THE DETECTION OF MALINGERED MENTAL RETARDATION IN HIGH- AND LOW-COGNITIVE ABILITY INDIVIDUALS

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
Daniel Anthony Proto
B.S., University of Georgia, 2004
May, 2008

TABLE OF CONTENTS

List of Tables

Abstract

The detection of malingering is an area of research that has received increasing attention in recent decades. Neuropsychologists in particular are often asked to assess the validity of symptoms such as cognitive impairment due to brain injury or toxic chemical exposure. Additionally, given the decision of the U.S. Supreme court in *Atkins v. Virginia* (2002), incentive to feign mental retardation in order to avoid capital punishment has greatly increased. However, few measures of malingering detection have been thoroughly studied for their applicability to mentally retarded individuals, and for their ability to accurately distinguish between malingerers, normal controls, and individuals with mental retardation (MR). The current study explores the use of the Bender Gestalt Test (BGT), and of the Koppitz and Lacks scoring systems in particular, as a screening device to determine the validity of claimed intellectual disability. Additionally, high- and low-cognitive ability groups were compared on their capacity to successfully malinger on the BGT, the Test of Memory Malingering (TOMM), the Rey Memory for Fifteen Items test (MFIT), and the Rey Dot-Counting test (RDCT). Results showed that high- and low-cognitive ability malingerers were not significantly different in their malingering performances, and both groups performed similarly to effortful responders with mild MR on all measures other than RDCT total errors. Also, of the TOMM, MFIT, and RDCT, only the RDCT did not misclassify high proportions (>30%) of mild MR participants as malingerers.

Introduction

Malingering, as defined by the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., Text Revision), is "the intentional production of false or grossly exaggerated physical or psychological symptoms" in the presence of external incentive, such as monetary gain or the evasion of criminal prosecution, and not attributable to mental disorder (American Psychiatric Association [APA], 2000). With prevalence rate estimates ranging from 8% in medical contexts, to 19% in forensic settings, to upwards of 30% in chronic pain and fatigue cases (Mittenberg, Patton, Canyock, & Condit, 2002), along with the high economic and societal costs associated with successful efforts (Hayes, Hale, & Gouvier, 1998; Johnstone & Cooke, 2003; Rohling, Langhinrichsen-Rohling, & Miller, 2003), the need to identify episodes of malingering is great. Psychological professionals have long-noted the importance of the detection of feigned impairment (e.g., Bash & Alpert, 1980; Hunt & Older, 1943), while simultaneously recognizing the shortcomings inherent in these processes (Faust, 1991, 1995). Accordingly, hundreds of research articles have been authored in relation to various aspects of malingering (Rogers, 1984), leading to the development of specific scoring criteria on existing measures such as the Minnesota Multiphasic Personality Inventory (MMPI and MMPI-2; Schretlen & Arkowitz, 1990; Shores & Carstairs, 1998; Walters, White, & Greene, 1988), the Halstead-Reitan Neuropsychological Test Battery (Goebel, 1983; Heaton, Smith, Jr., Lehman, & Vogt, 1978), the Personality Assessment Inventory (Wang, Rogers, Giles, Diamond, Herrington-Wang, & Taylor, 1997), the Rorschach (Bash & Alpert, 1980; Schretlen, 1988), the California Verbal Learning Test (Slick, Iverson, & Green, 2000), and versions of the Wechsler Adult Intelligence Scales (WAIS; Bash & Alpert, 1980; Johnstone & Cooke, 2003; Killgore & DellaPietra, 2000;

Langeluddecke & Lucas, 2004; Miller, Ryan, Carruthers, & Cluff, 2004). Additionally, entirely new measures, such as the Rey Memory for Fifteen Items Test (MFIT; Rey, 1964; in Lezak, 1995) and Rey Dot Counting Test (Rey, 1941; in Lezak, 1995), the Structured Interview of Reported Symptoms (SIRS; Rogers, Bagby, & Dickens, 1992), the Test of Memory Malingering (TOMM; Tombaugh, 1996), and the Green Word Memory Test (WMT; Green, Allen, & Astner, 1996), have been created for the explicit purpose of detecting feigned impairment. With such continuous advances in both theory and instrumentation, academic and practical interests in malingering appear to be steadily increasing, much as they did in the previous decade (Sweet, King, Malina, Bergman, & Simmons, 2002).

Table 1

Abbreviations Frequently Used in This Thesis

Acronym	Full Name/Description
BGT	Bender Gestalt Test
CST	Competency to Stand Trial
MR	Mental Retardation
MFIT	Rey Memory for Fifteen Items Test
MMPI	Minnesota Multiphasic Personality Inventory
PPVT-IIIB	Peabody Picture Vocabulary Test—3rd Edition, Form B
RDCT	Rey Dot Counting Test
RMI	Rarely Missed Index (derived from the WMS-III)
SVT	Symptom Validity Test
TBI	Traumatic Brain Injury
TOMM	Test of Memory Malingering
WAIS	Wechsler Intelligence Scales
WMS	Wechsler Memory Scales
WMT	Green's Word Memory Test

Literature Review

Literature searches of such databases as PsycINFO and MEDLINE return results related to malingering numbering in the hundreds, if not thousands. Topics range from the malingering of pain (Robinson, Myers, Sadler, Riley, III, Kvaal, & Geisser, 1997) to amnesia (Greiffenstein, Baker, & Gola, 1994) to hysterical blindness (Theodor & Mandelcorn, 1973). Reviewing the findings of such a large body of work would be beyond the scope of this paper, but a brief exploration of some of the common themes of malingering research would be useful. As previously noted, types of malingering can vary greatly, and so focusing on a specific context becomes necessary. This paper, then, will concern itself with the feigning of deficits in a forensic setting, and specifically with criminal forensic malingering.

Criminal Forensic Malingering. Malingering and litigation are linked in numerous circumstances, such as in worker's compensation hearings related to traumatic brain injury (TBI; e.g., Barth, Gideon, Sciara, Hulsey, & Anchor, 1986; Bianchini, Greve, & Love, 2003), toxin exposure (Bianchini et al., 2003), and many other conditions. However, in the criminal forensic domain, malingering and its research appear largely confined to the areas of competence to stand trial (CST), the feigning of mental disorders, and the feigning of cognitive deficits (Rogers & Bender, 2003). CST, an idea founded in English common law, has been defined by the U.S. Supreme Court in *Dusky v. United States* (1960) as the defendant having "sufficient and present ability to consult with his lawyer with a reasonable degree of rational understanding—and whether he has a rational as well as factual understanding of the proceedings against him" (p. 789; in Stafford, 2003). It is one of the more frequently-assessed aspects of mental health in a forensic setting, and various landmark legal decisions, such as *Jackson v. Indiana* (1972) have since been made in relation to it (Stafford, 2003). If found incompetent to stand trial, the individual is put into a treatment program to attempt to restore competence. Should such restoration be deemed impossible given the guidelines espoused by *Jackson v. Indiana* (1972), individuals are generally committed to civil facilities for treatment and observation, and, if not found to be a danger to themselves or others, are released (Stafford, 2003). Establishment of such dangers, and of CST in general, is intrinsically linked to the presence or absence of cognitive deficits and/or mental disorders.

Mental disorders in and of themselves do not automatically alleviate criminal culpability; however, they are often a key component of being found incompetent to stand trial, and of obtaining a verdict of not guilty by reason of insanity, hence the incentive to malinger their presence. However, unlike CST evaluations, in which a specific mental disorder or cognitive deficit is not necessary to be deemed incompetent a defendant and his or her attorney must prove the existence of a mental disorder or defect in order to establish legal insanity. This defect or disorder must then directly lead to an excusing condition, such as the “incapacity to appreciate the wrongfulness of one’s actions or to control one’s conduct” (p. 385, Goldstein, Morse, & Shapiro, 2003). In such cases, malingerers would be those individuals who fabricate psychopathology so as to appear insane at the time of the crime’s commission, thereby escaping criminal prosecution (Rogers & Bender, 2003). Should the insanity defense be successful, the individual is freed of all responsibility for the act in question, although he or she might still be committed to a hospital or mental institution for treatment and monitoring (Goldstein, Morse, & Shapiro, 2003).

As has been previously stated, the feigning of cognitive deficits can be used as a way of establishing a mental defect in hopes of receiving a not guilty by reason of insanity verdict, a ruling of incompetent to stand trial, or in an attempt to meet partial criteria for a diagnosis of mental retardation. Given that the U.S. Supreme court declared, via *Atkins v. Virginia* (2002), the execution of mentally retarded individuals to be cruel and unusual punishment, and therefore unconstitutional, such a diagnosis of retardation would provide powerful incentive for individuals charged with capital offenses punishable by death (Brodsky & Galloway, 2003). As pointed out by Rogers and Bender (2003), the malingering of cognitive deficits is fundamentally different from the malingering of mental disorders in that the individual, rather than fabricate psychopathology, must instead exhibit "effortful failure," thereby convincing the examiner that his or her deficits are genuine (p. 119). This paper, while recognizing the importance of CST and psychopathological research, will focus instead on this malingering of cognitive deficits.

Forensic Evaluation Perspectives. In an effort to examine the different ways in which professionals approach examinations of malingering in a forensic setting, Rogers and Bender proposed three commonly-adopted perspectives (Rogers & Bender, 2003). The first of these perspectives, the intuitional perspective, holds that professionals feel malingering and other simulation/dissimulation, such as defensiveness, will be readily identified by clinical judgment alone. However, given that clinical judgment has repeatedly been found insufficiently accurate in identifying malingerers (Faust, 1991; Guilmette & Giuliano, 1991; Heaton, Smith, Lehman, & Vogt, 1978), such a perspective is seen as lacking.

The second perspective, the standard perspective, has the examiner using traditional, well-validated psychological measures such as personality inventories and ability tests in a dualpurpose role—the measurement of the intended characteristic(s), and the identification of possible feigning (Rogers & Bender, 2003). Much research exists in this perspective, as using existent tests eliminates the need to develop entirely new, malingering-specific tools.

