Louisiana State University
LSU Digital Commons

2011

Simulated subaverage performance on the Block Span task of the Stanford-Binet Intelligence Scales-Fifth Edition

Alyse Ann Barker
Louisiana State University and Agricultural and Mechanical College, abarke1@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses

Part of the Psychology Commons

SIMULATED SUBAVERAGE PERFORMANCE ON THE BLOCK SPAN TASK OF THE
STANFORD-BINET INTELLIGENCE SCALES-FIFTH EDITION

A Thesis
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Psychology

by
Alyse A. Barker
B.S., Louisiana State University, 2007
May, 2011

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

Acronym or Abbreviation	Complete Term
BSRI	Bem Sex-Role Inventory
DCT	Dot Counting Test
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition, Text Revision
FCT	Forced-Choice Test
FSIQ	Full Scale Intelligence Quotient
IQ	Intelligence Quotient
MND	Malingered Neurocognitive Dysfunction
NGRI	Not Guilty by Reason of Insanity
PDRT	Portland Digit Recognition Test
RDS	Reliable Digit Span
SB-V	Stanford-Binet Intelligence Scales-Fifth Edition
SILS	Shipley Institute of Living Scale
TBI	Traumatic Brain Injury
TOMM	Test of Memory Malingering
VIP	Validity Indicator Profile
VSVT	Victoria Symptom Validity Test
WAIS	Wechsler Adult Intelligence Scale
WAIS-III	Wechsler Adult Intelligence Scale-Third Edition
WAIS-IV	Wechsler Adult Intelligence Scale-Fourth Edition
WAIS-R	Wechsler Adult Intelligence Scale-Revised
WMS-III	Wechsler Memory Scale-Third Edition
WMT	Word Memory Test

ABSTRACT

As clinical psychologists and neuropsychologists routinely assess individuals in medicolegal and criminal forensic settings, they are faced with the challenge of evaluating and testifying on the validity of these psychological and neuropsychological assessments. Individuals possess various motives for manipulating their responses or performance on psychological and neuropsychological assessment instruments. Malingering refers to poor effort on psychological and neuropsychological tests when an external incentive is present to reward poor performance. Malingering can be assessed by stand-alone measures of effort or measures derived from the response profiles of traditionally administered neuropsychological and psychological tests. Using a dataset from the Stanford-Binet Intelligence Scales-Fifth Edition (SB-V; Roid, 2003a) validation, the present study derived an embedded validity index from the SB-V to be used in the clinical detection of feigned mental retardation. The author explored the utility of this index in discriminating analog malingerers from individuals with genuine FSIQ scores in the mild mental retardation range (i.e., FSIQ scores from 50 to 75). The newly developed Block Span validity index demonstrated a sensitivity rate of 52% and a specificity rate of 100% in discriminating analog malingerers from individuals with genuine SB-V FSIQ scores in the mild mental retardation range. Analog malingerers in the aforementioned analysis had SB-V FSIQ scores ranging from 40 to 110. When analog malingerers with SB-V FSIQ scores above 85 were excluded from analyses, the sensitivity rate of the Block Span validity index was 63%, and the specificity rate was 100% in detecting feigned mental retardation from genuine impairment.

INTRODUCTION AND LITERATURE REVIEW

Clinical psychologists and neuropsychologists routinely evaluate individuals in a variety of settings including vocational rehabilitation and the forensic arenas. In these settings, psychologists and neuropsychologists are faced with the challenge of assessing neuropsychological and psychological functioning using standardized measures and further evaluating the validity of these assessments. Determining the validity of an assessment becomes particularly important in circumstances where individuals might purposefully exaggerate or feign deficits in cognitive functioning in order to obtain external incentives. Individuals may feign cognitive dysfunction and/or intellectual functioning consistent with mental retardation for a variety of reasons, for example, to obtain damage awards, disability determinations, or reduced penalties in the criminal justice system. One of the most notable circumstances of great incentive to feign mental retardation was born of the U.S. Supreme Court case of Atkins v. Virginia (2002) where the court ruled it was unconstitutional to execute persons with mental retardation. Under the U.S. Supreme Court ruling of Atkins v. Virginia (2002), criminal defendants in any jurisdiction can evade the death penalty by obtaining a sustainable legal determination of mental retardation.

<u>Defining Malingering</u>

"Malingering," as described by the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR), is "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). In the DSM-IV-TR (2000), malingering is not classified as a mental disorder, rather

it is classified as a condition that may be the focus of clinical attention, particularly if two or more of the following factors are present: a medicolegal setting, antisocial personality disorder, inconsistency between objective findings and subjective symptoms, or insufficient cooperation during the assessment and noncompliance with treatment (American Psychiatric Association, 2000). The DSM-IV-TR (2000) clearly differentiates malingering from other psychological disorders, including factitious disorder and conversion disorder, which are characterized by similar symptoms but are maintained by different functions. Factitious disorder involves a conscious exaggeration or feigning of psychological symptoms, but unlike malingering, individuals with factitious disorder have motives that are psychological, not external. Conversion disorder also involves psychological incentives, and, unlike malingering, this disorder does not involve a deliberate exaggeration or feigning of deficits. While malingering and factitious disorder are under volitional control, conversion disorder is not. According to the DSM-IV-TR's (2000) diagnostic criteria for malingering, individuals diagnosed with factitious disorder or conversion disorder cannot be diagnosed as malingerers.

In the DSM-IV-TR (2000), psychological disorders, such as conversion disorder and factitious disorder, have explicit diagnostic criteria; however, no explicit criteria exist for the identification of malingering per se. In response to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, (DSM-IV; American Psychiatric Association, 1994) and DSM-IV-TR's (2000) lack of specific diagnostic criteria of malingering, Slick, Sherman, and Iverson (1999) proposed a more objective and replicable set of standards to aid clinicians in diagnosing malingering. Their criteria involve three levels of certainty: possible, probable, and definite malingering of neurocognitive dysfunction. According to the Slick et al. (1999)

criteria, classification in the diagnostic category of "definite malingering of neurocognitive dysfunction" (p. 552) is appropriate if the client meets three criteria. First, a significant external incentive must be present. Secondly, a client must display a "definite negative response bias," (p. 552) which is operationally defined as performance below chance (p<0.05) on a forced-choice cognitive test, and lastly, malingering behaviors or poor performance on the forced-choice test cannot be better explained by psychiatric, neurological, or developmental issues (Slick, Sherman, & Iverson, 1999). The Slick et al. (1999) criteria for a diagnosis of "probable malingering of neurocognitive dysfunction" (p. 552) include a significant external incentive and neuropsychological testing that presents two forms of evidence of malingering or neuropsychological testing that presents one form of evidence of malingering and a self-report that presents another form of evidence of malingering. Furthermore, diagnosis of "probable malingering of neurocognitive dysfunction" (p. 552) is only warranted if a client does not exhibit a "definite negative response bias" (p. 552) and if the malingering behaviors cannot be better explained by psychological, developmental, or neurological issues (Slick, et al., 1999). A diagnosis of "possible malingering of neurocognitive dysfunction" (p. 553) is appropriate if a client does not warrant a diagnosis of definite or probable malingering but a significant external incentive is present and a client's self-report contains discrepancies consistent with malingering that are not better explained by psychological, neurological, or developmental issues or if a client would warrant a diagnosis of definite or probable malingering but explanations based on psychological, neurological, or developmental issues are possible (Slick, et al., 1999). These criteria are quite effective in detecting feigned impairment of neuropsychological dysfunction; however, there is limited research on the utility of the Slick

criteria in psychiatric settings and forensic settings where feigned mental retardation or psychosis is suspected.

<u>Prevalence of Malingering</u>

The prevalence of malingering is difficult to estimate because malingers would presumably exert considerable effort to conceal their deceptive behavior. Research has demonstrated base rates of malingered cognitive dysfunction vary by degree of external incentive, setting, referral source, and diagnosis or symptoms being exaggerated.

With regards to external incentives, Bianchini, Curtis, and Greve (2006) found higher base rates of malingering among individuals with mild traumatic brain injuries who had larger financial incentives. Their research indicated a "dose-response relationship" (p. 843); that is, the prevalence of malingering increases as the opportunity for financial gains increases (Bianchini, et al., 2006).

Medicolegal and criminal forensic arenas often provide external incentives for exaggeration or feigning of cognitive dysfunction. These incentives may take the form of increased financial compensation and reduced or easier sentences in medicolegal and criminal forensic settings, respectively. Consequently, base rates of malingering have been investigated in these settings. Larrabee (2003) compiled the results of eleven published studies and found that the prevalence rates of malingering in civil litigation ranged from 15% to 64%, with an average base rate of 40% (548/1363). Ardolf, Denney, and Houston (2007) evaluated the base rate of malingering in a sample of 105 pretrial, male criminal defendants. Using Slick et al. (1999) criteria, 32.4% of the sample were identified as probable malingerers, while 21.9% were classified as definite malingerers. A study by Weinborn, Orr, Woods, Conover, and Feix (2003) found that the presence of malingering was higher among

defendants awaiting competency to stand trial hearings than individuals adjudicated not guilty by reason of insanity (NGRI) or those civilly committed. When surveyed, neuropsychologists belonging to the American Board of Clinical Neuropsychology estimated the prevalence of probable malingering to be 30% and 21% for civil and criminal cases, respectively (Mittenberg, Patton, Canyock, & Condit, 2002). In the absence of litigation or compensation claims, base rates of malingered cognitive dysfunction ranged from 7-12% (Mittenberg, et al., 2002).

