the

paper and pencil version and captured response time to the millisecond.

Results indicated sensitivity and specificity on the basis of number of correct responses

were each 100%. With regards to response times (RTs), results indicated that response times

among both groups were comparable for Trial 1, but that the malingering group demonstrated

significantly longer response latencies on Trial 2. Researchers hypothesized that this may be

due to the fact that both groups were acclimating to the task during Trial 1. However, whereas

the control group simply responded truthfully and with increased ease on Trial 2, those in the

malingering group would have to first assess which picture is the correct answer and then decide

whether to answer truthfully or falsify their response for that given item. Incorporating these

results in routine administration of the TOMM is problematic, however, as some researchers

have demonstrated that brain-damaged individuals also have longer response times, thus making

it difficult to differentiate between malingering and genuine impairment (Rees et al., 1998).

Numerous other studies have gone on to validate the TOMM in a number of clinical

populations and circumstances. Overall, studies demonstrate the TOMM to have excellent

specificity and moderate to high sensitivity. That is, the TOMM can easily distinguish individuals exaggerating or feigning memory impairment from those with genuine memory impairment in most cases. Research also indicates that the TOMM is generally not sensitive to extraneous variables such as age, education, psychiatric symptoms, TBI, or other cognitive impairments. Some results do suggest slight sensitivity to dementia and advise caution in using the TOMM with moderately to severely demented individuals. Research has consistently supported the use of a cutoff score of 45 in lieu of the below-chance guideline. These robust findings have encouraged widespread use of the TOMM. Slick, Tan, Strauss, and Hultsch (2004) surveyed clinical neuropsychologists concerning their means of assessing for suboptimal effort and malingering. Of 24 total respondents, 21% reported that they often use the TOMM and 25% reported that they always use TOMM to assess effort.

Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)

This study explores the evaluation of malingering through the use of the floor effect strategy employed by the TOMM as well as by examining patterns of performance on the RBANS, a brief screening measure of neuropsychological functioning. Researchers have demonstrated that it is far more difficult for clients to successfully feign impairment throughout an entire battery of tests than on a single measure of effort (Heaton et al, 1978; Nies & Sweet, 1994). The problem with this approach, however, is that most neuropsychological batteries are quite lengthy, therefore requiring more time and resources than many treatment settings are able to allocate for each individual case. Thus, the present study will use the RBANS, as it typically requires less than 30 minutes for administration and is a widely used and empirically validated instrument for neuropsychological assessment (Randolph, Tierney, Mohr, & Chae, 1998; Strauss, Sherman, & Spreen, 2006).

The RBANS was produced by Christopher Randolph in 1998 for the dual purposes of identifying and characterizing cognitive decline in older adults and as a neuropsychological screening battery for a broader population including younger patients. The battery consists of 12 subtests which contribute to five indices: Immediate Memory (subtests are List Learning and Story Memory), Visuospatial/Constructional index (Figure Copy, Line Orientation), Language index (Picture Naming, Semantic Fluency), Attention index (Digit Span, Coding), and a Delayed Memory index (List Recall, List Recognition, Story Recall, and Figure Recall) (Randolph, 1998). Each index is recorded as a standard score with a mean of 100 and a standard deviation of 15 (Lezak, 2004). A Total Score is also computed by summing the index scores and converting them through the table provided in Appendix 2 of the manual. Scores are corrected for age but not for education (Randolph, 1998). However, the manual does provide means and standard deviations for the five indexes which are broken down according to age and education (Randolph, 1998). Duff, Patton, Schoenberg, Mold, Scott, and Adams (2003) expanded the normative information on the RBANS by calculating age- and education-corrected scaled scores for the subtests, index scores, and total score of the RBANS. The study used data from a group of 718 community-dwelling adults ages 65 and over who were recruited in primary care settings. Unfortunately, the current study uses a more heterogeneous sample with regards to age and as such will not be able to make sure of these expanded norms.

All of the subtests on the RBANS are variations of "tried-and-true" neuropsychological measures (Silverberg, 2007). In his initial standardization of the RBANS, Randolph (1998) examined the construct validity of the test using a clinical sample consisting largely of dementia patients. Results demonstrated that the RBANS has strong convergent validity with the neuropsychological measures it mirrors. The Visuospatial/Constructional Index score was highly

correlated with the Rey Complex Figure Test as well as the Judgment of Line Orientation test. The Language Index was highly correlated with the Boston Naming Test and the Controlled Oral World Association Test. The Total Scale score was also highly correlated with Full Scale IQ from the WAIS-R (Randolph, 1998).

