2013

Detecting malingered ADHD using the personality assessment inventory : an exploratory analysis in college students

Mandi Wilkes Musso
Louisiana State University and Agricultural and Mechanical College, mmusso4@tigers.lsu.edu

Recommended Citation

DETECTING MALINGERED ADHD USING THE PERSONALITY
ASSESSMENT INVENTORY: AN EXPLORATORY ANALYSIS IN
COLLEGE STUDENTS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

by
Mandi Wilkes Musso
B.S., University of Louisiana at Lafayette, 2006
M.A., Louisiana State University, 2011
August 2013

ACKNOWLEDGEMENTS

I would like to thank Dr. Benjamin Hill and Ms. Anneliese Boettcher for their assistance in providing archival data from their University's Psychology Clinic. I would like to express my appreciation to my dissertation committee, Dr. Alex Cohen, Dr. Mary Lou Kelley, Dr. Glenn Jones, and Dr. Shaomian Yao. I would also like offer my gratitude to Dr. Gouvier for his guidance, patience, supervision, and insight, not only on this project but throughout my training at Louisiana State University. Finally, I would like to thank my husband and my family for their unending love and support throughout graduate school and internship.

TABLE OF CONTENTS

ABSTRACT

Background: There has been a growing interest in assessment of effort during psychoeducational evaluations, where students may feign symptoms of ADHD to obtain academic accommodations or stimulant medications. Current research suggests most ADHD questionnaires and neuropsychological tests do not adequately distinguish clinical ADHD from simulated ADHD.

Objective: The purpose of the current study is to develop an embedded malingering index in the Personality Assessment Inventory (PAI) specifically for detecting feigned ADHD in college students.

Method: A sample of 310 undergraduate students were separated into three groups, ADHD Simulators, Prospective ADHD, and College controls. In addition, this study used archival data from individuals diagnosed with Clinical ADHD, No Diagnosis, Psychopathology, Comorbid ADHD-Psychopathology, and Suspect Effort. ADHD Simulators' scores on the items of the Personality Assessment Inventory were compared to the Clinical ADHD group. The item pool was narrowed by selecting the 40 items with the largest effect sizes. A discriminant function analysis was then used to select the items that discriminate best between the two groups. The items were weighted and summed into a scale. Next, logistic regression analyses and ROC curves were used to determine an appropriate cutoff score.

Results: Fourteen items were summed into a scale. When various cutoff scores were examined, a score of > 16 yielded specificity of .95 and .96 for the Clinical ADHD groups and .98 for the No

Diagnosis group and sensitivity of .64 for the ADHD Simulator groups. However, it did not yield adequate specificity for Psychopathology or Comorbid ADHD-Psychopathology groups. A cutoff score of greater than > 22 yielded specificity > .90 for all groups but sensitivity of .44 for the ADHD Simulators.

Conclusion: The use of a cutoff score of > 16 is recommended when individuals do not complain of comorbid psychopathology but a cutoff score of > 22 when comorbid psychopathology is in question. The newly developed scale of the PAI shows promise in identifying college students malingering ADHD symptoms.

1. INTRODUCTION

There is some debate over the verisimilitude of the construct of Attention Deficit Hyperactivity Disorder (ADHD) in adulthood. Initially, symptoms of ADHD were thought to subside as children matured and reached puberty (DuPaul, Guevermont, & Barkley, 1991), but more recent evidence suggests ADHD symptoms persist into adulthood for many individuals (Barkley, Fischer, Smallish, & Fletcher, 2002; Resnick, 2005). Despite a growing body of literature that supports the presence of adult ADHD, it has also been recognized that there are numerous incentives for some individuals to feign ADHD. Clinicians and researchers have recently begun to examine malingering in the context of psychoeducational evaluations where individuals might be tempted to feign ADHD. Estimated base rates of noncredible performance in college students range from 8.3% (Harrison, Rosenblum, & Currie, 2011) to 47.6% (Sullivan May, & Galbally, 2007) depending on criteria and measures used. There is a paucity of research on the detection of feigned ADHD, but current studies suggest that college students are successfully able to simulate ADHD on clinical measures designed to assess the presence of self-reported ADHD symptoms in a manner similar to their simulations of post concussion symptom reports (Martin, Hayes, & Gouvier, 1996).

This project will first examine the literature on adult ADHD. Next, a brief overview of malingering and the detection of malingered neurocognitive dysfunction will be presented. The readers will find a synopsis of the literature on the detection of malingered ADHD followed by a rationale for the present study. This study will focus on the development of an embedded effort index that offers carefully calibrated psychometric operating characteristics designed specifically for the purpose of detecting malingered ADHD in college students.

1

1.1. Attention Deficit/Hyperactivity Disorder

Diagnostic Criteria. The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Revised (*DSM-IV-TR*; American Psychological Association, 2000) defines ADHD as a disorder comprising one or more of several clusters of symptoms that are marked either by excessive symptoms of inattention and/or hyperactivity-impulsivity. In order to receive a diagnosis of ADHD, symptoms must cause significant impairment in at least two domains of one's life. The *DSM-IV-TR* requires that the symptoms be present before age seven but notes individuals may be diagnosed later if the symptoms have been present but were undiagnosed before age seven. Parenthetically, it should be noted that proposed changes for the 2013 Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (*DSM-V)* include a change in age of onset up to 12 years old (APA, 2012). Table 1 lists the *DSM-IV-TR* diagnostic criteria for ADHD. A diagnosis of ADHD, Combined Type is made if criteria A1 and A2 are met, ADHD. ADHD, Predominantly Inattentive Type, is diagnosed if only criterion A1 is met, and ADHD, Predominantly Hyperactive Type, is diagnosed if only criterion A2 is met. Criterion symptoms must be present for at least six months. Finally, ADHD Not Otherwise Specified is diagnosed if individuals currently meet criteria, but onset of symptoms is age seven or later or if individuals experience significant impairment but their symptom pattern does not meet full criteria for one of the more specific subspecialties (APA, 2000).

Some researchers believe that the *DSM* criteria are not appropriate for diagnosing ADHD in adulthood because the criteria were selected for their application with children (Barkley, Murphy, & Fischer, 2008). For example, Wender and colleagues developed the Utah Criteria for diagnosis of ADHD that include establishing impairment in childhood through interview of the patient and an informant as well as indication of current impairment including seven symptoms:

2

Table 1. DSM-IV-TR criteria for ADHD.

Criterion A1: *six of more symptoms of inattention persisting for at least six months*

1. often fails to give close attention to details or makes careless mistakes in schoolwork, work,
   or other activities

2. often has difficulty sustaining attention in tasks or play activities

3. often does not seem to listen when spoken to directly

4. often does not follow through on instructions and fails to finish schoolwork, chores, or
   duties in the workplace (not due to oppositional behavior or failure to understand
   instructions)

5. often has difficulty organizing tasks and activities

6. often avoids, dislike, or is reluctant to engage in tasks that require sustained mental effort
   (such as schoolwork or homework)
7. often loses things necessary for tasks or activities (e.g. toys, school assignments, pencils
   books, or tools)

8. is often easily distracted by extraneous stimuli

9. is often forgetful in daily activities

Criterion A2: *six or more hyperactivity-impulsivity symptoms persisting for at least 6 months*

1. often fidgets with hands or feet or squirms in seat

2. often leaves seat in classroom or in other situations in which remaining seated is
   expected
3. often runs about or climbs excessively in situations in which it is inappropriate (in
   adolescents or adults, may be limited to subjective feelings or restlessness)

4. often has difficulty playing or engaging in leisure activities quietly

5. is often "on the go" or often acts as if "driven by a motor"

6. often talks excessively

7. often blurts out answers before questions have been completed

8. often has difficulty awaiting turn

9. often interrupts or intrudes on others (e.g. butts into conversations or games)

3

inattentiveness, hyperactivity, mood lability, irritability and hot temper, impaired stress tolerance, disorganization, and impulsivity.   They introduced the Wender Utah Rating Scale (WURS; Ward, Wender, & Reimherr, 1993) to assess retrospective report of childhood ADHD symptoms. However, Barkley et al. (2008) argue that allowing symptoms such as irritability, hot temper, and mood lability confound the delineation between ADHD and other psychiatric disorders in adolescents and young adults. In fact, Hill and colleagues reported WURS scores were more significantly related to dysfunctional personality traits than performance on neuropsychological measures of attention in 522 college students self-referred for a psychoeducational evaluation (Hill, Pella, Singh, Jones, & Gouvier, 2009).

