2006

DETECTION OF MALINGERED MENTAL RETARDATION

Lili Odom Graue
University of Kentucky, lili.graue@uky.edu

ABSTRACT OF THESIS

DETECTION OF MALINGERED MENTAL RETARDATION

The 2002 Supreme Court decision (*Atkins vs. Virginia,* 536 U. S. 304*)*, prohibiting the execution of mentally retarded persons, may potentially increase malingering of mental retardation (MR). There is limited research addressing the detection of feigned MR. The present study compared results from tests of intelligence, adaptive functioning, legal/courtroom knowledge, and psychiatric and neurocognitive feigning to determine how effectively these instruments discriminate between MR participants and community volunteers asked to either approach the test honestly (CVH group) or feign, or malinger, MR (CVM group). CVMs suppressed their IQ scores sufficiently to appear MR. CVMs overestimated deficits on individuals with genuine MR on tests of adaptive functioning and courtroom knowledge. Psychiatric feigning instruments did not discriminate between MR and CVM groups. Neurocognitive feigning instruments discriminated between groups, however specificity and Positive Predictive Power were unacceptably low. Revising cutting scores to hold specificity at .95 improved PPP significantly, suggesting the potential utility of these instruments to detect feigned mental retardation. Results from this study suggest that applying published decision rules to MR populations on tests commonly used in forensic neuropsychological evaluations will likely result in a high rate of false positive errors. Given the high stakes associated with classification errors in capital cases involving MR defendants, alternative cutting scores appropriate for this population should be determined.

KEYWORDS: Malingering, Mental Retardation, Neuropsychological Assessment, Feigned Neurocognitive Impairment, Malingering Instruments

Lili Odom Graue

August 28, 2006

DETECTION OF MALINGERED MENTAL RETARDATION

By

Lili Odom Graue

David T. R. Berry, Director of Thesis

David T. R. Berry, Director of Graduate Studies

August 28, 2006

RULES FOR THE USE OF THESES

Unpublished theses submitted for the Master's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgements.

Extensive copying or publication of the thesis in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this thesis for use by its patrons is expected to secure the signature of each user.

THESIS

Lili Odom Graue

The Graduate School

University of Kentucky

2006

DETECTION OF MALINGERED MENTAL RETARDATION

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
College of Fine Arts at the University of Kentucky

By

Lili Odom Graue

Lexington, Kentucky

Director: Dr. David T. R. Berry, Professor of Psychology

Lexington, Kentucky

2006

ACKNOWLEDGMENTS

The following thesis, while an individual work, benefited from the insights and direction of several people. First, my thesis chair, Dr. David T. R. Berry, who provided encouragement, support, and instructive comments throughout every stage of the thesis process. Dr. Berry exemplifies high quality scholarship as demonstrated by his academic record. Perhaps more importantly he is an exceptional mentor who takes a personal interest in the success of his students. I have benefited considerably from his tutelage. Next, I wish to thank the other faculty on my Thesis Committee: Dr. Ruth A. Baer and Dr. Gregory T. Smith. Each provided insights that helped to guide the project to its completion.

In addition to the technical and instrumental assistance above, I received equally important assistance and support from family and friends. To my husband, Jefferson Graue, thank you for your continuous support and encouragement throughout the thesis process. Thanks also to my children, Ashley, Josh, and Kayla, for their patience with my busy schedule. I would also like to thank my colleagues who volunteered their time to help with data collection : Jessica Clark, Michelle Cardi, Myriam Sollman, Lindsey Schipper, Jaclyn Hopkins, and Anni Shandera. Finally, I wish to thank Dellynda Werline and Cindy Faulkner, of the Bluegrass Mental Health/Mental Retardation Board, for their assistance with participant recruitment, and the staff at Connections and the Berea and Richmond MARC Centers for facilitating us in the data collection of our developmentally disabled participants.

TABLE OF CONTENTS

LIST OF TABLES

Chapter One

Introduction

Neuropsychologists are frequently called upon to evaluate individuals involved in litigation or criminal proceedings to resolve questions regarding mental competence and cognitive ability. The results of psychological evaluations have become more widely accepted in courtrooms, often determining whether an individual claiming personal injury is entitled to considerable benefits such as workers' compensation, disability pensions, and other monetary awards (Berry, Baer, Rinaldo, & Wetter, 2002; Tombaugh, 2002; Vickery, Berry, Inman, Harris, & Orey, 2001). Such evaluations may also influence a court's decision regarding whether an individual should be held accountable for or be absolved of responsibility for criminal activity. In light of the substantial awards and benefits one stands to gain in these cases, as well as the costs to society as a result of successful false presentation, the issue of malingering must be addressed. Consequently, malingering has become a topic of considerable interest in recent years, particularly as the number of forensic evaluations has increased (Etcoff & Kampfer, 1996; Inman & Berry, 2002; Inman, et al., 1998; Nies & Sweet, 1994; Vickery et al., 2002; Zielinski, 1994).

In the last two decades, malingering research has focused primarily on psychiatric and head-injured populations. However, a recent decision by the U.S. Supreme Court may inadvertently encourage a new type of malingering that, prior to this event, had not garnered much attention from researchers in the field. In 2002, the Court ruled that executing people with mental retardation was cruel and unusual punishment, violating the Eighth Amendment of the Constitution (Atkins vs. Virginia, 2002). Although it is too soon for reliable statistics to have been gathered, it seems likely that defense attorneys will pursue this issue more aggressively than in the past. Because feigning of mental retardation has been virtually ignored in the literature, it is not known whether standard psychiatric and neurocognitive instruments for the assessment of feigning are appropriate for use with this population. Furthermore, even if it is determined that these standard measures can be effectively applied to this population, it may be necessary to adjust cutting scores derived from other patient populations to distinguish between honest and malingered performance in possibly mentally retarded defendants.

The goal of the present study was to assess the efficacy of neurocognitive and psychiatric feigning measures in detecting feigned mental retardation. To provide important background information, this review will begin with a brief overview of malingering, including definitions of malingering and its prevalence in medicolegal settings. Methodological designs most commonly used in malingering research will then be summarized, followed by a detailed discussion of a number of psychiatric and neurocognitive instruments that are frequently used to detect dissimulation. Next, definitions of mental retardation will be presented and the legal issues surrounding malingered mental retardation will be introduced. Then, the scant existing body of published literature addressing this issue will be summarized with a focus on the

following questions: (a) which instruments have been used to detect malingered MR, (b) how accurate are they, and (c) what methodological weaknesses are present in published studies? Finally, results from a study comparing mental retardation and community volunteers feigning MR will be presented.

Malingering

Malingering has been described as an attempt at impression management whereby an individual deliberately exaggerates or fabricates symptoms to appear mentally and/or cognitively impaired for the purpose of achieving some external goal. Malingering has been conceptualized by psychologists as one of several "response sets," a form of non-veridical responding commonly referred to as "faking bad" (Berry, et al., 2002).

Malingering is generally believed to take at least two forms relevant to psychological evaluation (Rogers, 1997). The first, malingered psychopathology, involves the exaggeration or fabrication of psychiatric symptoms to mimic global mental illness or a specific disorder. Individuals facing criminal prosecution and jail time or a call to military duty, for example, may seek a diagnosis of psychiatric illness to escape such undesirable conditions. Others may fabricate or exaggerate psychiatric symptoms in order to obtain desirable institutional placements or maintain their patient status in mental health facilities.

The second form, malingered neurocognitive impairment, involves providing inadequate effort or deliberately choosing incorrect answers in pursuit of a substandard performance on measures designed to tap cognitive abilities such as vocabulary, verbal reasoning, and memory. Compensation-seeking individuals may thus attempt to document false claims of neuropsychological impairment for the purpose of obtaining benefits such as workers' compensation, disability pensions, or monetary awards from responsible parties.

DSM-IV-TR Definition of Malingering

Malingering is formally defined in the DSM-IV-TR as "the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" (APA, 2000, p. 739).

The DSM-IV-TR provides no formal criteria to aid clinicians in determining whether a person is malingering (Berry, et al., 2002). Instead, it simply states that malingering should be "strongly suspected" when any combination of the following is observed: (a) the clinical assessment occurs within a medico-legal context, (b) there is a discrepancy between the person's stated impairment and the objective findings, (c) there is a lack of client cooperation either during the evaluation or with the recommended course of treatment, and (d) there is evidence of Antisocial Personality Disorder. Rogers (1997) argues that these criteria are insufficient for determining a diagnosis of malingering and suggests they should be

viewed merely as factors that increase the probability of malingering. Thus, without established criteria for "diagnosing" malingering, clinicians may inadvertently rely on methods and techniques that lack sufficient reliability and validity to accurately detect feigning.

Unfortunately, research in psychological assessment suggests that clinicians are poor judges of malingering on the basis of standard test results alone (Garb, 1998; Heaton, Smith, Lehman, & Vogt, 1978; Inman et al., 1998; Ziskin, 1984). Another important issue is that clinicians must themselves determine which feigning instruments to use when conducting a forensic evaluation as there are no consensus standards. Therefore, clinicians in this field should be familiar with the current research which provides guidelines for identifying malingering in the medico-legal context. Clinicians should also be familiar with a broad range of assessment tools and consider the research base that supports each instrument to best evaluate whether it can appropriately address the issue in question. Furthermore, a determination of malingering should never be made on the basis of a single measure. Rather, the use of diverse methods and multiple data sources with convergent findings is recommended for maximizing confidence in conclusions regarding malingered performance (Rogers, 1990; Slick, Sherman, & Iverson, 1999; Sweet, 1999).

Research published in the past decade (Faust & Ackley, 1998; Greiffenstein, Baker, & Gola, 1994; Greiffenstein, Gola, and Baker, 1995; Pankratz & Binder, 1997), which included modifications of the DSM definition of malingering and more specific diagnostic criteria, helped to guide research efforts in this area and provided clinicians with important information to improve diagnostic accuracy in forensic evaluations. Although these advances were useful, there remained a need for a more comprehensive model of malingering of cognitive deficits to increase our understanding of malingering and enhance current detection strategies. Slick, Sherman, and Iverson (1999) provided such a model.

Slick et al. (1999) Definition of Malingering

Slick et al. (1999) addressed perceived shortcomings in the literature by developing a more explicit definition and set of diagnostic criteria for identifying neurocognitive malingering that could be utilized by neuropsychologists. Building on previous models (DSM-IV; APA, 1994; Faust & Ackley, 1998; Greiffenstein, Baker, & Gola, 1994; Greiffenstein, Gola, and Baker, 1995; Pankratz & Binder, 1997; Rogers, 1990), Slick et al. provided a set of formal criteria for diagnosing possible, probable, and definite malingering of cognitive functioning to aid clinicians in decision-making in the context of neuropsychological evaluations. These criteria may also serve as guides to define populations for clinical research on malingering prevalence and detection. Slick et al. proposed the following definition (p. 552):

> Malingered Neurocognitive Dysfunction (MNCD) is the volitional exaggeration or fabrication of cognitive dysfunction for the purpose of obtaining substantial material gain, or avoiding or escaping formal duty or responsibility. Substantial gain includes money, goods, or services of

nontrivial value (e.g., financial compensation for personal injury). Formal duties are actions that people are legally obligated to perform (e.g., prison, military, or public service, or child support payments or other financial obligation). Formal responsibilities are those that involve accountability or liability in legal proceedings (e.g., competency to stand trial).

These formal diagnostic criteria were designed to draw a distinction between possible, probable, and definite malingering, thereby increasing diagnostic certainty. The criteria specify that Definite MNCD could be established when "clear and compelling evidence of volitional exaggeration or fabrication of cognitive dysfunction" exists without "plausible alternative explanations" (p. 552). Probable MNCD would be indicated "by the presence of evidence strongly suggesting volitional exaggeration or fabrication of cognitive dysfunction and the absence of plausible alternative explanations" (p. 552). Possible MNCD would be indicated "by the presence of evidence suggesting volitional exaggeration or fabrication of cognitive dysfunction and the absence of plausible alternative explanations" (p. 553). This comprehensive set of criteria provides detailed explanations of each criterion, with additional points for consideration. In addition, caveats and recommendations for clinicians are outlined in detail. A number of reputable researchers who study the problem of malingering have recommended using the Slick et al. (1999) criteria in investigations of malingering in neurological populations (Bianchini et al., 2001; Greve & Bianchini, 2004; Thompson, 2002).

Prevalence of Malingering

Although exact prevalence data are unavailable, research spanning the past ten years suggests that feigning on psychological measures is more common in forensic settings than in other contexts (Miller, 2001). The data suggest that the base rate of malingering rises roughly in relation to increases in the potential award involved (Berry, et al., 2002; Rogers, 1997; Rogers & Cruise, 2000; Sweet, 1999). Consequently, mental health professionals cannot assume the accuracy of information presented by clients on self-report measures during forensic evaluations, particularly when there is an incentive to fabricate psychiatric symptoms or cognitive impairment.

Rogers, Sewell, and Goldstein (1994) surveyed 320 forensic psychologists who provided malingering estimates of 15.7% and 7.4% in forensic and non-forensic settings, respectively. In a similar finding, the average estimates of malingering were 17.4% for forensic and 7.2% for non-forensic patients (Rogers, Salekin, Sewell, Goldstein, & Leonard, 1998). Other studies (Binder, Villaneuva, Howieson, & Moore, 1993; Frederick, Crosby, & Wynkoop, 2000; Guilemette, Whelihan, Sparadeo, & Buongiorno, 1994; Norris & May, 1998; Strong, Green, & Schinka, 2000; Trueblood & Schmidt, 1993; Youngjohn, Burrow, & Erdal, 1995) also reported evidence of poor motivation or exaggerated symptoms among a relatively substantial proportion of evaluees involved in litigation or criminal proceedings. In a recent survey of the literature yielding base rate estimates of malingering, Berry, et al. (2002) found that the

reported prevalence of malingered performance ranged from 7.3 to 27 percent in non-forensic settings, whereas that in forensic settings was between 31 and 45 percent. This substantial increase in feigning in forensic settings underscores the importance of evaluating for malingering in these evaluations.

Assessment of malingering requires the use of valid, reliable measures with cutoff scores set to minimize diagnostic errors. Cutoff scores are established in controlled validation studies by comparing test results from groups of individuals with the condition of interest (i.e., feigning) to those without the condition (patients answering honestly). When determining which measures to use in a forensic evaluation, it is important to consider the experimental design of validation studies and the quality of the research support for tests under consideration, including the operating characteristics of each.