The RBANS was standardized utilizing a stratified, nationally representative sample of 540 healthy adults broken down according to the following age groups: 20-39, 40-49, 50-59, 60-69, 70-79, and 80-89 (Randolph, 1998). The ratios of males to females, levels of education attained, and racial/ethnic background were comparable to those observed in U.S. Census data. Subjects represented various regions from across the United States including the Northeast, North Central, South, and West regions (Randolph, 1998). All members of the normative sample were screened both medically and psychiatrically. Those subjects with a history of head injury, epilepsy, stroke, infections of the central nervous system or other CNS disease, uncorrected hearing or vision problems, major psychiatric illness, drug or alcohol dependence, those using antipsychotic or antidepressant medications, and those demonstrating evidence of recent cognitive decline were excluded. As such, the norms provided in the RBANS manual are based on a diverse, nationally representative sample of healthy adults.

Randolph reported test-retest and split half reliability estimates. Split half reliability coefficients for Total Scale and all five indices were calculated for normal and clinical populations by age group and ranged from .75 to .95, with the majority of coefficients falling at approximately .80. Test-retest stability was initially evaluated in a sample of 40 older adults with a mean age of 70.7 years with repeat administrations of Form A at approximately 39 weeks. Reliability coefficients were found to be high for the Total Scale score (.88) and somewhat lower for individuals indices (.55 for Language index, .78 for Immediate Memory) (Randolph, 1998).

Randolph (1998) also performed a test-retest analysis of administration of the alternate forms (Form A-Form B) at intervals of one to seven days and again found high stability coefficient for the Total Scores (.82) and somewhat lower coefficient for index scores (.80 for Attention, .46 for Language). Duff et al. (2005) retested older adults (65 and older) at a one-year interval and found stability coefficients ranging from .58 to .83 for index scores and .51 to .83 for subtest scores. Practice effects were largely absent, with most scores slightly decreasing at retest.

Convergent and divergent validity for the RBANS were obtained by comparing scores for a mixed clinical sample (Alzheimer's Disease, traumatic brain injury, Parkinson's Disease, Huntington's Disease) with scores on measures assessing similar domains (Randolph, 1998). Total Scale scores demonstrated a strong correlation to FSIQ on the WAIS-R short form, although means on FSIQ were significantly higher than Total Score. These results demonstrate good convergent validity but greater ability on the part of the RBANS to detect acquired deficits. Scores on the RBANS Visuospatial Index were strongly correlated with established measures of visual discrimination and figure copy. Similarly, scores on the Immediate and Delayed Memory Indices were highly correlated with established tests of memory functioning (Randolph, 1998).

RBANS is still a relatively new instrument in the field of neuropsychology, and much of the research performed to date has focused on discrete populations such as older adults (Duff et al., 2003; Duff et al., 2005), patients with Alzheimer's disease (Randolph, 1998), Huntington's disease (Randolph, 1998), bipolar disorder (Hobart, Goldberg, Bartko, & Gold, 1999), and schizophrenia (Gold, Queern, Iannone, & Buchanan, 1999; Randolph, 1998). Much of this research has focused on the RBANS as a dementia battery and less so on its use as a general cognitive screening battery. The current literature review will not discuss these studies at length as their results are not generalizable to the current study, which will use the RBANS as a

63

screener among a more heterogeneous sample in terms of age and medical condition in order to maximize generalizability.

RBANS and Malingering

Only two studies to date have explored the issue of suboptimal effort on the RBANS. Silverberg, Wertheimer, & Fichtenberg (2007) endeavored to develop an internal validity indicator for the RBANS. Researchers noted that recognition memory tasks embedded in tests such as the California Auditory Verbal Learning Test-II (CVLT-II), the Rey Auditory Verbal Learning Test (RAVLT), and Word Lists subtests for the Wechsler Memory Scales-Third Edition have demonstrated the ability to discriminate between genuine and exaggerated memory problems. Several studies have demonstrated the tendency for recognition memory to remain intact even in the face of neurological injury (Bigler et al, 1996; Freed, Corkin, Growden, & Nissen, 1989; Millis & Dijkers, 1993 as cited in Larrabee, 2007). Silverberg and colleagues hypothesized that the recognition memory items on the List Recogntion subtest of the RBANS would be particularly sensitive to malingering given that its distracter items are not semantically or phonemically related to the target items and are never presented previously, thus making it more inconspicuously easy. The List Recognition subtest requires examinees to discriminate between ten semantically unrelated target words (presented in an earlier list learning task) and distracters from a list of 20 words by asking for each word, was ____on the list? (Randolph, 1998).