Alternatively, Barkley et al. (2008) attempted to establish age appropriate criteria for ADHD in adults by examining both *DSM-IV-TR* criteria and an item pool of 87 variables designed to tap self-reported neurocognitive symptoms such as inattention, working memory, and self-regulation. Nine items emerged as variables that significantly aided in the diagnosis of ADHD. Three items were from the *DSM-IV-TR* and six items were from the item pool of 87 variables. Symptoms include: makes decisions impulsively, has difficulty stopping activities or behaviors when he/ she should, starts a project without reading or listening to directions, shows poor follow-through on promises or commitments that he/she may make to others, has trouble doing things in their proper order or sequence, is more likely to drive a motor vehicle much faster than others (excessive speeding), prone to daydreaming when should be concentrating, trouble planning ahead or preparing for upcoming events, and can't seem to persist at things that he/she doesn't find interesting. These nine items were placed into a scale, and Barkley et al. (2008) recommended that individuals be diagnosed with ADHD if they endorsed six or more items "Often" or "Very Often". A factor analysis revealed that these nine symptoms loaded onto

4

three factors, consistent with *DSM-IV-TR* criteria: inattention and working memory, hyperactive-impulsive, and verbal impulsivity.

Fedel and colleagues (2010) were unable to replicate these findings when they asked a group of college students with ADHD and a group of controls to complete self-report questionnaires. Individuals were assigned to the ADHD group if they reported a history of diagnosis of ADHD or met current *DSM-IV* criteria for ADHD. The students completed 18 items from the Current Symptoms Scale-Self Report Form (Barkley & Murphy, 2006) as well as the 87 items used by Barkley et al. (2008). In their study, they found two factors, Disinhibition and Cognitive Inflexibility. These two factors were comprised of 17 items that distinguished adults with ADHD from those without.

Overall, there is considerable debate about the construct, relevant symptoms, and diagnostic criteria of ADHD. Though multiple studies have been conducted, replication studies have generally not been favorable. The debate continues, making the diagnosis of ADHD in adulthood difficult for clinicians and researchers.

Prevalence of Adult ADHD. Estimates of the prevalence of ADHD in college students vary. Significant differences in opinions of experts, regarding appropriate diagnostic criteria and approach to measuring symptoms, further affects estimates of prevalence of adult ADHD. For example, McKee (2008) reported approximately 20% of 1,096 college students met criterion threshold (97th percentile) for ADHD on the ADHD Response Evaluation (Glutting et al., 2002) when a norm-based approach was used. In the same study only 7.48% of individuals met *DSM-IV* criteria. McKee suggested this discrepancy in prevalence of ADHD, when using these two methods, questions developmental appropriateness of diagnostic criteria in college students.

5

Another challenge in assessing prevalence of ADHD symptoms includes inadequate reliability and convergent validity in measures of adult and childhood symptoms of ADHD. For example, Weyandt, Linterman, and Rice (1995) surveyed of 770 college students and reported that 7% of students endorsed significant symptoms (defined by 1.5 standard deviations above the mean) on the Adult Rating Scale, a measure they introduced in their study. In the same study, 8.7% of students endorsed significant childhood symptoms of ADHD on the WURS, but only 2.5% of students endorsed significant symptoms on both measures. When more stringent criteria were used (2 standard deviations), 4% and 3.8% of students met criteria for the Adult Rating Scale and WURS, respectively. However, only 0.5% met criteria when both measures were used. Further, the Adult Rating Scale and WURS correlated .54, illustrating relatively poor convergent validity in current adult and childhood measures of ADHD, which may account, in part, for discrepancies in research findings. Alternatively, the findings of this study may reflect the resolution of ADHD in some individuals diagnosed in childhood, persistence in others, and a group of college students who are self-referred for attentional impairments as adults but were not diagnosed with ADHD in childhood.

In their study, Heiligenstein, Conyers, Berns, & Smith (1998) found that 4% of the 448 college students surveyed endorsed symptoms consistent with a diagnosis of ADHD. Of this 4%, 56% endorsed inattentive symptoms while 22% endorsed symptoms consistent with the hyperactive-impulsive and combined types, respectively. DuPaul and colleagues (2001) reported that 2.9% of American men endorsed criteria for ADHD with a majority (2%) endorsing symptoms of hyperactivity-impulsivity. In addition, 3.9% of American women endorsed clinically significant symptoms of ADHD, with a majority (2.3%) endorsing the hyperactivity-impulsivity subtype. The National Comorbidity Survey Replication estimated the prevalence of

ADHD to be 4.4% (Kessler, et al., 2006). In their review of 23 studies, Weyandt and DuPaul (2006) found the estimated prevalence of adult ADHD ranged from 2 to 8%.

Impact of Adult ADHD. Farone and colleagues (2000) found that clinical correlates of ADHD are similar for children and adults; namely, adults with ADHD are impulsive, inattentive, and restless. Additional evidence suggests that clinic-referred young adults with ADHD, who were not diagnosed in childhood, may differ in symptoms and impairments when compared to adults that were diagnosed with ADHD as children (Barkley et al., 2008). For example, investigators argue that older individuals with ADHD might demonstrate a number of symptoms that are difficult to measure prior to age seven such as deficits in executive functioning, poor planning, forgetfulness, problems involving delay of gratification, self-control problems, and difficulty dividing and focusing one's attention (Wasserstein, 2005; Barkley et al., 2008). The *DSM-IV-TR* notes symptoms of excessive gross motor activity are less common in adults, and hyperactive symptoms in adults may be limited to feeling fidgety and internal restlessness that may interfere with occupational and social functioning.

In regards to functional impairment, adult ADHD is thought to cause substantial impairments in economic, academic, social, and occupational functioning (Barkley et al., 2008; Barkley & Murphy, 2010; Faraone, et al., 2000) including meeting deadlines, completing tasks, planning ahead, and poor sense of time (Riccio, et al., 2005). In their review of the literature, Weyandt and DuPaul (2008) found that ADHD results in poorer academic outcomes and increased psychological difficulties for college students. College students with self-reported ADHD symptoms also endorse significantly more depressive symptoms, emotional instability, substance abuse, and concerns about social functioning and academic performance (Blasé et al., 2009). There is some evidence to suggest that inattentive symptoms related to ADHD may

negatively impact college students' study skills, academic adjustment, and GPA (Norwalk, Norvilitis, & MacLean, 2009).

In their book, Barkley et al. (2008) report data from two large studies. The University of Massachusetts (UMASS) study examined adults who did not receive formal ADHD diagnoses as children but were self-referred for an ADHD evaluation. The Milwaukee study was a longitudinal study that collected data from children diagnosed with ADHD as they grew into adulthood. When data from both studies were compared, it was found the adults from the UMASS study endorsed significant impairment in most domains (with the exception of dating and marriage), with most the impairment in domain of education. Fewer adults with ADHD from the Milwaukee study reported problems with education. Similarly, when retrospective childhood reports were examined, the UMASS participants were more likely to endorse problems in school, and the Milwaukee adults were most likely to endorse problems with peer interaction. Interestingly, educational attainment was higher in individuals in the UMASS study compared to those in the Milwaukee study. These results may reflect differences in the educational systems of these two states, or they may illustrate differences between individuals were diagnosed with ADHD in childhood from self-referred adults who were not diagnosed with ADHD in childhood. While both groups endorsed significant occupational problems, the individuals in the UMASS study were more likely to have higher job status and work more hours per week. It is possible that adults presenting for ADHD evaluation for the first time may over-endorse academic problems despite relatively intact social and occupational functioning. See Barkley, Murphy, and Fischer (2008) for review.