Referral source is another variable that explains some of the variance in base rates of malingering in civil and criminal cases. Overall, base rates are higher when clients are referred by the opposing party as opposed to their own counsel (Mittenberg, et al., 2002). In civil cases, base rates of malingering are higher when plaintiffs are referred by defense attorneys and/or insurance companies. In criminal cases, base rates of malingering are higher when the defendant is referred by the prosecution as opposed to his or her own attorney.

Base rates of malingering also vary depending on the diagnoses and/or types of symptoms being exaggerated. Research reveals individuals with mild traumatic brain injuries perform significantly poorer on effort measures than individuals with moderate or severe traumatic brain injuries (Flaro, Green, & Roberson, 2007; Green, Iverson, & Allen, 1999). Even after adjusting for variance due to referral source, higher base rates of malingering have been reported among individuals with mild head injuries (41.24%) as opposed to moderate or serve head injuries (8.82%) (Mittenberg, et al., 2002). Diagnoses or symptoms associated with high malingering base rates include mild head injury (41.24%), fibromyalgia or chronic fatigue (38.61%), pain or somatoform disorders (33.51%), neurotoxic disorders (29.49%), and electrical injury (25.63%) (Mittenberg, et al., 2002).

Methods and Measures Used to Detect Malingering

Several different methods have been used to detect malingering of neurocognitive dysfunction. These methods involve analyzing atypical performance patterns and performance inconsistencies as well as using floor effects. Symptom validity tests and derived embedded validity indices are two types of measures used to detect malingering that may employ the aforementioned methods (Iverson & Binder, 2000).

Research reveals malingerers often demonstrate performance patterns that differ from individuals with genuine deficits. In a study by Heaton, Smith, Lehman, and Vogt (1978), malingerers obtained significantly lower scores than head-injured patients on sensory, psychomotor, and working memory measures, while head-injured patients displayed poorer performance on measures of executive functioning. Similarly, Mittenberg, Azrin, Millsaps, and Heilbronner (1993) found that while individuals with genuine brain injuries perform more poorly on memory measures than attention measures, malingerers demonstrate the opposite pattern, as they perform worse on attention measures than memory measures. Some assessment instruments detect malingering by analyzing inconsistencies in an individual's performance on easier test items compared to his or her performance on harder items of the same test. Generally, individuals are expected to perform more accurately and/or faster on easier items of tests than harder items. The Dot Counting Test (DCT; Rey, 1941) and Validity Indicator Profile (VIP; Frederick, 1997) are two maligning measures that use this approach of performance inconsistencies.

Malingering measures that use floor effect techniques are designed so that naïve individuals who wish to feign impairment perform at a level below individuals with substantial neuropsychological deficits. Standardized cutoff scores based on the lowest

performance levels seen among the impaired comparison groups are applied to distinguish individuals with true deficits from malingerers. The Word Memory Test (WMT, Green, 2003) and the Rey Fifteen-Item Test (Rey, 1964) are frequently employed malingering measures that use floor effect strategies (Slick, Tan, Strauss, & Hultsch, 2004).

Symptom validity tests are forced-choice measures specifically designed to detect insufficient effort. At their most basic level, these forced-choice tests (FCTs) operate on the premise that below chance scores result from individuals deliberately choosing incorrect responses (Bianchini, Mathias, & Greve, 2001). However, FCTs may also be failed by random response patterns or by scores that fall below empirically validated cutoff scores. As previously stated, a person must perform significantly below chance on a forced-choice effort test in order to meet Slick, et al. (1999) criteria for definite malingering, but even when performance surpasses chance levels, FCTs can be failed by unusually poor performance below empirically validated cutoff scores. The Test of Memory Malingering (TOMM; Tombaugh, 1996), Word Memory Test (WMT; Green, 2003), Victoria Symptom Validity Test (VSVT; Slick, Hopp, Strauss, & Spellacy, 1996), and the Portland Digit Recognition Test (PDRT; Binder, 1993) are just a few of the many symptom validity tests available for use in malingering detection. More detailed information about the Test of Memory Malingering (TOMM; Tombaugh, 1996) and Word Memory Test (WMT; Green, 2003) is included below because these two symptom validity tests were used in the present study.

Test of Memory Malingering

The Test of Memory Malingering (TOMM; Tombaugh, 1996) is a very popular symptom validity test (Slick, et al., 2004). A raw score below 45 on Trial 2 and/or the Retention Trial of the TOMM is suggestive of poor effort (Tombaugh, 1996). A study by

Tombaugh (1997) revealed that individuals with neurological impairments, including aphasia, traumatic brain injury (TBI), and cognitive impairment, obtain scores compatible with healthy individuals on Trial 2 and the Retention Trial of the TOMM. In this study, the performance of the individuals with neurological impairments was greater than 97% correct on Trial 2 and even higher on the Retention Trial. Even a group of individuals with dementia had average scores above 45 on Trial 2 and the Retention Trial of the TOMM. Research studies have substantiated the utility of the TOMM in the detection of malingered neurocognitive dysfunction (MND) in various populations and revealed the measure has acceptable sensitivity and specificity rates (Rees, Tombaugh, Gansler, & Moczynski, 1998; Greve, Bianchini, & Doane, 2006; Weinborn et al., 2003).

Word Memory Test

The Word Memory Test (WMT; Green, 2003) is another assessment instrument used to detect malingering. On two of the six subtests of the WMT, namely, the Immediate Recognition and the Delayed Recognition subtests, scores at or below 82.5% are indicative of poor effort, while scores at or above 90% suggest good effort. Scores falling between 82.5% and 90% are indicative of questionable effort (Green, 2003; Green, Allen, & Astner, 1996). Studies have demonstrated that external incentives have a greater effect on WMT performance than neurological impairment and the WMT is a useful measure in the detection of malingering (Green, 2003; Green, et al., 1996; Green, et al., 1999; Flaro, et al., 2007). Researchers purport the WMT demonstrates greater sensitivity than the TOMM in the detection of malingering, while maintaining adequate specificity (O'Bryant & Lucas, 2006).

Embedded Validity Measures

Like symptom validity tests, embedded validity indices also provide information about effort, but they are derived from standard neuropsychological tests. Embedded validity indices have some advantages over stand-alone effort measures. That is, they provide information about effort on specific tests of interest. Also, they do not require additional time or other resources to administer, and they are not as easily coached as stand-alone measures (Mathias, Greve, Bianchini, Houston, & Crouch, 2002). Embedded validity indices derived from the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997a) and Wechsler Memory Scale-Third Edition (WMS-III; Wechsler, 1997b) include the Digit Span age-corrected scaled score, the Reliable Digit Span (Greiffenstein, Baker, & Gola, 1994), the Digit Span Forward score, the Digit Span Backward score, Vocabulary minus Digit Span score, Mittenberg's discriminant function score (Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995), and the Rarely Missed Index (Killgore & DellaPietra, 2000). With exclusion of the Rarely Missed Index, all of the aforementioned indices use working memory scores in the detection of malingered neurocognitive dysfunction.

Working Memory Scores in the Detection of Malingering

Research suggests malingerers suppress their performance on working memory tasks. For example, a study by Mathias et al. (2002) showed probable malingerers obtained scores on the Working Memory Index of the WMS-III that were significantly below those of control participants. Consequently, measures of working memory have been used to distinguish malingers from individuals with true cognitive deficits.

Scores used to detect poor effort have been derived from the Digit Span subtest of the WAIS-III. These include the Digit Span age-corrected scaled score and the Reliable

Digit Span (Greiffenstein, et al., 1994). Iverson and Franzen (1994) used a Digit Span age-corrected scaled score cutoff of less than five to identify simulated student and inmate malingerers from patients with genuine head injuries. They correctly classified 90% of the student and inmate malingerers and 90% of the patients. Subsequent research by Iverson and Franzen (1996) using the same cutoff score revealed a sensitivity rate of .78 and a specificity rate of 1.00.

Greiffenstein, et al. (1994) developed the Reliable Digit Span to detect suspect effort. The Reliable Digit Span score is obtained by totaling the longest string of digits repeated correctly over two trials for both the forward and backward digit spans. Greiffenstein et al. (1994) obtained a sensitivity rate of .70 and a specificity rate of .73 using a Reliable Digit Span (RDS) cutoff score of seven to identify probable malingerers from non-malingering patients with TBI. Using data from 47 mild TBI litigants and 49 mild TBI non-litigants, Meyers and Volbrecht (1998) sought to cross-validate the results of Greiffenstein et al.'s (1994) study. A 95% sensitivity rate and 77.8% specificity rate was reported for a RDS cutoff score of seven when failure on a forced-choice test was used as the standard to identify poor effort. Meyers and Volbrecht (1998) found that seven of the nine mild TBI litigants who failed the forced-choice test were also identified as malingerers based on the RDS cutoff score of seven. In another study using a cutoff score of seven for RDS, Larrabee (2003) reported that 50% of the individuals in the definite malingered neurocognitive dysfunction (MND) group were correctly classified, while 93.5% of the non-malingering individuals with moderate-to-severe closed head injuries were properly identified.