Because the measurement of cognitive deficits often-times is delegated to neuropsychologists—indeed, by some estimates, more than half of psychologists offering neuropsychological services have testified at least once in a legal setting, with one in five having done so ten or more times (Guilmette, Faust, Hart, & Arkes, 1990)—it is no surprise that many existing neuropsychological measures have been examined for their utility in malingering detection. Lu, Boone, Cozolino, and Mitchell (2003) found that the Rey-Osterrieth Complex Figure Test, a commonly-used neuropsychological test of visuoconstructive abilities and memory, was able to provide a sensitivity of 74% when delineating participants with suspect effort from controls and those with various forms of memory impairment. Similarly, both Tenhula and Sweet (1996) and Forrest, Allen, and Goldstein (2004) found that the Halstead Category Test, a measure of abstract reasoning and concept formation, could accurately discriminate dissimulators from controls and individuals with brain damage when using certain malingering indices. Heaton, Smith, Jr., Lehman, and Vogt (1978) used the entire Halstead-Reitan battery, a popular neuropsychological tool, along with the MMPI and the WAIS, to examine the responses of uncoached/naïve fakers and nonlitigating head-injured patients. While malingerers were able to score in the impaired range as instructed, they were differentiated from impaired patients with 94-100% accuracy (Heaton et al., 1978). A potential flaw in this study was that it contained more predictor variables (37) than subjects (32), and did not include a control group, limiting its external validity (Forrest, Allen, & Goldstein, 2004). Goebel (1983) attempted to remedy these shortcomings, and was able to identify impaired versus unimpaired individuals with 94.9-97.2% accuracy, depending upon the base rate assumed. However, significantly less success was had in differentiating malingerers from non-malingerers.

Standard measures not specific to neuropsychology have also been adapted by mental health practitioners to screen for the simulation of cognitive abilities, foremost among them perhaps being the Wechsler series of tests. Johnstone and Cooke (2003) studied the ability of incarcerated adolescent offenders, postgraduate students, and research fellows to feign “mental handicap” on the WAIS-R, finding that all groups were successfully able to lower their scores to the mentally retarded range. However, participants’ results were inconsistent with qualitative observations such as the ability to understand complex directions. Killgore and DellaPietra (2000) developed a malingering scale for the WMS-III known as the Rarely Missed Index (RMI), which consists of six items from the Logical Memory Delayed Recognition subtest of the WMS-III that are answered correctly at better-than-chance levels by naïve subjects. Accuracy of classification of malingerers and neurologically-impaired patients using the RMI was 98%, with sensitivity of 97% and specificity of 100% (Killgore & DellaPietra, 2000). Miller, Ryan, Carruthers, and Cluff (2004) found similar accuracy results (95% overall) for the RMI when detecting symptom exaggeration in substance-abusing and head trauma groups, while also examining the accuracy of a separate index, Vocabulary minus Digit Span (V-DS) on the WAIS-III. Langeluddecke and Lucas (2004), however, reported accuracy rates of only 75% for the RMI when used with litigating head injury referrals in a 28% base rate condition.

As can be seen, then, the adaptation of standard measures to malingering detection has been met with mixed success, and Rogers and Bender (2003) have stated as much. These findings have lead to the more-widespread adoption of the third of Rogers and Bender's (2003) perspectives, the specialized perspective, in which standard testing is paired with measures specifically designed to detect feigning. This is the method recommended by Rogers and Bender (2003), and one which also adheres to the multi-test malingering detection method espoused by Heaton et al. (1978) among others.

Such malingering-specific tests often rely on three malingering detection strategies described by Rogers and Bender (2003)—floor effects, symptom validity testing, and performance curves. Measures such as the Test of Memory Malingering (TOMM) and Rey's Memory for Fifteen Items test (MFIT) make use of floor effects, in which a relatively simple task, completable by ≥90% of impaired individuals, is thought by malingerers to be of greater difficulty, thereby causing them to perform worse than would individuals with genuine deficits (Rogers & Bender, 2003). Limitations exist in that results are only valid for which norms of specific types of impairment are known and to which performance by suspected malingerers can be compared (Rogers & Bender, 2003). Additionally, as illustrated by Weinborn, Orr, Woods, Conover, and Feix (2003) in relation to the TOMM, when floor effects are used in an effort to create cut-off scores, specificity is somewhat sacrificed in order to obtain greater sensitivity.

The TOMM also makes use of a second detection strategy, symptom validity testing (SVT), as described by Brady and Lind (1961; in Rogers & Bender, 2003), and first named by Pankratz and colleagues (1975, 1983; in Rose, Hall, Szalda-Petree, & Bach, 1998) and Binder and Pankratz (1987; in Crawford, Stewart, & Moore, 1989). Essentially, SVT is based on the premise that should an examinee be given a multiple-choice task, he or she would be expected to perform at a certain level by chance alone. If found to be performing significantly below this chance level, it could be assumed that the individual knew the correct answer, and instead gave the incorrect response in an attempt to feign impairment (Rogers & Bender, 2003; Rose et al., 1998). However, given that such below-chance performance is uncommon in malingerers (Martin, Hayes, & Gouvier, 1996), SVT measures, while boasting high specificity (i.e., the ability to correctly classify individuals without the condition of interest; true negative rate), generally have only moderate sensitivity (i.e., the ability to correctly detect the condition of interest; true positive rate).

Performance curves, the third commonly-used feigning detection strategy, involve the concept that individuals should do well on simpler items, and less-well on more difficult items (Rogers & Bender, 2003). Malingerers would be thought either not to know of this concept, especially were they unaware of item difficulty, or to have trouble maintaining a "normal" consistency in their responses, thus producing an abnormal performance curve (Rogers & Bender, 2003). Rey's Dot Counting test (RDCT) makes use of such principles in its malingering detection paradigm, as individuals would be expected to produce more errors and require larger amounts of time when having to count greater amounts of dots, and when dealing with ungrouped as opposed to grouped patterns of dots (Lezak, 1995).

Expert Testimony Guidelines. Given the sometimes-limited accuracies of standard and malingering-specific measures alike, along with the fact that expert witnesses often play key roles in cases involving feigning in criminal settings (Goldstein, Morse, & Shapiro, 2003), guidelines have been developed concerning such expert testimony. The most binding of these guidelines is that espoused by the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993), in which the Court ruled that expert testimony should be permitted in legal proceedings only when “the reasoning or methodology underlying the testimony is scientifically valid and…properly can be applied to the facts in issue” (in Mossman, 2003). Mossman (2003) summarizes a response by Vallabhajosula and van Gorp’s (2001) to the proposed *Daubert* standard as suggesting that any measure used for malingering detection must yield a posterior probability of malingering (i.e., accuracy) of 80% when a base rate of 30% is assumed; however, Mossman (2003) criticizes this standard as not providing enough information, arguing that an infinite combination of sensitivity and specificity values could achieve such accuracy.

Additionally, studies have shown that base rates are fundamentally important in malingering detection, although this aspect has at times received little attention in both research and practical settings (Gouvier, Hayes, & Smiroldo, 1998). Thus, little progress has been made in reaching an agreement as to a global accuracy “standard” in terms of admissible expert witness testimony and malingering detection results. The only generally-accepted tenet is that, given constitutional protections of the rights of the accused, specificity (and thus the avoidance of Type I errors) is of greater importance than sensitivity (the avoidance of Type II errors), should a trade-off between the two be necessary (Hayes, Hale, & Gouvier, 1997; Reznek, 2005). Hence the frequent use of measures such as the TOMM and MFIT, which exhibit excellent specificity at the expense of sensitivity (Lezak, 1995; O’Bryant & Lucas, 2006; Reznek, 2005). However, continued efforts must be made to increase the sensitivity of current detection methods, or to develop new methods that display excellent accuracy in all respects.

Malingering in Special Populations. Given these guidelines regarding the accuracy of expert testimony, and the level of precision required by malingering detection strategies, it is of the utmost importance that the specific population of interest be as thoroughly identified and understood as possible. However, despite this importance, relatively little work has been done with special populations in regards to malingering measures. Indeed, while aspects such as race, level of education, age, and even psychopathology have been attended to, less work has been done in areas such as low intellectual functioning and mental retardation (MR); and that work which exists is not always promising. Studying these populations, especially individuals with mental retardation, is becoming increasingly important given such rulings as *Atkins v. Virginia* (2002).

Mental Retardation and Malingering. Mental retardation, defined by the DSM-IV-TR (2000) as possessing an IQ $\leq$70, deficits in at least two areas of adaptive functioning, and onset of before 18 years of age, affects an estimated 1-3% of the population (APA, 2000; Hodapp & Dykens, 2003). Additionally, an estimated 16% of murder/insanity defendants (Lanzkron, 1963; in Hayes, Hale, & Gouvier, 1997), and 2-10% of the prison population as a whole (Petersilia, 2000, in Davis, 2000) are thought to have intellectual disabilities (ID; a term now used synonymously with MR). These individuals pose unique challenges to mental health professionals assessing malingering in forensic populations. Many typical measures of malingering detection, such as the MMPI (Keyes, 2004) and the PAI (Morey, 1996), require reading ability and levels of insight beyond those possessed by the majority of individuals with MR. Other measures, such as the Rarely Missed Index (RMI), have little or no research exploring their use, and the development of appropriate norms, with intellectually-disabled examinees. As such, the development of measures that are suitable for intellectually-disabled populations, and that can also accurately discriminate between malingered and true cognitive deficits, are necessary additions to the existent body of feigning detection tools.

While commonly-used intelligence tests would be an obvious first choice, little work has been done regarding their use in the detection of malingered mental retardation, with interest instead focusing largely on feigned brain injury or psychosis (Schretlen, 1988). However, other measures have been studied in contexts related to the malingering of mental retardation. The MFIT is one example that is simple enough in design and administration to be used with MR individuals, and that has also been previously, albeit scarcely, used to examine faked intellectual disability. Hayes, Hale, and Gouvier (1998) found that when used with the SIRS, RDCT, and M-Test, the MFIT (with a cut-off of <9 items correct) lead to correct classification of all participants across three groups—pre-trial, not guilty by reason of insanity (NGBRI), and malingering. However, when the SIRS was removed from the discriminant function analysis, the remaining measures achieved an accuracy of only 59.5%, although this value was significant at the .05 level. Misclassification of pre-trial and NGBRI group members as malingerers occurred at rates of 23.1% and 27.8%, respectively. Hurley and Deal (2006) found that the SIRS, TOMM, and MFIT all misclassified participants of below-average intellectual functioning as malingerers; only the RDCT, using a cut-off of greater than 180 seconds total completion time, had a 0% false-positive rate. Thus, based on these results, it appears as though with a cut-off of <9 items, the MFIT might misclassify non-feigning individuals with MR as malingering at unacceptably high levels. However, with lower cut-off scores, it might in fact demonstrate high specificity if not sensitivity. Additionally, the previous studies documented shortcomings related to small sample size and subject homogeneity, indicating possible areas of future improvement and study.