Operating characteristics include sensitivity (SN), specificity (SP), positive predictive power (PPP), and negative predictive power (NPP). Sensitivity is the proportion of those individuals with a particular condition who are identified correctly by the test. Specificity is the proportion of individuals without that condition who are identified correctly by the test. Sensitivity and specificity vary according to the given cutting score used. Perhaps more useful to clinicians is predictive power which varies according to the base rate of a condition as well as SN and SP at the chosen cutting score. Predictive power is the probability that an individual has (or does not have) the condition, given that the test sign is positive (PPP) or negative (NPP). The predictive power of a given measure is an important index of the instrument's clinical validity. Most validation studies report predictive power along with other classification rates. Since actual base rates of malingering are unknown, researchers have used empirically based estimates to calculate the hypothetical predictive power of an instrument.

Research Designs

Research designs used in the published malingering literature include case studies, simulation (analog) studies, known-groups comparisons, and differential prevalence designs. Each design has methodological strengths and limitations. Case studies present detailed descriptions of individual cases of persons known to feign symptoms for some external gain. These may generate useful information to aid in the understanding of behavioral and cognitive components of malingering. However, the inability to generalize the results limits the utility of this design.

The simulation study has been used most frequently in the validation of instruments to detect feigned performance. Typically, this design compares simulators (participants instructed to feign symptoms or deficits associated with a particular condition) to a group of patients answering honestly. Ideally, the "honest" group is clinically relevant to the condition under examination and the "malingering" group is as similar to the criterion group as possible. The strength of this design lies in its experimental control as the feigning is induced by the experimenter. It is difficult, however, to determine

whether "real world" malingerers would perform similarly. Thus, simulation studies tend to have high internal, but limited external, validity.

In known-groups comparisons, individuals from a clinical or applied setting who are identified as malingerers or honest (using some "gold" standard or other criterion) are compared. The greatest strength of this design is that the results are likely to be more generalizable than those from analog designs. This design has weaknesses, however. First, the validity of the "gold" standard used for group memberships puts a ceiling on the potential observed validity of the measure being evaluated. Second, those individuals in the malingering group may not be representative of all malingerers, as it is possible that only the more blatant malingerers were identified while the more sophisticated malingerers escaped detection. Third, due to the lack of random assignment to groups and experimental manipulation, it cannot be assumed that the two groups differ only in the presence or absence of feigning symptoms and that all differences are due to the presence or absence of feigning. As noted below, support from both simulation studies and known-groups designs is ideal for validating a malingering instrument because the strengths and weaknesses of one design are complemented by those of the other.

Differential prevalence designs involve the formation of groups that differ on a characteristic believed to be related to the condition being evaluated. As an example, in malingering research, compensation-seeking participants might be compared to non-compensation-seeking individuals to determine how well a malingering measure differentiates the two groups. Since no effort is made to classify persons within each group as honest or malingering, results from differential prevalence designs cannot provide information regarding the prevalence rate of feigning in different referral groups. Thus, classification parameters of the instruments used in the study cannot be calculated. This design, although it provides indirect evidence for the construct validity of a measure, has little to offer in regards to criterion-related validity and clinical classification.

While each of these designs may provide useful information about the construct itself, only simulation studies and known-groups comparisons meet the standards necessary to adequately establish the validity and classification accuracy of a given measure. Rogers (1995) suggested using the combination of simulation studies and known-groups comparisons to address the need for both experimental rigor and generalizability. He recommended using the latter to cross-validate findings from simulation studies.

Detection of Malingering

Although the DSM-IV-TR (APA, 2000) treats malingering as a broad phenomenon that may include the exaggeration or fabrication of physical or psychological symptoms, malingering research in psychology has focused primarily on attempts to appear either mentally ill (malingered psychopathology) or cognitively impaired (malingered neurocognitive deficits). As noted earlier, both forms of malingering

are believed to be fairly common in medico-legal settings (Owens, 1995; Rogers, 1997). There are many empirically validated instruments that are widely used in forensic evaluations to assess both general and specific functioning. These include, but are not limited to, intelligence tests, personality inventories, neuropsychological tests, and both self-report and structured interview measures of psychopathology. A typical forensic evaluation involves the administration of a test battery covering most of these domains.

Instruments used to address the presence of malingering make use of several detection strategies for dissimulation, discussed below (Rogers, 1990). Validating these detection strategies is necessary for determining the usefulness of an instrument. Rogers proposed three strict criteria for validating detection strategies. The first is the convergence of observations across research designs (e.g., simulation and known groups). The second involves the convergence of findings across methods of assessment (e.g., self-report and cognitive tests). Finally, detection strategies should be cross-validated on clinically diverse samples. Strategies that meet all of these criteria may be used with a high degree of confidence.

Detection of Malingered Psychopathology

Consistent with other experts, Rogers (1997) described malingered psychopathology as the fabrication or exaggeration of psychiatric symptoms for some external gain. Possible motives for this type of dissimulation include the avoidance of prosecution, the prospect of receiving financial awards or benefits (e.g., insurance settlements, long-term disability), and obtaining institutionalized care. Psychiatric malingering has been addressed in a number of ways. Rogers (1990) described four strategies for detecting feigned psychopathology that have been cross-validated across research designs and methods of assessment. These strategies include (a) endorsement of rare symptoms, (b) indiscriminant symptom endorsement, (c) obvious symptoms, and (d) improbable symptoms (p. 374). Two others, symptom combination and symptoms of extreme severity, almost meet Rogers' criteria but currently lack sufficient validation.

A number of instruments incorporating these detection strategies have been validated for use in the identification and diagnosis of psychopathology in both research and clinical settings. A brief description of each, along with classification data obtained from validation studies will be presented next.

Minnesota Multiphasic Personality Inventory-2 (MMPI-2): The MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) is the most widely accepted self-report inventory of psychopathology today. It consists of 567 true/false questions which are scored on a number of clinical, content, validity, and supplementary scales. Validity scales detect random responding and other response sets such as nay-saying, yea-saying, faking good, and faking bad. Once random responding has been ruled out through referring to the Variable Response Inconsistency (VRIN) and True Response Inconsistency (TRIN) scales (Gallen & Berry, 1996; 1997; Wetter, Baer, Berry, Smith, & Larsen, 1992), the interpreter may evaluate several other scales to determine the presences of exaggerated symptoms.

The Infrequency (F) scale is elevated by highly atypical responses on items in the first part of the test (i.e., items are endorsed by ≤10% of the normative sample). The Back Infrequency (Fb) scale is similar to the F scale, but differs in that it identifies atypical responses in the latter part of the test. The fabrication or exaggeration of symptoms may also be detected by subtracting the raw K score (tapping defensiveness) from the raw F score (tapping symptom magnification). Thus a high, positive F minus K score indicates the tendency to over-report symptoms as opposed to under-report or deny symptoms (Gough, 1950, cited in Berry et al., 2002). The Infrequency-Psychopathology, F(p), scale was developed to help discriminate feigned symptoms from legitimate psychopathology (Arbisi & Ben-Porath, 1995), as an elevated F(p) scale is unlikely to be indicative of severe psychopathology due to inclusion of items rarely endorsed by psychiatric inpatients. The MMPI-2 also has several supplementary scales related to feigned symptomatology that may be useful to consider, although they are not used as frequently, nor are they as extensively supported, as the previously mentioned infrequency scales.

Berry et al. (2002) reviewed the literature on the detection of feigning using the MMPI-2. In this manuscript, data from numerous studies were collapsed to provide a rough picture of the operating characteristics of these scales since there was much variation in the findings from each study taken alone. A wide range of optimal cutting scores was reported for the 4 fake bad indices. The highest average sensitivity (.84) was obtained from the F – K index using a cutting score of 16, followed by the F scale (.83) with a cutting score of 106T. The highest average specificity was obtained on the F(p) scale (.92) with a cutting score of 96T. PPP and NPP were calculated at 2 base rates (.19 and .27). The F(p) scale had the highest PPP at both base rates. However, PPP at the latter base rate was only .77, which indicates that there is approximately a 1 in 4 chance that the test will falsely identify an honest respondent as feigning. Results suggested that these 3 scales would be best used to identifying honest responding rather than feigning since NPP for all 3 scales was at or above .90 at both base rates.

One caveat must be considered when using the MMPI-2 validity scales. Consideration of all available malingering scales, with any one beyond the cutting score taken, suggesting feigned performance, may result in a high incidence of false positives because the scales are not perfectly correlated (Greene, 2000, cited in Berry et al., 2002). Clinicians may minimize the probability of making diagnostic errors by determining beforehand 1 or 2 scales to evaluate. As noted earlier, diagnostic accuracy may also be increased by using a number of different instruments to address the issue since converging evidence from multiple sources will strengthen any conclusions made from test results.

Structured Interview of Reported Symptoms (SIRS): The SIRS (Rogers, 1986) has 172-items organized into 8 primary scales and 5 supplementary scales designed to assess feigning of psychological symptoms and other response sets. Three of the primary scales focus on atypical symptom reporting. Four assess symptom severity. The final primary scale addresses discrepancies between self-reported overt

symptoms and interviewer observations. The SIRS is presented in a structured interview format, which standardizes the test administrator's behavior and improves the reliability and validity of the measure.

Using empirically determined ranges, raw scores from each of the scales are classified as honest, indeterminate, probable feigning, or definite feigning. Both simulation studies and known-groups comparisons have been utilized to evaluate the SIRS. In two validity studies (Gothard, Viglione, Meloy, & Sherman, 1995; Rogers, et al., 1992) using a cutting score of ≥3 on primary scales in the probable feigning range, sensitivity was .48 and .97 and specificity was .99 and .97. The SIRS appears to be the best supported malingering instrument currently available. Inter-rater reliability has been strong (.96) and it has been extensively validated. The major limitation of the SIRS is the length of time required for clinicians to administer the test.

Structured Inventory of Malingered Symptomatology (SIMS): The SIMS (Widows & Smith, 1995) is a brief self-report measure designed to screen for malingering. Written at a 4th grade reading level, the SIMS consists of 75 true/false items and includes scales sensitive to different forms of malingering, including Psychosis (P), Amnesia (Am), Neurological Impairment (N), Affective Disorder (Af), and Low Intelligence (Li). The Total Score, which is a summation of all SIMS scales, is typically used to identify malingered performance.

Validation studies of the SIMS suggest its potential value in detecting dissimulation (Lewis, Simcox, & Berry, 2000; Rogers, Hinds, & Sewell, 1996; Smith and Burger, 1997). Available validation research included a psychiatric control group, simulation designs, known-groups comparisons, and the use of incentives (in the simulation study) to increase external validity. The cutting score of >16 was fairly consistent across studies for predicting malingering. This cutting scores yielded sensitivity rates between .44 and 1.00 and specificity rates ranging from .55 to .94. Limitations to the instrument includes the absence of a scale for addressing random responding, the lack of research on non-forensic psychiatric patients for comparison to those involved in litigation, and the failure to check understanding of and compliance with instructions in the simulation designs using honestly responding patient groups (Berry, et al., 2002).

Miller Forensic Assessment of Symptoms Test (M-FAST): Based on the SIRS, the M-FAST (Miller, 2001) is a brief structured interview consisting of 25 questions assigned to one of seven scales (Reported vs. Observed, Extreme Symptomatology, Rare Combinations, Unusual Hallucinations, Unusual Symptom Course, Negative Image, and Suggestibility) designed to screen for malingered psychopathology. The creation of test items was based on previous research that suggested response styles and interview strategies common to known malingerers. The test was validated using known-groups comparisons (clinical sample) and simulation designs (non-clinical samples) to maximize both external and internal validity (Miller, 2001).

The M-FAST manual suggests using a raw score of ≥6 to detect malingered psychopathology in both clinical and non-clinical samples. The reported classification accuracy of this instrument is encouraging. In clinical samples, using this cutting score, sensitivity was .93 and specificity was .83. Predictive power, calculated with a malingering base rate of .51, was equally encouraging. PPP and NPP were .68 and .97, respectively. In non-clinical samples sensitivity equaled .93, specificity equaled 1.00, and PPP and NPP were 1.00 and .94, respectively.

In a more recent validation study, Miller (2004) used a sample of criminal defendants deemed incompetent to stand trial because of mental illness and found that the same cutting score resulted in a sensitivity of .93, specificity of .83, and an overall hit rate of .86. Evidence for construct validity was demonstrated by convergent and discriminant data that showed correlations between individual scales and total scores on the SIRS, MMPI-2 and M-FAST.

Data from these validation studies suggested that the M-FAST was a good initial screening instrument for use in forensic evaluations. One of the strengths of the M-FAST is its brevity, taking approximately 5 minutes to administer. Also, the interview format allows for more interaction between the clinician and client and gives the clinician the opportunity to make behavioral observations during test administration. Limitations included those associated with the research design. In the known-groups design, participants were classified as honest or malingering according to SIRS test scores. The potential for incorrect group assignment makes it impossible to know whether test performance of the criterion group is representative of all malingerers. In addition, the validity of the criterion measure puts a ceiling on the validity of the instrument under examination. Therefore, it is possible that the actual validity of the M-FAST may be higher than was reported.

Detection of Feigned Neurocognitive Deficits

Malingering instruments are developed under the assumption that persons attempting to feign cognitive impairment lack understanding of key principles underlying cognitive functioning. Instruments designed to detect malingered cognitive impairment typically have a high face difficulty level, but a low true difficulty level. This allows those with legitimate deficits to perform well, while potentially "inviting" malingerers to perform poorly. When used in a test battery, malingering measures are perceived as having a high difficulty level and, therefore, may be valuable tools for discriminating malingerers from honest responders.

Several detection strategies for identifying feigned neurocognitive deficits are described by Rogers, Harrell, and Liff (1993). One is the floor effect. This is seen when the test-taker fails tasks that even those with severe impairments would be likely to perform successfully. Another detection strategy, the performance curve, was described by Goldstein (1945). It is based on the assumption that dissimulators will not consistently consider item difficulty when deciding which items to fail. Given this,

they may be detected by comparing their performance curve across item difficulty levels to that of persons with legitimate cognitive deficits.

The most popular strategy for detecting feigned cognitive impairment is the symptom validity test (SVT). This procedure, which was originally used to detect feigned sensory deficits, was named by Pankratz (1983). It has since been adapted for use in the detection of exaggerated memory deficits (Heubrock & Petermann, 1998). Most SVTs use forced-choice recognition across numerous trials to accomplish this goal. Forced-choice testing typically involves presentation of a stimulus followed by a delay period after which the original stimulus must be chosen from a paired foil. In two-alternative forced-choice procedures, the probability of chance performance can be obtained by referring to a table of binomial probabilities. A statistically significantly below-chance performance indicates that the correct stimulus was known and actively avoided. Some would-be malingerers tend to try too hard to appear impaired and thereby exceed the binomial probability for wrong responses that would be obtained by random responding. In other words, they achieve a score that is below that which would be obtained by guessing alone.