Silverberg et al. (2007) also elected to incorporate Digit Span in their development of an effort index for the RBANS. Digit span, which requires the examinee to immediately repeat back a string of numbers read by the clinician, is among the easiest of neuropsychological tasks. Several researchers have studied the Digit Span subtest of the Wechsler Adult Intelligence

Scales, which includes forward and backward trials, as a measure of poor effort. Digit Span has

been demonstrated to be insensitive to many neurological impairments but quite sensitive to poor

effort and symptoms exaggeration (Axelrod, Fichtenberg, Millis, & Wertheimer, 2006; Bernard,

1990; Griffenstein, Baker, & Gola, 1994; Iversen & Franzen, 1996; Mathias, Greve, Bianchini,

Houston & Crouch, 2002 as cited in Silverberg, 2007). Even individuals with severe amnestic

disorders typically do quite well on Digit Span tasks (Larrabee, 2007). The Digit Span subtest on

the RBANS, which includes only forward digits, should share these properties and thus be

indicative of poor effort on the RBANS. The Digit Span subtest of the RBANS requires

examinees to listen to digit spans of increasing length and immediately repeat them back to the

examiner (Randolph, 1998).

Silverberg and colleagues conducted 2 experiments in order to first establish and then test

a potential effort index for the RBANS. Participants in the first experiment included 103 patients

from a Midwestern outpatient neurorehabilitation clinic who had completed extensive

neuropsychological testing with no evidence of insufficient effort across several effort measures.

Participants in this sample presented with neurological impairments including traumatic brain

injury, cerebrovascular accident, dementia, epilepsy, multiple sclerosis, anoxia, and various

psychiatric diagnoses. The mean score on the RBANS Total Scale was 1.3 standard deviations

below the normative mean (M=81.47, SD=14.08). Researchers explored various cutoff scores for

Digit Span and List Recognition, but consistently produced false positives. They elected to

create an Effort Index (EI) by determining the subtest scores associated with the following

percentile ranges: 0, 0.1-1.9, 2-4.9, 5-8.9, 9-15.9, 16-24.9, and greater than or equal to 25. Raw

scores falling in these ranges were assigned weighted scores of 6, 5, 4, 3, 2, 1, and 0 respectively,

such that less frequently occurring scores were assigned higher weighted scores. Weighted

scores from each subtested were then summed to compute the EI, which may range from 0-12. Results suggest that EI scores greater than 3 should be considered suspicious in a general population referred for neuropsychological testing.

Silverberg and colleagues then conducted a second experiment to test the Effort Index with a sample of mild traumatic brain injury patients, a population commonly associated with high rates of malingering and thus commonly utilized in malingering research (Mittenberg et al, 2002). Participants were divided into five groups: clinical TBI group (n=32), clinical malingering group (n=15), simulated-naïve malingerers (n=28), simulated-coached malingerers (n=24), and a control group (n=24). Participants completed the RBANS and the TOMM as part of a comprehensive neuropsychological battery. Results revealed that patients with legitimate TBIs obtained higher scores (high scores indicate intact abilities) than the clinical, simulated-naïve, and simulated-coached malingerers on all RBANS indices. Overall classification accuracy for the EI was optimal (86.9%) at a cutoff score of > 0. However, researchers suggest a more liberal cutoff score of >1 in order to screen for inadequate effort in TBI populations. Overall, this experiment demonstrated that the EI discriminated well between genuine and feigned TBI. (Silverberg et al, 2007).