Bender Gestalt Test. Another measure that is appropriate for individuals with mental retardation, and that is also commonly-used by clinical psychologists (Groth-Marnat, 2003), but which has been infrequently applied to forensic settings, is the Bender Gestalt Test (BGT; Bender, 1938). Adapted by Bender from a portion of Wertheimer's (1923) original designs, the BGT is a measure of "visual-motor integration" (Koppitz, 1975) in that it requires examinees to visually perceive a series of stimuli, to accurately grasp their "gestalt" properties, and then to reproduce the figures. The measure itself consists of nine separate designs, each contained on a separate card. No standard instructions are provided or required (Bender, 1938), although generally examinees are told to reproduce each design as close to the original as possible; they are then provided with an 8 ½ x 11" sheet of blank paper, and shown the separate Bender cards one at a time. No time limit is imposed, and no specific corrections are to be given other than general statements such as "do your best," or, "do it the way you think best" (Lacks, 1984). Behavioral observations should be made during administration, such as the examinees' rotations of cards and/or paper, the general time required for reproduction of each design, any erasures or spontaneous corrections, etc. (Bender, 1938; Koppitz, 1975; Lacks, 1984).

While Bender (1938) did not advocate, and was in fact skeptical of, the use of standardized scoring systems, many have nonetheless been developed. Pascal and Suttell (1951), Canter (1968), and Lacks (1984) developed three of the more well-known scoring criteria intended for use with adult examinees, while Koppitz (1963, 1975) created the most widely-used developmental scoring system for children. All appear to be well-normed, valid, and reliable, although the Pascal-Suttell method has been criticized for its intricacy and the length of time required to examine profiles (Groth-Marnat, 2003). Additionally, performance on the BGT via many of these scoring systems has been correlated at least moderately with IQ, especially in the average and below ranges (Koppitz, 1975). However, while the BGT is still quite popular with clinical psychologists, it has fallen out of favor with many neuropsychologists due to the non-specific nature of results (Groth-Marnat, 2003), which generally are only able to differentiate "organic" from "functional" impairment. Nonetheless, its brief administration time, simple instructions, and non-traditional protocol might be useful to forensic screening applications in special populations.

Bruhn and Reed (1975) examined such applications in their study exploring the ability of the Pascal-Suttell and Canter scoring systems in differentiating college student controls, individuals told to feign "brain damage," and individuals pre-determined to have existing organic impairment. Their results suggested that neither the Pascal-Suttell nor the Canter method was able to successfully separate malingerers from impaired and unimpaired individuals. Bash and Alpert (1980) attempted to adapt the malingering of cognitive deficits guidelines first mentioned by Bender (1938) to quantitative use. These guidelines indicated that malingers would produce small, inhibited designs, would demonstrate uneven performance (i.e., higher-level performance on some designs, and lower-level performance on others), would change figure position while maintaining pattern properties (e.g., simplifying squares as circles), would maintain gestalt function of designs while altering relation or direction of individual parts, would retain the shape of a design while simplifying the symbols, and would add extraneous, complex elements to the designs. This scoring protocol was able to successfully discriminate between malingerers and three control groups, including psychotic (i.e., schizophrenic) and non-psychotic inpatients. Schretlen (1988, 1990) successfully revised and expanded the Bash and Alpert (1980) malingering criteria. While uniform discrepancies between genuinely-impaired and malingering subjects were not noted, the criteria—via a discriminant function analysis—were able to successfully differentiate between malingerers, controls (including non-malingering inmates with psychiatric illness), and individuals with mental retardation. False positive rates generally varied between 0% and 7.5% overall. However, it should be noted that the mean IQs of the mentally retarded groups involved were in the 40-50 range, suggesting moderate rather than mild retardation.

Neither the Koppitz nor Lacks method has been used in the detection of malingered mental retardation. However, both would be appropriate for MR populations. The Koppitz system, as previously mentioned, is developmental in nature, and thus intended for children. Its utility essentially disappears after age 10-12 years, at which point children tend to earn perfect scores, although it is useful after that point for individuals whose visuo-perceptual constructive abilities are below those of a typical nine year old (Koppitz, 1975). Total scores can range from zero to 30, with higher scores indicating greater numbers of errors. Scores are then compared to normative data divided into six-month age intervals to determine level of development. Test-retest reliabilities have ranged from .53 to .90 depending upon the interval between examinations, with Koppitz suggesting that some difference would be expected as children mature, and that individuals with “minimal brain dysfunction” would be expected to provide more reliable scores (Koppitz, 1975). Given that many individuals with intellectual disabilities are developmentally immature in relation to their chronological age, it is possible that a developmental, child-centered scoring strategy would be more appropriate than one intended only for adults when used with ID populations. Also, given that many malingerers might overestimate the difficulty of the test, they would obtain error scores significantly higher than those of individuals with MR, with both groups obtaining error scores greater than normal controls.

The Lacks system, conversely, is intended for use with adult participants, and is meant to screen for the existence of organic cerebral impairment. It is a restatement, refinement, and expansion of the scoring system originally proposed by Hutt and Briskin (1960). Key supports of its utility are that scores are obtained quickly, and that novice examiners are generally as accurate as experienced professionals when using it (Lacks, 1984). Interrater reliability for total error score ranged from .87 to .95, and intrarater reliability was .93. Total scores can range from zero to 12, with higher scores indicating more errors; any total score of five or greater is taken to indicate the presence of organic impairment or subpar interest and effort. As with the Koppitz method, it is possible that malingerers would overestimate the difficulty of the BGT, and thus would display significantly more errors than normal controls or MR subjects using the Lacks criteria.

Statement of Purpose

Given the relative recency of the *Atkins* (2002) decision, it should be expected that much work remains to be done in studying its repercussions in the legal and mental health communities. The ability to examine the effects of this judgment, especially in a malingering context, so close to its inception is an amazing opportunity. Additionally, given the need for more work in the area of malingering in mental retardation in general, it would behoove any researcher to attempt to combine these various aspects when possible.

The current study, then, plans to examine the performance of individuals with mild MR, normal controls, and analog (i.e., simulated) malingerers on common malingering measures, namely the TOMM, the MFIT, and the RDCT. Additionally, with the benefits offered by the BGT, and specifically by the Koppitz and Lacks scoring systems, in relation to malingering detection, the performance of individuals with mild MR in comparison to normal controls and analog malingerers will be explored on this measure. With the BGT having been at least moderately correlated with intelligence, and with this correlation increasing at the lower ranges of the IQ spectrum (Koppitz, 1975), this measure might provide a screening device uniquely appropriate for individuals with MR and/or below-average IQ in addition to those persons with average or above-average intelligence. Previous research (e.g., Bash & Alpert, 1980; Bruhn & Reed, 1975; Schretlen, 1988, 1990) has shown promising results, but replication and expansion are necessary, especially with a more-standardized and well-known scoring procedure, and in individuals with mild as opposed to more-severe MR.

Finally, given that higher intelligence has been correlated with improved ability to malinger psychopathology (Pelfrey, 2004), more work is needed in exploring the relationship between intelligence and performance on cognitive disability malingering measures. The present study, then, will attempt to examine the performances of high- versus low-cognitive ability analog malingerers on the TOMM, MFIT, RDCT, and BGT when instructed to feign mental retardation.

Research Questions and Hypotheses

Question 1: How will high-cognitive ability simulated malingerers perform compared to low-cognitive ability simulated malingerers?

Hypothesis 1a: It is hypothesized that high-cognitive ability malingerers will commit significantly fewer errors on the BGT under both the Koppitz and Lacks scoring systems than will low-cognitive ability malingerers.

Hypothesis 1b: It is hypothesized that high-cognitive ability malingerers will perform significantly better (i.e., produce fewer errors in Trial 2) on the TOMM than low-cognitive ability malingerers.

Hypothesis 1c: It is hypothesized that high-cognitive ability malingerers will perform significantly better (i.e., respond more quickly and produce fewer errors) on the RDCT than will low-cognitive ability malingerers.

Hypothesis 1d: It is hypothesized that high-cognitive ability malingerers will perform significantly better (i.e., remember more items) on the MFIT than will low-cognitive ability malingerers.

Question 2: How will analog malingerers perform on the Bender Gestalt Test when compared to normal controls and individuals with mild mental retardation?

Hypothesis 2a: It is hypothesized that analog malingerers will commit significantly more errors on the BGT under both the Koppitz and Lacks scoring systems than will individuals with mental retardation.

Hypothesis 2b: It is hypothesized that analog malingerers will commit significantly more errors on the BGT under both the Koppitz and Lacks scoring systems than will normal controls.

Method

Participants. A power analysis for alpha = .05 and power = .8 indicates that in order to obtain an effect size *d* = .5 with five groups and six response variables (Lacks score, Koppitz score, TOMM trial 2 score, RDCT total time and total errors, and MFIT score), a total sample size of 100 is recommended. Dividing this between the various conditions provides a cell size of n = 20 for each.

High- and Low-Cognitive Ability Malingerers and Controls. Individuals were recruited from undergraduate Psychology classes at Louisiana State University via the PSYC Experiments internet-based recruiting system. The experiment was open to all current students, and contained the following description:

> Eligible students (18+ years old, no current psychological diagnosis, and no pending major litigation/legal issues) will be asked to complete a variety of tasks examining areas such as memory, vocabulary, and perceptual abilities.
> The session will take place in the basement of Johnston Hall (room 33) at the Psychological Services Center, and should last approximately 45-60 minutes.

LSU student volunteers received two points of research credit for their participation. As is seen in the above recruitment message, all participants were screened for the exclusionary criteria of current psychological diagnosis, age less than 18 years, and current involvement in major (i.e., anything beyond a minor traffic offense such as a speeding ticket) litigation. Individuals in both the high- and low-cognitive ability groups were randomly divided into either the malingering or control condition before beginning the experiment.

In total, 79 participants were recruited in this manner, with six being dismissed due to existing psychological diagnosis, resulting in a final sample of 73 students. The sample consisted of 25 males and 48 females with the following self-reported ethnicities: 21 Caucasians; 9 African-Americans, 5 Asians, and 2 Hispanics. The mean cognitive ability level, as measured by PPVT-IIIB scores, was 108.4 (*SD* = 10.2). The mean age was 20.5 years (*SD* = 3.7) and the mean years of education completed was 13.2 (*SD* = 1.1).