Mathias et al. (2002) reported the sensitivity and specificity rates of various cutoff scores of the RDS in the identification of MND. In their study, a cutoff score of six had a

sensitivity rate of 38% and a specificity rate of 97%, whereas a cutoff score of seven had a sensitivity rate of 67% and a specificity rate of 93%. Cutoff scores of five and eight demonstrated less utility. Research by Heinly, Greve, Bianchini, Love, and Brennan (2005) provides further support for the use of the RDS cutoff score of seven in the detection of poor effort.

The RDS has been effective in identifying malingering in various populations. A RDS cutoff score of seven has shown utility in identifying malingerers among patients with chronic pain (Etherton, Bianchini, Greve, & Heinly, 2005). Greve et al. (2007) demonstrated the utility of the RDS in the detection of malingered cognitive deficits after toxic exposure. Duncan and Ausborn (2002) investigated the utility of the RDS in a criminal, forensic sample and found that a cutoff score of six had a sensitivity rate of 56.6% and a specificity rate of 90.3%, whereas a cutoff score of seven has a sensitivity rate of 67.9% and a specificity rate of 71.6%. Though the RDS is effective in detecting malingering in several samples, including TBI, chronic pain, toxic exposure, and criminal forensic populations, research by Graue et al. (2007) and Marshall and Happe (2007) reveals non-malingerers with mental retardation often perform below established RDS cutoff scores.

Babikian, Boone, Lu, and Arnold (2006) assessed the utility of six variables derived from the standard administration of the WAIS-III Digit Span subtest in the detection of malingering. These variables included the Digit Span age-corrected scaled score, the Reliable Digit Span score, the Digit Span Forward raw score, the total number of trials attempted, the longest correct Digit Span Forward string, and the longest correct Digit Span Forward string over both trials. They reported a sensitivity rate of 45% and a specificity rate of 93% using a RDS cutoff score of equal to or less than six. The RDS cutoff score of six

combined with a cutoff score equal to or less than four for the longest correct Digit Span Forward string yielded a sensitivity rate of 54% and specificity rate of 88%.

While the research on malingering indices derived from verbal working memory tasks is abundant, research on the effectiveness of malingering indices derived from visual-spatial working memory tasks is scarce. Among the few studies in this area, analog malingerers have been shown to perform significantly poorer than normal controls on visual working memory tasks (Beetar & Williams, 1995; Iverson & Franzen, 1994). Iverson and Franzen (1994) investigated the use of the Knox Cube Test total score and memory span score in discriminating analog malingers from patients with genuine head injuries. They found that the mean total score and mean memory span score of the malingerers were significantly lower than the mean total score and mean memory span score of the patients with genuine head injuries (Iverson & Franzen, 1994). A Knox Cube Test total score cutoff of less than four yielded a sensitivity rate of 80% and a specificity rate of 100%, while a Knox Cube Test memory span cutoff scores of less than four correctly classified 42% of the malingerers with a zero false positive rate (Iverson & Franzen, 1994). Research shows promising results with regards to the effectiveness of visual-spatial working memory tasks in the detection of malingered neurocognitive dysfunction.

Malingering and Mental Retardation

In the case of Atkins v. Virginia (2002), the U.S. Supreme Court ruled that the execution of persons with mental retardation was cruel and unusual punishment and in violation of the U.S. Constitution's Eighth Amendment. As a result of this ruling by the U.S. Supreme Court, any defendant, regardless of the state law, has an opportunity to avoid the

death penalty with a legal finding of mental retardation. This provides a powerful external

incentive to feign mental retardation.

Definitions of mental retardation vary across states and organizations; however, in

general, diagnostic criteria include deficits in intellectual and adaptive functioning with onset

before the age of 18. The DSM-IV-TR's (2000) diagnostic criteria for mental retardation are

an intelligence quotient (IQ) of 70 or below and adaptive functioning deficits in two or more

domains with onset preceding the age of 18. Also, the DSM-IV-TR (2000) recognizes the

issue of measurement error in intelligence scores and indicates that individuals with IQ

scores up to 75 are still eligible for the diagnosis if the other criteria are met.

Subaverage intellectual functioning, the first criterion in diagnosing mental

retardation, is typically evaluated through use of standardized intelligence tests. Although

standardized intelligence tests, including the WAIS-III, SB-V, and Kaufmann Adolescent

and Adult Intelligence Test, are the most widely accepted measures of intellectual

functioning in the judicial system (Everington & Olley, 2008), they are not free of

limitations. Measurement error, the Flynn effect, practice effects, and poor effort call into

question the veracity of IQ scores obtained from standardized intelligence tests.

Practice effects are often an issue in forensic cases, as defendants are routinely

assessed multiple times to allow opportunity for evaluation by the prosecution's and

defense's experts. In Atkins cases, practice effects are a serious concern because they may

mask the presence of cognitive deficits severe enough to warrant a diagnosis of mental

retardation. Consequently, individuals with mental retardation may go undiagnosed and

erroneously face the death penalty. Like practice effects, the possibility of malingering is

also a concern when standardized intelligence tests are used in Atkins cases. Malingering by

an individual desiring a diagnosis of mental retardation can result in an invalid diagnosis and avoidance of the death penalty.

Research shows available malingering measures that are effective in detecting malingered neurocognitive dysfunction are likely not appropriate for screening for malingered mental retardation. Generally, limited research with persons with mental retardation has demonstrated unacceptably high false positive rates for malingering when standard cutoffs are used (Graue et al., 2007; Marshall & Happe, 2007).

The embedded indices from the WAIS-III have not been shown to be particularly effective at detecting malingered mental retardation. Dean, Victor, Boone, and Arnold (2008) investigated the relationship between full scale intelligence quotient (FSIQ) scores measured by the WAIS-III and nine commonly used effort measures, including the WAIS-III Digit Span scaled score and Reliable Digit Span score. Lower FSIQ scores were associated with failing a greater number of effort measures (Dean et al., 2008). Adjusting the score cutoffs is one possible answer; for individuals with FSIQ scores below 70, cutoff scores less than 3 yielded specificity rates of 100% and 93% for the Digit Span scaled score and Reliable Digit Span score, respectively (Dean et al., 2008). Results of Graue et al.'s (2007) study also demonstrate the ineffectiveness of the embedded validity indices of the WAIS-III in discriminating individuals with mental retardation from analog malingerers using standard cutoffs. The Digit Span scaled score and Reliable Digit Span score yielded unacceptable specificity rates of .19 and .15, respectively. The specificity rate of the Mittenberg Discriminant Function index (.65) was higher, although still inadequate. Marshall and Happe's (2007) study also yielded an unacceptably low specificity rate for the Reliable Digit Span index, as only 31% of individuals with mental retardation scored above the established

cutoff of six. Marshall and Happe (2007) found that the Vocabulary minus Digit Span index produced a specificity rate of .98; however, the index's sensitivity was not investigated. Research by Graue et al. (2007) on the Vocabulary minus Digit Span index yielded a specificity rate of 1.00 but a sensitivity rate of 0.00.

Compared to the WAIS-III validity indices, stand-alone, symptom validity tests appear more promising in the detection of malingered mental retardation. Results regarding the effectiveness of the TOMM in the detection of malingered mental retardation are mixed. When standard cutoff scores for the TOMM (that is, Trial 2 and/or Retention Trial scores <45) were used by Graue et al. (2007) to discriminate analog malingerers from 25 individuals with mental retardation, the TOMM yielded a sensitivity rate of 80% and specificity rates of 69% and 81% for Trial 2 and the Retention Trial, respectively. Similarly, in Hurley and Deal's (2006) study, a TOMM Trial 2 cutoff score of <45 misclassified 41% of the individuals with subaverage intelligence as malingerers. In a study by Simon (2007), the TOMM was administered to 21 individuals with mild mental retardation (Mean FSIQ=60) who were housed at a forensic facility. Seventeen of the 21 participants in this study were diagnosed with Axis I disorders in addition to mild mental retardation (Simon, 2007). The legal charges of all participants had been resolved (Simon, 2007). Nineteen of the participants had been found not guilty by reason of insanity, while two of the participants were deemed not restorable with regards to the issue of competency (Simon, 2007). This study found that the standard cutoff score of <45 on Trial 2 and the Retention Trial of the TOMM produced acceptable specificity rates of 95% and 100%, respectively (Simon, 2007). More research is needed to determine the effectiveness of various effort measures in the detection of malingered mental retardation.

RATIONALE FOR THE PRESENT STUDY

In Atkins cases where practice effects are a concern and the WAIS-III or Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) has already been administered, the Stanford-Binet Intelligence Scales-Fifth Edition (SB-V; Roid, 2003a, 2003b, 2003c) may be the test of choice for many who wish to obtain an intelligence score untainted by prior experience. Presently, however, one drawback of using the SB-V is its lack of embedded validity indices. To address this shortcoming and increase the SB-V's utility, the current study aims to develop an embedded validity index for the SB-V.

Generally research aimed at detecting malingered mental retardation has produced dismal results. Limitations of the existing research will be discussed, as well as methods used by the present study to remedy these shortcomings. Currently the bulk of the research on the detection of malingered mental retardation has evaluated the utility of previously established cutoff scores. The ineffectiveness of these cutoff scores is not surprising because they were validated using inappropriate comparison groups. To address this limitation, the present study will attempt to derive a validity index using the SB-V data of individuals with mental retardation, the most relevant comparison group.