In a 2009 study, Hook, Marquine, and Hoelzle evaluated the RBANS Effort Index proposed by Silverberg et al. using archival data from 44 clinically referred, non-litigating older adults (over 60) referred to an urban medical center for neuropsychological evaluation. Most patients were referred for concerns related to global mental status changes or memory decline. Participants had completed the Geriatric Depression Scale, Mini-Mental Status Examination, RBANS, and the Wechsler Test of Adult Reading. EI scores were calculated according to the instructions provided in the Silverberg (2007) study. Hook et al. elected to use the cutoff score

of >3 proposed in Silverberg's first study. They found that 31% of their sample obtained

RBANS EI scores greater than 3. This suggests that scores of this magnitude are common in

clinical populations of medically ill older adults not engaged in litigation. MMSE scores and

RBANS total scores were significantly correlated with the RBANS EI score. Followup analyses

revealed that no one in a non-impaired group scored above 3 on the EI, but half of a cognitively

impaired group scored above 3. Results suggest that the EI index is confounded by cognitive

impairment and thus is not a useful indicator of effort in a medically ill older adult population

(Hook et al, 2009).

In a recent study, Duff et al. (in press) sought to further validate the RBANS Effort Index

using five independent samples of geriatric patients, one of which included community-dwelling

primary care patients as well as four clinical samples, for a total of 1,379 participants. EI scores

were again calculated according to the procedure described in Silverberg et al. (2007), and

compared to the suggested cutoff scores (e.g., >3, >0). Duff et al. also examined the

relationships between EI and demographic variables (age, education, gender) as well as between

EI scores and overall level of cognitive functioning. Given that EI scores are calculated using

raw scores which have not been corrected for age or education, researchers had concerns that

older adults, who typically obtain lower raw scores, may be unfairly penalized by the index. As

suspected, researchers found that even among a cognitively intact sample, older and less

educated patients were more likely to obtain high scores thought to suggest poor effort on the

index. Education effects were observed among three of the four clinical samples used in the

study. In addition, among more severely cognitively impaired geriatric patients, more than one

third scored in the elevated range on the index. These results suggest that, as was found with the

Hook et al. study, poor cognitive functioning is significantly associated with suspicion of poor

effort. Give that the EI does not appear to be a pure measure of effort, researchers suggest that age, education, and level of cognitive functioning must be taken into account when considering results on the index. Considerable caution is suggested in examining EI scores among cognitively impaired older adults.

Marker, Horner, and Bachman (2010) also sought to examine the validity of the Effort Index, in this case using a sample of 303 predominantly male geriatric veterans. The study utilized retrospective chart review from patients who sought services at a Veterans Affairs memory disorders clinic. Patients had completed the TOMM, RBANS, Wechsler Memory Scale –Third Edition Information and Orientation subtest, the Trail Making Test, and the Geriatric Depression Scale. Researchers classified subjects as demonstrating suspect effort (n=45) or probable good effort (n=258) on the basis of both clinical consensus (clinical interview and behavioral observations) and TOMM performance.   EI scores were calculated according to the instructions provided in the Silverberg study (2007) using a cutoff score of >3. Researchers also examined scores on the Digit Span, List Recognition, and Picture Naming subtests to assess their ability to predict malingering status. Researchers found that the Effort Index correctly classified 95.3% of subjects with probable good effort, but demonstrated substantially lower sensitivity, correctly classifying only 42.2% of those with suspect effort. The individual subtests, Digit Span and List Recognition demonstrated comparable classification rates, while Picture Naming showed modest sensitivity to detect suspect effort, but no incremental validity beyond the predictive utility of the Effort Index. These classification rates are considerably poorer than those demonstrated by the TOMM, again indicting the EI to be an inferior measure of effort as compared with the TOMM.

Summary

Malingering has been demonstrated to be a prevalent problem in the practice of neuropsychology. Clinicians cannot confidently interpret test results without measures of effort to indicate that the results at hand are an accurate representation of the examinee's full abilities. Only four studies to date have examined the issue of malingering as related to the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). Silverberg et al. (2007) endeavored to create an effort index to assess effort as part of the RBANS. However, a subsequent validation study by Hook et al. (2009) found that the index did not yield valid results. Similarly, Duff et al. (in press) found that the Effort Index is confounded by age, education, and cognitive ability. Conversely, Barker, Horner, and Bachman (2010) examined the validity of the Effort Index among a geriatric sample of predominantly male veterans and found that the EI was not significantly correlated with the EI, although researchers acknowledged that scores from the List Recognition subtest were significantly correlated with age (r= -.16, p<.01). No studies to date have examined all of the subtests and indices of the RBANS for patterns of performance indicative of malingering among a heterogeneous clinical sample. The current study is the first to examine such patterns as well as assessing effort through the Test of Memory Malingering.