One advantage of the SVT in detecting malingered performance is that it is less susceptible to base rate variations, which adversely affect norm-based decision rules (Bianchini et al., 2001; Bickart, Meyer, & Connell, 1991). The main limitation in using significantly below-chance cutoffs is that they may be sensitive only to the most blatant malingerers. Thus, only a relatively small percentage of individuals attempting to feign deficits score significantly below the chance level. Consequently, the low sensitivity of the SVT using the significantly below chance criterion increases the potential for making false negative classifications, meaning that the more sophisticated malingerers will go undetected (Bickart et al., 1991; Guilmette, Hart, & Giuliano, 1993).

In an attempt to improve sensitivity, two approaches to solving this problem have been investigated. These include modifying the cutting score to reflect a better-than-chance level and using a norm-based approach which identifies an "empirical floor" based on the lowest performance by some impaired but non-compensation-seeking comparison group (Bianchini, Mathias, & Greve, 2001). In general, these techniques have increased sensitivity without compromising specificity.

Other strategies to improve the sophistication of SVTs have been (a) developing procedures manipulating face difficulty levels (Hiscock & Hiscock, 1989); (b) using more than 2 alternatives (Rogers, 1997); (c) examining consistency of performance across time (Iverson, Franzen, & McCracken, 1991); and (d) measuring different abilities (Pritchard & Moses, 1991). The importance of incorporating into a test battery a number of SVTs using a variety of stimuli (e.g., pictures, letters, digits, etc.) was demonstrated by Gervias, Rohling, Green, and Ford (2003). This study found that the different measures varied substantially in sensitivity, challenging the assumption that all SVTs are equally sensitive to

response bias. The impressive overall results from validation studies of SVTs suggest that these instruments should be an important part of any neuropsychological evaluation, particularly within a medico-legal context.

The following SVT-based instruments have been validated extensively and are commonly used in forensic test batteries. Each instrument will be described briefly, followed by a summarization of empirical findings from validity studies that provide classification accuracy data and recommended cutting scores.

Digit Memory Test (DMT): The standard DMT (Hiscock & Hiscock, 1989), a refined SVT, is a forced-choice digit recognition test which involves identifying a previously presented 5-digit number after an increasing time delay. The first block of 24 trials has a 2.5-second delay which increases to 5 seconds during the second block, giving the illusion of increasing difficulty. During the third block the delay is increased to 10 seconds. In a study comparing analog malingerers with non-compensation seeking brain injured individuals, Guilmette et al. (1993) demonstrated that a cutting score of <90% correct on a 36-item short form of the DMT identified inadequate motivation with sensitivity and specificity at 90% each. Similar results were found in a cross-validation study using mildly head injured college students as analog malingerers (Inman and Berry, 2002). Internal consistency reliability for the DMT is adequate at .87. Overall, the DMT is one of the most extensively validated motivational tests currently available (Vickery et al., 2001).

Test of Memory Malingering (TOMM): The TOMM (Tombaugh, 1996), another extensively studied SVT, is a forced-choice recognition test that involves the recognition of simple line drawings of common objects in 2 learning trials which are followed by a retention trial. The learning trials have 2 phases: a study phase and a test phase. In the study phase, each of the 50 targets (line drawings) is presented for 3 seconds. In the test phase, each target is paired with a new drawing as a foil and the test-taker must choose the one that had been presented earlier, in the study phase. If an incorrect response is given, the error is pointed out. This procedure is repeated with novel foils in the second learning trial. After a 20-minute delay period, the retention trial, which consists only of a test phase with 50 new foils, is administered. The total administration time is 15 minutes or less, excluding the delay period.

The TOMM was designed to discriminate between genuine and feigned symptom reporting. Validation studies (Rees, Tombaugh, & Boulay 2001; Rees, Tombaugh, Gansler, & Moczynski, 1998; Tombaugh, 1997) demonstrated that the TOMM was insensitive to depression and brain injury. In these studies, cognitively intact adults and individuals with traumatic brain injury obtained mean scores of 99.8% and 99.2% correct, respectively, on the delayed retention trial. Also, it was found that highly depressed individuals rarely obtained scores below the recommended cutoff for valid responding. In

addition, individuals with dementia associated with Alzheimer's disease obtained a mean score of 94.0% correct.

Tombaugh (1997) recommended a cutting score of < 90% correct for identification of feigned memory dysfunction. This cutting score correctly classified 100% of community-dwelling participants and 95% of clinical patients without dementia. Tombaugh (2002) reported on 5 experiments that used clinically diverse samples to examine classification accuracy. Specificity was quite high (ranging from 90 to 100 percent) and sensitivity was adequate at >88%, using the recommended cutting score.

Letter Memory Test (LMT): The LMT (Inman et al., 1998) is a computer-administered 45-item forced-choice recognition test which uses consonant letters as stimuli. Face difficulty level is manipulated along two dimensions: number of letters to be remembered and number of letter choices presented with the target stimulus. The LMT offers several advantages over the TOMM. First, it uses less of the clinician's time because it is administered by computer. Second, it manipulates face difficulty level multiple times by changing the number of letters to be remembered and the number of choices from which the target stimulus must be chosen.

Validation studies of the LMT (Inman et al., 1998) demonstrated that cognitively impaired individuals and persons feigning cognitive deficits obtained mean scores of 99.3% and 66.1% correct, respectively. Using a cutting score of <93% correct, sensitivity and specificity were .84 and 1.00, respectively. The overall hit rate was 92%. Like the TOMM, the LMT is not affected by affective state, as specificity was maintained at 1.00 when used with individuals with moderate levels of depression (Inman et al., 1998). In a study that cross-validated previously published cutting scores of several malingering instruments, Inman and Berry (2002) evaluated how well LMT performed as a part of a psychological battery. Using the recommended cutting score of <93% correct the LMT resulted in a specificity of 1.00, a sensitivity of .73, and an overall hit rate of 87%. PPP and NPP were calculated using minimum (.15), median (.23), and maximum (.48) base rates and previously noted sensitivity and specificity figures. The resulting PPP was constant at 1.00 at each of the base rates. At the lowest base rate, NPP was .96. Although NPP decreased as base rates increased, it remained fairly strong (.83) at the highest base rate. In the original validation study, large effect sizes were reported for the LMT when head injured controls are compared to analogue malingerers (d = 2.0) or to forensic evaluees identified as probable malingerers (d = 3.5). Sensitivity ranged from between .83 and .95 and specificity was 1.00 (Inman et al., 1998).

Wechsler Adult Intelligence Scale-3rd Edition (WAIS-III): The WAIS-III is an individually administered intelligence scale for adults consisting of 14 subtests that may be combined to provide indices of verbal, performance, and full scale IQ. The Verbal IQ score is an overall measure of verbal comprehension and working memory. The Performance IQ is a general measure of perceptual organization and processing speed. The Full Scale IQ is a broad measure of intelligence which is derived

by averaging the Verbal and Performance IQ scores (Kaufman & Lichtenberger, 1999). The standardization sample for the WAIS-III was selected according to 1995 U.S. census data and was stratified according to age, sex, race/ethnicity, geographic region, and education level. IQ scores represent a person's performance compared to others of the same age. Thirteen age groups, ranging from ages 16 to 89, are represented in the sample.

The reliability and validity of this measure are supported by a significant body of peer-reviewed research. Split-half reliability coefficients were .97 for Verbal IQ, .94 for Performance IQ, and .98 for Full Scale IQ. Test-retest reliability was equally strong for Verbal (.96), Performance (.91), and Full Scale IQ (.96). Much of the validation research on this instrument was obtained from studies that used the previous version of the test (WAIS-R; Wechsler, 1987). However, results from the WAIS-R can be generalized to the WAIS-III since the two versions share relevant subtests, correlate highly, and produce almost identical scaled scores (Wechsler, 1997).

The WAIS-III (Wechsler, 1997) is used in approximately 90% of neuropsychological evaluations (Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995) and is frequently used in the assessment of traumatic brain injury (Millis, Ross, & Ricker, 1998; Rogers, Harrell, & Liff, 1993). In forensic cases involving claims of substantial cognitive impairment, such as that seen in mental retardation, the WAIS-III is likely to be employed. However, since research has shown that the WAIS-III is susceptible to feigning (Bernard, 1990; Bianchini et al., 2001; Iverson & Tulsky, 2003; Millis, Ross, & Ricker, 1998; Mittenberg et al., 1995), it will be critical to include feigning detection instruments in settings where malingering is probable.

Mental Retardation

Mental retardation is a developmental disorder that is characterized by low intelligence and poor adaptive abilities, with onset in infancy, childhood, or adolescence. Mental retardation may involve any number of etiological factors including heredity, events occurring during embryonic or prenatal development, environmental influences, mental disorders, and medical conditions acquired during infancy or childhood (DSM-IV-TR; APA, 2000).

According to the DSM-IV-TR, there are three key criteria that must be met for a proper diagnosis of mental retardation to be given (pp. 41-42). The essential feature of this condition is significant subaverage intellectual functioning. This must be accompanied by limitations in adaptive functioning in 2 or more of the following skill areas: communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health, and safety. Finally, the onset of the symptoms must occur prior to age 18. Significant subaverage intelligence is defined in the DSM as an IQ of $\leq$70 (approximately 2 standard deviations below the mean) obtained by evaluation on 1 or more standardized IQ tests. Measurement error on different IQ tests varies and must be

considered when making a diagnosis, particularly when the IQ score is close to 70. Adaptive functioning refers to the ability to cope with life's demands, and to establish personal independence appropriate to age level and sociocultural standards. Adaptive skills may be influenced by education, motivation, personality, opportunities for social and vocational progress, and other mental or physical health conditions that may coexist with mental retardation.

Levels of Severity

Four degrees of severity of mental retardation are specified by the DSM-IV-TR (APA, 2000, pp. 43-44): mild, moderate, severe, and profound. According to the DSM, individuals with Mild Mental Retardation make up 85% of those diagnosed as mentally retarded. These individuals have an IQ between 50 and 70, and are considered educable, which means that they develop social and communication skills early in life. During the early developmental period, mildly mentally retarded persons have minimal impairment in sensorimotor areas, and may not be distinguishable from children without the condition until a later age. As adults, these individuals are capable of learning vocational skills, seeking employment, and living independently with minimal assistance.

Moderately mentally retarded individuals have IQ levels that range from 35 to 55. These individuals make up about 10% of the population of persons diagnosed with mental retardation. They typically acquire communication skills early in life and can attend to their personal care with moderate supervision. They benefit from vocational training and training in social skills, but do not usually progress academically beyond a second-grade level.

Those with Severe Mental Retardation make up 3% - 4% of individuals with mental retardation. These individuals have IQ levels ranging from 20 to 40 and acquire only minimal communicative skills. As children, they may be able to learn some words and, as adults, they may be able to perform simple tasks under close supervision. Most adapt to life in group homes or with their family unless they have associated conditions that require specialized care.

Individuals diagnosed with Profound Mental Retardation usually have a neurological condition that explains their mental retardation. These individuals have IQ levels that fall below 20-25 and constitute 1% - 2% of those with mental retardation. In childhood they display considerable sensorimotor dysfunction and require constant supervision and individual assistance from a caregiver.

The major alternative to the DSM for diagnosing mental retardation is provided by the American Association on Mental Retardation (AAMR, 1992; 2002). The AAMR uses the criteria similar to the DSM to define mental retardation. The AAMR (2002, p. 8), in a refinement of an earlier definition, described individuals with mental retardation as having "significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills." Conceptual skills include receptive and expressive language, reading and writing, money concepts, and

self-direction. Social skills include interpersonal, responsibility, self-esteem, gullibility, naiveté, and the ability to follow rules, obey the law, and avoid victimization. Practical skills include personal activities of daily living, instrumental activities of daily living, occupational skills, and maintaining a safe environment. The definition also states that the disability originates prior to age 18.

The AAMR's expanded description emphasizes the coexistence of both strengths and weaknesses in individuals with mental retardation, related to psychological factors, health-related diagnoses, life experiences, opportunities, availability of services, and participation in enriching activities, which may explain some individual differences in persons within the same intellectual range of mental retardation (Fredericks & Williams, 1998; Luckasson et al, 2002). In addition, the AAMR recommends a multidimensional assessment similar to the DSM's multiaxial diagnosis to identify relevant intellectual, psychological, physical, and environmental factors.

Clinical Diagnostic Protocol

Two widely accepted and comprehensive sources outlining clinical diagnostic protocols for mental retardation are the DSM (DSM-IV-TR; APA, 2000) and the ICD-10 (World Health Organization, 1992). Since serious psychopathology is two to three times more common in mentally retarded children than in the general population (Corbett, 1979; Einfeld & Tonge, 1996), making a differential diagnosis may be particularly problematic. Clinical judgment becomes increasingly difficult when it is necessary to differentiate between various behaviors and traits that can be common in both mental retardation and psychological disorders. For example, behavioral patterns and disturbances often shared by both mentally retarded children and adults and those with disorders characterized by maladaptive behavior include impaired judgment, poor insight, disruptive conduct, aggressive and/or hyperactive behavior, difficulty with attention and impulse control, criminal activity and interpersonal difficulties (Anderson & Hewitt, 2002; Einfeld & Aman, 1995; Einfeld & Tonge, 1996).

Clinicians evaluating persons suspected of having developmental disabilities should use multiaxial classification systems of the DSM or ICD-10 as a basis for diagnosing psychiatric disorders in individuals with mental retardation. This will address medical conditions and psychological disorders that may contribute to or complicate existing emotional and behavioral problems. To establish whether an individual meets diagnostic criteria for intellectual and adaptive functioning, standardized IQ tests and measures of adaptive skills have generally been used. The third criteria, diagnosis prior to age 18, may seem quite simple. However, in some instances, individuals who otherwise meet the criteria for mild mental retardation were not identified prior to age 18. Under these circumstances, school records are particularly important for establishing the "before age 18" criterion set forth by the DSM and the AAMR. All records, including attendance, disciplinary actions, grades, test scores, and psychological records should be reviewed in this regard. In addition, it may be helpful to interview teachers or other individuals

whose names appear in school records for additional information about the person's school performance and conduct. A thorough assessment such as this is necessary to establish whether a person meets the criteria for a diagnosis of mental retardation and to confirm the presence of associated conditions.

Historically, intelligence testing was based on the belief that an important component of intelligence was the ability of an individual to learn from experience and adapt to his environment. Today, the most widely used IQ tests in assessing mental retardation are the Wechsler Intelligence Scale for Children (WISC-III; Wechsler, 1991) and the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997). The WISC-III manual suggests that the most accurate diagnosis of mental retardation will derive from multiple data sources, including assessment of the individual's functioning at home, in school, and in the community (Wechsler, 1991, p. 8). In other words, intellectual deficiency alone, as demonstrated by performance on an IQ measure, is not sufficient to establish a diagnosis. Therefore, even when mental retardation is suggested by an IQ score, a diagnosis of MR should not be given in the absence of additional data documenting significant adaptive impairment.