Challenges in Neuropsychological Assessment of Adult ADHD. As the previous sections have highlighted, diagnosing ADHD in adulthood is difficult. First, as demonstrated in

8

the previous sections, there is ongoing debate about the diagnosis of ADHD in adulthood. There also appears to be a lack of convergent validity in the construct of adult ADHD as studies have demonstrated significant differences between individuals diagnosed with ADHD in childhood and adults presenting for ADHD diagnoses who were not diagnosed in childhood (Barkley et al., 2008). In addition, research has failed to replicate data regarding adult ADHD, and some commonly used measures of self-reported ADHD symptoms do not demonstrate adequate convergent or construct validity. For example, the WURS has been found to measure ongoing dysfunctional personality traits rather than retrospective reports of childhood ADHD (Hill et al., 2009).

Adults are often inaccurate historians when recalling childhood symptoms of ADHD (Mannuzza, Klein, Klein, Bessler, & Shrout, 2002; American Psychiatric Association, 2000), raising questions about validity of self-reported symptom onset, severity, and duration that can preclude diagnostic accuracy. In addition, diagnosing ADHD in adulthood is difficult because many individuals with ADHD also experience impairments related to other Axis I diagnoses (Kessler et al., 2006; Sobanski, et al., 2007; Wilens et al., 2009) including anxiety, depression, substance use disorders (Kessler et al., 2006), and eating disorders (Sobanski et al., 2007). Adult ADHD has also been associated with personality disorders (Williams et al., 2010). Some symptoms of Axis I and Axis II disorders overlap with symptoms of ADHD. For example, problems with attention and concentration are included in *DSM-IV-TR* diagnostic criteria for anxiety and depressive disorders. In addition, students may mistake other symptoms of anxiety and depression, such as feeling restless and psychomotor agitation, as symptoms of hyperactivity.

9

Further, there is debate about number of symptoms required to diagnose ADHD in adults. Some researchers suggest that college students, as a whole, endorse fewer symptoms of inattention and hyperactivity than children (Heiligenstein, et al., 1998; Barkley et al., 2008). Heiligenstein and colleagues (1998) recommend that endorsement of only four or more symptoms should be sufficient to classify college students as ADHD. In the Milwaukee Study, a cutoff of four or more *DSM* symptoms was also recommended (Barkley et al., 2008). Despite the problem of "opening up the floodgates" by lowering the diagnostic requirements for meeting the threshold, there remains controversy in proposed changes to the *DSM-V* and it has not yet been decided whether the threshold of symptoms should be lowered for adults and adolescents over the age of 17 (APA, 2012).

Another challenge in diagnosing ADHD in adults is the lack of agreement on a neuropsychological profile that is characteristic of ADHD (Wasserstein, 2005). Several authors have reported that individuals who complained of ADHD symptoms did not perform differently on neuropsychological measures compared to individuals who made no such complaints (Rosselli et al., 2000; Riccio et al., 2005). Other authors suggest self-reported executive functioning deficits play a larger role in impairments in occupational functioning than performance on neuropsychological measures (Barkley & Murphy, 2006). However, there is some evidence that adults with ADHD perform significantly poorer on measures of executive functioning that require response inhibition (Rapport, VanVoorhis, Tzelepis, & Friedman, 2001) and also show impairments in vigilance, selective attention, divided attention, and cognitive flexibility (Tucha et al., 2008). In addition, there is evidence that individuals with ADHD demonstrate impairments on neuropsychological tests of attention, behavioral inhibition (Hervey, Epstein, & Curry, 2004), working memory (Hervey et al., 2004; Marchetta, Hurks, Jolles, &

Krabbendam, 2008), and set shifting (Marchetta et al., 2008). A meta-analysis of neuropsychological performance in adults with ADHD found moderate effect sizes for impaired performance in domains of complex attention and verbal memory in adults with ADHD (Schoechlin & Engel, 2005). Woods and colleagues (2002) propose using discrepancy analyses produces increased sensitivity in detecting ADHD in a battery of neuropsychological tests.

These challenges in the diagnosis of adult ADHD are particularly troublesome when one considers an alternative explanation for adults presenting with self-reported ADHD symptoms. For example, Diller (2010) expressed concern about initial diagnosis of ADHD in late adolescence, suggesting this subgroup may not actually "have" ADHD. Rather, they may belong to a cohort of under-performing adolescents who are not motivated or academically prepared for higher education standards. Diller suggests that their symptoms do not generalize to every aspect of their lives and tend to be context specific, arising when goals are not easily met. He encourages clinicians to consider this explanation and assess this possibility as part of the diagnostic process.

This section highlights the significant challenges in diagnosing adult ADHD. Despite these challenges, release of the *DSM-V* in May 2013 may make it easier for to acquire a diagnosis of ADHD by lowering the symptom threshold for adults. Issues such as disagreement regarding the construct of adult ADHD, the lack of a neuropsychological profile, relying on self-reported childhood symptoms, and questionable validity of commonly used measures also present challenges in detecting malingering in the context of ADHD evaluations. At the current time, adult ADHD is so poorly understood, that clinicians are frequently reluctant to diagnose malingered ADHD except in the most obvious cases. Furthermore, the current literature on malingered ADHD suggests college students easily simulate retrospective and current self-

11

reported symptoms and neuropsychological tests, successfully feigning current diagnostic criteria. This becomes more alarming when one considers rewards for successfully feigning ADHD include academic accommodations or obtaining schedule II narcotics (e.g., Adderall) that are frequently abused. The next section will briefly review the literature regarding malingering in general which will be followed by a literature review of malingered ADHD, specifically.

1.2. Overview of Malingering

Defining Malingering. Response distortion is not unique to psychological testing and is often seen in daily life. Individuals have myriad reasons to distort their responses and portray themselves in a certain manner. They may engage in positive impression management by responding in a defensive manner, denying or minimizing symptoms of psychopathology. Alternatively, they may choose to respond in a manner that exaggerates any current symptoms they may actually have or fabricate symptoms altogether. These two broad categories of response sets may be referred to as simulation (attempting to feign or exaggerate symptoms of psychopathology) and dissimulation (attempting to feign good health; Price, 1995 cited in Hayes et al., 1998). It should be noted that a there is a lack of consensus in definitions and terms used for simulation and dissimulation. In fact, Iverson (2006) states that, through the evolution of research in malingering, terms have been blurred, and now, terms that were once used to define distinct response styles are used interchangeably. For example, Bush (2005) and Rogers (2008) define dissimulation as deliberate distortion of symptoms (without a connotation of defensiveness).

The *DSM-IV-TR* currently classifies malingering as a V-code (V65.2) under conditions that should be the focus of additional clinical attention. The *DSM-IV-TR* states malingering should be suspected (1) in a medicolegal context, (2) if an individual is not cooperative during an

12

evaluation, (3) if performance is markedly different from what would be expected based on person's reported disability, or (4) if the person has a history of Antisocial Personality Disorder. These current guidelines fall short of actual diagnostic criteria for nearly all genuine Axis I disorders, excepting that of Factitious Disorder which is essentially the same as malingering except for its absence of obvious and tangible external incentive (other than secondary gain in the context of the factitious experience) to reward the simulator.

The *DSM-IV-TR* conceptualization of malingering is based on remnants of what has been called the Puritanical Model, which assumes individuals who malinger are morally "bad" (Rogers, 1990). In this article, Rogers also championed movement towards an empirical view of detecting malingering. Resnick (1997) proposed there are three distinct types of malingering, with increasing degree of conscious manipulation: "pure malingering", "partial malingering", and "false imputation." In his definition, "pure malingering" is conceptualized as a complete falsification of symptoms. He classified partial malingering as exaggeration of actual symptoms, and "false imputation" as intentional misattribution of actual symptoms for means of compensation. While several criteria have been proposed, clinicians and researchers most often refer to empirical criteria proposed by Slick, Sherman, and Iverson (1999) for classifying an individual as malingering (see Table 2). In order to diagnose malingered neurocognitive dysfunction (MND), an external incentive must be present. In addition, Slick et al. (1999) propose that MND is classified based on evidence from objective test data and/or self-reported symptoms. The Slick et al (1999) criteria allow clinicians to assert the degree of diagnostic accuracy (i.e. definite, probable, possible) of MND.

Rogers (2008) stated that, in his learned opinion, the most common error made in malingering detection is the overly precise attribution of antisocial labels to persons who display

13

Table 2. Operational criteria for malingering proposed by Slick, Sherman, & Iverson (1999).