A shortcoming of the existing research on the detection of malingered mental retardation using working memory measures is that this research has focused almost exclusively on verbal working memory tasks. Individuals with mild mental retardation display capacity deficits in both verbal and visual-spatial working memory tasks; however, research by Rosenquist and colleagues suggests functional impairments may be limited to verbal working memory components, involving the phonological loop (Rosenquist, Conners, & Roskos-Ewoldsen, 2003). The weakened word length effect in individuals with mental

retardation suggests they have problems rehearsing verbal information in the phonological loop (Rosenquist, et al., 2003). Conversely, the presence of an equivalent visual complexity effect in individuals with and without mental retardation suggests rehearsal problems in individuals with mental retardation are found exclusively in the phonological loop and not in the visual-spatial sketchpad (Rosenquist, et al., 2003). No studies known to date have investigated the utility of visual-spatial working memory tasks in the detection of malingered mental retardation. Also, another potential benefit of using a visual-spatial working memory task to detect malingered mental retardation is that one would expect analog malingerers to have less exposure to the visual-spatial working memory capacities of individuals with mild mental retardation compared to their verbal working memory capacities. This reduced exposure may make it more difficult for analog malingerers to successfully feign visual-spatial working memory deficits of individuals with mental retardation compared to verbal working memory deficits.

Block Span Approach

The SB-V includes a nonverbal working memory subtest that is divided into six testlets of various difficulty levels. Five of the six nonverbal working memory testlets involve the Block Span task. The Block Span task is a visual-spatial working memory measure based on the Knox Cube Test. A Block Span score can be derived from an individual's performance on the Block Span task. The Block Span score is the total number of raw points obtained on the five Block Span testlets. All five Block Span testlets of the SB-V contain six Block Span items each worth one point; therefore, Block Span scores can range from zero to thirty.

Study 1

With aims at developing a validity index for the SB-V, the author explored the use of a Block Span cutoff score in Study 1. Dr. Gale Roid, the author of the SB-V, provided the author of this study with standardization sample data from healthy control individuals and individuals with mental retardation. Using individuals whose SB-V FSIQ scores fell in the mild range of mental retardation (FSIQ=50-75), a Block Span cutoff score that satisfactorily described individuals with genuine intellectual functioning in the mild mental retardation range as non-malingerers was determined using frequency distributions.

Study 2

The effectiveness of the Block Span validity index was evaluated in Study 2. A group of college students were randomly assigned to one of two conditions, normal controls or analog malingerers. The analog malingerers and normal controls were administered the SB-V, along with the Word Memory Test, the TOMM, the Digit Span subtest of the WAIS-III, and the Bem Sex-Role Inventory. As a manipulation check, the performance of analog malingerers and normal controls was compared on previously established malingering indices, including the TOMM Trial 2 score, scores on the Immediate Recognition and Delayed Recognition subtests of the Word Memory Test, and the Reliable Digit Span score from the WAIS-III. Analog malingerers were expected to obtain significantly lower scores than normal controls on the aforementioned indices. The Bem Sex-Role Inventory (BSRI) was included in this study as a measure of discriminant validity. Analog malingerers and normal controls were not expected to obtain significantly different scores on the BSRI. Of interest to this project was whether analog malingerers would obtain significantly lower Block Span scores than non-malingerers (i.e., normal controls and individuals with mental

retardation). Also this study examined the effectiveness of the Block Span cutoff score in identifying individuals malingering mental retardation from individuals with genuine intellectual functioning in the mild mental retardation range.

Research Questions and Hypotheses

Research Question 1 (Manipulation Check):

How will the analog malingerers perform on previously established malingering indices and a measure of discriminant validity compared to normal controls?

Hypothesis 1:

It is hypothesized that the analog malingerers will have significantly lower scores than normal controls on Trial 2 of the TOMM, the Immediate Recognition and Delayed Recognition subtests of the Word Memory Test, and the Reliable Digit Span from the WAIS-III Digit Span subtest. It is hypothesized that the analog malingerers and the normal controls will not obtain significantly different scores on the Bem Sex-Role Inventory, an index of discriminant validity.

Research Question 2:

How will analog malingers perform on the Block Span task compared to normal controls?

Hypothesis 2:

It is hypothesized that analog malingerers will have significantly lower Block Span scores than normal controls.

Research Question 3:

How will the analog malingerers perform on the Block Span task compared to individuals with genuine intellectual functioning in the mild mental retardation range?

Hypothesis 3:

It is hypothesized analog malingerers will have significantly lower Block Span scores than individuals with genuine intellectual functioning in the mild mental retardation range.

Research Question 4:

How effective is the Block Span cutoff score in discriminating analog malingerers from individuals with genuine intellectual functioning in the mild mental retardation range?

Hypothesis 4:

It is hypothesized that there exists a Block Span cutoff score that will demonstrate a specificity rate equal to or greater than 95% in discriminating individuals with genuine intellectual functioning in the mild mental retardation range from analog malingerers. Sensitivity rates are an exploratory matter.

METHODS

<u>Participants</u>

Power Considerations

A power analysis conducted using G*Power 3.1.0 revealed that for a Mann-Whitney U one-tailed test to detect a medium effect size (d=.5) when alpha equals 0.05 and power equals 0.8, a total sample size of 106 is needed (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007).

Stanford-Binet Intelligence Scales-Fifth Edition Standardization Sample

The norming sample of the SB-V (Roid, 2003c) was collected during 2001 and 2002 by approximately 100 examiners from urban and rural areas of the Northeast, Midwest, South, and West. The norming sample included 4,800 individuals. None of these individuals had severe medical conditions, severe communication or sensory impairments, and/or severe emotional or behavioral problems. Standard administration of the test was used. The norming sample was stratified by age, sex, race/ethnicity, education level, and geographic area to match the U.S. Census data.

The SB-V (Roid, 2003c) was also administered to special populations, including 119 individuals with documented diagnoses of mental retardation. These individuals ranged in age from 2 to 25. Fifty-six percent of the sample was male, and forty-four percent was female. Forty-two percent of the sample was white, while 39% was African-American, 12% was Hispanic, and 7% was of another race/ethnicity.

Dr. Roid, the author and publisher of the SB-V, provided standardization data of 38 individuals diagnosed with mental retardation aged 2 to 20 and 343 control individuals aged 17 to 35. The control standardization sample used in the present study included 307

individuals. Twenty-two individuals were excluded because their data contained missing

data points, and 14 individuals in the control standardization sample provided by Dr. Roid

had SB-V FSIQ scores of 70 or below and were, therefore, grouped in the mental retardation

sample. The control standardization sample used in this study had a mean age of 24.11 years

(SD = 5.19) and a mean SB-V FSIQ score of 102.48 (SD = 14.18) with SB-V FSIQ scores

ranging from 72 to 146.

Only individuals at least ten years of age with SB-V FSIQ scores in the mild mental

retardation range (IQ ranging from 50-75) were included in the mental retardation group in

this study (n = 29). These inclusion criteria excluded 23 individuals from the original dataset

of individuals with mental retardation obtained from Dr. Roid. Fourteen individuals were

excluded because they were under the age of ten, and eleven individuals were excluded

because they had SB-V FSIQ scores outside the range of 50-75. The 29 individuals in the

mental retardation group ranged in age from 10 to 34 years with a mean age of 18.86 years

(SD = 6.59). The SB-V FSIQ scores of these individuals ranged from 52 to 75 with a mean

SB-V FSIQ score of 62.17 (SD = 5.87). The sample was composed of 14 males and 15

females. Also, 11 individuals in the sample were African-American, 9 were Caucasian, 3

were Asian, 2 were Hispanic, and 4 were of another race/ethnicity. The educational level of

the individuals with mental retardation was not provided by Dr. Roid.

College Sample

The college sample consisted of Louisiana State University undergraduate students.

Students who participated in the study were at least 18 years of age with no previous

neurological or psychological diagnoses. Participants received extra credit or research credit

in a psychology course for participating in this experiment. One hundred and ten participants

were recruited for the study, and data collected from two of the participants were excluded

because of examiners' errors on the routing tests of the SB-V. Individuals in the college

sample ranged from 18 to 48 years of age with a mean age of 20.48 (SD = 4.11). The sample

consisted of 69 males and 39 females. Approximately three-fourths of the sample was

Caucasian (75.9%), 14.8% was African-American, 5.6% was Asian, 1.9% was Hispanic, and

2% was of another race/ethnicity. Regarding education, 42.6% of the sample was college

freshman, 14.8% was college sophomores, 20.4% was college juniors, and 22.2% was

college seniors.

Materials

Consent Form and Demographic Questionnaire

Individuals in the college sample were read the consent form (See Appendix A) and

any questions or concerns were addressed prior to obtaining their signatures. After obtaining

informed consent, participants were asked to complete a demographic questionnaire (See

Appendix B) and provide the following information: gender, age, race/ethnicity, as well as

psychological and neurological diagnoses.

Shipley Institute of Living Scale

The Shipley Institute of Living Scale (SILS; Shipley, 1940; Zachary, 1986) assesses

intellectual abilities. This self-administered instrument consists of two subtests. The

vocabulary subtest includes 40 multiple-choice vocabulary items, while the abstraction

subtest includes 20 fill-in-the-blank abstraction items. The test takes about twenty minutes to

complete. This measure can provide a rapid estimate of FSIQ.