There are a number of valid measures currently available that may be used to assess adaptive functioning. Most rely on third-party observations (i.e., caregivers, teachers, etc.), resulting in limited objectivity and reliability and the potential for response biases (Giller et al., 1986). The two most widely used examples of this type of measure are the Vineland Adaptive Behavior Scales (VABS; Sparrow, Balla, & Cicchetti, 1984) and the AAMD Adaptive Behavior Scales (ABS; Nihira, Foster, Shellhaas, & Leland, 1974). These tests were normed on children but rely heavily on reports by parents and other caregivers. They are most effective in evaluating behaviors characteristic of moderate and severe, rather than mild, retardation.

The Street Survival Skills Test (SSSQ; Linkenhoker & McCarron, 1983) is a measure of adaptive functioning that is gaining attention as a viable alternative to the VBS and ABS (Janniro, Sapp, & Kohler, 1994; Reynolds, Price, & Niland, 2003). The SSSQ is a self-report instrument developed to assess adaptive functioning (i.e., knowledge and skills relevant to community adjustment) in adults with mild to moderate mental retardation. The test is appropriate for ages 9 and older with administration time at approximately 30-60 minutes. This 216-item test is comprised of 9 subscales: (1) Basic Concepts, (2) Functional Signs, (3) Tools, (4) Domestic Management, (5) Health and Safety, (6) Public Services, (7) Time, (8) Money, and (9) Measurement. Items are presented orally and the examinee responds by pointing to 1 of 4 pictures. Each subscale yields a raw score and the raw scores of each subscale are summed to determine the total raw score. The total raw score is converted to a standard score, called the Survival Skills Quotient (SSQ). The SSQ is scaled to have a mean of 100 and a standard deviation of 15 so SSSQ scores may be compared to intelligence.

The SSSQ is supported by research demonstrating its reliability and validity (Giller, Dial, & Chan, 1986; Janniro, Sapp, & Kohler, 1994). These studies reported overall test-retest reliability (r = .99) and internal consistency (r = .97) and test-retest reliability of the subtests ranging from .87 to .95. Content validity was established through item analysis, using correlations between the SSSQ and various measures of intelligence and achievement scores. Correlation coefficients were obtained for receptive language, using the Peabody Picture Vocabulary Test (r = .64), and for reading, using the Wide Range Achievement Test (r = .21). The SSSQ was also significantly correlated with the public school version of the Adaptive Behavior Scale (r = .74), the WISC-R Full Scale IQ (r = .84), and the Daily Living subscale of the Vineland Adaptive Behavior Scale.

The SSSQ has several advantages over other tests of adaptive functioning that are based on third-party reporting. Advantages include objectivity, relatively short administration time, and self-report format. Also, it is unique in that it may be used with individuals with special needs. Its large print and graphic format are designed to accommodate those with visual acuity up to 20/200 in either eye. It is appropriate for those with motor impairments and poor verbal communication skills since no movement or language is required. A sign language version is also available.

A disadvantage of the SSSQ is that it taps knowledge relevant to adaptive skills rather than measuring the typical adaptive behavior of the individual. Thus, a person may be able to demonstrate knowledge of adaptive behaviors, yet be unable to apply that knowledge to real-life situations. Although preliminary studies provide support for the use of the test, the authors recommend that the SSSQ be included in a comprehensive test battery rather than independently until additional a larger body of independent studies on the instrument can be conducted (Giller, et al., 1986; Janniro, et al., 1994).

Although a number of objective measures, as just reviewed, are available for clinicians in the assessment of mental retardation, making a diagnosis may be challenging. Clinicians should keep in mind that test performance may be lowered as a result of concentration problems, test anxiety, fatigue, lack of interest or motivation, effects of medication, and so forth. Thus, poor performance on psychological measures may be erroneously interpreted as evidence for psychopathology or intellectual impairment. Likewise, any number of factors may affect a person's ability to adapt to his environment, live independently, and maintain a steady job. Therefore, gathering corroborative evidence from multiple data sources, including documentation of problems beginning in childhood, is necessary. The responsibility of the clinician, therefore, is to conduct a thorough and careful evaluation and consider all available data bearing on diagnostic conclusions.

Finally, in forensic settings, where clinical recommendations may influence important decisions regarding eligibility for benefits, financial compensation, and various forms of legal protection, the possibility of malingering of mental retardation must not be overlooked. The issues surrounding a recent

Supreme Court decision and the resulting implications illustrate this point and will be discussed at length throughout the remainder of this manuscript.

Legal Issues and Mental Retardation

Atkins vs. Virginia

In June 2002, the U.S. Supreme Court decided, in the case of *Atkins vs. Virginia,* that executing persons with mental retardation was unconstitutional. In this case, Daryl Renard Atkins, who was mentally retarded, and an accomplice committed armed robbery at a convenience store while under the influence of alcohol and marijuana. They then abducted a U.S. Air Force airman and forced him to withdraw money from an ATM, after which they drove him to a remote location where Atkins killed him by shooting him 8 times. At that time, Atkins had a criminal history of no less than 16 felony convictions for violent offenses. Despite the brutality of the crime and Atkins' extensive history of prior offenses, the Court found his history of mental retardation to be a crucial factor in this capital case.

The Court's ruling in *Atkins vs. Virginia* reversed a precedent set 13 years earlier in *Penry vs. Lynaugh* (1989) which held that it was within the bounds of the Constitution to execute mentally retarded individuals. John Paul Penry, whose mental retardation was documented in early childhood, was tried 3 times for the same capital murder charge because of questions surrounding the inadmissibility of evidence for his mental disabilities and evidence that his Fifth Amendment rights had been violated. Each time, however, the jury reached the same decision and Penry was sentenced to death. The Penry case brought national attention to the issue of executing persons with mental retardation, however, it was not until the *Atkins* decision that persons with mental retardation obtained federal protection from execution.

The Court supported its *Atkins* decision on the basis of 3 factors: (1) changes in society's prevailing attitudes about decency with regards to legislation pertaining to punishment, (2) legislation passed by a growing number of states prohibiting the execution of mentally retarded persons, suggesting that these individuals are less culpable, and (3) the diminished mental capacities of mentally retarded individuals (Reynolds, Price, & Niland, 2003).

Judge Stevens, writing for the majority of the Court, outlined 6 areas where diminished mental capacity is demonstrated: (1) understanding and processing information, (2) communication, (3) abstract learning and learning from experience, (4) logical reasoning, (5) impulse control, and (6) understanding the reactions of others. In light of these deficiencies, the Court acknowledged that mentally retarded persons are less culpable for the crimes they commit, although these mental deficiencies did not warrant an exemption from justice. Other potentially serious problems associated with these individuals' limitations are as follows: they have a higher incidence of making false confessions (Brodsky & Galloway, 2003), they make poor witnesses (Reynolds et al., 2003), and they fail to convey remorse to

juries (Reynolds, et al., 2003). As a result, they are more likely to be sentenced to death in spite of evidence that may justify a lesser sentence (Brodsky and Galloway, 2003).

Finally, the Court justified its decision on the grounds that executing such individuals would not noticeably contribute to the purposes of justice, including the two most often cited justifications for the death penalty, namely deterrence and retribution (Annas, 2002; Brodsky & Galloway, 2003; Reynolds, et al., 2003). This is because mentally retarded individuals have difficulty learning from mistakes and cannot readily apply consequences from past mistakes to new situations.

Competency to Stand Trial

Despite the mental incapacities noted above, mentally retarded individuals frequently understand the difference between right and wrong and may give the erroneous impression that they understand and agree to the proceedings taking place. Most, in fact, are determined by the Court to be competent to stand trial (Smith & Broughton, 1994). Competency to stand trial (CST) is an important issue to address, one that is critical for mentally retarded defendants. An estimated 14,000–20,000 offenders with mental retardation currently reside in state and federal prisons (Anderson & Hewitt, 2002), with many others likely residing in residential and mental health facilities and local jails (Smith & Hudson, 1995). In the criminal justice system, CST assessment is not routinely performed. Those most likely to be selected for CST evaluation are persons with suspected psychiatric conditions. Most criminal defendants, including those with mental retardation, do not receive CST evaluations (Applebaum & Applebaum, 1994).

To be deemed competent to stand trial, one must have an understanding of all aspects of the criminal justice system, including arrest, Miranda Rights, confessions, trial and sentencing procedures, and the roles of those involved in the proceedings (Brodsky & Galloway, 2003; Fulero & Everington, 2004). Since many cases are won on the basis of the confession, it is crucial that the defendant understand his Miranda Rights and the consequences of waiving them. He must also understand the nature and purpose of the proceedings against him, understand the roles of the key personnel in the courtroom, be able to recall and communicate information pertaining to the alleged offense to his attorney, possess the ability to appraise the potential outcomes of the proceedings, comprehend instructions and advice from his attorney, and conduct himself appropriately during the trial (Bonnie, 1990; Reynolds et al., 2003). Given the vulnerabilities of mentally retarded individuals and the potential impact on their lives when their rights are violated, the issue of competency to stand trial becomes the cornerstone for ensuring a fair trial.

CAST-MR

The Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST-MR; Everington & Luckasson, 1992) was the first validated measure of competency to stand trial for mentally retarded adults. The CAST-MR consists of 25 multiple-choice and 10 open-ended questions that

are administered in interview style. Section I measures knowledge of basic legal concepts, the meaning of the trial process, the roles and functions of key courtroom personnel and the jury. Section II assesses the defendant's understanding of the client-attorney relationship. Section III addresses the defendant's understanding of facts surrounding the case and the charges against him. Administration time is approximately 30-45 minutes.

Validation studies of this instrument (Everington, 1990; Everington & Dunn, 1995; Everington & Luckasson, 1992) reported similar findings which provided strong support for its usefulness. Findings suggest a high degree of internal reliability across different samples, with reliability coefficient ranges reported as follows: internal reliability for the CAST-MR total score (r = .92-.93), test-retest reliability (r = .89-.90). Internal reliability for individual sections as follows: Section I (r = .84- .91), Section II (r =.79-.76), and Section III (r = .83). Interrater reliability was adequate (r = .80-.87). Discriminant function analyses demonstrated the instrument's ability to discriminate between groups at cutting scores established for the total score and each of the 3 sections (Everington & Dunn, 1995). Classification rates were shown to be similar to other measures designed to address this issue. Overall, these results indicated that the CAST-MR can be a valuable tool for determining whether a mentally retarded individual is competent to stand trial.

Malingering Mental Retardation

A final issue relevant to defendants with mental retardation involves malingering. Interestingly, a recent (2003) report from Louisiana's Chapter of AAMR stated that it is virtually impossible to fake mental retardation. This claim was based on the belief that well-established, scientifically accepted, and valid measurement tools have safeguards (i.e., validity scales) built in to alert experts to malingered performance. Furthermore, the article purports that the real problem with dishonest responding is that mentally retarded individuals may try to fake having normal cognitive abilities. Others have shared this observation, citing examples from research showing the tendency of the mentally retarded towards acquiescence and the strong desire to please authority figures (AAMR, 1992; Schlesinger, 2003; Shaw & Budd, 1982; Sigelman et al., 1982). However, this assertion is controversial, as will be reviewed below.

As a result of the *Atkins vs. Virginia* ruling, the criminal justice system may have a larger number of capital defendants claiming mental retardation as a defense to avoid the possibility of execution. Rather than settling the issue with a single ruling, the Court's decision generated new debate, primarily because it failed to establish classification criteria for determining who qualifies for this status. The Court made it the responsibility of each individual state to determine whether a person is mentally retarded, although it recommended using the clinical definitions set forth by the American Psychiatric Association (APA) and the American Association on Mental Retardation (AAMR). However, by not establishing standard

criteria, the Court increases the potential for states to adopt different standards which then creates inequity within the nation in making execution exemption decisions (Greenspan & Switzky, 2003).

Addressing this controversial ruling, Supreme Court Justice Antonin Scalia argued that the symptoms of mental retardation could be readily feigned with little to no risk on the part of the capital defendant (Annas, 2002). Scalia pointed out that one who feigns psychiatric illness risks being institutionalized until he can be rehabilitated, after which he may be tried and executed. The offender who successfully feigns mental retardation, on the other hand, is not suffering from an illness that can be treated. Therefore, it would seem that rehabilitation is not a viable option. A recent study (Anderson & Hewitt, 2002) investigated the effects of competency restoration training on mentally retarded defendants who were found NGBRI by the court. The training program was designed to educate defendants on the types of information they would be expected to understand in order to be declared competent to proceed. Results indicated that such training may not be effective. These findings supported previous research (Daniel & Menninger, 1983; Ellis & Luckasson, 1985) suggesting that MR is a chronic condition that is largely unresponsive to treatment. Perhaps not surprisingly, Anderson and Hewitt found that IQ was related to treatment efficacy such that those with higher IQs (in the upper limits of mild MR) were able to derive some benefit from the training.

The Clinical Challenge

The preceding issues underscore the complexity of the criminal justice system, particularly in dealing with special populations. On the one hand is the responsibility to punish those who commit wrongful acts and prevent them from causing further harm. On the other hand, it is necessary to protect individuals who lack the capacity to understand the implications of their behavior by ensuring that they receive a fair trial and any protections to which they are entitled by law. In order to serve justice, then, it is imperative to devise effective methods to accurately distinguish between those who are mentally retarded and deserve protection and those who may be feigning impairment. The most common method to distinguish between the two involves an extensive evaluation that includes tests designed to detect malingering. Although many of the tests routinely used in forensic evaluations have been validated on various clinical samples (i.e., psychiatric patients and head-injured patients) few studies have looked at malingered mental retardation. Therefore, it has not yet been determined how well these tests perform when used with mentally retarded samples.

Research on the Detection of Malingered Mental Retardation

The first study that attempted to validate a test battery to detect malingered mental retardation used prison inmates as simulators (Schretlen and Arkowitz, 1990). Two groups of inmates, one instructed to feign psychiatric illness, and the other instructed to feign mental retardation, were compared to two criterion groups responding honestly and a control group comprised of prison inmates. The criterion

groups included psychiatric inpatients and mentally retarded adults. The inclusion of both clinical and control groups, the administration of a battery of tests rather than a single measure, and provision of monetary incentive to comply with instructions to feign believably, increased laboratory control and enhanced the generalizability of the study.

The test battery consisted of the MMPI (Hathaway & McKinley, 1943), Bender-Gestalt (Bruhn & Reed, 1975), and the Malingering Scale (a 90-item, pencil and paper instrument developed for this study). The MMPI-2's F and F-K raw scores, 5 of the 11 Bender-Gestalt scales, and the 4 subtests of the Malingering Scale (Information, Arithmetic, Vocabulary, and Abstraction) were used to detect malingered performance. The mentally retarded participants did not take the MMPI-2 due to its high reading level. The test battery correctly classified 85% of subjects who faked insanity with no false positives and 80% of subjects who faked mental retardation with 7.5% false positives. The classification accuracy of the battery was better than that of any single instrument in the battery taken alone. Unfortunately, those instructed to feign mental retardation scored significantly lower on the Malingering Scale than inmate controls but not as low as the mentally retarded criterion group.