Criterion A: One identifiable, substantial *external incentive*

Criterion B: *Evidence from test data*

1. Definite Response Bias: below chance (P < .05) performance on at least one FCT

2. Probable Response Bias: performance consistent with malingering on at least one validated measure or index of malingering

3. Discrepancy between performance and known patterns of functioning: must be consistent with known pattern of exaggeration

4. Discrepancy between test data and observed behavior: data from two or more tests within one domain are discrepant with observed functioning

5. Discrepancy between test data and reliable collateral reports: data from tests of at least one domain are discrepant with day-to-day functioning

6. Discrepancy between test data and background/history: data from two or more tests of a domain are discrepant with reported neurological history

Criterion C: *Evidence from self-report*

Discrepancy between self-report and documented history: consistent with attempt to exaggerate deficits

Discrepancy between self-report and known patterns of brain functioning: unlikely in number, severity, or pattern or inconsistent with known functioning

Discrepancy between self-report and observed behavior: self-report is discrepant with observed functioning

Discrepancy between self-report and reliable collateral reports: self-reported symptoms are discrepant with day-to-day functioning reported by reliable informants

Evidence of exaggerated or fabricated dysfunction: evidence from self-report and test data suggest exaggeration or malingering

Criterion D: *Behavior not fully accounted for by psychological, neurological, or developmental factors*

Classification Criteria

Definite MND: Criteria A, B1, and D

Probable MND: Criteria A, B2-6 OR C1-5, and D

Possible MND: either Criteria A, C1-5, and D OR criteria for Probable MND except for criteria D

14

deviant response styles, meaning that clinicians tend to automatically and pejoratively label these response styles as malingering even when insufficient motivational evidence exists. Rogers (2008) further recommends practitioners make decisions to favor using nonspecific descriptions (unreliable, deception, biased responding) of performance and self-report rather than merely inferring and pejoratively labeling specific responses styles (i.e. malingering, defensiveness, etc.). Terms such as "suboptimal effort" and "overreporting" are sometimes equated with malingering; however, he notes these terms lack precision and clarity and could be used to describe most clients. Clinicians and researchers should be judicious in labeling someone as malingering and may, instead, wish to use less value-laden terms to describe performance. Rogers (2008) differentiates "malingering" from "feigning" because feigning includes symptom fabrication without assumptions about its goals. For example, the diagnosis of Factitious Disorder is appropriately used if an individual feigns symptoms of psychopathology in order to establish and maintain a "sick role" (APA, 2000).

Similarly, Iverson (2006) states that, while there is overlap between the constructs of exaggeration, poor effort, and malingering, these terms are not synonymous. Poor effort, as he defines it, refers to a behavior during testing while exaggeration describes reporting of symptoms. In his conceptualization, malingering may involve one or both of these constructs with the intention of influencing scores for secondary gain. This conceptualization highlights the two components of malingering: effort and intent. Individuals who do not perform adequately during an assessment may purposefully and intentionally attempt to distort their responses (e.g. simulation, malingering, Factitious Disorder). However, other individuals may perform poorly due to lack of effort or motivation (e.g. fatigue, boredom etc.), but these individuals may not be intentionally attempting to portray deficits in order to obtain an external incentive. Despite the

potential for misuse, the term "malingering" is most commonly used in the literature regarding clinical populations to describe suboptimal performance on neuropsychological tests, without differentiation between poor effort, obvious non-compliance during examination, and exaggeration of symptoms in the presence of an external incentive (Rogers, 2008; APA, 2000; Slick, Sherman, & Iverson, 1999).

Construct of Malingering. As the field matures, more and more fine-grained delineations are being drawn. For example, researchers have begun to investigate whether there is overlap in malingering constructs, specifically malingered neurocognitive dysfunction versus malingered psychopathology. Larrabee (2012a; 2012b) and Bigler (2012a; 2012b) recently published a dialogue in the Journal of the International Neuropsychological Society where they suggest using the terms "performance validity" and "symptom validity" in the place of more general terms such as "effort" and "response bias." Larrabee (2012) defines performance validity as validity of test performance (e.g., what have traditionally been referred to as symptom validity tests) and symptom validity as validity of an individual's symptomatic complaints (e.g., MMPI-2) and argues these are two distinct constructs.

Nelson and colleagues performed an exploratory factor analysis using cognitive and psychological effort measures and reported that four distinct constructs emerged (Nelson, Sweet, Berry, Bryant, & Granacher, 2007). The first factor, "underreporting of psychological symptoms", involved the Minnesota Multiphasic Personality Inventory, Second Edition (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) scales used to detect minimization of psychopathology: Superlative (S; Butcher & Han, 1995), Defensiveness scale (K; Butcher et al., 1989) and the Lie Scale (L; Butcher et al., 1989). Factor 2, "overreporting of neurotic symptoms", included the MMPI-2 scales Response Bias Scale (RBS; Gervais et al., 2005), Fake

16

Bad Scale (FBS; Lees-Haley, English, & Glenn, 1991), Malingered Depression scale (Md; Steffan, Clopton, & Morgan, 2003). Factor 3, "insufficient cognitive effort", was comprised of stand-alone, neurocognitive effort measures (Letter Memory Test, Victoria Symptom Validity Test, and Test of Memory Malingering (TOMM). The MMPI-2 scales Infrequency (F), Infrequency Psychopathology Scale (Fp,; Arbisi & Ben-Porath, 1995) and Dissimulation Scale loaded onto Factor 4, "overreporting of psychotic/rarely endorsed symptoms." Nelson et al. (2007) suggest that malingered psychopathology differs significantly from malingered neurocognitive disorders because cognitive and psychological effort measures loaded onto distinct factors.

Other research indicates it is not uncommon for individuals to respond in ways that indicate they are suffering from multiple types of disability (neurocognitive, physical/somatic, and/or psychiatric) as the result of an illness or injury (Henry, Heilbronner, Mittenberg, Enders, & Stanczak, 2008). Investigators have found that measures designed to detect feigned psychopathology may be useful in neuropsychological evaluations where malingered cognitive dysfunction is suspected (Youngjohn, Wershba, Stevenson, Sturgeon, & Thomas, 2011; Henry et al., 2008; Whiteside, Dunbar-Mayer, & Water, 2009). In fact, some MMPI-2 scales have specifically been designed to detect malingering in head injury litigation (Lees-Haley Fake Bad Scale: FBS; Lees-Haley, 1991) and disability cases (e.g. Henry-Heilbronner Index: HHI; Henry, Heilbronner, Mittenberg, & Enders, 2006). Such scales show promise in detecting MND, but have been criticized as methods whose design and initial validity depended upon highly skewed and generally non-representative participant groups (Hielbronner, Sweet, Morgan, Larrabee, & Millis, 2009; Bigler, 2012a; Bigler, 2012b; Larrabee, 2012a; Larrabee, 2012b).

17

A recent study used taxometric procedures to examine the construct of malingering (Walters et al., 2008). The authors reported malingering has a dimensional structure with various levels of feigning/exaggeration as opposed to a dichotomous, malingering-honest or "good" vs. "bad" construct. This study highlights the need for empirically derived cutoffs using agreed upon criteria for defining unacceptable effort.

As this brief overview of malingering indicates, most research conducted on malingering involves cases of neuropsychological and psychiatric conditions. Malingering research has made significant progress in this past two decades, but it is still relatively in its infancy. Future research is necessary to clarify the construct of malingering as well as the most appropriate means of assessing effort throughout evaluations.

Detection of Malingering. Individuals may choose a variety of strategies when they attempt to feign a disability, so effort indices use a variety of methods to detect feigning. Greene (1997) proposed a bipolarity hypothesis in which malingering and defensiveness are polar opposites, and he suggests that malingering can be detected by the absence of defensiveness. Rogers (1997) outlined a number of methods that are commonly used to detect response bias. These include: using items that are rarely endorsed by individuals with genuine deficits (rare symptoms), comparing symptom severity endorsed by malingerers to actual patients (symptom severity), and examining subtle items that are not generally recognized by nonprofessionals compared to obvious items that are well known symptoms of the disorder in question (obvious and subtle symptoms). Another detection strategy, "erroneous stereotypes", detects response bias by identifying item responses that reflect common misperceptions about disorders (Rogers, 2003).