Because of the difficulty experienced in recruiting individuals with a below-average PPVT-IIIB score, which was used as an estimate of overall cognitive ability, a median split was conducted in order to determine group membership. The median PPVT-IIIB score was determined to be 109; thus, individuals with scores of 109 and above were designated as part of the high-ability condition, and individuals with scores of 108 and below were designated as part of the low-ability condition. This resulted in group sizes of 18 for low-ability malingerers, 20 for high-ability malingerers, 16 for low-ability controls, and 19 for high-ability controls. See Table 1 for the number of males and females, ehtnicities, mean ages, mean years of education completed, and mean PPVT-IIIB scores for each of the high- and low-ability groups.

Individuals with Mental Retardation. Individuals with mild mental retardation were recruited from community housing on the campus of a large civil and forensic facility located in eastern Louisiana. Eligible participants (i.e., those diagnosed with mild mental retardation) were identified by staff, and were approached by the principle investigator in the presence of at least one staff member. Participation was on a voluntary basis, and those individuals who completed testing were paid $10 for their time. In total, 11 participants were selected, none of whom were dismissed or excluded. Brief records reviews revealed a mean WAIS-III IQ of 61.1 (*SD* = 3.6), mean age of 38 (*SD* = 7.9), and mean years of education completed of 10.8 (*SD* = 1.9). The mean PPVT-IIIB score for this group was 48.4 (*SD* = 11.8).

Of the 11 individuals, four were female and seven were male, with ten being African-American and one being Caucasian.

Measures. Written informed consent was obtained from all participants per one of two standardized consent forms prior to inclusion in the study, with one form used for LSU students (see Appendix A), and the other for individuals with mild mental retardation (see Appendix B). For individuals with mental retardation, additional steps were taken to ensure that proper consent was obtained. These steps included additional explanation of the consenting procedure, reading aloud of the consent form, and verification through staff that each individual was a competent major. The following tests were administered to all participants:

Structured Interview: A brief structured interview was developed and administered in order to obtain demographic information regarding age, race, years of education completed, current psychological diagnoses, and current involvement in major litigation. For individuals with mild MR, this information was verified via records review to ensure correctness. See Appendix C.

Bender Gestalt Test (BGT): The Bender Gestalt Test, as previously described, was administered to all participants. Because the administration instructions are not thoroughly or rigorously standardized, and are similar in both the Lacks and Kopptiz systems, those procedures detailed by Lacks (1984) were used in presentation of the nine individual stimulus cards. To aid in standardization, no behavioral observations were recorded, as neither system officially scores such observations. Total error scores using both the Lacks and Koppitz criteria were coded and recorded, along with the time required to complete each design.

Test of Memory Malingering (TOMM): The TOMM consists of 50 line drawings of common objects. Respondents are shown each object for three seconds, and are then presented with 50 forced-choice recognition items consisting of one correct response and one foil. After each item, respondents are provided with feedback regarding the correctness of their responses. Once the first trial is finished, a second is administered immediately afterwards; an optional retention trial can also be given after a 15-minute delay. A cutoff of < 45 correct responses on either Trial 2 or the delayed retention trial is recommended as an indication of suspect effort (Weinborn et al., 2003; Tombaugh, 1996). The TOMM has been shown to be resistant to effects of age, education, depression, and many neurological conditions sans a moderate or greater level of dementia (Weinborn et al., 2003). Vallabhajosula and van Gorp (2001) have also specifically mentioned the TOMM as adequately meeting the demands of the *Daubert* standard (in Weinborn et al., 2003). Its effectiveness has been supported in psychiatric forensic settings (Weinborn et al., 2003), and its sensitivity and specificity have been compared to that of the Word Memory Test (WMT; Green, Allen, & Astner, 1996), with findings declaring it to be highly specific if only moderately sensitive at base rates of malingering ranging from 10% to 70% (O'Bryant & Lucas, 2006). Additionally, in a search of published state and federal case law resources, the TOMM was one of only four tests of cognitive malingering mentioned specifically by name (Mossman, 2003). The TOMM has also been used in MR samples, with varying degrees of success (Hurley & Deal, 2006; Simon, 2007). Thus, its use in a variety of settings seems both valid and acceptable.

In this study, the TOMM was administered per standard instructions (Tombaugh, 1996). Only Trials 1 and 2 were given; the retention trial was not be used. Total number of correct responses on Trials 1 and 2 were recorded for further analysis. However, as the manual only recommends the use of Trial 2 in determining suspect effort, only Trial 2 scores will be analyzed in this study.

Rey Dot-Counting Test (RDCT): The RDCT, as described by Lezak (1995), is a simple, straight-forward task designed to assess the validity of "general cognitive impairment or specific visuoperceptual defects." It consists of 12 cards, each containing different numbers of dots. The first six cards contain ungrouped dots, and the final six contain dots grouped in various ways. Participants are asked to count the number of dots on each card as quickly and accurately as possible, with the examiner recording response time and number of dots counted for each card. Lezak (1983) proposed cut-offs of 4.8 seconds for mean grouped dot-counting time, and a mean grouped dot counting time that was not at least twice as fast as mean ungrouped dot counting time (in Boone, Savodnik, Ghaffarian, Lee, Freeman, & Berman, 1995). Boone, Lu, Back, King, Lee, Philpott, Shamieh, & Warner-Chacon (2002, in Nelson et al., 2003) suggested the use of a combination score consisting of mean grouped time, mean ungrouped time, and number of errors. Sensitivity and specificity for this method were reported to be 75-100% and ≥90%, respectively (Boone, et al., 2002; in Nelson et al., 2003). Paul, Franzen, Cohen, and Fremouw (1992; in Hurley & Deal, 2006) put forth a cut-off of greater than 180 seconds total time required as indicative of inadequate effort, which was later validated in a population of individuals with mental retardation (Hurley & Deal, 2006). Indirect support for use of time rather than errors made in certain circumstances also exists, as mentally retarded individuals attempting to malinger actually produced fewer errors than did non-malingering mentally retarded individuals (Hayes, Hale, & Gouvier, 1997). Finally, the RDCT has been shown to be resistant to the effects of depression (Lee et al., 2000), and to generally provide acceptable levels of incremental validity in a malingering detection battery (Nelson et al., 2003).

In this study, the RDCT was administered to participants per standard instructions (Lezak, 1995). The number of dots counted per card, as well as total time required per card, were recorded. From these, a total error score (i.e., number of cards incorrectly counted) and an overall total time were determined and used in further analyses. See Appendix D.

Memory for Fifteen Items Test (MFIT): The MFIT consists of a single stimulus card on which fifteen items (five rows of three items each) are printed. Examinees are told that they will be shown the card for ten seconds, after which it will be withdrawn and they must reproduce as many of the items as possible from memory. They are then given a blank piece of 8 ½” x 11” paper to make their reproduction, and the test is begun.

Like the TOMM, the MFIT is based on the premise of appearing very difficult while in actuality being quite easy to complete (Lezak, 1995). However, despite its simple, straight-forward administration and scoring (generally only the raw number of missed or incorrectly recalled characters is recorded), there has been controversy regarding an appropriate cut-off score. Lezak (1983), in her review of MFIT literature, suggested a cut-off of recalling fewer than three of the five complete character sets (nine items total). This cut-off was later supported, after modification to fewer than nine total items recalled rather than three complete character sets, when applied to psychiatric and mentally retarded inpatients (Goldberg & Miller, 1986), brain-damaged and comparison groups (Bernard & Fowler, 1990), criminal defendants referred for inpatient forensic evaluation (Simon, 1994), and in differentiating TBI subjects from probable malingerers (Greiffenstein, Baker, & Gola, 1994), and has been found resistant to the effects of depression (Lee, Boone, Lesser, Wohl, Wilkins, & Parks, 2000). A meta-analysis of 13 studies reporting MFIT data suggested that with a cut-off of nine, the measure has low sensitivity but high specificity (90%) when excluding individuals with mental retardation (Reznek, 2005). On occasion, cut-off scores above eight items correct have been suggested, although these results were reported in only small sample sizes (Taylor, Kreutzer, & West, 2003). The MFIT has also been shown to provide incremental validity when included among a battery designed for malingering detection (Nelson, Boone, Dueck, Wagener, Lu, & Grills, 2003), which is important considering the breadth of literature suggesting the use of multiple tests rather than a single measure for such tasks (Bash & Alpert, 1980; Heaton et al., 1978). While different versions of the measure have been developed (e.g., Fisher & Rose, 2004; Griffin, Glassmire, Henderson, & McCann, 1997), the original MFIT still appears to be the most popular variant (Kreznek, 2005; Taylor, Kreutzer, & West, 2003).

In this study, the MFIT was administered per standard instructions (Lezak, 1995). The only value recorded was the total number of figures correctly recalled. As per previous research and recommendations (e.g., Hayes, Hale, & Gouvier, 1997; Lezak, 1995), a cut-off score of <8 rather than <9 total correctly-recalled items will be used as an identifier for suspect effort. See Appendix E.

Peabody Picture Vocabulary Test—3rd Edition (PPVT-III): The PPVT-III is a brief test of "receptive vocabulary over a wide age range using a very nonthreatening approach," consisting of two alternate and parallel forms (Form IIIA and Form IIIB) of 204 items each (Williams & Wang, 1997). The forms displayed excellent coefficient alpha (.92 to .98, median .95) and split-half (.86 to .97, median .94) reliabilities across 25 different age group categories.

Coefficients of equivalence between the two forms were also high, ranging from .88 to .96, with a median of .94 across age groups. Test-retest stability coefficients ranged from .91 to .94 across a delay interval of between 8 and 203 days, mean 42 days. Correlations in adults between the PPVT-III and the Kaufman Adolescent & Adult Intelligence Test (KAIT, Kaufman & Kaufman, 1993), a measure of general intelligence, were .76 (Form IIIA) and .85 (Form IIIB) for Fluid IQ, .87 (Form IIIA) and .91 (Form IIIB) for Crystallized IQ, and .85 (Form IIIA) and .91 (Form IIIB) for Composite IQ.

The standardization sample for the PPVT-III was demographically balanced across a variety of categories (age, race/ethnicity, gender, socioeconomic status, geographic region, and level of education) to match, as closely as possible, data from the *Current Population Survey, March 1994* conducted by the Bureau of the Census for the Bureau of Labor Statistics (Williams & Wang, 1997). Additionally, the PPVT-III was validated with a variety of special populations, including children/adolescents (N=44) and adults (N=41) with mild mental retardation. Average scores for these groups ranged from 59.3 (Form IIIA) and 58.5 (Form IIIB) in adults to 75.2 (Form IIIA) and 74.4 (Form IIIB) in children/adolescents.