It is important to note that the test battery, the Malingering Scale in particular, had not been systematically validated previously. Ideally, the group faking mental retardation would have performed more poorly than the actual mentally retarded group on the Malingering Scale. This unexpected finding may point to a potential shortcoming of the Malingering Scale's ability to accurately classify malingered performance. Interestingly, those feigning insanity scored lower than the psychiatric inpatient group, but not as low as those feigning mental retardation. This suggests that even psychiatric malingerers at times intentionally present themselves as intellectually impaired.

The next major relevant study (Hayes and Hale, & Gouvier, 1997) investigated whether several brief tests commonly used to identify malingered psychiatric illness and neurocognitve impairment could accurately classify mentally retarded individuals as feigning or honest. The test battery consisted of the Dot Counting Test (Lezak, 1995), Memory for Fifteen Items Test (MFIT; Lezak, 1995), and the M-Test (Beaber, Marston, Michelli, & Millis, 1985). The subjects were recruited from a state facility for the criminally insane and comprised three groups: (a) pretrial patients who were not malingering, (b) patients found by the court to be not guilty by reason of insanity (NGBRI), and (c) individuals previously identified as malingering. Results ran counter to expectations. The malingering patients performed better than the non-malingering group on all of these measures. Discriminant function analyses resulted in the correct classification of only 59.5%. Consequently, the authors suggested that these tests were not appropriate for identifying malingered performance among mentally retarded defendants.

The following year, Hayes, Hale, and Gouvier (1998) expanded the previous study by including the administration of the Structured Interview of Reported Symptoms (SIRS; Rogers, et al., 1992) with

the three malingering tests used previously (Dot Counting Test, M-Test, and MFIT) to a sample of thirty-nine patients in a forensic mental hospital. Diagnoses of mental retardation and malingering were independently made by an interdisciplinary team consisting of psychologists, social workers, correction officers, and nurses. The 3 experimental groups included: (1) a nonmalingering group made up of individuals with a pretrial status who were considered not competent to stand trial by the courts, (2) a malingering group made up of pretrial patients who met DSM criteria for malingering and had been observed engaging in behavior inconsistent with their stated complaints, and (3) a group of individuals found by the courts to be NGBRI.

Upon completion of testing, Hayes, et al. (1998) assessed the classification accuracy of the test battery by running three separate multiple discriminant function analyses (DFA). The first, using the raw scores from the eight primary scales of the SIRS, found that the SIRS correctly classified 94.9% of cases. Eighty-nine percent of malingers, 94.4% of NGBRI patients, and 100% of the nonmalingering pretrial patients were correctly classified. The second DFA took into account the data from the three malingering measures and found that, when added to the SIRS, all cases were correctly identified. The third analysis omitted the SIRS and found that only 59.9% of cases were correctly classified. False positives for each group ranged from 23.1% (nonmalingerers) to 33.3% (malingerers).

Unfortunately, results from these studies are somewhat conflicting. While the Schretlen and Arkowitz (1990) research suggested the utility of cognitive feigning instruments for identifying malingered MR, Hayes et al. (1998) reported the superiority of psychiatric feigning indices in detecting feigning. However, this is not necessarily surprising given that the latter study was designed to investigate feigned psychiatric disorder, not MR.

Although there were several methodological weaknesses in the studies that limited both internal and external validity, these were the first published studies to use real-world mentally retarded malingerers. They clearly demonstrated the utility of using validated feigning indices, such as the SIRS, for detecting malingering and indicated the importance of including such instruments in neuropsychological test batteries.

These early studies potentially paved the course for new research and suggested ways to improve methodologies used to investigate malingered mental retardation. However, in the decade following Schretlen and Arkowitz's (1990) attempt to validate a test battery for this purpose, there was not a great deal of activity in this area of research. Since the *Atkins vs. Virginia* ruling, however, many are likely to turn their interest to this topic as it has become increasingly important for forensic neuropsychologists to understand and document prototypical patterns of malingered mental retardation and improve on existing measures for detecting feigned performance.

Statement of the Problem

This small body of existing research provides some information about current malingering measures applied to mentally retarded samples. However, there remains a growing need for further research in this area. A probable implication of *Atkins vs. Virginia* is an increase in the number of capital defendants who will claim mental retardation as a defense against the death penalty. Should this increase occur, psychologists will need to have available a repertoire of empirically supported, psychometrically sound instruments that can be applied to the detection of feigned mental retardation.

The problems associated with misdiagnosing a psychiatric illness or falsely classifying individuals presenting legal claims affect not only those individuals, but also society, collectively, as costs incurred by employers, insurance companies, disability programs are passed along to consumers and taxpayers or when criminals are released into their former communities, threatening the security of its citizens. Other damages resulting from false classification include failure to provide services for individuals with legitimate claims and reduced credibility of psychological reports and testimony in the courtroom (Bianchini, et al., 2001). Because the costs of malingering can be quite high, it is important for clinicians to make accurate decisions about the extent of psychiatric and neurological impairment the client may be exhibiting (Owens, 1995; Sweet, 1999).

These circumstances suggest that clinicians evaluating criminal defendants for this condition would benefit from the availability of objective techniques for identifying attempts at feigning mental retardation. However, it has not been determined whether validated cutting scores for standard psychiatric and neuropsychological instruments, established on other impaired populations, can be effectively applied to mentally retarded individuals to detect malingered performance. There is a need for empirical research in this area to determine (a) how the "typical" mildly mentally retarded person will perform on these measures, (b) which tests are most appropriate for the detection of malingering in this population, (c) how the "normal" group instructed to dissimulate will approach the tests (i.e., what types of deficits they will elect to feign), and (d) how cutting scores may need to be adjusted to maximize classification accuracy.

Purpose of the Study

The purpose of the present study is to evaluate standard psychological and neurocognitive feigning instruments to determine how effectively they distinguish between persons with mild mental retardation and community volunteers instructed to feign symptoms to appear mildly mentally retarded. Since more severe levels of the condition should be well documented and undisputed, a criterion group limited to individuals diagnosed with mild mental retardation will be selected for participation.

Eight instruments will be evaluated independently and in combination for the purpose of constructing a test battery that will potentially increase the classification accuracy in forensic evaluations of individuals claiming mild mental retardation. It will be important to determine what measurable

differences exist between the honestly responding mentally retarded criterion group and “normals” instructed to malinger on these measures. If the measures used in this study prove to be appropriate for this population, the results may require adjustments to the previously established cutting scores to distinguish between honest and dissimulated performance and improve the classification accuracy of these measures.

Chapter Two

Methodology

Participants

Participants included both mentally retarded (MR) and community volunteers. The MR group consisted of 26 mildly mentally retarded adults who were their own legal guardians, recruited from various local day-treatment centers for developmentally disabled adults, who are served by the Bluegrass Regional Mental Health/Mental Retardation Board. Community volunteers (CV) were 35 individuals recruited through fliers posted in the local vocational assistance office and medical clinics associated with University of Kentucky. Community volunteers were required to have completed no more than the 11th grade of high school. This was to ensure a sample that might plausibly claim intellectual deficits. The community volunteers were screened for substance abuse, neurological conditions, and major psychiatric illness. The community sample was randomly assigned to either the honest (CVH; n=10) or malingering (CVM; n=25) group.

Materials

Materials included an approved IRB informed consent agreement signed by all participants, a release form signed by MR participants permitting their medical records to be examined for diagnostic information, a self-report demographics questionnaire (Appendix A), instruction sets (Appendices B & C) partly adapted from Schretlen and Arkowitz (1990), the Wechsler Test of Adult Reading (WTAR: Psychological Corporation, 2001) which uses irregular word reading ability and demographic information to estimate premorbid FSIQ, and a battery of tests which included a test of intelligence (WAIS-III), 2 psychiatric malingering instruments (SIMS and M-FAST), 3 motivational tests for detecting malingered neurocognitive deficits (DMT, LMT, and TOMM), 1 test of adaptive functioning (SSSQ), and 1 test for competency to stand trial (CAST-MR). Community volunteers assigned to the malingering group also received information about MR (Appendix D), and a MR quiz (Appendix E) to assess understanding of MR facts.

Following test administration, all CV participants completed a post-test debriefing form (Appendix F) on which they were asked to reproduce their initial instructions and rate, on a scale of 0 (low) to 10 (high), how well they understood the instructions and how successful they believed themselves to have been in fulfilling their assigned role. Participants in the malingering (CVM) group received an additional debriefing statement (Appendix G) explaining the reason for deception in the study and thanking them for their participation.

Procedure

After obtaining IRB approval from both the University of Kentucky and the Bluegrass MH/MR Board, the first author contacted facility case managers and explained the study and the recruitment criteria. These managers then approached day MR program participants who met diagnostic and legal criteria about participation in the study. Those who indicated willingness to participate met with either the first author or an IRB approved research assistant who explained the study, obtained consent and written permission to review medical records, presented the demographic questionnaire orally, and recorded participants' answers. Following this, the researcher administered the test battery in counterbalanced order using standard instructions, in a private room within the facility serving the participant, allowing for breaks as needed. Upon completion of the testing, participants were paid and thanked.

Community volunteers contacted the first author by phone or email to indicate interest in participation. After screening the potential participant for years of education, substance abuse, and major neurological or psychological conditions, a meeting date/time was scheduled. Testing was conducted in the Psychology Department's training clinic on the University of Kentucky campus. Testing of all CV participants required 2 research assistants, hereafter referred to as RA1 and RA2. CV participants were met by RA1 who obtained consent, supervised completion of the demographics questionnaire, administered the WTAR under standard instructions, and provided the participant with a sealed envelope containing the instruction assigning him or her to either the honest (CVH) or malingering (CVM) group.

If assigned to the CVM group, the basis of the study was briefly explained and a fact sheet containing information about MR was presented. This was followed by the MR quiz. Given the potential implications associated with successful dissimilation (i.e., that the participant will be judged as stupid by the examiner) RA1 reassured the participant that the examiner (RA2) would be told afterwards what the participant's instructions were. This was done in an attempt to minimize potential anxiety associated with participants' instructions to feign and appear unintelligent. If assigned to the CVH group, participants did not receive additional information about the basis of the study or MR.

Following these procedures, RA1 left the room and RA2 entered. RA2, who was blind to participants' group assignment, administered tests in counterbalanced order, allowing for breaks as needed. Upon completion of testing, RA1 returned and instructed participants to complete the debriefing questionnaire to assess for compliance with instructions. Next, in the CVM condition, RA1 debriefed the participant and revealed his/her instructions to feign MR to RA2. Finally, RA1 paid and thanked all participants.

Chapter Three

Results

Demographic and background characteristics are presented in Table 3.1. Two of the 3 possible comparisons were of most interest to the present study. Comparison of the 2 Community Volunteer participant groups (CVH vs. CVM) provided an experimental manipulation check, whereas contrasts between the CVM and the Mentally Retarded participants (CVM vs. MR) discriminate between feigned and genuine MR. Continuous variables were analyzed using t-tests, whereas chi-square procedures were used to analyze categorical variables. These analyses revealed that the groups were comparable on all variables except years of education, special education, and developmental disability status. MR participants had significantly more years of education than the CVM group ($t(49) = 5.43$, $p < .001$), all MR participants (but no CVM) received special education ($t(49) = -51.00$, $p < .001$), and all MR participants carried a diagnosis of developmental disability ($t(49) = -51.00$, $p < .001$). Results showed that the CVH and CVM groups were comparable on the WTAR estimated FSIQ ($M = 80.2$, $SD = 8.89$ and $M = 82.12$, $SD = 9.26$, respectively), which was administered under standard instructions prior to group assignment. This finding suggests that the intellectual level of the CV groups were similar prior to experimental manipulation. Also worth noting is the finding that 43% of the community volunteers reported a criminal history which involved either a felony or misdemeanor offense. This finding suggests that this sample is ecologically valid for investigating feigning to evade criminal punishment.

Results from the WAIS-III IQ and Index Score are presented in Table 3.2. These data were analyzed using one-way ANOVAs and Tukey's HSD for post-hoc comparisons with an alpha level of .05. Due to the varying numbers of participants in the two contrasts, the magnitude of the group differences was standardized using Cohen's $\underline{d}$. Main effects of group across all variables were significant. Follow-up comparisons reveal group differences as significant only in the CVH vs. CVM comparison. Despite being comparable on the WTAR estimated FSIQ, the CVM group produced FSIQ scores approximately 2 standard deviations below the CVH group ($\underline{d} = 2.11$). This indicates that the experimental manipulation (i.e., instruction to feign MR) resulted in a substantial suppression of performance on this measure. This pattern was consistent across all IQ and Index scores. The other comparison of interest, CVM vs. MR, was not statistically significant. The CVM group produced FSIQ scores quite comparable to the MR group ($\underline{d} = -.24$), suggesting that the CVM group was able to feign cognitive deficits similar to those demonstrated by individuals with genuine MR. Again, this pattern was consistent across all WAIS-III comparisons.

Table 3.3 illustrates results from measures of adaptive functioning and knowledge of legal concepts. Similar to the previous analyses, these data were analyzed using one-way ANOVAs followed by Tukey's HSD comparisons, with an alpha level of .05. Analyses revealed significant main effects (*F*

(2, 58) = 15.53, p < .001) of group on the Street Survival Skills Questionnaire (SSSQ). The CVM group suppressed their scores significantly compared to the CVH group ($\underline{d}$ = 1.80). Compared to the MR group, however, the CVM seemed to overestimate the level of impairment in adaptive functioning in individuals with genuine MR ($\underline{d}$ = 1.02). Two separate analyses were conducted to examine group differences on the Competency Assessment to Stand Trial for Defendants with Mental Retardation (CAST-MR). The first analysis, which looked at group differences on the CAST-MR Part I (assessing knowledge of courtroom personnel and procedures) was significant (F (2, 58) = 19.23, p < .001). The second analysis, examining group differences on the CAST-MR Part II (assessing skills to assist in defense) was also significant (F (2, 58) = 14.74, p < .001). On both Parts I and II, the performance of the CVM group was suppressed relative to the CVH group ($\underline{d}$ = 2.27 and 2.02, respectively). Once again, however, on both scales, the CVM group overestimated the level of impairment that persons with genuine MR are likely to demonstrate (Part I, $\underline{d}$ = .90; Part II, $\underline{d}$ = .68). Results of these analyses suggest that the CVM group had more difficulty portraying realistic deficits in adaptive functioning and knowledge of legal concepts than they did on the WAIS-III.