18

There are myriad tests and indices designed for the detection of MND and malingered psychopathology. Many stand-alone tests, also known as symptom validity tests (SVT), have been developed specifically for the detection of sub-optimal effort or intentional distortion of responses. These tests are administered for the sole purpose of assessing malingered deficits during a neuropsychological evaluation. Another strategy for detecting malingering is developing embedded effort indices within popular neuropsychological measures. Embedded indices are designed to detect poor effort on specific psychological tests such as the Wechsler Adult Intelligence Scales, Third Edition (WAIS-III; Wechsler, 1997a) and Wechsler Memory Scale, Third Edition (WMS-III; Wechsler, 1997b). Advantages of embedded effort indices include: providing information about effort on specific psychological and neuropsychological tests and obtaining information about effort without extending the length/time of the evaluation. Within each of these two broad categories, there are multiple methods for inferring the presence of response bias.

Forced-choice Effort Tests (FCT). Forced-choice tests provide respondents with a dichotomous-choice response format. Usually, one choice is a target that was previously presented, and the other choice is a foil. Some forced choice tests present dichotomous yes/no response set. Statistically, the probability of guessing correctly, by chance alone, would approximate 50%. The initial intent of Forced-choice Tests was to detect below chance levels of responding; however, such response sets are rare. For many Forced-choice Tests, cutoff scores have been developed and validated. These cutoff scores are usually set below the level attained from the most impaired patient sample appropriate for the test. These have proven useful in detecting individuals who are exhibiting suboptimal performance. Forced-choice methodologies

19

may be found in stand-alone SVTs or in embedded malingering indices (e.g. WMS-III Rarely Missed Items Index; Killgore & DellaPietra, 2000).

Floor Effect Strategies. Normative floor effect strategies compare an individual's performance to performance of persons with known impairment (Frederick, 2000). Floor effect strategies are used to determine whether an individual's performance is at or below the level that would be expected for the disability in question. Many floor effect strategies rely on previously established cutoff scores. The rationale behind cutoff scores is that an individual need not demonstrate below-chance performance if the vast majority of individuals with a given disorder perform above the cutoff score. It should be noted this approach has limitations. For example, actual cognitive impairments can interfere with effort testing in individuals putting forth their best effort (Merten, Bossink, & Schmand, 2007). Interpretation of response profiles that fall below empirically derived cutoff scores for neurologically impaired or intellectually disabled individuals is that the individual in question is may not be providing sufficient effort by responding carelessly (effort) and/or may be intentionally distorting his/her responses (intent). Floor effect strategies may be used in either symptom validity tests or embedded effort indices.

Pattern Analysis. Meehl (1956) purported that a methodology based on multivariate statistical analyses can, almost always, outperform clinical judgment in detecting a diagnosis of interest. In addition, there is evidence that malingerers are able to successfully generate scores that are characteristic of individuals who have a specified disorder (Heaton, Smith, Lehman, & Vogt, 1978). In such cases, it is necessary to examine malingering based on patterns of performance on individual subtests in neuropsychological testing. Heaton et al. (1978) first proposed that malingerers could be identified by their performance patterns on psychological and neuropsychological tests. Today, such measures include the Minnesota Multiphasic Personality

Inventory, Second Edition (MMPI-II; Butche et al., 1989), WAIS-III, and Halstead Reitan battery (Reitan & Wolfson, 1985). Similarly, Mittenberg and colleagues used a discriminant function analysis to identify patterns of performance on the WAIS-R that successfully differentiated malingerers from individuals with traumatic brain injuries (Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995).

Rare or Improbable Symptoms. Malingerers will often endorse symptoms or fail test items that are rarely endorsed/missed by individuals with actual impairments (Rogers, 2008). The Miller Forensic Assessment of Symptoms Test (M-FAST; Miller, 2001), PAI Negative Impression Management scale, and Structured Interview of Reported Symptoms (SIRS; Rogers, Bagby, & Dickens, 1992) rely on this methodology for detecting feigned psychopathology. The WMS-III Rarely Missed Items Index (Killgore & DellaPietra, 2000) and the Stanford Binet-5 Rarely Missed Items-Nonverbal Index (Musso, Barker, Jones, Roid, & Gouvier, 2011) capitalize on this principle for detecting intentionally suppressed cognitive functions. Each of these scales include items that are rarely endorsed or missed by individuals with the disorder in question but are often endorsed or missed by individuals attempting to malinger.

Of all of the measures previously discussed, the Personality Assessment Inventory (PAI) will be used for the current study. The PAI employs several methods previously mentioned. For example, the Infrequency scale is based on items that are rarely endorsed by most individuals, and the Rogers Discriminant Function and Malingering Indices were derived from pattern analyses. The current study will employ a pattern analysis approach to developing a scale embedded in the PAI designed to detect malingered ADHD. The next section reviews broad literature regarding validity indices of the PAI. In addition, the author will present a brief review of utility of the PAI in detecting feigned disorders in a variety of settings.

1.3. Malingering on the Personality Assessment Inventory

There are eight PAI indices designed specifically to detect response distortion. Three indices were developed to detect positive impression management: Positive Impression Management scale (PIM; Morey, 1991), the Defensiveness Index (Morey 1993, 1996), the Cashel Discriminant Function (CDF; Cashel, Rogers, Sewell, & Martin-Cannici, 1995). There are also three indices designed to detect malingered symptoms: Negative Impression Management scale, the Malingering Index (Morey 1993, 1996), and Rogers Discriminant Function (Rogers, Sewell, Morey, & Ustad, 1996). Two additional scales identify profiles fraught with rarely endorsed or inconsistent responses (Infrequency and Inconsistency; Morey, 1991).

The Negative Impression Management scale is composed of eight items that do not overlap with the clinical scales. Its purpose is to detect negative response bias (i.e., responses that present an overly unfavorable impression of the individual). While the Negative Impression Management scale can be elevated in outright malingering, overly negative self-evaluative styles, indicating pessimism or low self-esteem, can also cause elevations on this scale (Morey & Lanier, 1998). A relatively high cutoff t-score of 92 is used for distinguishing malingered profiles from genuinely pathological ones. The Malingering Index is composed of eight "configural features" of the PAI that are elevated significantly more often in simulated psychopathology than in actual psychopathology. The Rogers Discriminant Function was developed to differentiate individuals feigning psychopathology from individuals with actual psychological disorders. Thus, the Rogers Discriminant Function is an example of pattern analysis and is based on the premise that individuals feigning psychopathology often have difficulty simulating profiles of actual patients.

22

A recent meta-analysis examined the utility of the PAI validity scales (Negative Impression Management scale, Malingering Index, and Rogers Discriminant Function) for detecting feigned psychopathology and found that all three scales were able to discriminate malingerers and simulators from honest respondents and patients (Haws & Boccaccini, 2009). It should be noted, however, that effect sizes were smaller when trials were conducted using criterion groups compared to simulation designs. Another study reported that, when psychiatric inpatients were asked to fake-bad on the PAI, they obtained elevated scores on the Negative Impression Management scale, Rogers Discriminant Function, Malingering Index and Cashel Discriminant Function (Baity, Siefert, Chambers, & Blais, 2007). In the same study, discriminant function analyses indicated that the Negative Impression Management scale, Rogers Discriminant Function, and Defensiveness Index made significant contributions in distinguishing exaggerated psychopathology from actual psychopathology.

There is some debate of the usefulness of the PAI validity indices in criminal forensic settings. One study compared responses of forensic psychiatric patients asked to either malinger psychopathology or respond honestly on the Structured Inventory of Malingering Symptomatology, Structured Interview of Reported Symptoms, and PAI validity scales: Negative Impression Management scale, Rogers Discriminant Function, and Malingering Index (Edens, Watkins-clay & Poythress, 2007). The authors reported that only the Rogers Discriminant Function and Malingering Index scales performed adequately in differentiating individuals from the psychiatric unit who were asked to feign from those judged to be genuinely mentally ill. However, other authors reported only the Negative Impression Management scale of the PAI performed adequately in differentiating a group of criminal defendants, who were deemed to be malingering, from a group determined to be mentally ill (Kucharski, Toomey, Fila,

& Duncan, 2007). The Negative Distortion Scale is a new validity scale, derived for the PAI, designed specifically for detecting malingered psychopathology in forensic settings, where there are expected to be higher levels of psychopathology (Mogge, Lepage, Bell, & Ragatz, 2010). The authors report that initial validation was promising, but further research is needed.