Shipley Institute of Living Scale scores demonstrate acceptable psychometric

properties. Total scores yielded split-half and test-retest reliability estimates of .92 and .78,

respectively (Zachary, 1986). Correlations between the Shipley total score and FSIQ scores from the Wechsler-Bellevue, Wechsler Adult Intelligence Scale (WAIS), and Wechsler Adult Intelligence Scale-Revised (WAIS-R) were high, ranging from .68 to .90 (Zachary, 1986). Consequently the measure has been used as a brief estimate of intellectual functioning.

Stanford-Binet Intelligence Scales-Fifth Edition (SB-V)

The Stanford-Binet Intelligence Scales-Fifth Edition (Roid, 2003b, 2003c) is a standardized intelligence test. The test assesses five factors of intellectual functioning: Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing, and Working Memory. Each of these five factors is measured in both the verbal and nonverbal domains yielding ten subtests.   The first subtests administered are the nonverbal Fluid Reasoning routing subtest and the verbal Knowledge routing subtest. Scores on the nonverbal and verbal routing subtests determine the starting level for the nonverbal and verbal domains, respectively. Excluding the routing subtests, the remaining eight subtests are further divided into five or six testlets that span five or six difficulty levels. Basal and ceiling rules applied to each testlet determine the difficulty of items administered for each subtest. Examinees can usually receive a maximum of six points for each testlet. After the basal has been established, if an individual obtains two or fewer points on any testlet, the subtest is discontinued. Five of the six nonverbal working memory testlets use the Block Span task. The SB-V takes about 45 to 75 minutes to administer.

Scores from the SB-V (Roid, 2003b, 2003c) demonstrate excellent psychometric properties. FSIQ scores demonstrate very high internal reliability estimates (.97-.98) across all ages (Roid, 2003c). Test-retest reliability and inter-rater reliability are also satisfactory

(Roid, 2003c). With regards to convergent validity, SB-V FSIQ scores and WAIS-III FSIQ scores are highly correlated (.82) (Roid, 2003c).

Test of Memory Malingering (TOMM)

The TOMM (Tombaugh, 1996) is a forced-choice effort test. Individuals are shown 50 line drawings, one at a time for three seconds each. On Trial 1, individuals are presented with two line drawings and are asked to indicate which one they saw previously. Individuals are then given feedback regarding the correctness of their responses. Scores on Trial 1 can range from 0 to 50 and are usually not interpreted; however, scores below chance are indicative of insufficient effort. Following Trial 1, individuals are again shown the same 50 line drawings, one at a time. They are then presented with two line drawings (a correct response and a novel line drawing) and asked to indicate which drawing they viewed previously. Trials 1 and 2 of the TOMM take approximately 15 minutes to administer.

Word Memory Test (WMT)

The Word Memory Test (WMT; Green, 2003) is a test of verbal learning and effort during which examinees are presented with a list of 20 pairs of words. They are then asked in the Immediate Recognition subtest to choose the word they have previously viewed from a new set of 40 pairs consisting of a previously seen word and a foil. The examinee is told whether his or her response is correct or incorrect. After thirty minutes, the examinee is again asked to choose the previously learned word from a set of 40 pairs that include a previously seen word and a different foil in the Delayed Recognition subtest. Four other memory subtests follow: Multiple Choice subtest, Paired Associates subtest, Free Recall subtest, and an optional Long Delayed Free Recall subtest. During the Multiple Choice subtest, the first word of a learned pair is presented, and the examinee picks the second word

in the pair from a list of eight choices. During the Paired Associates subtest, the examinee is asked to recall the second word of a learned pair when presented with the first word. During the Free Recall subtest, the examinee recalls any word remembered from the learned list in any order. The instructions for the optional Long Delayed Free Recall subtest are the same as those for the Delayed Free Recall subtest; however, the subtest follows a 20-minute delay.

Digit Span subtest of WAIS-III

The Digit Span subtest of the WAIS-III is a measure of verbal working memory. In the Digit Span forward task, the examiner dictates a string of digits at a one-second interval, and the examinee is asked to repeat the digits in the same order. In the Digit Span backward task, the examiner again dictates a string of digits at a one-second interval; however, in this task, the examinee is asked to repeat the digits in the reverse order. As described earlier, several variables derived from the Digit Span subtest of the WAIS-III have been used to detect poor effort.

Bem Sex-Role Inventory (BSRI)

The Bem Sex-Role Inventory (BSRI; Bem, 1974) is a self-report measure assessing masculine and feminine personality traits. The inventory has a total of 60 items, including 20 masculine personality traits, 20 feminine personality traits, and 20 neutral personality traits. Examinees are asked to use a Likert scale ranging from 1 ("never or almost never true") to 7 ("always or almost always true") to rate how much each of the 60 personality traits describes him or her. Averaging the 20 scores for the masculine personality traits and the 20 scores for the feminine personality traits yields a masculinity scale score and femininity scale score, respectively. Notably, masculinity and femininity are scored as two independent dimensions, not on a continuum. The masculinity and femininity scales yield high internal consistency

estimates ranging from .80 to .82 for the femininity scale and .86 for the masculinity scale (Bem, 1974). The test-retest reliability of the masculinity and femininity scales is also high with estimates of .90 for both scales (Bem, 1974).

Participant Effort Rating Scale

A Participant Effort Rating Scale (See Appendix C) was administered to participants in the analog malingering condition. The scale includes two questions: number one, "How much effort did you put into performing as you think someone who is mentally retarded would?" and, number two, "How successful do you think you were at performing as someone who is mentally retarded would?" Scores ranged from 1 to 5 on a continuum (1="not much/not at all"; 3="average"; 5="very much/very successful").

Procedure

Standardization sample

Dr. Roid, the author and publisher of the SB-V, provided SB-V standardization data which included individuals with mental retardation and control individuals.

College sample

College students enrolled in psychology classes at Louisiana State University were recruited to participate in the present study in order to earn research credit or extra credit. Students interested in participating in this study were first given an informed consent form (Appendix A). The consent form was read aloud to the student, and then any questions and/or concerns were addressed. Students agreeing to participate in the study signed the consent form and were given a copy of the informed consent form for their records. Participants' names only appeared on their consent forms. All other data were labeled with a subject number randomly assigned to the participant. Participants' test data were in no way

linked to their consent forms, names, or other identifying information. Consent forms were

kept in a separate folder from test data. All participants were asked to complete a

demographic questionnaire (See Appendix B), as well as the Shipley Institute of Living Scale

to obtain estimates of participants' intellectual functioning. Participants were randomly

assigned to one of two conditions, the control condition or the analog malingering condition.

Participants assigned to the control condition were given the following instructions:

> "As a child, you were able to complete schoolwork and read, write, spell, and
> do math as well as other children. In addition, you were able to adequately perform
> daily functions related to interpersonal relationships, monetary transactions,
> communication, and self-care.
> As you grew older, you perceived yourself as "normal" and were never
> diagnosed with an intellectual disability. You graduated from high school and were
> accepted to Louisiana State University. You began hanging around people who
> convinced you that obtaining a college education was the best way to support
> yourself. One night you decided to sign up for a psychology course at LSU. The
> psychology course offered you an opportunity for research credit or extra credit by
> participating in psychology experiments. You chose to sign up for this psychology
> experiment.
> Now you are sitting in the LSU Psychological Services Center participating in
> this study for research credit or extra credit. You are asked to take some tests as part
> of a neuropsychological evaluation. The examiner explains that, if you perform with
> your best effort, you will contribute to psychological research. The examiner further
> explains that psychological experiments are crucial to testing hypotheses and
> developing theories.
> The validity of this research study depends on the effort you put forth on these
> tests. You are instructed to respond to the tests with your very best effort" (M. W.
> Musso, personal communication, October 30, 2009).

Participants assigned to the analog malingering condition were given these alternative

instructions:

> "As a child, you struggled continuously with schoolwork and were never able
> to read, write, spell, or do math as well as other children. In addition, you struggled
> with daily functions related to interpersonal relationships, monetary transactions,
> communication, and self-care.
> As you grew older, you knew 'something was wrong.' You either 'squeaked
> by' though school, got socially promoted, or dropped out. You continued to struggle
> with basic skills. You began hanging around people who convinced you that stealing

was the best way to support yourself. One night you broke into a house to steal some money, and the homeowner confronted you with a gun. In a split-second reaction, you shot him fatally.

Now you are in prison, facing the death penalty. You are asked to take some tests as part of a neuropsychological evaluation. Your lawyer explains that, if the tests show your IQ is less than 70, you can be found mentally retarded and escape the death penalty, as mentally retarded persons cannot be executed. Your lawyer further explains, however, that, if the examiner discovers you are faking, you will likely receive the death penalty.

Your life is now depending on the scores of these tests. You are instructed to respond to the tests in a manner that ensures you will be found mentally retarded without being detected as a faker" (M. W. Musso, personal communication, October 30, 2009).

After receiving instructions on how to perform, both groups of participants were administered the following measures: WMT, TOMM, WAIS-III Digit Span subtest, BSRI, and the SB-V. The administration of the WMT, TOMM, WAIS-III Digit Span subtest, and BSRI was counterbalanced with the administration of the SB-V. After completing these measures, participants in the analog malingering condition completed the Participant Effort Rating Scale. At the conclusion of each testing session, the participant was debriefed and thanked for his or her time. Data were concurrently collected for this present study and a companion thesis project entitled "Simulated subnormal performance on the Stanford Binet-V: An exploratory investigation of the rarely missed items index."