Results from the measures of psychiatric feigning are presented in Table 3.4. One-way ANOVAs, followed by Tukey's HSD comparisons, were used in these analyses. Significant main effects of groups were found on each of the 3 indices. However, follow-up tests revealed that only the CVH vs. CVM comparisons were statistically significant. The CVM group elevated scores relative to the CVH group by about 1 standard deviation on all 3 indices (MFAST, $\underline{d}$ = -1.03; SIMS, $\underline{d}$ = - 1.19; and SIMS LI, $\underline{d}$ = -1.09). The CVM group also elevated scores relative to the MR group, however, with only a modest difference ($\underline{d}$ = -.37; SIMS; $\underline{d}$ = -.51; and SIMS LI, $\underline{d}$ = -.56). These results suggest that these instruments would be unlikely to allow for reliable classification of individuals with genuine vs. feigned MR.

Results from the neurocognitive feigning instruments are presented in Table 3.5. Main effects were significant across all measures. Effect sizes (Cohen's $\underline{d}$) for the CVH vs. CVM group comparison were quite high, ranging from 1.94 to 2.24. For the more clinically relevant comparison between CVM and MR groups, effect sizes remained high, ranging from 1.59 to 2.29. It is noteworthy that the MR group means for these indices all fall very close to recommended cutting scores, which range from .90 to .93. Given that traumatic brain injury (TBI) populations usually score close to .99 correct, this finding suggests that individuals with MR have mild deficits on neurocognitive feigning tests.

Table 3.6 presents the test operating characteristics for all psychiatric and neurocognitive feigning indices using the recommended cutting scores for each test. Overall hit rates for psychiatric feigning measures are near chance level. Positive and negative predictive powers (PPP and NPP) for these 2 measures are poor, even at the rather high base rate (.49) in the present sample. Consistent with $\underline{d}$ scores reviewed previously (see Table 3.4), these measures are unlikely to be useful, at the recommended cutting

scores, in detecting feigned MR. Area under the curve (AUC) data, which represent the accuracy of a test at all possible cutting scores, point to the superiority of the neurocognitve indices over the psychiatric feigning indices for detecting feigned MR. Results from the neurocognitive feigning instruments, however, are more promising. The overall hit rates at recommended cutting scores range from .73 (LMT) to .81 (DMT). Sensitivity (SN) data for these measures are moderately strong and acceptable, ranging from .76 (DMT) to .88 (LMT). However, specificity (SP) is lower than those reported in other populations (e.g., TBI), ranging from .57 (LMT) .85 (DMT). This is likely due to MR participants' lower mean scores on these tests. Given the poorer performance on neurocognitive feigning tests by persons with MR and the increased likelihood of execution associated with false positive errors, it may be useful to adjust the cutting scores to reflect the MR group's lower scores.

Given the modest specificity rates discussed above and the high stakes associated with false positive errors, cutting scores on the neurocognitive feigning instruments were revised such that specificity would be maintained at approximately .95 in the present sample. Positive and negative predictive powers were calculated with a malingering base rate of .39, which is an intermediate between estimates of feigned psychiatric disorder (.37) and feigned neurocognitve impairment (.41) in forensic settings. The proposed revisions to recommended cutting scores are presented in Table 3.7. Not surprisingly, adjusting the cutting scores improved SP considerably with only a mild decline in SN. Changing the recommended cutting scores also increased PPP across all measures, ranging from .90 to .92, although with some decline in NPP (ranging from .77.to 86). These results suggest adequate confidence for predicting feigned performance with a positive test sign.

Considering the malingered neurocognitive deficit (MNCD) tests as a battery, Table 3.8 explores the effect of increasingly stringent criteria for predicting malingered performance using the 4 neurocognitive feigning indices (TOMM has two indices: T2 and RT) and revised cutting scores. These analyses suggest that using a threshold of ≥ 2 or more failed tests results in an acceptable PPP (.95) and moderately strong NPP (.84). Although these results are promising, cross-validation will be necessary to determine whether these improved test operating characteristics can be maintained across new samples.

Table 3.1

Demographic and Background Characteristics of Groups

Variable		CVH (N=10)	CVM (N=25)	MR (N=26)	CVH vs. CVM t / χ^2	CVM vs. MR t / χ^2
Age	M	32.30	34.24	37.12	-.34	7.94
	SD	14.51	15.52	9.51		
Education	M	10.30	9.72	11.42	1.77	5.43 ***
	SD	0.68	1.24	.99		
Special Ed.	%	0%	0%	100%	---	-51.00 ***
Dev. Disability	%	0%	0%	100%	---	-51.00 ***
Male	%	50%	56%	65%	-.10	-.47
Caucasian	%	50%	40%	57%	-.61	-2.34
Right-handed	%	100%	100%	76%	---	-6.54
Live Alone	%	30%	48%	31%	-.95	-1.59
Learning Disability	%	2%	4%	1%	-2.33	-.31
Psychiatric Hx	%	30%	12%	19%	-1.63	-.50
Arrest Hx	%	60%	36%	23%	-1.68	-1.03
Head Injury Hx	%	30%	28%	27%	-.01	-.01
WTAR Predicted IQ	M	80.22	82.12	---	-.53	---
	SD	8.89	9.26	---		

Note: CVH = Community Volunteer Honesr; CVM = Community Volunteer Malingering; MR = Mentally Retarded. t = t-test value; χ^2 = Chi-square value; * = $p < .05$; ** = $p < .01$; *** = $p<.001$. Education = years of education completed; Male = percent male; Special Ed. = special education program while in school; Live Alone = currently living alone; Dev. Disability = Mental Retardation; Psychiatric Hx = History of Psychiatric Diagnosis; WTAR Predicted IQ = Wechsler Test of Adult Reading predicted IQ score.

Table 3.2

WAIS-III IQ and Index Scores from Community Volunteers and MR Participants

Variable		CVH	CVM	MR	F (2, 58)	CVH vs. CVM d	MR vs. CVM d
WAIS FSIQ	M	80.70	61.72	60.00	28.20 ***	2.11 ***	-.24
	SD	9.14	8.94	5.51			
WAIS VIQ	M	80.10	64.96	62.50	22.83 ***	1.80 ***	-.37
	SD	7.92	8.62	4.90			
WAIS PIQ	M	84.40	64.84	64.23	22.78 ***	2.01 ***	-.08
	SD	11.33	9.16	6.45			
WAIS VCI	M	83.10	67.64	65.23	17.84 ***	1.84 ***	-.29
	SD	7.52	8.73	7.92			
WAIS POI	M	90.20	66.68	67.69	33.98 ***	2.52 ***	.16
	SD	13.82	7.65	5.15			
WAIS WMI	M	82.50	63.40	57.96	26.17 ***	1.75 ***	-.72
	SD	14.15	9.73	5.57			
WAIS PSI	M	85.10	66.60	63.35	23.68 ***	1.77 ***	-.40
	SD	9.42	10.82	5.37			

Note: CVH = Community Volunteer Honest; CVM = Community Volunteer Malingering; MR = Mentally Retarded; VCI = Verbal Comprehension Index; POI = Perceptual Organizational Index; WMI = Working Memory Index; PSI = Processing Speed Index. d = Cohen's d. * = $p<.05$; ** = $p<.01$; *** = $p<.001$ using ANOVA or Tukey's HSD.

Table 3.3

Results from Street Survival Skills Questionnaire and Competency Assessment to Stand Trial for MR Defendants

Variable		CVH	CVM	MR	F (2, 58)	CVH vs. CVM d	MR vs. CVM d
SSSQ Total	M	192.00	108.52	150.31	15.53 ***	1.80 ***	1.02 **
	SD	7.02	61.29	21.29			
SSSQ DIS	M	19.90	9.08	13.81	16.47 ***	2.03 ***	.89 **
	SD	.32	7.20	3.45			
SSQ	M	94.70	46.40	67.31	18.47 ***	2.06 ***	.94 **
	SD	4.74	30.29	13.95			
CAST MR 1	M	93.60	40.80	62.77	19.23 ***	2.27 ***	.90 **
	SD	3.86	30.46	18.62			
CAST MR 2	M	87.90	38.12	56.00	14.74 ***	2.02 ***	.68 *
	SD	9.66	30.24	22.31			

Note: CVH = Community Volunteer Honest group (N = 10); CVM = Community Volunteer Malingering group (N = 25); MR = Mentally Retarded group (N = 26); SSSQ DIS = Street Survival Skills Questionnaire (Disabled Norms); SSQ = Survival Skills Quotient; CAST MR = Competency to Stand Trial for MR Defendants (parts 1 and 2). d = Cohen's d. * = $p<.05$; ** = $p<.01$; *** = $p<.001$ based on ANOVA or Tukey's HSD.

Table 3.4

Results from Psychiatric Feigning Measures

Variable		CVH	CVM	MR	F (2, 58)	CVH vs. CVM d	MR vs. CVM d
MFAST Total	M	3.10	8.92	6.73	3.81 *	-1.03 *	-.37
	SD	3.35	6.52	5.49			
SIMS LI	M	4.20	6.72	5.58	5.73 **	-1.09 **	-.56
	SD	1.93	2.46	1.60			
SIMS Total	M	18.30	36.00	28.65	5.94 **	-1.19 **	-.51
	SD	11.03	16.25	12.45			

Note: CVH = Community Volunteer Honest group; CVM = Community Volunteer Malingering group; MR = Mentally Retarded group; MFAST = Miller Forensic Assessment of Symptoms Test; SIMS = Structured Interview of Malingered Symptomatology; SIMS LI = Low Intelligence Scale from SIMS; d = Cohen's d. * = p<.05; ** = p<.01; *** = p<.001 based on ANOVA or Tukey's HSD.

Table 3.5

Results from Neurocognitive Feigning Measures

Variable		CVH	CVM	MR	F (2, 58)	CVH vs. CVM d	MR vs. CVM d
DMT % Correct	M	100.00	58.36	94.93	28.50 ***	1.94 ***	2.07 ***
	SD	.00	29.49	6.24			
LMT % Correct	M	99.20	48.84	91.22	39.06 ***	2.34 ***	2.29 ***
	SD	1.40	29.13	8.37			
TOMM T2 % Correct	M	100.00	57.40	91.40	23.32 ***	2.04 ***	1.59 ***
	SD	.00	14.34	7.18			
TOMM RT % Correct	M	99.60	55.20	93.20	26.61 ***	2.03 ***	1.77 ***
	SD	0.42	14.86	6.73			

Note: CVH = Community Volunteer Honest group; CVM = Community Volunteer Malingering group; MR = Mentally Retarded group. DMT = Digit Memory Test; LMT = Letter Memory Test; TOMMT2 = Test of Memory Malingering 2nd Learning Trial; TOMMRT = Test of Memory Malingering Retention Trial; d = Cohen's d; * = p<.05; ** = p<.01; *** = p<.001 based on ANOVA or Tukey's HSD.

Table 3.6

Test Operating Characteristics for Discriminating Community Volunteer Malingerers and MR Participants Using Recommended Cutting Scores

Test	CUT SCORE	SN	SP	HR	PPP	NPP	AUC
MFAST Total	≥ 6	.68	.50	.59	.57	.62	.61
SIMS Total	> 16	.88	.23	.55	.52	.67	.65
DMT % Correct	< 90%	.76	.85	.81	.83	.79	.83
LMT % Correct	< 93%	.88	.58	.73	.67	.83	.87
TOMM T2 % Correct	< 90%	.80	.69	.75	.71	.78	.85
TOMM RT % Correct	< 90%	.80	.81	.80	.80	.81	.86

Note: MFAST = Miller Forensic Assessment of Symptoms Test; SIMS = Structured Interview of Malingered Symptomatology; DMT = Digit Memory Test; LMT = Letter Memory Test; TOMMT2 = Test of Memory Malingering, 2nd Learning Trial; TOMMRT = Test of Memory Malingering, Retention Trial; SN = Sensitivity; SP = Specificity; HR = Percent Correct Classification (Hit Rate); PPP = Positive Predictive Power; NPP = Negative Predictive Power; AUC = Area Under the Curve.

Table 3.7

Neurocognitive Malingering Test Operating Characteristics for Discriminating Community Volunteer Malingerers and MR Participants Using Cutting Scores Adjusted to Maintain Specificity of Approximately .95 and Malingering Base Rate of .39

Test	CUT SCORE	SN	SP	HR	PPP	NPP
DMT % Correct	< 80 %	.72	.96	.87	.92	.84
LMT % Correct	< 70 %	.76	.96	.88	.92	.86
TOMM T2 % Correct	< 60 %	.56	.96	.80	.90	.77
TOMM RT % Correct	< 60 %	.60	.96	.82	.91	.77

Note: MFAST = Miller Forensic Assessment of Symptoms Test; SIMS = Structured Interview of Malingered Symptomatology; DMT = Digit Memory Test; LMT = Letter Memory Test; TOMMT2 = Test of Memory Malingering 2nd Learning Trial; TOMM RT = Test of Memory Malingering Retention Trial; SN = Sensitivity; SP = Specificity; HR = Percent Correct Classification (Hit Rate); PPP = Positive Predictive Power; NPP = Negative Predictive Power.

Table 3.8

Test Operating Characteristics for Using Increasingly Stringent Criteria for Predicting Feigned MR Using Neurocognitive Malingering Tests and Cutting Scores Maintaining Specificity of Approximately .95 and Malingering Base Rate of .39

Test Failures	SN	SP	HR	PPP	NPP
≥ 1	.84	.89	.87	.83	.90
≥ 2	.72	.96	.87	.92	.84
≥ 3	.60	1.00	.84	1.00	.80
≥ 4	.48	1.00	.80	1.00	.75

Note: Failures = number of neurocognitive malingering tests falling below recommended cutting scores; SN = Sensitivity; SP = Specificity; HR = Percent Correct Classification (Hit Rate); PPP = Positive Predictive Power; NPP = Negative Predictive Power.

Chapter Four

Discussion

Detection of malingering has become increasingly more important in recent years, particularly as results from forensic evaluations have become more widely accepted in courtrooms. Rising costs to society when malingering goes undetected are quite substantial. Conversely, erroneous classification of an honest person as malingering may be equally devastating on a more personal level in terms of the individual's reputation as well as the consequences of being denied compensation or services that are needed and justified. Further, in high stakes situations, such as those involving mentally retarded defendants convicted of a capital crime, such decisions may literally mean life or death.

Although exact prevalence data are unavailable, research suggests that feigning on psychological measures is more common in forensic settings than in other contexts (Miller, 2001) and that base rates rise in relation to increases in the potential award involved (Berry et al., 2002). In a recent survey of malingering literature, malingering base rates in forensic settings were estimated to be between 31% and 45% (Berry et al., 2002). Therefore, in any context where one stands to gain some benefit from presenting as impaired, malingering must be considered.

Malingering research in the past several decades has focused primarily on psychiatric and head-injured populations. However, a recent Supreme Court decision preventing the execution of capital defendants with mental retardation may have increased the probability that certain individuals may elect to falsely claim mental retardation as a defense to minimize serious judicial consequences if convicted. Unfortunately, little research into the detection of feigned MR has been published and the few available studies addressing this issue report conflicting results regarding the superiority of psychiatric vs. neurocognitive instruments in identifying malingering in MR populations.