As previously noted, the PPVT-III consists of 204 test items, each with its own accompanying illustration page in the test booklet. The illustration pages are divided into quarters, with a separate line drawing portrayed in each quadrant. The test items are read aloud to the examinee one at a time, and after each item the examinee is asked to choose, via pointing to the design or saying aloud its number, which illustration best defines or shows the meaning of the test item. Research has shown that the PPVT-III has been effectively used with children with Down Syndrome (Fidler, Most, & Guiberson, 2005) and Autism (Condouris, Meyer, & Tager-

Flusberg, 2003). Additionally, while the PPVT-III tends to underestimate full-scale WAIS-III IQ scores above the High Average range, it accurately estimates full-scale IQ in the Average and High Average ranges and is related to full-scale and verbal IQ overall (Bell, Lassiter, Matthews, & Hutchinson, 2001). Finally, the PPVT-III has successfully been used to screen cognitive abilities in adults with developmental disabilities residing in state care facilities, showing strong correlations with both the Kaufman Brief Intelligence Test and the Wide Range Achievement Test—3rd Edition, with correlations ranging from .72 to .92 (Powell, Plamaondon, & Retzlaff, 2002). It was felt that the PPVT-III would be the most easily-administered and appropriate brief screening device to use in this study, given that many of the individuals with mild MR might be illiterate, or have difficulty with other, popular screening devices, such as the Wechsler Test of Adult Reading, that require reading words aloud.

In this study, the PPVT-III, form B, was administered and scored per standard instructions to produce a scaled score that was used as an estimate of cognitive ability.

Subject Rating Scale: A subject rating scale, adapted from that used by Tombaugh (1996), was administered to all participants in the two malingering conditions. It asked the participant how much effort they put into their faking attempt, and how successful they felt they were in faking mental retardation. Both questions were rated on a 5-point Likert scale. See Appendix F.

Design and Procedure

High- and Low-Cognitive Ability Malingerers: All malingerers were first given the informed consent form to explain the confidentiality of their participation and responses. Participants' names were recorded on this form only, after which they were randomly assigned an identification number for use on all subsequent records. Participants were then administered the structured interview, and any individuals meeting exclusion criteria were thanked for their time and dismissed, while also receiving research credit for signing up. For individuals who met inclusion criteria, the PPVT-III form B was administered per standard instructions. Participants then received the following set of instructions, adapted from those used by Tombaugh (1996), regarding the tests that are to come, and how they are to perform:

> In this study, you will be asked to complete a variety of tasks. These tasks are often included in test batteries that examine whether or not a person suffers from some level of mental impairment.
>
> As you take each test, I would like you to assume the role of someone who has been convicted of first-degree murder, and that in your case, the death penalty has been recommended. Your lawyer has advised you of a recent Supreme Court decision that if you are mentally retarded, the court cannot sentence you to death.
>
> Your job today is to try to respond on the tests in a way that will convince the tester that you have mild mental retardation. Keep in mind that if you do not succeed in this deception, you will likely receive the death penalty. Also, be aware that with such a serious charge, there will be the suspicion that people might try to fake intellectual disability to avoid the death penalty. This means that your responses must be believable. Major exaggerations, such as not being able to do anything, remembering absolutely nothing, or completely failing to respond, are easy to detect.

Following these instructions, participants were administered the Bender Gestalt Test, the MFIT, the first two trials of the TOMM, and finally the RDCT, in fixed order. Participants were then given the subject rating scale, thanked for their time, allowed the opportunity to ask questions regarding the experiment, and dismissed. Results of the various assessments were then calculated. Additionally, it was planned that any individual indicating, via the participant rating scale, to have not put forth at least an “average” amount of effort (score of three or greater on question two of the rating scale) would not have their results included in the final analyses. However, no participants were excluded based on this effort factor.

High- and Low-Cognitive Ability Controls: The same procedures as those used for high- and low-ability malingerers were followed for high- and low-ability controls, with the exception of the instructions given. The following, non-malingering instructions were used instead:

> In this study, you will be asked to complete a variety of tasks. These tasks are often included in test batteries that examine whether or not a person suffers from some level of mental impairment.
>
> As you take each test, I would like you to perform to the best of your ability. That is, just try your hardest on each of the different tasks that you are given.

Individuals with Mental Retardation: The same procedures and instructions used for the two control groups were used for individuals with mental retardation, with the exception of testing location. For high- and low-ability participants, all testing occurred at the Psychological Services Center of LSU in Johnston Hall. For mild MR participants, testing occurred on their homes/units, which were located on the campus of a large civil and forensic state-run mental health facility in eastern Louisiana. Additional explanation of the instructions and the various tasks were given as necessary. Only those individuals who had previous documentation of having met DSM-IV-TR diagnostic criteria standards for mild Mental Retardation were included in this sample. Consent was obtained from the participant directly in all instances, as all MR participants were competent majors. At the conclusion of the testing, as with the high- and low-ability controls, participants were allowed to asked questions, provided with their $10 compensation, thanked for their time, and dismissed.

All other measures of this study (BGT, MFIT, TOMM, RDCT, and subject rating scale) were given as per the high- and low-ability conditions. The instructions used, as noted, were identical to those received by the control groups.

Results

Demographic Analysis: To examine equivalence across groups on demographic variables (i.e., age, education, sex, and ethnicity) and PPVT-IIIB scores, various analyses were run. A one-way, between-groups ANOVA revealed significant differences on age, $F(4,79) = 38.6$, $p<.001$, years of education completed, $F(4,79) = 9.3$, $p<.001$, and PPVT-IIIB scores, $F(4,79) = 180.7$, $p<.001$. Post-hoc testing in the form of Tukey's HSD revealed that for both age and years of education, the mild MR group differed significantly from all other groups (low- and high-ability malingerers and low- and high-ability controls), while none of the non-MR groups significantly differed from one another. In relation to PPVT-IIIB scores, the two low-ability groups (malingering and control) did not differ from one another, nor did the two high-ability groups (malingering and control) differ from one another. However, both low-ability groups differed significant from both high-ability groups, and all groups differed significantly from the mild MR group. See Table 2.

Multiple independent samples t-tests, with alpha set at .005 to control for experiment-wise error rates (.05/10 = .005), were run to determine if groups differed in terms of sex and ethnicity. The results revealed that no groups differed from one another significantly in relation to gender, and the high- and low-congitive ability malingerers and controls did not differ significantly from one another on either characteristic. However, the mild MR group did significantly differ in terms of race from high-ability malingerers, $t\ (29) = -3.12$, $p = .003$, and high-ability controls, $t\ (28) = 3.96$, $p < .001$. See Table 2.

Table 2

Demographic Characteristics of Study Participants

Demographic Characteristic	Low-Ability Malingerers (N=18)	High-Ability Malingerers (N=20)	Low-Ability Controls (N=16)	High-Ability Controls (N=19)	Mild Mental Retardation (N=11)
Mean Age, Years (SD)	21.06 (2.56)	19.50 (1.10)	21.94 (7.06)	19.84 (1.50)	38.09 (7.92)*
Mean Education, Years (SD)	13.39 (1.09)	13.00(.92)	13.19 (1.10)	13.21 (1.23)	10.82 (1.89)*
Sex	M = 5 F = 13	M = 7 F = 13	M = 6 F = 10	M = 7 F = 12	M = 7 F = 4
Ethnicity	Cauc. = 14 Afr. Am. = 3 Asian = 0 Hispanic = 1	Cauc. = 18 Afr. Am. = 1 Asian = 0 Hispanic = 1	Cauc. = 9 Afr. Am. = 3 Asian = 4 Hispanic = 0	Cauc. = 16 Afr. Am. = 2 Asian = 1 Hispanic = 0	Cauc. = 1* Afr. Am. = 10* Asian = 0 Hispanic = 0
Mean PPVT-IIIB Score (SD)	100.67 (5.56)[a]	115.40 (6.02)[b]	98.88 (6.45)[a]	116.21 (7.61)[b]	48.36(11.79)[c]

*Indicates a value that is significantly different ($p < .005$) from every other value in that row
[abc] Values with identical superscript characters are not significantly different ($p < .005$) from one another

Hypothesis 1: To test hypothesis 1 (1a through 1d), a between-subjects MANOVA using group membership (high-cognitive ability malingerer and low-cognitive ability malingerer) as the independent variable, and Koppitz score, Lacks score, TOMM raw score, MFIT number correct, RDCT total time, and RDCT number of errors as the dependent variables, was initially proposed. However, upon further analysis and review of the data, it was determined that a MANOVA would be inappropriate, as none of the variables included was normally-distributed.

Thus, the nonparametric procedure of a Mann-Whitney *U* test was chosen to examine differences in performances on these measures between independent groups.

While the hypothesis itself would suggest the use of 1-tailed significance testing, based on the exploratory nature of this study, it was decided to instead examine the 2-tailed results. In contrast to the literature, it was felt that low-ability malingerers might actually have been more successful at malingering, given that they function intellectually in a range closer to MR than do high-ability malingerers. However, the tests revealed no significant differences between high- and low-ability malingerers on any of the experimental measures. See Table 3.

Table 3

Mann-Whitney U Test Results for Malingerers on the BGT, TOMM, RDCT, and MFIT (N = 38)

	Koppitz Score	Lacks Score	MFIT Score	TOMM Trial 2 Score	RDCT Total Time	RDCT Total Errors
Z value	-1.18	-.93	-.33	-.75	-.95	-.97
Significance	.24	.35	.75	.45	.34	.33
	Low-Cognitive Ability Malingerers (n = 18)					
Mean Rank	21.72	21.22	18.89	18.08	17.69	21.33
	High-Cognitive Ability Malingerers (n = 20)					
Mean Rank	17.50	17.95	20.05	20.78	21.13	17.85

Hypothesis 2: To test hypothesis 2 (both 2a and 2b), a between-subjects MANOVA, using collapsed group membership (malingerer, control, mild MR) as the independent variable

and the Koppitz and Lacks scoring system results for these groups as the dependent variables. The collapsing of the high- and low-malingering groups was not felt to confound the results, as the two groups—as previously mentioned—did not differ significantly in their malingering performances.