RESULTS

Statistical Analyses

Study 1

Based on the performance of individuals with mental retardation and control individuals from the SB-V standardization sample, a cutoff score that demonstrated adequate specificity was derived from the Block Span task using frequency distributions. The cutoff score of less than 10 demonstrated 100% specificity and did not misclassify any of the individuals in the mental retardation group or any control individuals as malingerers.

Study 2

Research Question 1 (Manipulation Check)

Independent sample t tests were conducted to assess whether college participants in the two conditions, control subjects and analog malingerers, differed in age, years of education, and SILS estimated WAIS-R IQ scores. No significant differences between the analog malingerers and control subjects were found for age, $t(106) = 1.22$, $p = .270$, education, $t(106) = .79$, $p = .432$, or SILS estimated WAIS-R IQ scores, $t(106) = -.94$, $p = .351$. Also, chi-square tests revealed the control subjects and analog malingerers did not differ significantly in race/ethnicity, $\chi^2(4, N = 108) = 1.11$, $p = .892$, or gender, $\chi^2(1, N = 108) = .361$, $p = .548$.

As a manipulation check, the analog malingerers' SB-V FSIQ scores were compared to the control samples' SB-V FSIQ scores. The author hypothesized the analog malingerers would have significantly lower SB-V FSIQ scores than the control sample. An independent samples t test supported this hypothesis, $t(81) = -12.11$, $p < .001$. Analog malingerers had significantly lower SB-V FSIQ scores than the control sample and were successful in

obtaining SB-V FSIQ scores in the range of mild mental retardation lending support for the

effectiveness of the experimental manipulation in this study. See Table 1.

Table 1.

Demographic Characteristics and SB-V FSIQ scores of the College Sample

	Analog Malingerers	Control Subjects
	M (SD)	*M (SD)*
Age in Years	20.96 (5.28)	20.00 (2.41)
Education (Years completed)	13.31 (1.21)	13.13 (1.23)
Estimated WAIS-R IQ scores	66.13 (20.56)	104.48 (10.91)
SB-V FSIQ scores	66.13 (20.56)	104.48 (10.91)
	% (n)	*% (n)*
Gender		
Males	61.1 (33)	66.7 (36)
Females	38.9 (21)	33.3 (18)
Ethnicity/race		
Caucasian	79.6 (43)	72.2 (39)
African-American	13 (7)	16.7 (9)
Asian	3.7 (2)	7.4 (4)
Hispanic	1.9 (1)	1.9 (1)
Other	1.9 (1)	1.9 (1)

Note. Estimated WAIS-R IQ scores = estimated Wechsler Adult Intelligence Scale-Revised IQ scores from the Shipley Institute of Living Scale; SB-V FSIQ scores = Stanford-Binet Intelligence Scales-Fifth Edition full scale IQ scores

As a manipulation check, TOMM Trial 2 scores, WMT Immediate Recognition and

Delayed Recognition scores, and the Reliable Digits Span scores of the analog malingerers

and control sample were compared. Analog malingerers were hypothesized to have

significantly lower scores on all of these effort measures. Shapiro-Wilk tests of normality

revealed the scores obtained by the analog malingerers and control subjects on the Trial 2 of

the TOMM, $W(106) = .71$, $p < .001$, the Immediate Recognition subtest, $W(106) = .80$, $p <$

.001, Delayed Recognition subtest, $W(106) = .78$, $p < .001$, and the Reliable Digit Span,

$W(106) = .97$, $p < .05$, were all non-normal. Therefore, instead of using a between-subjects

MANOVA, nonparametric Mann-Whitney U tests were conducted to analyze differences

between groups on the aforementioned scores. The analog malingers obtained significantly

lower scores than the control college subjects on Trial 2 of the TOMM, $z = -6.96$, $p < .001$,

the Immediate Recognition subtest of the WMT, $z = -7.32$, $p < .001$, the Delayed

Recognition subtest of the WMT, $z = -7.83$, $p < .001$, and the Reliable Digit Span, $z = -5.41$,

$p < .001$. The mean rank scores of the analog malingerers on Trial 2 of the TOMM, the

WMT Immediate Recognition subtest, the WMT Delayed Recognition subtest, and the

Reliable Digit Span were 35.96, 32.81, 31.20, and 38.31, respectively. The mean rank scores

of the control college students on Trial 2 of the TOMM, the WMT Immediate Recognition

subtest, the WMT Delayed Recognition subtest, and the Reliable Digit Span were 73.04,

75.59, 76.65, and 70.69, respectively. Large effect sizes were noted for group membership

on the TOMM Trial 2 scores, $r_{pb} = -.67$, WMT Immediate Recognition scores, $r_{pb} = -.71$,

WMT Delayed Recognition scores, $r_{pb} = -.76$, and Reliable Digit Span scores, $r_{pb} = -.52$.

The Bem Sex-Role Inventory (BSRI) was used as a discriminant validity measure.

The BSRI masculinity and femininity scale scores of the analog malingerers and control

subjects were not expected to differ significantly. Shapiro-Wilk tests of normality indicated

scores of the analog malingerers and control subjects on the BSRI masculinity scale, $W(106)$

$= .95$, $p < .05$, and BSRI femininity scale, $W(106) = .97$, $p < .05$, were non-normal.

Therefore, Mann-Whitney U tests were conducted instead of a between-subjects MANOVA.

Results of two Mann-Whitney U tests revealed the analog malingerers had significantly

lower scores than the control college students on the BSRI masculinity scale, $z = -2.50$, $p =$

.012 and the BSRI femininity scale, $z = -2.61$, $p = .009$. For the BSRI masculinity scale, the mean rank of the analog malingerers was 46.96, and the mean rank of the control college students was 62.04. For the BSRI femininity scale, the mean rank of the analog malingers was 46.63, and the mean rank for the control college students was 62.37. Notably, while the previously established effort measures demonstrated large effect sizes across groups, only medium effect sizes were noted for the masculinity, $r_{pb} = -.24$, and femininity scale scores, $r_{pb} = -.25$, reflecting the expected trend towards discriminant validity.

The Participant Effort Rating Scale was administered to only participants in the analog malingering condition. Question 1 of the Participant Effort Rating Scale asked participants how much effort they put into performing as they thought someone who was mentally retarded would. There was a significant negative correlation between self-reported effort and participants' SB-V FSIQ scores, $r = -.35$, $p = .012$. More self-reported effort by individuals into performing as they thought someone who is mentally retarded would was associated with lower SB-V FSIQ scores. Self-reported effort scores were also significantly correlated with TOMM Trial 2 scores, $r = -.37$, $p = .007$, and Block Span scores, $r = -.32$, $p = .020$, with lower TOMM Trial 2 scores and lower Block Span scores associated with individuals self-reporting putting more effort into faking mental retardation. Self-reported effort was not correlated with WMT Immediate Memory scores, $r = -.16$, $p = .256$, WMT Delayed Memory scores, $r = -.20$, $p = .148$, Reliable Digit Span scores, $r = -.18$, $p = .215$, BSRI masculinity scores, $r = -.03$, $p = .832$, or BSRI femininity scores, $r = 0.01$, $p = .949$.

The second question of the Participant Effort Rating Scale asked participants how successful they thought they were at performing as a person with mental retardation would. There were no significant correlations between individuals self-reported judgment of success

at malingering and SB-V FSIQ scores, $r = -.06$, $p = .660$, Block Span scores, $r = 0.01$, $p = .950$, WMT Immediate Recall scores, $r = -.23$, $p = .095$, Reliable Digits scores, $r = -.20$, $p = .150$, BSRI masculinity scores, $r = .11$, $p = .443$, or BSRI femininity scores, $r = .14$, $p = .321$. However, individuals self-reported judgment of success at malingering was correlated with WMT Delayed Recall scores, $r = -.32$, $p = .022$, and TOMM Trial 2 scores, $r = -.40$, $p = .003$, with higher levels of perceived success associated with lower scores on these measures.

Research Question 2

The Block Span scores of the analog malingerers and control college students were non-normal, as indicated by a Shapiro-Wilk test of normality, $W(106) = .95$, $p < .001$. A Mann-Whitney U test revealed Block Span scores of the analog malingerers were significantly lower than the Block Span scores of the normal controls, $z = -7.33$, $p < 0.001$. A medium effect size, $r_{pb} = .30$, was noted for the effect of group membership on Block Span scores. The mean rank of the analog malingerers' Block Span scores was 32.45, and the mean rank of the college control subjects' Block Span scores was 76.55.

Research Question 3

A Mann-Whitney U test also showed the Block Span scores of the analog malingerers were significantly lower than the Block Span scores of individuals in the mental retardation group, $z = -2.76$, $p = 0.003$. The mean rank of the analog malingerers' Block Span scores was 36.68, while the mean rank of the mental retardation groups' Block Span scores was 51.91.

The analog malingerers and mental retardation group were compared along several demographic variables including age, gender, and race, as well as SB-V FSIQ scores. An independent sample t-test revealed the individuals in the mental retardation group did not

differ significantly from the analog malingerers in age, $t(48) = 1.48$, $p = .145$. Likewise, a chi-square test revealed no significant differences in gender between the analog malingerers and individuals in the mental retardation group, $\chi^2(1, N = 83) = 1.27$, $p = .261$. However, according to a chi-square test, the race of the analog malingers and individuals in the mental retardation group differed significantly, $\chi^2(4, N = 83) = 19.711$, $p = 0.001$. A Mann-Whitney U test revealed the analog malingerers and mental retardation sample did not have significantly different SB-V FSIQ scores, $z = -.081$, $p = .935$. The mean rank of the analog malingerers' SB-V FSIQ scores was 42.16, and the mean rank of the mental retardation samples' SB-V FSIQ scores was 41.71. See Table 2.