Clinicians involved in forensic cases where a diagnosis of MR has not been previously made may be called upon to determine whether the defendant meets the diagnostic criteria for the condition and is therefore entitled to protection under the law. In such evaluations, it would be necessary to determine the individual's intellectual capacity and adaptive functioning and, if possible, compare this information with existing corroborative data such as school records. A typical forensic evaluation involves the administration of a test battery covering all relevant domains, including malingering. Assessment of malingering requires the use of valid, reliable measures with cutting scores established in controlled validation studies using relevant criterion and control groups. Unfortunately, no such validation studies have been successful in establishing which tests are appropriate to use for detecting malingered MR.

The present study attempted to address these issues by comparing participants with mild MR to demographically comparable community volunteers instructed to feign the condition. Two out of 3 possible comparisons were of interest in the present study. The CVH vs. CVM comparison served as a

manipulation check, whereas the CVM vs. MR contrast demonstrated how well the tests discriminated genuine from feigned MR. In addition to measures of psychiatric and neurocognitive feigning, tests of intelligence, adaptive functioning, legal knowledge were administered by research assistants who were blind to CV group assignment.

Demographic and background data suggested that the groups were comparable on most relevant variables. Results demonstrated that community volunteers feigning MR (CVM) were able to suppress their performance on the WAIS-III to produce IQ and Index Scores comparable to those in the MR group. The CVM group, however, scored significantly lower than the CVH group on this measure, despite being relatively identical on estimated premorbid IQ prior to the instruction set, indicating successful experimental manipulation. Performance on measures of adaptive functioning and knowledge of courtroom personnel and procedures suggested that the CVM group overestimated MR deficits in these areas.

Psychiatric feigning instruments were relatively ineffective in distinguishing genuine and feigned MR using recommended cutting scores. Neurocognitive feigning measures showed more promise, but specificity values at the recommended cutting scores were lower than reported in other populations and unacceptable in light of the high costs involved with false positive errors. This was likely due to mild deficits in MR individuals on these tests. Positive Predictive Power (PPP) was modest, and again unacceptable for making predictions in such a high stakes situation. Since recommended cutting scores were derived from non-MR patient populations, alternative cutting scores based on the present samples were developed. Revised cutting scores were selected to maintain a specificity of approximately .95 and predictive power was calculated using a modified malingering base rate of .39. Using revised cutting scores, PPP ranged from .90 - .92, which is acceptable for such important decisions.

Using the neurocognitive feigning measures as a battery may be useful to increase certainty of predictions based on positive test results. At the revised cutting scores, a criterion of ≥ 2 failed tests yielded a sensitivity of .72 and specificity of .96. Using the modified base rate of .39, PPP was strong at .92 and NPP was adequate at .84.

Overall, results from this study suggest that detecting malingered MR might be addressed most effectively using tests designed to identify neurocognitive, as opposed to psychiatric, feigning. Perhaps the ineffectiveness of psychiatric feigning instruments should not be surprising since these measures focus on rarely reported psychological symptoms. It is possible that the community sample did not have a clear understanding of the primary characteristics associated with MR and elected to endorse unusual psychiatric symptoms in addition to cognitive deficits. It is also likely, based on examiner observation, that individuals with mild MR were unable to fully understand the items on these measures and as a result tended towards acquiescence.

Conversely, it makes intuitive sense that tests designed to detect neurocognitive feigning would be more effective, given that cognitive impairment is a core feature of MR. Perhaps the relative success of the SIRS, a psychiatric feigning measure, in the Hayes et al. (1998) study was related to the research focus being on malingered insanity, which may have elicited from participants more emphasis on psychological symptoms than cognitive impairment. The conflicting reports from the 2 major publications in this area (Schretlen & Arkowitz, 1990 and Hayes et al., 1998) may have also been a result of using a test that had not been previously validated (i.e., the Malingering Scale) and older, less sophisticated psychiatric and cognitive malingering tests. It is possible that using more sophisticated feigning measures may have resulted in better classification accuracy in those studies.

Perhaps the most important finding of the present study was that feigning instruments developed and cross-validated on different cognitively impaired populations may not be directly applicable to individuals with MR. Data from this study suggests that malingering tests that are appropriate for TBI patients will not necessarily generalize to MR populations without a revision of the recommended cutting scores. Therefore, clinicians involved in forensic evaluations where a diagnosis of MR is suspect should refrain from using decision rules adapted for use in these other populations.

Limitations of the present study include the following: MR participants and community volunteers were tested in different locations, test administrators were not blind to MR participants' diagnostic status, no debriefing information was obtained from MR participants (i.e., how well they understood and complied with instructions), modest sample size, restricted geographical representation in sample (i.e., all participants were from central Kentucky), lack of observer information on measures of adaptive functioning, and the likelihood that the malingering base rate in the present study (49%) was unrealistically high.

The present study addressed a need for additional research on the detection of malingered MR and provided preliminary data suggesting the importance of future research in this area. Despite the limitations noted above, results demonstrated the potential utility of several commonly used neurocognitive feigning instruments to detect malingered MR. Although results appear promising, these data also suggest that applying published decision rules to MR populations will likely result in high false positive errors. Given the high stakes associated with inaccurate classification of capital defendants with suspected MR, it is imperative that accurate cutting scores appropriate for this under-researched population be determined. The cutting scores derived from the present sample should be cross-validated in new samples before being applied in forensic evaluations. Presuming successful replication of these findings, these measures should next be validated using Known-Groups methodologies (Rogers, 1997).

Appendix A

Self-report Questionnaire

Subject number_____ **Date**_____

Sex: Male Female

Race: Black Hispanic Asian White Other _____

Are you left-handed or right-handed? _____ Right-handed _____ Left-handed

Date of Birth _____ **Age** _____

Education (highest grade completed) _____

Did you repeat any grades? Which one(s)? _____ How many times?_____

Do you have a GED? _____ Yes _____ No

What was your most recent job?_____

List your last 3 jobs. (1)_____

(2)_____

(3)_____

Do you live alone? Yes No **If no, with whom do you live?**_____

Have you ever been diagnosed with a learning disability? _____ Yes _____ No

What type? _____

Have you ever been diagnosed with a psychiatric disorder? _____ Yes _____ No

If yes, what was your diagnosis? List all if more than one. _____

What was your age at the time of diagnosis? _____

Have you ever been diagnosed with mental retardation or other developmental disorder?

_____ Yes _____ No **If yes, what was your diagnosis?** _____

What was your age at time of diagnosis?_____

Please list all medications you are currently taking._____

_____

Have you ever been convicted of a misdemeanor or felony? _____ Yes _____ No

If yes, please list here. _____

Have you ever had a head injury? ___Yes ___No **Your age at the time of injury?** _____

If yes, please describe the injury. _____

Appendix B

Instructions for the CVH Group

In this study, you will be asked to take several tests. It is important that you try your best on each of these tests. The testing process will take some time, so be patient and let the test administrator know if you need a break. Some of the questions you will be asked are hard and you may not know the answer. This is to be expected. Just try your hardest and move on if you do not know the answer.

Do not tell the person giving you the tests what your instructions are. Just answer all of the questions to the best of your ability. You will receive $75.00 for your participation in the study.

Appendix C

Instructions for the CVM Group

In this study, you will be asked to take several tests. Although you probably typically try your best on tests, that is not what we want you to do today. Instead, we would like you to pretend that you are mentally retarded. People who are mentally retarded are not smart. They have very low intelligence and have a lot trouble understanding things that you learn in school. They also have trouble understanding how to take care of themselves and get along on their own.

As you take these tests, think about how someone who is mentally retarded might respond and try to answer the questions this way. Do not let the examiner know your instructions to pretend you are mentally retarded. If you can perform on these tests so that your answers look like those of a mentally retarded person, you will receive an extra $20.00 as a bonus. That means you will get $95.00. Keep in mind that if you make it too obvious that you are trying to fake mental retardation when you answer these questions, the examiner will know that you are faking and you will not get the bonus.

Think of it like a game. It might help to pretend you have committed a crime and you can avoid going to jail or the death penalty by pretending you are mentally retarded. A doctor would give you some tests to try to prove it. You will give answers that will make the doctor think you really are mentally retarded. If you can convince the doctor you are mentally retarded, then you might go free. You should respond on all of these tests in the way you think someone with mental retardation would.

Appendix D

Facts about Mental Retardation

- People with Mental Retardation generally have very low intelligence and struggle in school. Although some are able to read at a low level, many are unable to read.

- In addition to difficulty reading, people with Mental Retardation often have trouble communicating and only understand things that are said very simply.

- People with Mental Retardation often have difficulty taking care of themselves, frequently including the following:

    - need help keeping clean and dressing appropriately
    - need help taking care of their health
    - need help handling money
    - need help reading signs
    - need help staying safe at home or in public
    - need help understanding consequences of doing bad things
    - need help in avoiding being easily mislead by others
    - need help with social interactions

Appendix E

Quiz

Please circle T or F for each statement about people with Mental Retardation (MR):

T F 1. A person with MR is very smart.

T F 2. A person with MR does poorly on tests that measure things like word knowledge, problem-solving, and puzzles.

T F 3. A person with MR can get along well on their own.

T F 4. A person with MR knows right from wrong and is not easily persuaded by others.

T F 5. A person with MR has trouble understanding language unless it is very simple.

T F 6. A person with MR can function well in public and understands how to read signs and follow safety instructions.

Appendix F

Debriefing Questionnaire

Subject Number _____ Date_____

Please answer the following questions regarding the tests that you took. Your answers will help us to improve the tests.

1. Please write the instructions that you were given at the beginning of the test briefly and in your own words below (the instructions that were contained in the first envelope):

2. How well did you understand the instructions (the role that you were to follow) that you were given before taking the tests?

Didn't Understand 0 1 2 3 4 5 6 7 8 9 10 Fully Understood At all

3. How hard did you try to follow these instructions as you took the tests?

No Effort 0 1 2 3 4 5 6 7 8 9 10 Maximum Effort

4. Please rate how successful you were in following the instructions as you took the tests:

Not Successful 0 1 2 3 4 5 6 7 8 9 10 Very Successful

5. What strategies did you use to follow these instructions?

Appendix G

Debriefing Statement

When you were given the instruction set, it stated that you would receive an additional $20.00 if you were successful at pretending to be mentally retarded. This reward was offered so that you would feel similar to those who fake mental retardation in the real world and hope to receive a reward for successfully faking. In other words, they hope to receive a lesser sentence than the death penalty.

Deception about the reward was necessary to create a situation more like the real world and give you a reason to try hard to be successful in your faking attempt. In reality, all of the participants in the faking group of this study will receive the $20.00.

This is a new area of research and we are just discovering how to successfully identify those who feign mental retardation. Your participation in this study will add to our understanding of how those who attempt to fake mental retardation respond to these tests.

Thank you for your participation.

References

American Association on Mental Retardation. (1992). *Mental retardation: Definition, classification and systems of support* (9th ed.). Washington, DC: Author.

American Association on Mental Retardation. (2002). *Mental retardation: Definition, classification and systems of support* (10th ed.). Washington, DC: Author.

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

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

Anderson, S. D. & Hewitt, J. (2002). The effect of competency restoration training on defendants with mental retardation found not competent to proceed. *Law and Human Behavior*, 26(3), 343-351.

Annas, G. J. (2002). Moral progress, mental retardation, and the death penalty. *The New England Journal of Medicine*, 347(22), 1814-1818.

Applebaum, K. L., & Applebaum, P. S. (1994). Criminal-justice-related competencies in defendants with mental retardation. *Journal of Psychiatry and the Law*, 3, 483-501.

Arbisi, P. A., & Ben-Porath, Y. S. (1995). On MMPI-2 infrequency response scale for use with psychopathological populations: The Infrequency Psychopathology scale F(p). *Psychological Assessment*, 7, 424-431.

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

Beaber, R. J., Marston, A,. Michelli, J., & Mills, M. J. (1985). A brief test for measuring malingering in schizophrenic individuals. *American Journal of Psychiatry*, 142, 1478-1481.

Bernard, L. (1990). Prospects for faking believable memory deficits on neuropsychological tests and the use of incentives in simulation research. *Journal of Clinical and Experimental Neuropsychology*, 12, 715-728.

Berry, D. T. R., Baer, R. A., Rinaldo, J. C., & Wetter, M. W. (2002). Assessment of Malingering. In J. Butcher (Ed,), Clinical personality assessment: Practical approaches (2nd Ed.). New York: Oxford University Press.

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

Bickart, W.T., Meyer, R.G. & Connell, D. (1991). The symptom validity technique as a measure of feigned short-term memory deficit. *American Journal of Forensic Psychology*, 9, 3-11.

Binder, L.M., Villaneuva, M. R., Howieson, D., & Moore, R. T. (1993). The Rey AVLT recognition memory task measures motivational impairment after mild head trauma. *Archives of Clinical Neuropsychology*, 8, 137-147.

Bonnie, R. J. (1990). The competency of criminal defendants with mental retardation to participate in his own defense. *The Journal of Criminal Law and Criminology*, 81, 419-446.

Brodsky, S. L, & Galloway, V.A. (2003). Ethical and professional demands for forensic mental health professionals in the post-Atkins era. *Ethics & Behavior*, 13(1), 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(3), 244-255.

Butcher, J., Dahlstrom, W., Graham, J., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasis Personality Inventory-2: Manual for administration and scoring*. Minneapolis: University of Minnesota Press.

Corbett, J. A. (1979). Psychiatric morbidity and mental retardation. In *Psychiatric illnesses and mental handicap* (ed. F. E. James and R. P. Snaith), pp. 11-25. London: Gaskell Press.

Daniel, A. E., & Menninger, K. (1983). Mentally retarded defendants: Competency and criminal responsibility. *American Journal of Forensic Psychiatry*, 4, 145-156.

Einfeld, S. L., & Aman, M. (1995). Issues in the taxonomy of psychopathology in mental retardation. *Journal of Autism and Developmental Disorders*, 25, 143-167.

Einfeld, S. L., & Tonge, B. J. (1996). Population prevalence of psychopathology in children and adolescents with intellectual disability. II: Epidemiological findings. *Journal of Intellectual Disability Research*, 40, 99-109.

Ellis, J. W., & Luckasson, R. A. (1985). Mentally retarded criminal defendants. *George Washington Law Review*, 53, 414-493.

Etcoff, L., & Kampfer, K. (1996). Practical guidelines in the use of symptom validity and other psychological tests to measure malingering and symptom exaggeration in traumatic brain injury cases. *Neurolopsychology Review*, 6, 171-201.

Everington, C. T. (1990). The Competency Assessment for Standing Trial for Defendants with Mental Retardation (CAST:MR): A validation study. *Criminal Justice & Behavior*, 17(2), 147-168.