However, as with hypothesis 1, it was determined that a MANOVA would be inappropriate due to the non-normality of the dependent variables. As such, a nonparametric, K-samples Kurskal-Wallis Test was run. The test was significant for both the Koppitz score, $\chi^2(2, N = 84) = 57.95$, $p < .001$, and the Lacks score, $\chi^2(2, N = 84) = 64.75$, $p < .001$.

Follow-up tests in the form of Mann-Whitney *U* tests, were conducted to examine the pairwise differences between groups. See Table 4.

Hypothesis 2a: The results of the pairwise tests indicated that mild MR participants performed significantly worse (i.e., produced more errors) than did malingerers on the Lacks score, $z = -3.57$, $p < .001$, but not on the Koppitz score, $z = -1.61$, $p = .108$. Mean ranks for mild MR participants on the Lacks and Koppitz scores were 38.27 and 31.05, respectively. Mean ranks for malingerers were 21.16 on the Lacks scores and 23.25 on the Koppitz score.

Hypothesis 2b: According to the Mann-Witney *U* Test, malingerers performed significantly worse than did controls on both the Lacks score, $z = -7.33$, $p < .001$, and the Koppitz score, $z = -7.06$, $p < .001$. Mean ranks for the controls were 18.34 on the Lacks scores and 19.00 on the Koppitz score. Mean ranks for malingerers were 54.18 on the Lacks score and 53.58 on the Koppitz score.

Table 4

Mann-Whitney U Test Results for the TOMM, MFIT, and RDCT

	MFIT Score	TOMM Trial 2 Score	RDCT Total Time	RDCT Total Errors
	Malingerers vs. Controls (N = 75)			
Z value	-6.32**	-6.34**	-6.52**	-5.98**
Significance	.000	.000	.000	.000
	Malingerers vs. Mild MR (N = 46)			
Z value	-1.90	-1.72	-.53	-2.59*
Significance	.057	.086	.598	.010
	Controls vs. Mild MR (N = 49)			
Z value	-6.42**	-5.46**	.4.92**	-2.08**
Significance	.000	.000	.000	.000

* Denotes a difference significant at the $p < .05$ level
** Denotes a difference significant at the $p < .001$ level

Unproposed Analyses*:* As an additional, exploratory analysis not listed in the initial thesis proposal, results were obtained in order to compare the performance of mild MR participants with controls and malingerers on the TOMM trial 2, RDCT, and MFIT. As with the previous analyses, parametric tests were deemed inappropriate due to the nonnormality of the dependent variables. Instead, a Kurskal-Wallis Test was run to explore for significant differences between the three participant groups. Results for the test were significant for the TOMM trial 2, $\chi^2(2, N = 84) = 49.08$, $p < .001$, the MFIT, $\chi^2(2, N = 84) = 43.25$, $p < .001$, the RDCT total time,

$\chi^2(2, N = 84) = 50.51$, $p < .001$, and the RDCT total number incorrect, $\chi^2(2, N = 84) = 37.30$, $p < .001$.

Follow-up tests in the form of Mann-Whitney *U* Tests were conducted to examine pairwise differences. Results indicated significantly worse performance by malingerers compared to controls on all four measures (TOMM trial 2, MFIT, RDCT total time, and RDCT total number incorrect), significantly worse performance by malingerers compared to individuals with mild MR on RDCT total number of errors, and significantly worse performance by individuals with mild MR compared to controls on the TOMM trial 2, the MFIT, and RDCT total time. Results are summarized in Table 5.

Table 5

Identifications and Misidentifications of Participants Using Recommended Cut-Scores for the TOMM, MFIT, and RDCT

Experimental Group Membership	Classification Based on Measure		
	Malingerer	Non-malingerer	% Misidentified
	MFIT < 8 Recalled (Sensitivity = 36%, Specificity = 83%)		
Malingerer (n = 38)	14	24	63%
Control (n = 35)	0	35	0%
Mild MR (n = 11)	8	4	73%
	TOMM Trial 2 Score < 45 (Sensitivity = 68%, Specificity = 91%)		
Malingerer (n = 38)	26	12	32%
Control (n = 35)	0	35	0%
Mild MR (n = 11)	4	8	36%
	RDCT Total Time > 180 seconds (Sensitivity = 18%; Specificity = 83%)		
Malingerer (n = 38)	7	31	82%
Control (n = 35)	0	35	0%
Mild MR (n = 11)	0	11	0%

Discussion

The malingering of cognitive disabilities and psychopathology in a forensic setting carries with it significant financial and societal costs. Successful malingerers might be found incompetent to stand trial or not guilty by reason of insanity, thereby escaping significant punishment. Additionally, those individuals found incompetent would be further evaluated in an effort to restore them to competence, and thus would artificially inflate costs required for such programs. Given the recent Supreme Court decision in *Atkins v. Virginia* (2002), the need to identify potential malingerers, especially in relation to mental retardation, is greater now than ever. However, along with this need for the identification of malingering comes the need to ascertain the appropriateness of effort tests in the mentally retarded population. With previous research having identified potential problems in using such popular measures as the MMPI-2 (Keyes, 2004) and the TOMM (Hurley & Deal, 2006) in the assessment of individuals with intellectual disabilities, much work remains to be done to determine which tools are, in fact, suited for this population. The aim of the current study was to examine the performance of individuals with mild mental retardation, in comparison to analog malingerers and effortful responders, on many popular malingering scales. An additional goal was to explore the effects of high- versus low-cognitive ability as a factor of performance on cognitive ability effort tests, given that previous research has shown higher intelligence to be associated with more success in malingering psychopathology (Pelfrey, 2004).

This study made use of the simulation design of malingering, as explained by Rogers (1997), in which individuals are instructed to purposefully feign a specific impairment—in this case mild mental retardation—and are compared to effortful responders (i.e., controls) and clinical groups with genuine disorders. This type of design allows for maximum experimental control, while decreasing generalizability of results. Given the exploratory nature of the study, and of this pilot use of the Koppitz and Lacks scoring systems for the Bender Gestalt in the assessment of malingering, the control-for-generalizability trade-off was deemed appropriate and necessary.

Internal validity was felt to have been achieved by randomly assigning non-MR participants to the various malingering and control groups. The statistical demographic analysis supported this idea, as there were no significant differences between the non-MR groups in relation to any of the demographic factors. However, given that all non-MR participants were students attending LSU, they likely would not be representative of many other groups, such as older individuals, or people with fewer years of education. Additionally, the mild MR group was significantly different from high- and low-cognitive ability malingerers in terms of race, with the MR group having a significantly larger proportion of African-Americans, and a significantly smaller proportion of Caucasians. Given this finding, comparisons between individuals with mental retardation and high-ability malingerers and controls cannot be fully attributed to the experimental conditions (i.e., effortful malingering vs. effortful responding), and is instead confounded by race.

Also, it was originally hoped to have recruited a mild MR sample in which no individuals carried comorbid psychiatric or personality disorders. However, securing such a sample proved to be exceedingly difficult, and resulted in ten of the eleven mild MR participants being accepted despite dually-diagnosed psychological conditions. Thus, internal validity is further compromised with respect to the attribution of causality of performance (i.e., it is difficult to attribute the performance of mild MR individuals to the fact that they are MR as opposed to the effects other psychological disorders), although external validity and the applicability of these results to real-world settings might consequently be improved.

Examining the results of the study first in relation to hypothesis 1, it was found that there were no significant differences in performance on any of the included effort measures between malingerers with high vs. low cognitive ability. Thus, hypothesis 1 (parts 1a through 1d) was proven false. It must be noted that the small difference in estimated cognitive abilities (PPVT-IIIB scores) between these two groups, due to the use of a median split and the difficulty in recruiting low-ability participants, could be at least partially responsible for this lack of a difference in malingering performance. While the low-ability groups were significantly below the high-ability groups in terms of cognitive ability, the split between them was not as large as had originally been hoped for. However, these results did allow for the collapsing of the high- and low-ability malingerers into a single malingering group during subsequent analyses.

Exploring the results of hypothesis 2, it was seen that malingerers performed significantly worse on the Koppitz and Lacks scores for the BGT than did controls, thus supporting hypothesis 2a. However, with respect to hypothesis 2b, mild MR participants performed similarly to malingerers on the Koppitz scores, while actually performing worse on the Lacks scores. MR participants thus exhibited results opposite to what were expected, thereby providing no support for hypothesis 2b.

In terms of the exploratory, unproposed and unhypothesized analyses, more-promising outcomes were existent. As would have been expected, malingerers performed significantly worse than controls on all of the common malingering measures (TOMM trial 2, MFIT, and

RDCT). Of greater interest is the fact that malingerers also performed significantly worse than individuals with mild mental retardation on one measure—the total number of incorrect responses to the RDCT. Also of note would be the finding that individuals with mild MR performed significantly worse than controls on all measures other than RDCT total number of incorrect responses. These results would suggest that total number of errors on the RDCT might prove to be a useful indicator of suspect effort in individuals thought to be malingering mild mental retardation. Additionally, the data caution that individuals with MR perform worse on common effort tests than do healthy controls.

In order to identify the extent to which this lowered performance by individuals with mild MR might lead to misclassification using the TOMM, MFIT, and RDCT, standard cut-offs for suspect effort (fewer than 45 correct responses on TOMM trial 2, 8 or fewer items remembered on the MFIT, and greater than 180 seconds total to finish the RDCT), as identified in prior research and testing manuals, were used to determine sensitivity (true positives/[true positives + false negatives]) and specificity (true negatives/[true negatives + false positives]) (Hayes, Hale, & Gouvier, 1997; Paul et al., 1992; Tommbaugh, 1996) values. Manual examination of the data yielded the results shown in Table 6. At first glance, all three measures would appear acceptable or near-acceptable in terms of sensitivity and specificity, with only the MFIT crossing below 90% specificity, and the TOMM trial 2 having a rather strong sensitivity of 68%. However, given that all false positives occurred in the mild MR sample, the outcome is a bit more sobering, with the TOMM having misidentified 36% and the MFIT having misidentified 72% of effortful mild MR responders as potentially malingering. Commensurate with previous research, only the RDCT total time variable provided an acceptable level of specificity (100%) in the mild MR group, while conversely displaying the lowest sensitivity (Hurley & Deal, 2006). Thus, the results of this study support the notion that only RDCT total time is appropriate for use in mild MR populations, with MFIT and TOMM trial 2 cut-off scores misclassifying unacceptably high levels of intellectually disabled individuals.