The Negative Impression Management scale, Malingering Index, and Rogers Discriminant Function indices of the PAI have not provided adequate psychometric properties for the detection of malingered pain-related disabilities (Hopwood, Orlando, & Clark, 2010). Therefore, Hopwood et al. (2010) used discriminant function analyses to develop a new validity index of the PAI that demonstrates adequate sensitivity and specificity for the detection of malingered pain-related disorders. However, this index was developed using undergraduate simulators and needs further validation in actual clinical samples where malingered pain related disabilities are more likely.

A recent study found the PAI Negative Impression Management and Infrequency scales (designed to assess careless responding) are correlated with Test of Memory Malingering scores (Whiteside, Dunbar-Mayer, & Water, 2009). This suggests that individuals who fail these measures may also malinger cognitive impairments. A subsequent study indicated the Somatization clinical scale was significantly predictive of Test of Memory Malingering-Trial Two failure in individuals completing a neuropsychological evaluation (Whiteside, et al., 2010), suggesting that individuals who malinger cognitive effort may be inclined to malinger physical/somatic symptoms on self-report measures of psychopathology.

The PAI was chosen for the current study, in part, because it has proven useful in detecting malingering in a variety of disorders, including feigned cognitive disorders. Wasserstein (2005) reported that measures of psychopathology should be employed in ADHD

evaluations because psychiatric symptoms may mimic inattention and poor concentration and be confused with ADHD among naïve college students. Therefore, an embedded index in the PAI could potentially be useful for detecting response bias in this population. The next section will review the current literature on malingered ADHD in college students.

1.4. Malingered ADHD

In order for malingering to occur, an external incentive must be present. Clinicians and researchers have begun to recognize that college students have two primary external incentives to feign ADHD: the legal acquisition of stimulant medications and receipt of academic accommodations that may improve their grades. Little research has been conducted in this area, but at this point, it appears that individuals are easily able to feign ADHD symptoms and few tests are able to adequately distinguish individuals feigning ADHD from individuals who actually have neurocognitive impairments that interfere with their functioning. This section will review the literature relevant to malingered ADHD.

Incentives for Malingering ADHD Symptoms. College students have a number of incentives for feigning ADHD deficits. Individuals diagnosed with ADHD may receive federally regulated stimulant medications that are frequently abused. These medications are commonly used recreationally and are a valuable commodity in the illicit marketplace, selling for over $5 per pill in some instances (Garnier-Dykstra, Caldeira, Vincent, O'Grady, & Arria, 2012). In addition, students with ADHD are typically eligible for services such as academic accommodations that may further improve performance in school. External incentives create an atmosphere that increases the possibility that some individuals may malinger deficits. Effort must be considered in addition to observed deficits during testing prior to diagnosing ADHD, or in the words of our belated President, Ronald Reagan, "Trust but verify."

25

Stimulant abuse. College students abuse stimulant medications for a number of reasons including: desire to improve attention, reduce hyperactive symptoms, improve grades, and recreational intoxication (White, Becker-Blease, & Grace-Bishop, 2006; Teter et al., 2005). In one study of individuals who were prescribed stimulant medications for ADHD, 31% reported misuse of their medications (Rabiner, Anastopoulos, Costello, McCabe, & Swartzwelder, 2009). Of those students, 8% reported intranasal ingestions of the stimulant medication, and 26% of students who had been prescribed stimulant medications for ADHD reported they sometimes diverted their medications to peers. In a survey of a northeastern college, 9.8% of students prescribed stimulant medications for ADHD reported they misused their medications (White et al., 2006). In a different study that surveyed 390 northeastern college students, 7.5% reported they engaged in nonmedical use of stimulant medication, and 60% reported that they knew individuals who misused stimulant medications (Weyandt, et al., 2009). Students in this study who endorsed nonmedical stimulant use also endorsed increased psychological distress. In another northeastern study, 16.2% of college students with no diagnosis of ADHD reported misusing or abusing stimulant medications (White et al., 2006). Similar results were found when 21,294 students from a large Midwestern university were surveyed (McCabe, Teter, & Boyd, 2006). Of the sample, 8.1% of students reported using illicit stimulant medications in their lifetime while 5.4% reported illicit use in the past year. The majority of students in this survey reported they obtained the illicit stimulant medication from a peer. Rates of illicit use of stimulant medication may be much higher. For example, Advokat and colleagues found that 43% of the 1,550 college students surveyed at a large Southern university reported they have used illicit stimulant medications without having an ADHD diagnosis (Advokat, Guidry, & Martino, 2010). Another survey of 483 college students prescribed prescription medications found that

26

35.8% diverted their medications at least once. The rates of diversion of stimulant medications were much higher, as 61.7% of college who were prescribed stimulant medications reported diverting them at least once. Of the sample, 9.8% of the respondents reported they sold their prescription medication (Garnier et al., 2010).

Academic Accommodations. In addition to access to stimulant medications, students diagnosed with ADHD are offered myriad opportunities for academic assistance. The United States government has passed several laws (i.e. the Individuals with Disabilities Act of 1975, Section 504 of the Rehabilitation Act of 1973, and the Americans with Disabilities Act of 1990) that enable individuals with disabilities to receive accommodations (Latham, 1996). Weydandt and DuPaul (2006) note the American Council of Education (1995) reported the percentage of college students receiving academic accommodations rose from 2.2% in 1978 to 8.8% in 1991. Depending on the area of impairment, usually demonstrated in a neuropsychological evaluation, students might receive extended time, quiet testing environments, note takers, readers, or even alternative courses (Sullivan et al., 2007).

This brief review of the literature highlights the incentives for college students to feign ADHD. The ongoing debate about the construct and diagnosis of adult ADHD and different assessment practices by different professionals creates an atmosphere ripe for malingering by college students. This is a relatively new domain in malingering and the next section will review the current literature regarding malingered ADHD by college students in psychological and neuropsychological assessment.

Research Involving Malingered ADHD. Typically, researchers investigate malingered ADHD by either asking analog subjects to perform as though they were intentionally faking ADHD (simulation studies) or by classifying individuals from clinical samples that fail either

gold-standard effort indices or that meet Slick et al. (1999) criteria as malingering (clinical data). The literature to date has examined malingered ADHD on self-report questionnaires, broad personality inventories, neuropsychological tests, and symptom validity tests (for review see Musso & Gouvier, 2012).

Malingering ADHD on self-report questionnaires. There have been a number of studies that have investigated college students' abilities to feign childhood symptoms of ADHD on self-report questionnaires. Quinn (2003) conducted one of the first studies on malingered ADHD, using self-report questionnaires, and found that college students easily faked symptoms of ADHD on the ADHD Behavior Checklist-Retrospective (Murphy & Barkley, 1996a). Two studies have also examined college students' abilities to feign ADHD symptoms on the Wender Utah Rating Scale and found that college students asked to distort their responses are able to produce profiles consistent with ADHD (Jachimowicz & Geiselman, 2004; Booksh et al., 2010). Similarly, individuals suspected of noncredible effort (failed effort indices) produced scores on the WURS that were indistinguishable from a clinical sample that received ADHD diagnoses and did not fail any effort indices (Suhr et al., 2008).