Everington, C. T., & Dunn, C. (1995). A second validation study of the Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST:MR). *Criminal Justice & Behavior*, 22(1), 44-59.

Everington, C. T., & Luckasson, R. (1992). Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST:MR). Worthington, OH: International Diagnostic Systems.

Faust, D., & Ackley, M. A. (1998). Did you think it was going to be easy? Some methodological suggestions for the investigation and development of malingering detection techniques. In C. R. Reynolds (Ed.) *Detection of malingering during head injury litigation* (pp. 261-286). New York: Plenum Press.

Frederick, R. I., Crosby, R. D., & Wynkoop, T. F. (2000). Performance curve classification of invalid responding on the Validity Index Profile. *Archives of Clinical Neuropsychology*, 15, 281-300.

Fredericks, D. W., & Williams, W. L (1998). New definition of mental retardation for the American Association of Mental Retardation. *Image: Journal of Nursing Scholarship*, 30(1), 53-56.

Fulero, S. M., & Everington, C. (2004). Assessing the capacity of persons with mental retardation to waive Miranda Rights: A jurisprudent therapy perspective. *Law & Psychology Review*,28, 53-59.

Gallen, R. T., & Berry, D. T. R. (1996). Detection of random responding in MMPI-2 protocols. *Assessment*, 3(2), 171-178.

Gallen, R. T., & Berry, D. T. R. (1997). Partially random MMPI-2 protocols: When are they interpretable? *Assessment*, 4(1), 61-68.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington DC: American Psychological Association.

Gervias, R. O., Rohling, M. L., Green, P., and Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. *Archives of Clinical Neuropsychology*, 19(4), 475-488.

Giller, V., Dial, J. G., & Chan, F. (1986). The Street Survival Skills Questionnaire: A correlational study, *American Journal of Mental Deficiency*, 91(1), 67-71.

Goldstein, H. (1945). A malingering key for mental tests. *Psychological Bulletin*, 42, 104-118.

Gothard, S., Viglione, D. J., Meloy, J. R., & Sherman, M. (1995). Detection of malingering in competency to stand trial evaluations. *Law & Human Behavior*, 19(5), 493-505.

Gough (1950). The F minus K dissimulation index for the Minnesota Multiphasic Personality Inventory. *Journal of Consulting Psychology*, 14, 408-413.

Greene, R. L. (2000). *The MMPI-2: An interpretive manual* (2nd ed.). Boston: Allyn & Bacon.

Greenspan, S., & Switzky, H. N. (2003). Execution exemption should be based on actual vulnerability, not disability label. *Ethics & Behavior*, 13(1), 19-27.

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

Greiffenstein, M. F., Gola, T., & Baker, J. W. (1995). MMPI-2 validity scales versus domain specific measures in detection of factitious traumatic brain injury. *Clinical Neuropsychologist*, 9, 230-240.

Greve, K. W., Bianchini, K. B., Mathias, C. W., Houston, R. J., & Crouch, J. A. (2003). Detecting malingered performance on the Wechsler Adult Intelligence Scale: Validation of Mittenberg's approach in traumatic brain injury. *Archives in Clinical Neuropsychology*, 18, 245-260.

Guilmette, T. J., Hart, K. J. & Giuliano, A. J. (1993). Malingering detection: The use of a forced-choice method in identifying organic versus simulated memory impairment. *TheClinical Neuropsychologist*, 7 (1), 59-69.

Guilemette, T. J., Whelihan, W., Sparadeo, F. R., & Buongiorno, G. (1994). Validity of neuropsychological test results in disability evaluations. *Perceptual and Motor Skills*, 78, 1179-1186.

Hathaway, S. R., & McKinley, J. C. (1943). The Minnesota Multiphasic Personality Inventory (revised ed.). Minneapolis, MN: University of Minnesota Press.

Hayes, J. S., Hale, D. B., & Gouvier, W. D. (1997). Do tests predict malingering in defendants with mental retardation? *Journal of Psychology*, 131(5), 575-577.

Hayes, J. S., Hale, D. B., & Gouvier, W. D. (1998). Malingering detection in a mentally retarded forensic population. *Applied Neuropsychology*, 5(1), 33-36.

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

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

Huebrock, D., & Petermann, F. (1998). Neuropscychological assessment of suspected malingering: Research results, evaluation techniques, and further directions of research and application. *European Journal of Psychological Assessment*, 14, 211-225.

Inman, T. H., & Berry, D. T. R. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology*, 17, 1-23.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Edwards, C. L., Lamb, D. G., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort given during neuropsychological testing: The Letter Memory Test. *Psychological Assessment*, 10, 128-139.

Iverson, G. L., & Tulsky, D. S. (2003). Detecting malingering on the WAIS-III unusual digit span performance patterns in the normal population and in clinical groups. *Archives of Clinical Neuropsychology*, 18, 1-9.

Iverson, G. L., Franzen, M. D., & McCracken, L. M. (1991). Application of a forced-choice memory procedure designed to detect experimental malingering. *Archives of Clinical Neuropsychology*, 15(6), 667-676.

Janniro, F., Sapp, G. L., & Kohler, M. P. (1994).Validating the Street Survival Skills Questionnaire. *Psychological Reports*, 74, 191-194.

Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment.* New York: John Wiley & Sons, Inc.

Lewis, J. L., Simcox, A. M., & Berry, D. T. R. (2000). Screening for feigned psychiatric symptoms in a forensic sample by using the MMPI-2 and the Structured Inventory of Malingered Symptomatology. *Psychological Assessment*, 14(2), 170-177.

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

Linkenhoker, D. & McCarron, L. (1993). *The Street Survival Skills Questionnaire*. Dallas, TX: McCarron-Dial Systems.

Luckasson, R., Borthwick-Duffy, S., Buntinx, W. H. E., Coulter, D. L., Craig, E. M., Reeve, A., Schalock, R. L., Snell, M. E., Spitalnik, D. M., Spreat, S. Tasse, M. J. (2002). *Mental retardation: Definition, classification, and systems of supports (10th ed.).* Washington, DC: American Association on Mental Retardation.

Miller, H. A. (2001). *Miller-Forensic Assessment of Symptoms Test (M-FAST): Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Miller, H. A. (2004). Examining the use of the M-FAST with criminal defendants incompetent to stand trial. *International Journal of the Offender Therapy and Comparative Criminology*, 48(3), 268-280.

Millis, S. R., Ross, S. R., & Ricker, J. H. (1998). Detection of incomplete effort on the Wechsler Adult Intelligence Scale-Revised: A cross-validation. *Journal of Clinical and Experimental Neuropsychology*, 20(2), 167-173.

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

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

Nihira, K., Foster, R., Shellhaas, M., & Leland, H. (1974). *AAMD Adaptive Behavior Scale (1974 rev.)*. Washington, DC: American Association on Mental Deficiency.

Norris, M. P., & May, M. C. (1998). Screening for malingering in a correctional setting. *Law & Human Behavior*, 22(3), 315-323.

Owens, R. G. (1995). The psychological signatures of malingering: Assessing the legitimacy of claims. *American Jornal of Forensic Psychology*, 13, 61-75.

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

Pankratz, L., & Binder, L. M. (1997). Malingering on intellectual and neuropsychological measures. In R. Rogers (Ed.), *Clinical Assessment of Malingering and Deception* (2nd ed., pp. 2230236). New York: Guilford Press.

Penry v. Lynaugh (1989). 492 U.S. 302.

Pritchard, D., & Moses, J. (1991). Tests of neuropsychological malingering. *Forensic Reports*, 5, 287-290.

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

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Psychology*, 16, 501-506.

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

Reynolds, C. R., Price, J. R., & Niland, J. (2003). Applications of neuropsychology in capital felony (death penalty) defense. *Journal of Forensic Neuropsychology*, 3(4), 89-123.

Rogers, R. (1986). *Structured Interview of Reported Symptoms (SIRS).* Odessa, FL: Psychological Assessment Resources.

Rogers, R. (1990). Development of a new classificatory model of malingering. *Bulletin of the American Academy of Psychiatry and Law*, 18, 232-333.

Rogers, R. (1995). *Diagnostic and structured interviewing: A handbook for psychologists.* Psychological Assessment Resources: Odessa, FL.

Rogers, R. (1997). *Clinical assessment of malingering and deception.* New York: Guilford Press.

Rogers, R. & Cruise, K. R. (2000). Malingering and deception among psychopaths. In C. Gacono (Ed.), *The clinical and forensic assessment of psychopathology: A practitioner's guide* (pp. 269-284). New York: Lawrence Erlbaum.

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). The SIRS test manual. Tampla, FL: Psychological Assessment Resources.

Rogers, R., Harrell, E. H., & Liff, C. D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review*, 13, 255-274.

Rogers, R., Hinds, J. D., & Sewell, K. W. (1996). Feigning psychopathology among adolescent offenders: Validation of the SIRS, MMPI-A, and SIMS. *Journal of Personality Assessment*, 67(2), 244-258.

Rogers, R., Salekin, R. T., Sewell, K. W., Goldstein, A., & Leonard, K. (1998). A comparison of forensic and nonforensic malingerers: A prototypical analysis of explanatory models. *Law & Human Behavior*, 22(4), 353-367.

Rogers, R., Sewell, K. W., & Goldstein, A. (1994). Explanatory models of malingering: A prototypical analysis, *Law and Human Behavior*, 18, 543-552.

Schlesinger, L. B. (2003). A case study involving competency to stand trial: Incompetent defendant, incompetent examiner, or "malingered by proxy"? *Psychology, Public Policy, and Law*, 9(3-4), 381-399.

Schretlen, D., & Arkowitz, H. (1990). A psychological test battery to detect prison inmates who fake insanity or mental retardation. *Behavioral Sciences and the Law*, 8, 75-84.

Shaw, J. A. & Budd, E. C. (1982). Determinants of acquiescence and nay saying of mentally retarded persons. *American Journal of Mental Deficiency*, 87(1), 108-110.

Sigelman, C. K., Winer, J. L., & Schoenrock, C. J. (1982). The responsiveness of mentally retarded persons to questions. *Education & Training of the Mentally Retarded*, 17(2),120-124.

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

Smith, G. P., & Burger, G. O. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology. *Journal of the American Academy of Psychiatry and Law*, 25, 183-189.

Smith, S. A., & Broughton, S. F. (1994). Competency to stand trial and criminal responsibility: An analysis in South Carolina. *Mental Retardation*, 32, 281-287.

Smith, S. A., & Hudson, R. L. (1995). A quick screening test of competency to stand trial for defendants with mental retardation. *Psychological Reports,* 76(1), 91-98.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales: Expanded form manual* (interview ed.). Circle Pines, MN: American Guidance Service.

Strong, D. R., Green, R. L., & Schinka, J. A. (2000). A taxometric analyis of MMPI-2 infrequency scales [F and F(p)]. *Psychological Assessment*, 12, 166-173.

Sweet, J. J. (Ed.). (1999). *Forensic neuropsychology: Fundamentals and practice.* Exton, PA: Swets & Zeitlinger.

Thompson, G. B. (2002). The Victoria Symptom Validity Test: An enhanced test of symptom validity. *Journal of Forensic Neuropsychology*, 2(3/4), 43-67.

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

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

Tombaugh, T. N. (2002). The Test of Memory Malingering (TOMM) in forensic psychology. *Journal of Forensic Neuropsychology*, 2, 69-96.

Trueblood, W., & Schmidt, M. (1993). Malingering and other validity considerations in the neuropsychological evaluation of mild head injury. *Journal of Clinical and Experimental Neuropsychology*, 15, 578-590.

Vickery, C. D., Berry, D. T. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Crager, D. E., & Orey, S. A. (2002). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology*, 599 (2002), 1-12.

Vickery, C. D., Berry, D. T. R., Inman, T. H., Harris, M. J., & Orey, S. A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology*, 16, 45-73.

Wechsler, D. (1987). *The Wechsler Adult Intelligence Scale-Revised: manual*. New York: The Psychological Corporation.

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

Wechsler, D. A. (1991). *Manual for the Wechsler Intelligence Scale for Children-Third Edition.* San Antonio, TX: The Psychological Corporation.
Wechsler, D.A. (1997). *The Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.
Wetter, M. W., Baer, R. A., Berry, D. T. R., Smith, G. T., & Larsen, L. H. (1992). Sensitivity of MMPI-2 validity scales to random responding and malingering. *Psychological Assessment*, 4, 369-374.
Widows, M. R. & Smith, G. P. (2005). Structured Inventory of Malingered Symptomatology: Professional Manual: Odessa, FL: Psychological Assessment Resources.
World Health Organization (1992). *International statistical classification of diseases and related health problems, 10th revision*. WHO, Geneva.
Youngjohn, J. R., Burrows, L., & Erdal, K. (1995).Brain damage or compensation neurosis? The controversial post-concussive syndrome. *The Clinical Neuropsychologist*, 9, 112-123.
Zielinski, J. (1994). Malingering and defensiveness in the neuropsychological assessment of mild traumatic brain injury. *Clinical Psychology: Science and Practice*, 1, 169-184.
Ziskin (1984). Malingering of psychological disorders. *Behavioral Sciences and the Law*, 2, 39-49.

Vita

Lili Odom Graue

Date of Birth: June 5, 1964

Place of Birth: Beaumont, Texas

Education:
Salem College, Winston-Salem, North Carolina
BA in Psychology, 2001

Honors:
Magna Cum Laude, 2001
President's Prize, 2001
Sophisteia Award, 2001
Salem Honor Society, 1999-2001
Who's Who Among Students in American Universities & Colleges, 2000, 2001
Dean's List, 1998, 1999, 2000

Professional Positions:

University of Kentucky, Lexington, KY
Research assistant, July 2005 – present.

University of Kentucky, , Orofacial Pain Center, Lexington, KY
Psychology resident, July 2004 – present.

University of Kentucky, Jesse G. Harris Psychological Services Center
Therapist, August 2004 – present.

University of Kentucky, Testing and Counseling Center, Lexington, KY
Process observer in group therapy, January 2005 – May 2005.

Kentucky Clinic, Department of Neuropsychology, Lexington, Kentucky
Test administrator, July 2005 – December 2005.

University of Kentucky, Lexington, KY
Research assistant, March 2004 – November 2004.

University of Kentucky, Lexington, KY
Teaching Assistant, August 2002 – May 2003, January 2005 – May 05.

Wake Forest University School of Medicine, Winston-Salem, North Carolina, Neuropsychology Dept.
Research assistant, July 2001 – July 2002

Professional Publications:
Graue, L. O., Berry, D. T. R., Clark, J. A., Sollman, M. J., Cardi, M., Hopkins, J., & Werline, D. (in press). Identification of feigned mental retardation using the new generation of malingering detection instruments: Preliminary Findings. *The Clinical Neuropsychologist.*