Further review of these data would suggest that the usefulness of the Koppitz and Lacks systems in differentiating genuine mental retardation from malingered cognitive deficit to be very limited. Poor performance on either measure would, based upon the results of this study, suggest only that an individual—in comparison to normal controls—is either potentially mildly mentally retarded or malingering, and as such would be of little use in the differentiation of malingered and genuine intellectual disability.

Future directions to be taken could explore the results of this study further, especially in relation to the trends uncovered in the unproposed/additional analyses. Given that malingerers produced significantly greater numbers of errors than individuals with mild MR, developing a cut-off score for these errors as an identifier of suspect effort might be a possibility.

Also, given the success had by Bash and Alpert (1980), and later by Schretlen (1988, 1990) in using their set of malingering criteria on the Bender Gestalt, the use of this scoring methodology on the present study's data set might yield promising results. Such an approach was initially ruled against in favor of the wider use of, familiarity with, and research base held by the Koppitz and Lacks systems. In light of the findings of the current study, though, the examination of the Schretlen (1990) criteria would now seem to be the next logical step.

Finally, as the sample sizes for nearly all groups were lower than the 20 individuals initially proposed and expected, future work could aim at replication and expansion—such as with the aforementioned Schretlen malingering criteria and the RDCT total errors cut-score—of the current design in a larger sample. Obtaining more voluntary mild MR participants, especially those who are community-dwelling and without comorbid psychiatric diagnoses, might particularly help to increase internal validity and possibly produce more positive findings. Given that the current study's participants with mild MR were housed in assisted-living facilities, their performances might not be representative of the mild MR population as a whole. This facet might also explain the poor specificity of the TOMM in relation to misclassifying mild MR participants as malingerers. Indeed, Simon (2007) specifically mentions community-dwelling versus assisted-living status as a potential reason for the disparity between his results, which supported the use of the TOMM in MR individuals, and those of Hurley & Dean (2006), whose assisted-living participants performed more similarly to those examined in the current study.

In conclusion, then, the Koppitz and Lacks scoring systems proved to be unsuccessful in potentially differentiating malingered from genuine mental retardation. However, promising findings were reported with respect to RDCT total time, with malingerers performing significantly worse than both controls and individuals with mild MR. Future studies should focus on increasing sample size, gathering a more-representative mild MR group, and examining the possible use of the Schretlen (1990) modification of Bash & Alpert's (1980) BGT malingering score and establishing an RDCT cut-off value.

References

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

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

Barth, J., Gideon, D., Sciara, A., Hulsey, P., & Anchor, K. (1986). Forensic aspects of mild head trauma. *Journal of Head Trauma Rehabilitation, 1,* 63-70.

Bash, I. & Alpert, M. (1980). The determination of malingering. *Annuals of the New York Academy of Sciences, 347,* 86-99.

Bell, N. L., Lassiter, K. S., Matthews, T. D., & Hutchinson, M. B. (2001). Comparison of the Peabody Picture Vocabulary Test—Third Edition and Wechsler Adult Intelligence Scale—Third Edition with University Students. *Journal of Clinical Psychology, 57,* 417-422.

Bender, L. (1938). *A Visual-Motor Gestalt Test and its Clinical Use.* New York: The American Orthopsychiatric Association.

Bernard, L. C. & Fowler, W. (1990). Assessing the validity of memory complaints: Performance of brain-damaged and normal individuals on Rey's task to detect malingering. *Journal of Clinical Psychology, 46,* 432-436.

Bianchini, K. J., Greve, K. W., & Love, J. M. (2003). Definite malingered neurocognitive dysfunction in moderate/severe traumatic brain injury. *Clinical Neuropsychologist, 17,* 574-580.

Bianchini, K. J., Houston, R. J., Greve, K. W., Irvin, T. R., Black, F. W., Swift, D. A., & Tamimie, R. J. (2003). Malingered neurocognitive dysfunction in neurotoxic exposure: An application of the Slick criteria. *Journal of Occupational and Environmental Medicine, 45,* 1087-1099.

Binder, L. & Pankratz, L. (1987). Neuropsychological evidence of a factitious memory complaint. *Journal of Clinical and Experimental Neuropsychology, 9,* 167-171.

Boone, K. B., Savodnik, I., Ghaffarian, S., Lee, A., Freeman, D., & Berman, N. G. (1995). Rey 15-item memorization and dot counting scores in a "stress" claim worker's compensation population: Relationship to personality (MCMI) scores. *Journal of Clinical Psychology, 51,* 457-463.

Boone, K. B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., Shamieh, E., & Warner-Chacon, K. (2002). Sensitivity and specificity of the Rey dot counting test in patients with suspect effort and various clinical samples. *Archives of Clinical Neuropsychology, 17,* 1-19.

Brady, J., & Lind, D. (1961). Experimental analysis of hysterical blindness. *Archives of General Psychiatry, 4,* 331-339.

Brodsky, S. L., & Galloway, V. A. (2003). Ethical and professional demands for forensic mental health professionals in the post-*Atkins* era. *Ethics & Behavior, 13,* 3-9.

Bruhn, A. R. & Reed, M. R. (1975). Simulation of brain damage on the bender-gestalt test by college subjects. *Journal of Personality Assessment, 39,* 244-255.

Canter, A. (1968). BIP Bender test for the detection of organic brain disorder: Modified scoring method and replication. *Journal of Consulting and Clinical Psychology, 32,* 522-526.

Condouris, K., Meyer, E., & Tager-Flusberg, H. (2003). The relationship between standardized measures of language and measures of spontaneous speech in children with autism. *American Journal of Speech-Language Pathology, 12,* 349-358.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1993).

Davis, L. A. (2000). *People with mental retardation in the criminal justice system.* Retrieved January 22, 2006, from http://www.thearc.org/faqs/crimjustice.doc.

Dusky v. United States, 362 U.S. 402, 80 S. Ct. 788 (1960).

Faust, D. (1991). Forensic neuropsychology: The art of practicing a science that does not yet exist. *Neuropsychology Review, 2,* 205-231.

Faust, D. (1995). The detection of deception. *Neurologic Clinics, 13,* 255-265.

Fidler, D. J., Most, D. E., & Guiberson, M. M. (2005). Neuropsychological correlates of word identification in Down syndrome. *Research in Developmental Disabilities, 26,* 487-501.

Fisher, H. L. & Rose, D. (2004). Comparison of the effectiveness of two versions of the Rey memory test in discriminating between actual and simulated memory impairment, with and without the addition of a standard memory test. *Journal of Clinical and Experimental Neuropsychology, 27,* 840-858.

Forrest, T., Allen, D., & Goldstein, G. (2004). Malingering indexes for the Halstead category test. *The Clinical Neuropsychologist, 18*, 334-347.

Goebel, R. (1983). Detection of faking on the Halstead-Reitan neuropsychological test battery. *Journal of Clinical Psychology, 39,* 731-742.

Goldberg, J. O. & Miller, H. R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology, 42,* 792-795.

Gouvier, W. D., Hayes, J., & Smiroldo, B. (1998). The significance of base rates, test sensitivity, test specificity, and subjects' knowledge of symptoms in assessing TBI sequelae and malingering. In C. Reynolds (Ed.), *Detection of Malingering during Head Injury Litigation* (pp. 55-79). New York: Plenum Press.

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.

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

Griffin, G. A., Glassmire, D. M., Henderson, E. A., & McCann, C. (1997). Rey II: Redesigning the Rey screening test of malingering. *Journal of Clinical Psychology, 53,* 757-766.

Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4th ed.). Hoboken, NJ: John Wiley & Sons.

Guilmette, T., Faust, D., Hart, K., & Arkes, H. (1990). A national survey of psychologists who offer neuropsychological services. *Archives of Clinical Neuropsychology, 5,* 373-392.

Guilmette, T. & Giuliano, A. (1991). Taking the stand: Issues and strategies in forensic neuropsychology. *The Clinical Neuropsychologist, 5,* 197-219.

Hayes, J., Hale, D., & Gouvier, W. D. (1997). Do tests predict malingering in defendants with mental retardation? *The Journal of Psychology, 131,* 575-576.

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

Hiscock, M. & Hiscock, C. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology, 11,* 967-974.

Hodapp, R. M. & Dykens, E. M. (2003). Mental retardation (intellectual disabilities). In E. J. Mash & R. A. Barkley (Eds.) *Child Psychopathology (2nd ed.)* (pp. 486-519). New York: The Guilford Press.

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.

Hutt, M. L. & Briskin, G. J. (1960). *The Clinical Use of the Revised Bender Gestalt Test,* New York: Grune & Stratton.

Jackson v. Indiana, 406 U.S. 715, 92 S. ct. 1845 (1972).

Johnstone, L. & Cooke, D. (2003). Feigned intellectual deficits on the Wechsler Adult Intelligence Scale-Revised. *British Journal of Clinical Psychology, 42,* 303-318.

Keyes, D. W. (2004). Use of the Minnesota Multiphasic Personality Inventory (MMPI) to identify malingering mental retardation. *Mental Retardation, 42,* 151-153.

Koppitz, E. M. (1963). *The Bender Gestalt Test for Young Children.* New York: Grune & Stratton.

Koppitz, E. M. (1975). *The Bender Gestalt Test for Young Children, Volume II, research and application, 1963-1973.* New York: Grune & Stratton.

Lacks, P. (1984). *Bender Gestalt Screening for Brain Dysfunction.* New York: John Wiley & Sons.

Lanzkron, J. (1963). Murder and insanity: A survey. *American Journal of Psychiatry, February,* 754-758.

Lee, A., Boone, K. B., Lesser, I., Wohl, M., Wilkins, S., & Parks, C. (2000). Performance of older depressed patients on two cognitive malingering tests: False positive rates for the Rey 15-item memorization and dot counting tests. *The Clinical Neuropsychologist, 14,* 303-308.

Lezak, M. (1983). *Neuropsychological Assessment* (2nd ed.). New York: Oxford University Press.

Lezak, M. (1995). *Neuropsychological Assessment* (3rd ed.). New York: Oxford University Press.

Martin, R., Hayes, J., & Gouvier, W. D. (1996). Differential vulnerability between postconcussion self-report and objective malingering tests in identifying simulated mild head injury. *Journal of Clinical and Experimental Neuropsychology, 18,* 265-275.

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

Morey, L. C. (1996). *An Interpretive Guide to the Personality Assessment Inventory (PAI).* Odessa, FL: Psychological Assessment Resources.