Results from this study contribute to the study of malingering by providing a fast, efficient means through which to assess neuropsychological status as well as assess the potential for malingering from two vantage points. Through identifying patterns of performance on the RBANS which are associated with malingering, clinicians have a means by which to identify individuals suspected of symptom exaggeration on this brief cognitive screener. The TOMM, which typically requires fewer than fifteen minutes to administer, may also be utilized to evaluate effort in conjunction with the pattern analysis approach to the RBANS. Many clinical

settings which do not have time or resources to administer lengthy neuropsychological batteries

and multiple effort tests may benefit from this efficient approach which provides two

checkpoints for the assessment of malingering.

VITA

Neila Renae Sackett

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE RELATIONSHIP BETWEEN SCORES ON THE REPEATABLE BATTERY FOR THE ASSESSMENT OF NEUROPSYCHOLOGICAL STATUS AND THE TEST OF MEMORY MALINGERING

Major Field: Counseling Psychology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Counseling Psychology at Oklahoma State University, Stillwater, OK, in July, 2011

Completed the requirements for the Master of Science in Educational Psychology at Oklahoma State University, Stillwater, OK in 2007

Completed the requirements for the Bachelor of Arts in English at University of Missouri, Kansas City, MO, in 2004

Experience:
Practicum Counselor - Veterans Affairs Medical Center, Neuropsychology Lab
Oklahoma City, OK, July 2009-April 2010

Practicum Counselor - Oklahoma Department of Corrections
Tulsa, OK, February 2008-June 2010

Practicum Counselor - Associated Centers for Therapy
Tulsa, OK, August 2007-May 2008

Practicum Counselor - Oklahoma University Health Sciences Center
Oklahoma City, OK, June 2006-May 2007

Practicum Counselor - Oklahoma State University Counseling Psychology Clinic
Stillwater, OK, September 2005-May 2005

Professional Memberships:
American Psychological Association, Associate Member

Name: Neila Sackett   Date of Degree: July, 2011

Institution: Oklahoma State University   Location: Stillwater, Oklahoma

Title of Study: THE RELATIONSHIP BETWEEN SCORES ON THE REPEATABLE BATTERY FOR THE ASSESSMENT OF NEUROPSYCHOLOGICAL STATUS AND THE TEST OF MEMORY MALINGERING

Pages in Study: 72   Candidate for the Degree of Doctor of Philosophy

Major Field: Counseling Psychology

Scope and Method of Study: The purpose of the current study was to establish a means of assessing and confirming malingering or symptom exaggeration in the context of a brief neuropsychological screening, specifically, the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). This study used results from the Test of Memory Malingering (TOMM), a well-established measure of malingering on neuropsychological evaluations, to identify subjects suspected of symptom exaggeration. Scores on the TOMM were compared with scores on the RBANS to assess for potential patterns indicative of malingering or symptom exaggeration on the RBANS. Statistical analyses included correlation analysis, multiple regression, and binary logistic regression.

Findings and Conclusions: Correlation analyses revealed that the majority of scores on the RBANS were found to be significantly correlated with effort as determined by the TOMM. Additionally, subtest and index scores on the RBANS were examined to determine their ability to account for variation in TOMM scores. Consistent with the hypothesized relationship, Digit Span and List Recognition subtest scores were found to be predictive of TOMM scores. These two scores, in conjunction with Coding subtest scores, were found to account for 21% of variance in TOMM Scores. The predictive utility of RBANS index scores were examined in a separate regression equation and results indicated that the Attention Index and Delayed Memory Index accounted for 20% of variance in TOMM scores. A third analysis explored the extent to which RBANS subtest and index scores were predictive of group status (i.e., suspected of malingering versus not suspected of malingering.) Two of the RBANS subtest scores, list learning and coding, were found to predict group membership. With regards to index scores, the Attention Index and Delayed Memory Index were found to account for a significant amount of variance in group membership. This study also examined the relationship between the RBANS Effort Index and the TOMM, a well established measure of malingering on neuropsychological examinations. Results showed scores on the two measures to be significantly correlated. Overall, results from this study supported three of the four hypotheses and established several significant relationships between variables on the RBANS and the TOMM.

ADVISER'S APPROVAL: <u>Al Carlozzi, Ed.D.</u>