When studies examined the ADHD Behavior Checklist-Current (Murphy & Barkley, 1996b), it was found ADHD Simulators did not differ significantly in their responses compared to clinical ADHD groups (Quinn 2003; Fisher & Watkins, 2008). Fisher and Watkins (2008) also examined coached simulators' abilities to feign symptoms on the College ADHD Response Evaluation (CARE; Glutting, Sheslow, & Adams, 2002) and reported 93% of simulators obtained believable, clinically elevated scores on the CARE. Other measures that have proved susceptible to malingered ADHD include the ADHD Current and Childhood Symptoms Scales-Self-Report Forms (Barkley & Murphy, 2006; Young & Gross, 2011), the Attention Rating Scale

Current and Childhood Symptom Checklists (Barkley & Murphy, 2006; Sollman, Ranseen, & Berry, 2010), the ADHD Rating Scale (ARS; DuPaul, Power, Anastopoulos, & Reid, 1998; Jachimowicz & Geiselman, 2004), and the Brown Adult ADHD Scale (BAAS; Brown, 1996; Jachimowicz & Geiselman, 2004), as each of these measures has been successfully manipulated by individuals attempting to feign ADHD. Students asked to simulate ADHD endorsed more symptoms on the Attention Deficit Scales for Adults (ADSA; Triolo & Murphy, 1996) and the Brown Attention Deficit Disorder Scale (BADDS; Brown, 1996) compared to controls and a clinical ADHD sample; however, their scores were not statistically or clinically useful in differentiating feigned from actual ADHD (Booksh et al., 2010 and Tucha, Sontag, Wlitza, & Lange, 2009, respectively). Marshall et al. (2010) found individuals who met Slick criteria for malingering (i.e. non-credible group) performed similarly to an honest, clinical ADHD group on the Barkley ADHD Self-report subscales. Marshall et al. (2010) reported a noncredible ADHD group endorsed significantly more inattentive behaviors compared to controls, a clinical ADHD sample, and a sample of college students who met Slick criteria (but were not diagnosed with ADHD). In addition, Marshall et al. (2010) examined performance on the Clinical Assessment of Attention Deficit-Adult Infrequency Scale (Bracken & Boatwright, 2005). Overall, they reported that recommended cutoffs did not offer an acceptable balance of sensitivity and specificity for detecting malingered ADHD.

Of all the self-report measures, the ability of the Conners' Adult ADHD Rating Scale (CAARS; Conner, Erhardt, & Sparrow, 1999) to detect malingered ADHD has been most extensively studied. Studies that examined analog ADHD malingerers (Harrison, Edwards, & Parker, 2007; Jachimowicz & Geiselman, 2004; Harp et al., 2011) or clinical groups suspected of malingered ADHD (Suhr et al., 2008; Harrison & Edwards, 2010; Sollman et al., 2010; Jasinski

29

et al., 2011) have found that symptoms of ADHD are easily feigned on the CAARS. While no subscales have proven clinically useful in distinguishing feigned ADHD from clinical ADHD, several studies found that individuals feigning ADHD tended to endorse more symptoms of hyperactivity/restlessness on the CAARS (Harrison & Edwards, 2010; Harp et al., 2011). Similarly, Suhr et al. (2008) reported that individuals suspected of noncredible performance during psychoeducational evaluations endorsed significantly more symptoms on the hyperactivity/restlessness subscales of the CAARS compared to a psychological symptom (no ADHD) group, but their scores did not differ significantly from a clinical ADHD group. Jachimowicz and Geiselman (2004) reported that hyperactivity accounted for 83% of the variance in positive diagnoses of ADHD in ADHD Simulators, suggesting the students may associate hyperactivity with ADHD to a greater degree than inattention.

Interestingly, none of the studies that examined the CAARS found the inconsistency index, designed to detect invalid responses styles, was useful in detecting feigned ADHD. In 2011, Suhr, Buelow, and Riddle developed an Infrequency Index for the CAARS (CII). They choose twelve items that were endorsed "pretty much, often" to "very much, very frequently" by fewer than ten percent individuals in their sample. They summed the scores of these items to create a scale. The authors reported that a cutoff score of ≥ 20 achieved specificity greater than 90%. When they performed a validity study using a hold-out sample, they found that the CII demonstrated acceptable sensitivity and excellent specificity for failure of Word Memory Test (WMT, Green, 2003).

Overall, it appears that self-report questionnaires designed to measure childhood or current symptoms of ADHD are easily faked by college students. The CAARS and the Clinical Assessment of Attention Deficit-Adult Infrequency Scale have validity indices, but these indices

were not useful in distinguishing individuals suspected of feigning ADHD from individuals that responded honestly and met diagnostic criteria for ADHD. Suhr et al. (2011) developed an additional validity index for the CAARS, the Infrequency Index, which demonstrated some promise in detecting feigned ADHD; however, further research is needed to validate the scale.

Malingering ADHD on Broad, Objective Personality Inventories. The MMPI-2 and PAI have numerous validity indices that are able to detect a variety of response biases. Few studies have examined the utility of these indices in detecting malingered ADHD in college students. Young and Gross (2011) reported that individuals asked to simulate ADHD on the MMPI-2 scored higher than the clinical ADHD on a number of scales including: the Infrequency (F; Butcher et al., 1989), Infrequency Psychopathology Scale (Fp,; Arbisi & Ben-Porath, 1995), Back-Infrequency Scale (Fb; Butcher et al., 1989), Response Bias Scale (RBS; Gervais, Ben-Porath, Wygant, & Green, 2007), and the Henry-Heilbronner Index (HHI; Henry et al., 2006); however, few scales yield adequate sensitivity and specificity. The most useful scales were the Fp ($\geq$ 5) and the HHI ($\geq$ 9), which obtained sensitivities of .59 and .47 and specificities of .94 and .89, respectively. The elevations of these scales suggest individuals attempting to feign ADHD may endorse symptoms consistent with externalizing complaints rather than cognitive or somatic symptoms. Harp and colleagues (2011) examined the MMPI-2-RF and found that, for the most part, simulators produce profiles that were consistent with a clinical ADHD sample. The ADHD simulators did obtain significantly higher scores on the RC1 subscale and the F-r subscale. Sullivan et al. (2007) examined the validity indices of the PAI by comparing college students that failed at least one index on the WMT to individuals that did not. They found that few individuals scored above the cutoffs for the PAI validity scales. Overall, it appears that the

31

current embedded validity indices of the MMPI-2 and PAI may be useful in detecting malingered ADHD, but further research is needed.

Malingering ADHD on Neuropsychological Tests. Many studies have sought to determine whether individuals asked to feign ADHD could be differentiated from individuals that were diagnosed with ADHD on standardized neuropsychological tests. The most widely investigated measures in the literature are Continuous Performance Tests (CPT) that measure of attention, concentration, reaction time, impulsivity, and inhibition. Quinn (2003) compared performance of ADHD Simulators to a clinical ADHD sample and found that the analog sample scored significantly lower on most subtests. She proposed an impairment index for the Integrated Visual and Auditory Continuous Performance Test (IVA CPT) that is composed of a combination of the Auditory Attention Quotients (< 44) and Response Control Score (< 74). She reported that this index yielded a sensitivity of .94 and specificity of .91. This index has not been investigated further. Booksh et al. (2010) found that ADHD Simulators scored significantly lower than controls on most variables of the Conners CPT (Conners, Erhart, & Sparrow, 1999); however, their scores differed from the Clinical ADHD sample only on the Total Score and the number of clinically elevated scales. Sollman et al. (2010) reported that, while individuals asked to feign ADHD obtained significantly different scores compared to a clinical ADHD sample, their scores were in the believable range. Marshall et al. (2010) reported that individuals that met Slick et al. (1999) criteria for poor effort did not obtain significantly different scores on the CPT compared to a clinical ADHD sample. Suhr et al. (2011) reported similar findings in a clinical sample of individuals suspected of giving poor effort based on failure of the WMT compared to a clinical ADHD sample, stating that only the CPT reaction time variability and reaction time change over interstimulus intervals variables differed between groups. Overall, the literature

suggests that CPT measures could be feigned, as, even when there are significant differences between groups, the scores are not clinically significant because scores from both groups remain in a believable range.