Mossman, D. (2003). *Daubert*, cognitive malingering, and test accuracy. *Law and Human Behavior, 27*, 229-249.

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

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.

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

Pankratz, L., Fausti, A., & Peed, S. (1975). A forced-choice technique to evaluate deafness in the hysterical or malingering patient. *Journal of Consulting and Clinical Psychology, 43,* 421-422.

Pascal, G. R. & Suttell, B. J. (1951). *The bender-Gestalt Test: Its quantification and validity for adults.* New York: Grune & Stratton.

Paul, D., Franzen, M. D., Cohen, S. H., & Fremouw, W. (1992). An investigation into the reliability and validity of two tests used in the detection of dissimulation. *International Journal of Neuropsychology, 14,* 1-9.

Pelfrey, W. V. (2004). The relationship between malingerers' intelligence and MMPI-2 knowledge and their ability to avoid detection. *International Journal of Offender Therapy and Comparative Criminology, 48(6),* 649-663.

Petersilia, J. (2000). Doing justice? Criminal offenders with developmental disabilities. *CPRC Brief, 12,* California Policy Research Center, University of California.

Powell, S., Plamondon, R., & Retzlaff, P. (2002). Screening cognitive abilities in adults with developmental disabilities: Correlations of the K-BIT, PPVT-3, WRAT-3, and CVLT. *Journal of Developmental and Physical Disabilities, 14*, 239-246.

The Psychological Corporation. (2001). *Wechsler Test of Adult Reading Manual*. San Antonio, TX: Author.

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.

Reznek, L. (2005). The Rey 15-item memory test for malingering: A meta-analysis. *Brain Injury, 19,* 539-543.

Robinson, M., Myers, C., Sadler, I., Riley, III, J., Kvaal, S., & Geisser, M. (1997). Bias effects in three common self-report pain assessment measures. *The Clinical Journal of Pain, 13,* 74-81.

Rogers, R. (1984). Towards an empirical model of malingering and deception. *Behavioral Sciences & The Law, 2,* 93-111.

Rogers, R. (1997). Researching dissimulation. In R. Rogers (Ed.), *Clinical Assessment of Malingering and Deception* (2nd ed., pp. 398-426). New York: Guilford Press.

Rohling, M. L., Langhinrichsen-Rohling, J., & Miller, L. S. (2003). Actuarial assessment of malingering: Rohling's interpretive method. In Franklin, R. D. (Ed.), *Prediction in forensic and neuropsychology: Sound statistical practices* (pp. 171-207). Mahwah, NJ: Lawrence Erlbaum Associates.

Rose, F., Hall, S., Szalda-Petree, A., & Bach, P. (1998). A comparison of four tests of malingering and the effects of coaching. *Archives of Clinical Neuropsychology, 13,* 349-363.

Schretlen, D. J. (1988). The use of psychological tests to identify malingered symptoms of mental disorder. *Clinical Psychology Review, 8,* 451-476.

Simon, M. J. (1994). The use of the Rey memory test to assess malingering in criminal defendants. *Journal of Clinical Psychology, 50,* 913-917.

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

Tenhula, W. & Sweet, J. (1996). Double cross-validation of the booklet category test in detecting malingered traumatic brain injury. *The Clinical Neuropsychologist, 10,* 104-116.

Theodor, L. & Mandelcorn, M. (1973). Hysterical blindness: A case report and study using a modern psychophysical technique. *Journal of Abnormal Psychology, 82,* 552-553.

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

Vallabhajosula, B. & van Gorp, W. G. (2001). Post-Daubert admissibility of scientific evidence on malingering of cognitive deficits. *Journal of the American Academy of Psychiatry and the Law, 29,* 207-215.

Walters, G. D., White, T. W., & Greene, R. L. (1988). Use of the MMPI to identify malingering and exaggeration of psychiatric symptomatology in male prison inmates. *Journal of Consulting and Clinical Psychology, 56,* 111-117.

Wang, E. W., Rogers, R., Giles, C. L., Diamond, P. M., Herrington-Wang, L. E., & Taylor, E. R. (1997). A pilot study of the Personality Assessment Inventory (PAI) in corrections: Assessment of malingering, suicide risk, and aggression in male inmates. *Behavioral Sciences and the Law, 15,* 469-482.

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,* 979-990.

Wertheimer, M. (1923). Studies in the Theory of Gestalt Psychology. *Psychol. Forsch., 4,* 300.

Williams, K. T., & Wang, J.-J. (1997). *Technical References to the Peabody Picture Vocabulary Test--Third Edition (PPVT-III)*. Circle Pines, MN: American Guidance Service, Inc.

Appendix A

Louisiana State University
Psychology
Consent Form

<u>The Effects of Effort on Cognitive Testing Performance</u>

This study is held at the LSU Psychological Services Center and Audubon Hall on the LSU campus, and involves looking at how people perform on cognitive tests. Today we will ask you to complete a set of tasks that are often used to measure a variety of abilities in people with and without intellectual disabilities. By participating in this study, you will be one of about 140 other participants providing valuable information about how people approach these testing exercises.

You will be given four separate tasks that will include tests of memory, perception, and mental speed, such as counting dots, remember letters, numbers, and shapes, and drawing designs. Additionally, you will be asked to fill out a questionnaire about general medical history, any past head injuries you might have had, and any current psychological difficulties you are experiencing. The testing itself should take about thirty to forty-five minutes and will be completed in a single session.

Any individual who is at least 18 years old, has no significant history of neurological disease or seizure disorder, no current psychological disorder, and no pending criminal charges may take part in this study. LSU students who participate will be given extra credit in their psychology class. All other volunteers will receive $10 for their time. **Participation is voluntary**. While there are no risks foreseen by taking part in this study, you may choose not to participate or to withdraw from the study at any time without penalty.

The researchers conducting this study have been properly trained to administer all of the tests you are about to take. All information will remain anonymous. Any forms containing your name or phone number will be kept separately from your testing data.

If you have any questions please contact:

Name:	Daniel Proto
Department:	Psychology
Phone:	(678) 480-7092
Available:	Mon. & Wed.: 1:00pm – 3:30pm

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. 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

Appendix B

Louisiana State University
Psychology
Consent Form

<u>The Effects of Effort on Cognitive Testing Performance</u>

This study is held at the LSU Psychological Services Center and Audubon Hall on the LSU campus, and will look at how people perform on different types of mental tests. Today we will ask you to do tasks that look at how people think. If you join this study, you will be one of about 140 other people helping us learn about how people do these types of tests.

We will ask you to do four tasks today that will look at how well you can remember things, how you see and draw shapes, and how fast you can count. This should take about an hour and we will finish everything today.

Anyone who is 18 or older, and who does not have any neurological condition, psychological diagnosis or pending criminal charges can be in this study, and we will pay you $10. **You do not have to do these things if you do not want to**. You will not be harmed in this study, and you can choose not to do it or to stop at anytime without punishment.

The people testing you know how to give all of the tests you are about to take. We will not put your name on any of your answers, so no one will be able to tell how you did if you do not want them to. We will also keep all of the forms with your name on them separate from your answers.

If you have any questions please contact:

Name:	Daniel Proto
Department:	Psychology
Phone:	(678) 480-7092
Available:	Mon. & Wed.: 1:00pm – 3:30pm

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. 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

The study subject has indicated to me that he/she is unable to read. I certify that I have read this consent form to the subject and explained that by completing the signature line above, the subject has agreed to participate.

__________________________________ _____________
Signature of Reader Date

(Appendix B cont.)

I certify that by signing below, I am indicating that as acting guardian for the subject, I consent to their participation in this study based on the information described in the form above. Additionally, I am indicating that assent to participate as been given by the subject.

____________________________________ ________________

Signature of Guardian/Staff Member Date

Appendix C

Demographic Questionnaire

Subject #:______ Examiner: ____________________

Age: _________ (circle one): MR / Non-MR

Race: ________

Gender: _______

Highest grade completed: ________

Do you currently have, or have you previously had, any type of neurological disorder, for example epilepsy?: Y / N
*If yes, please explain:

____
____
____
____

Have you ever been hit on the head so hard that you blacked out?: Y / N
*If yes, please explain:

____
____
____
____

Are you currently diagnosed with any psychological disorder(s), including ADHD or learning disorders?: Y / N

Are there any pending criminal charges against you?: Y / N

Appendix D

Memory for Fifteen Items Test

Subject #:_______   Examiner:______________________

Column 1 Correct: _________

Column 2 Correct: _________

Column 3 Correct: _________

Column 4 Correct: _________

Column 5 Correct: _________

Total # Correct : _____________

Appendix E

Dot Counting Test

Subject #: ______ Examiner: ____________________

Card 1: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 2: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 3: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 4: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 5: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 6: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 7: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 8: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 9: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 10: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 11: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Card 12: Number of Dots Counted: ______ Correct: Y / N Time (sec): ______

Total Number of Errors: __________

Appendix F

Participant Rating Scale

Subject #:______ Examiner: ____________________

1.) How successful do you think you were in trying to portray someone with a brain-injury? (Circle one)

Not at all		An average amount		Very
1	2	3	4	5

2.) How hard did you try? (Circle one)

Not at all		An average amount		Very
1	2	3	4	5

Vita

Daniel Proto was born in Plano, Texas in January, 1980 to mother Denise Proto, and father Vincent Proto. He spent his early years in a variety of states, moving between Texas, New York, and Massachusetts, before finally settling in Norcross, Georgia in the early Fall of 1985. His parents divorced soon afterwards, and while Daniel and his older brother, Michael Proto, chose to live with their mother, they both maintained a positive relationship with their father, whom they visited bi-weekly.

Educationally, Daniel generally performed well, entering the gifted program during the 2^{nd} or 3^{rd} grade, and continuing with advanced courses through middle and high schools. English eventually arose as a favorite subject, with Daniel earning a handful of awards in this area. His first long-standing career goal, formed in late junior high, centered around astronomy and astrophysics. However, this aspiration began shifting toward psychology by the time he had entered college at the University of Georgia in 2001, and each subsequent course taken in the area further piqued his interest.

After obtaining his Bachelor of Science degree in 2004, Daniel sought full-time employment the following year as he began preparing his graduate school applications. After attending interview day at Louisiana State University, and meeting Dr. William Drew Gouvier, he decided that the program in Baton Rouge was where he wished to enroll. Daniel is currently a third-year doctoral student in the aforementioned clinical psychology program at Louisiana State University, where his research interests focus on neuropsychology and malingering assessment.