Table 2.

Demographic Characteristics of Analog Malingers and Individuals in the Mental Retardation Sample

	Analog Malingerers	Mental Retardation Sample
	M (SD)	*M (SD)*
Age in Years	20.96 (5.28)	18.86 (6.59)
SB-V FSIQ scores	66.13 (20.56)	62.17 (5.87)
	% (*n*)	**% (*n*)**
Gender		
Males	61.1 (33)	48.3 (14)
Females	38.9 (21)	51.7 (15)
Ethnicity/race		
Caucasian	79.6 (43)	31 (9)
African-American	13 (7)	37.9 (11)
Asian	3.7 (2)	10.3 (3)
Hispanic	1.9 (1)	6.9 (2)
Other	1.9 (1)	13.8 (4)

Note. SB-V FSIQ scores = Stanford-Binet Intelligence Scales-Fifth Edition full scale IQ scores

Research Question 4

Analog malingerers were found to have significantly lower Block Span scores than individuals in the mental retardation group; therefore, the cutoff score of less than 10 derived in Study 1 was used to determine the effectiveness of the Block Span score in discriminating the two groups. A Block Span cutoff score of <10 yielded a sensitivity rate of 0.52 and a specificity rate of 1.00 in discriminating analog malingerers from individuals with genuine IQ scores in the mild mental retardation range. That is, 52% of the individuals faking mental retardation were correctly classified as malingerers, and no individual with a genuine IQ score in the mild mental retardation range was misclassified. Sensitivity and specificity rates of other potential cutoff scores are presented in Table 3; however, a cutoff score of <10 appears to be the most clinically useful.

Table 3.

Sensitivity and Specificity Rates of Various Cutoff Scores in Distinguishing Analog Malingerers from Individual with Genuine FSIQ Scores in the Mild Mental Retardation Range

Block Span Cutoff Score	Sensitivity Rate	Specificity Rate
<5	.26	1.0
<6	.33	1.0
<7	.35	1.0
<8	.48	1.0
<9	.48	1.0
<10	.52	1.0
<11	.56	.76
<12	.68	.62
<13	.72	.41
<14	.78	.28
<15	.78	.14

In this study, the positive predictive power of a cutoff score of <10 on the Block Span validity index was 1.00, and the negative predictive power was .53. That is, 100% of the individuals obtaining Block Span validity index scores below 10 were malingerers, and 53% of individuals obtaining Block Span validity index scores of 10 or above were non-malingerers.

The SB-V FSIQ scores of the 54 analog malingerers in this study ranged from 40 to 110. In real word settings, individuals who obtain FSIQ scores within one standard deviation of the population mean are unlikely to be eligible for a diagnosis of mental retardation, as most definitions of mental retardation require that an individual's FSIQ score is at least two standard deviations below the population mean. Ten analog malingerers in the current study failed to obtain SB-V FSIQ scores below 86 and therefore would presumably be ineligible for a diagnosis of mental retardation in real world settings. An additional analysis exploring the sensitivity and specificity rates of the Block Span cutoff score of less than 10 was conducted excluding the ten analog malingerers with SB-V FSIQ scores above 85. This analysis yielded a sensitivity rate of .63 and specificity rate of 1.0 in discriminating analog malingerers with SB-V FSIQ scores below 86 from individuals with genuine intellectual impairment in the mild mental retardation range.

DISCUSSION

The 2002 U.S. Supreme Court ruling in the case of Atkins v. Virginia provides incentive for defendants to fake or malinger cognitive impairment consistent with mental retardation, as the Court ruled individuals with mental retardation are not eligible for the death penalty. Clinicians use stand alone symptom validity tests and embedded validity indices to aid them in the identification of individuals feigning impairment. Research evaluating the appropriateness of using currently published symptom validity tests to identify individuals feigning mental retardation provides mixed results at best, with most studies demonstrating unacceptably high rates of individuals with genuine mental retardation identified as malingerers. Embedded validity indices with clinical utility in identifying individuals feigning mental retardation are also lacking. The purpose of the present study was to develop a SB-V embedded validity index which would be useful in identifying individuals malingering mental retardation.

A malingering index was derived from the Block Span testlets of the SB-V. A Block Span score was calculated by summing an individual's raw scores on the five Block Span testlets. Analog malingerers were found to have significantly lower Block Span scores than individuals with genuine intellectual impairment in the mild mental retardation range. Using a Block Span cutoff score of less than 10, 52% of the analog malingerers were correctly identified as fakers, and none of the individuals with genuine impairment were misclassified. When only analog malingerers with SB-V FSIQ scores of less than 86 were included in analyses, 63% of the analog malingerers and 100% of the individuals with genuine intellectual impairment in the mild mental retardation range were correctly classified.

The present study yields promising results regarding the effectiveness of a Block span cutoff score of less than 10 in identifying individuals feigning mental retardation. However, several limitations of the present study are also noted. First, the Block Span validity index was evaluated using a simulation study design. This design produces high internal validity; however, generalizability of these results to actual malingerers is compromised. Further research is needed to investigate the utility of the Block Span validity index using a known-groups comparison design. The generalizability of the results of the present study is also restricted with regards to age, as the analog malingerers ranged in age from 18 to 48 with only two individuals above the age of 29 and individuals in the mental retardation sample ranged in age from 10 to 34. More research is needed to evaluate whether the results of the present study will generalize to middle-aged and older adults, although those groups are less likely than young adults to have the question of mental retardation raised.

A manipulation check was conducted in which the scores of analog malingerers were compared to college controls on previously established malingering measures. As hypothesized, the analog malingering group obtained significantly lower scores on all of the malingering measures when compared to controls. There were no significant differences between analog malingerers and controls on demographic variables or measures administered before the experimental manipulation. However, the analog malingerers obtained significantly lower scores than controls on the BSRI masculinity and femininity scales. These scales were included in the study as measures of discriminant validity, and the two groups were not expected to have significantly different scores. One possible explanation for the groups' significantly different scores may be the fact that the BSRI self-report measure included an intellectual component. That is, individuals were instructed to read and

presumably comprehend a list of 60 personality traits and then rate themselves on these characteristics. Participants in the analog malingering condition may have perceived this self-report measure as possessing an intellectual component and responded with a random response pattern consisting of lower scores to demonstrate they were incapable of understanding the items and/or completing the task.

Another limitation of the present study is that the mental retardation sample had a significantly higher percentage of African-Americans than the analog malingering sample, while the analog malingering sample had a significantly higher percentage of Caucasians than the mental retardation sample. Further research is needed to rule out race as a contributing factor to the differences in Block Span scores between these two groups. The sample size of the mental retardation group in this study was smaller than expected. Replication of this study using a larger sample of individuals with mental retardation would be helpful to rule out the possibility that these findings are an artifact of this particular mental retardation sample. Along this vein, it is unclear how representative the mental retardation sample used in this study is to the general population of individuals with mild mental retardation. Notably, the mental retardation sample used in this study included 15 females and 14 males, when the ratio of males to females with mental retardation in the general populations is estimated to be 1.5:1 (American Psychiatric Association, 2000). When this study was proposed it was anticipated that only individuals with previous diagnoses of mental retardation would be included in the mental retardation sample. However, due to the small sample size resulting from this initially proposed criterion, individuals who scored in the range of mild mental retardation on the SB-V but had no previous diagnoses of mental retardation were included in the mental retardation group.

Although the present study has some limitations, it is an improvement upon the existing mental retardation malingering literature. The Block Span validity index was developed using an appropriate comparison group, and it demonstrates excellent specificity (100%) as well as adequate sensitivity (52% to 63%) in this study. The sensitivity and specificity rates obtained in this study are superior to those obtained in the literature using the WAIS-III embedded validity indices to detect feigned mental retardation (Graue et al., 2007; Marshall & Happe, 2007). In Atkins cases where defendants' efforts are suspect, clinicians may benefit from using the SB-V over the WAIS-III or WAIS-IV because the SB-V now contains a validity index which can aid clinicians in detecting feigned mental retardation.

REFERENCES

Ardolf, B. R., Denney, R. L., & Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist, 21(6),* 899-916.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th ed.).* Washington, DC: Author.

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

Atkins v. Virginia. (2002), 260 Va. 375, 534S. E. 2d 312.

Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspect effort. *Clinical Neuropsychologist, 20(1),* 145-159.

Beetar, J. T., & Williams, J. M. (1995). Malingering response styles on the memory assessment scales and symptoms validity tests. *Archives of Clinical Neuropsychology, 10(1),* 52-72.

Bem, S. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42(2),* 155-162.

Bianchini, K. J., Curtis, K. L., & Greve, K. W. (2006). Compensation and malingering in traumatic brain injury: a dose-response relationship? *Clinical Neuropsychologist, 20,* 831-847.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist, 15(1),* 19-45.

Binder, L. M. (1993). Assessment of malingering after mild head trauma with the Portland Digit Recognition Test. *Journal of Clinical and Experimental Neuropsychology, 15(2),* 170-182.

Dean, A. C., Victor, T. L., Boone, K. B., & Arnold, G. (2008). The relationship of IQ to effort test performance. *The Clinical Neuropsychologist, 22,* 705-722.

Duncan, S. A., & Ausborn, D. L. (2002). The use of reliable digits to detect malingering in a criminal forensic pretrial population. *Assessment, 9(1),* 56-61.

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

Everington, C., & Olley, J. G. (2008). Implication of Atkins v. Virginia: Issues in Defining and Diagnosing Mental Retardation. *Journal of Forensic Psychology Practice, 8(1),* 1-23.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

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

Flaro, L., Green, P., & Robertson, E. (2007). Word Memory Test failure 23 times higher in mild brain injury than in parents seeking custody: the power of external incentives. *Brain Injury, 21(4),* 373-383.

Frederick, R. I. (1997). *Validity Indicator Profile manual*. Minnetonka, MN: NCS Assessments.

Graue, L. O., Berry, D. T. R., Clark, J. A., Sollman, M. J., Cardi, M., Hopkins, J., & Werline, D. (2007). Identification of feigned mental retardation using the new generation of malingering detection instruments: preliminary findings. *The Clinical Neuropsychologist, 21(6),* 929-942.

Green, P. (2003). *Green's Word Memory Test for Microsoft Windows*. Edmonton, Alberta: Green's Publishing Inc.

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

Green, P., Iverson, G. L., & Allen, L. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury, 13(10),* 813-819.

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

Greve, K. W., Bianchini, K. J., & Doane, B. M. (2006). Classification accuracy of the Test of Memory Malingering in traumatic brain injury: results of a known-group analysis. *Journal of Clinical and Experimental Neuropsychology, 28,* 1176-1190.

Greve, K. W., Springer, S., Bianchini, K. J., Black, F. W., Heinly, M. T., Love, J. M., Swift, D. A., & Ciota, M. A. (2007). Malingering in toxic exposure: classification accuracy of reliable digit span and WAIS-III digit span scaled scores. *Assessment. 14(1),* 12-21.

Heaton, R. K., Smith, Jr., H. H., Lehman, R. A. W., & Vogt, A. T. (1978). Prospects of faking believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 46(5),* 892-900.

Heinly, M. T., Greve, K. W., Bianchini, K. J., Love, J. M., & Brennan, A. (2005). WAIS Digit Span-based indicators of malingered neurocognitive dysfunction: Classification accuracy in traumatic brain injury. *Assessment, 12(4),* 429-444.

Hurley, K. E., & Deal, W. P. (2006). Assessment instruments measuring malingering used with individuals who have mental retardation: Potential problems and issues. *Mental Retardation*, *44,* 112-119.

Iverson, G. L., & Binder, L. M. (2000). Detecting exaggeration and malingering in neuropsychological assessment. *Journal of Head Trauma Rehabilitation, 15(2),* 829-858.

Iverson, G. L., & Franzen, M. D. (1994). The Recognition Memory Test, digit span, and Knox Cube Test as markers of malingered memory impairment. *Assessment, 1*(4), 323-334.

Iverson, G. L., & Franzen, M. D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology, 18(1),* 38-51.

Killgore, W. D. S., & DellaPietra, L. (2000). Using the WMS-III to detect malingering: empirical validation of the rarely missed index (RMI). *Journal of Clinical and Experimental Neuropsychology, 22(6),* 761-771.

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

Marshall, P., & Happe, M. (2007). The performance of individuals with mental retardation on cognitive tests assessing effort and motivation. *The Clinical Neurologist, 21,* 826-840.

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

Meyers, J. E., & Volbrecht, M. (1998). Validation of reliable digits for detection of malingering. *Assessment, 5(3),* 303-307.

Mittenberg, W., Azrin, R., Millsaps, C., & Heilbronner, R. (1993). Identification of Malingered Head Injury on the Wechsler Memory Scale-Revised. *Psychological Assessment, 5(1),* 34-40.

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptoms exaggeration. *Journal of Clinical & Experimental Neuropsychology, 24(8*), 1094-1102.

Mittenberg, W., Theroux-Fichera, S., Zielinski, R. E., & Heilbronner, R. L. (1995). Identification of malingered head-injury on the Wechsler Adult Intelligence Scale-Revised. *Professional Psychology: Research and Practice, 26(5)*, 491-498.

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

Rees, L. M., Tombaugh, T. N., Gansler, D. A., & Moczynski, N. P. (1998). Five validation experiments of the Test of Memory Malingering (TOMM). *Psychological Assessment, 10(1),* 10-20.

Rey, A. (1941). L'examen psychologie dans les cas d'encephalopathie tramatique. *Archives de Psychologie, 28,* 286-340.

Rey, A. (1964). *L'examen Clinique en psychologie.* Paris: Presses Universitaires de France.

Roid, G. H. (2003a). Stanford-Binet Intelligence Scales: Fifth Edition. Itasca, IL: Riverside.

Roid, G. H. (2003b). Stanford-Binet Intelligence Scales: Fifth Edition, Examiner's Manual. Itasca, IL: Riverside.

Roid, G. H. (2003c). Stanford-Binet Intelligence Scales: Fifth Edition, Technical Manual. Itasca, IL: Riverside.

Rosenquist, C., Conners, F. A., & Roskos-Ewoldsen, B. (2003). Phonological and visuo-spatial working memory in individuals with intellectual disability. *American Journal of Mental Retardation, 108(6), 403*-413.

Shipley, W. C. (1940). A self-administered scale for measuring intellectual impairment and deterioration. *Journal of Psychology, 9,* 371-377.

Simon, M. J. (2007). Performance of mentally retarded forensic patients on the Test of Memory Malingering. *Journal of Clinical Psychology, 63*, 339-344.

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(4),* 545-561.

Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: a survey of experts' practices. *Archives of Clinical Neuropsychology, 19(4),* 465-473.

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

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

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale (3rd ed.)*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale (3rd ed.)*. San Antonio, TX: The Psychological Corporation.

Weinborn, M., Orr, T., Woods, S. P., Conover, E., & Feix, J. (2003). A Validation of the Test of Memory Malingering in a forensic psychiatric setting. *Journal of Clinical and Experimental Neuropsychology, 25(7),* 979-990.

Zachary, R. A. (1986). *Shipley Institute of Living Scale: Revised manual.* Los Angeles: Western Psychological Services.

Louisiana State University
Psychology
Consent Form

Study Title: Simulated sub-normal performance on the Stanford Binet-V: An exploratory investigation.

Performance Site: Louisiana State University Psychological Services Center and Audubon Hall

Contact Information:
 Wm. Drew Gouvier, Ph.D.   Mandi Musso, B. S.   Alyse Barker, B.S.
wgouvie@tigers.lsu.edu   mmusso4@tigers.lsu.edu   abarke1@tigers.lsu.edu

This study will be looking at whether test scores can distinguish individuals faking mild mental retardation from individuals giving their best effort. Today, you will be asked to complete a demographic questionnaire as well as several activities, including providing word definitions, identifying patterns, solving number problems, providing answers to questions, repeating information, tapping blocks, and identifying previously seen objects. These tasks will take approximately three hours to complete.

Individuals at least 18 years of age who have no previous neurological or psychological diagnoses are eligible to participate in this study. Participation is voluntary. You may at any time withdraw from the study without penalty. LSU students who participate in this study will receive extra credit in their undergraduate psychology course. There are no foreseeable risks in participating in this study. In addition, standards of psychological practice in forensic and general clinical settings may benefit from the development of indices that can be used to identify people who are faking impairment in order to obtain financial compensation or other secondary gain.

All data collected will be anonymous. Your name and identifying information will in no way be linked to your test scores.

The study has been discussed with me and all my questions have been answered. I may direct additional questions regarding study specifics to the investigators. If I have questions about subjects' rights or other concerns, I can contact Robert C. Mathews, Chairman, LSU Institutional Review Board, (225)578-8692, irb@lsu.edu, www.lsu.edu/irb. I agree to participate in the study described above and acknowledge the researchers' obligation to provide me with a copy of this consent form if signed by me.

Subject Signature:_____   Date:_____
Witness Signature: _____   Date: _____

APPENDIX B: DEMOGRAPHIC QUESTIONNAIRE

Subject Number: _____

Examiner: _____

Please complete the following information.

Age: _____

Race/Ethnicity: _____

Please circle one.

Gender: M   F

Classification in College: Freshman   Sophomore   Junior   Senior   Graduate Student

Have you ever been diagnosed with a neurological disorder (e.g., epilepsy, traumatic brain injury, meningitis, encephalitis, extreme fever, stroke, hematoma)?

Yes   No

If yes, please explain _____

Have you ever been diagnosed with a psychological disorder (e.g., depression, anxiety, ADHD, learning disability, OCD, schizophrenia, bipolar disorder, alcohol or substance abuse or dependence)?

Yes   No

If yes, please explain _____

APPENDIX C: PARTICIPANT EFFORT RATING SCALE

Subject Number: _____

Examiner: _____

How much effort did you put into performing as you think someone who is mentally retarded would? (Please circle one number.)

1	2	3	4	5
Not much		Average		Very much

How successful do you think you were at performing as someone who is mentally retarded would? (Please circle one number.)

1	2	3	4	5
Not at all		Average		Very Successful

VITA

Alyse A. Barker was born and reared in southern Louisiana. She was valedictorian of E.D. White Catholic High School's class of 2004. In December of 2007, she earned a Bachelor of Science degree in psychology from Louisiana State University. Currently, she is enrolled in Louisiana State University's clinical psychology doctoral program.