Multiple measures of attention have been examined in order to determine whether feigned ADHD could be detected. In one of the first published studies on malingered ADHD in college students, Leark and colleagues (2002) reported that ADHD Simulators obtained significantly different scores on numerous variables of the Test of Variables of Attention (TOVA; Greenberg et al., 1996). However, no further studies have examined the utility of the TOVA at this time. Studies of the Trail Making Test Parts A and B (Reitan, 1955) were not consistent. Surh et al (2008) reported individuals with suspect effort, when compared to clinical ADHD group who did not fail any effort indices, obtained similar scores on Trails A but lower scores on Trails B. Booksh et al. (2010) reported ADHD Simulators obtained similar scores on Trails B but lower scores on Trails A compared to a clinical ADHD group. Two studies (Harrison et al., 2007; Sollman et al., 2010) examined the Stroop Color and Word Test (Golden, 1978; Golden & Freshwater, 2002). Harrison et al. (2007) reported significant differences between ADHD Simulators and clinical ADHD samples only on the Interference score, and Sollman and colleagues (2010) reported that, while simulators scored statistically lower on the Color Word subtests, scores were in the believable range. Marshall et al. (2010) examined the Delis-Kaplan Executive Function System Color-Word Interference Test (Delis, Kaplan, & Kramer, 2001) and reported individuals with suspect effort performed worse, but the measure did not offer adequate sensitivity for detection of malingering. These authors also found individuals with suspect effort performed similarly to a clinical ADHD group on the NAB Numbers and

33

Letters Test (White & Stern, 2003), a measure of processing speed. At this time, it appears these measures of attention and processing speed are easily faked.

Several studies have examined performance on measures of academic achievement. Sollman et al. (2010) reported that ADHD Simulators' scores did not differ significantly from a clinical ADHD group on the Nelson-Denny Word Reading Test (Brown, Fishco, & Hanna, 1993). Frazier, Frazier, Busch, Kerwood, and Demaree (2008) reported that ADHD Simulators scored significantly lower than a control group on the reading subtest of the Wide Range Achievement Test, Third Edition (Wilkinson, 1993). Studies of the Woodcock-Johnson Psychoeducational Battery (Woodcock et al., 2001) found that ADHD Simulators (Harrison et al., 2007) and individuals with suspect effort (Harrison et al., 2010) scored lower on the Decision Making Speed subtest, but their scores were not clinically useful in detecting malingering. Harrison et al. (2007) stated that ADHD Simulators also scored lower than a control group and the clinical group on the Reading Fluency, Visual Matching, and Processing speed subtests of the Woodcock Johnson Psychoeducational Battery. They attempted to use discriminant function analysis with the CAARS and Woodcock Johnson scores in order to classify malingerers, but found that the best function produced a combined 25% error rate. These studies suggest that, while ADHD Simulators may score lower on measures of academic achievement, differences do not offer clinical utility in discriminating ADHD Simulators from Clinical groups.

Measures of memory for word lists produced conflicting data. Sollman et al. (2010) found ADHD Simulators performed similarly compared to the clinical ADHD sample on the Word Lists Subtest of the WMS-III. Harrison et al. (2010) reported their suspect effort group's scores were significantly lower on the Logical Memory and Family Pictures Immediate and Delayed subtest mean, but scores remained in the believable range. However, Suhr et al. (2008)

34

reported a group of individuals who failed at least one index of the WMT obtained significantly lower scores on the Auditory Verbal Learning Test compared to clinical ADHD and psychological symptoms samples, and Sullivan et al. (2007) reported that WMT failure correlated with scores on the California Verbal Learning Test, Second Edition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000). It appears individuals suspected of feigning ADHD are able to suppress their scores on memory tests; however, scores fall within believable ranges, and these individuals would not be easily identified as malingering.

Finally, multiple studies have examined performance on the WAIS-III in the detection of malingered ADHD. Sullivan et al. (2007) reported that failure of the WMT correlated with WAIS-III FSIQ scores. Harrison and Edwards (2010) reported that there were significant differences in WAIS-III FSIQ scores between individuals with suspect effort and a clinical ADHD sample. Malingerers scored lower than a clinical sample on the Performance IQ subtest in two studies (Sullivan et al., 2007; Harrison & Edwards, 2010). Sullivan et al (2007) did not find significant differences between groups for the Verbal IQ index; however, Harrison and Edwards reported that their groups differed significantly on VIQ, but the scores were believable. Investigation of the processing speed index yielded no clinically significant differences in three studies (Suhr et al., 2008; Booksh et al., 2010; Harrison & Edwards, 2010). However, Marshall et al. (2010) reported that a suspect effort group differed significantly from the clinical group on the processing speed subtests of the WAIS-III. These authors pointed out that the Suspect Effort group obtained scores that would be believable to naïve clinicians, limiting the clinical utility of these subtests as effort indexes. Frazier et al. (2008) reported ADHD Simulators scored significantly lower on the Digit Symbol Coding Subtest of the WAIS-III when compared to a control group. On the Working Memory Index, Suhr et al. (2008) reported significant differences

between groups while Harrison and Edwards (2010) did not find significant differences. Booksh et al (2010) examined WAIS-III Working Memory subtests (Digit Span and Letter Number Sequencing) and found ADHD Simulators did not obtain significantly different scores on these measures when compared to clinical groups.

Based on this literature, it is evident that neuropsychological test data alone is not sufficient to detect poor effort during psychoeducational evaluations. Studies that use data from credible and noncredible patients highlight the necessity of effort assessment during ADHD evaluations (Sullivan et al., 2007; Suhr et al., 2008; Marshall et al., 2010) because individuals who fail gold-standard malingering indices such as the WMT produce similar neuropsychological profiles compared to individuals who do not fail any effort measures. As with all other areas of malingering, it is best to collect data from multiple sources for determination of ADHD (Booksh et al., 2010). Many of the studies that examined the utility of neurocognitive validity indices in detecting malingered ADHD found Simulators and individuals suspected of putting forth poor effort easily suppressed scores to appear impaired though not suspiciously so.

Malingering ADHD on Effort Measures. Stand-alone measures designed to detect malingered cognitive dysfunction have been examined. Frazier et al. (2008) found significant differences between an ADHD simulator group and a control group for the Rey Fifteen Item Test (Rey, 1964), Victoria Symptom Validity Test (VSVT; Slick, Hopp, Strauss, & Thompson, 1997) and most indices of the Validity Indicator Profile (VIP; Fredrick, 2003) except the verbal slope. On the VIP, a cutoff score of 75 yielded .79 sensitivity and .96 specificity. Sensitivity of .71 and specificity of .93 were found when a cutoff score of 12 was used for the MFIT. The best

discrimination between ADHD simulators and the control group was found for the VSVT hard items yielding a sensitivity of .80 and specificity of 1.00.

Marshall and colleagues (2010) considered performance non-credible if either two separate SVT measures were failed, if one SVT and unusually impaired performance on one cognitive test was noted, or when one SVT or unusually impaired performance of one cognitive test combined with invalid completion of a behavioral rating scale. In their study, Dot Counting Test (Boone, Lu, & Herzberg, 2002a) produced acceptable specificity but poor sensitivity, and the WMT immediate recall, consistency, and pass/fail cutoffs of 82.5% yielded adequate sensitivities of 63.64%, respectively, and specificity above 90%. Booksh et al. (2010) also reported using the WMT to classify malingering improved classification accuracy over clinical judgment alone.

Sollman et al. (2010) investigated the utility of several neurocognitive effort indices for detecting feigned ADHD. They also examined the usefulness of the Miller Forensic Assessment of Symptoms Test (M-FAST; Miller, 2001), a measure of feigned psychopathology. They found the M-FAST yielded excellent specificity but inadequate sensitivity. In addition, Sollman et al. (2010) examined the utility of the Digit Memory Test (DMT; Hiscock & Hiscock, 1989), Letter Memory Test, Card Version (LMT; Inman et al., 1998; Schipper, Berry, Coen, & Clark, 2008), and Green's Nonverbal-Medical Symptom Validity Test (NV-MVST; Green, 2006) for distinguishing ADHD simulators from the clinical ADHD group. They reported that the Test of Memory Malingering (TOMM; Tombaugh, 1996) demonstrated the best balance between sensitivity and specificity. All other measures examined demonstrated excellent specificity but less sensitivity. Jasinski et al. (2011) administered the Digit Memory Test, Letter Memory Test, Test of Memory Malingering, b Test error score, and Nonverbal Medical Symptom Validity

37

Test. They found the ADHD Simulators scored significantly worse on all measures compared to a clinical ADHD group. When individuals with ADHD were asked to exaggerate their symptoms, they scored significantly lower compared to the clinical ADHD group on most measures with the exception of the b Test and the Nonverbal Medical Symptom Validity Test. All symptom validity tests administered were more sensitive to a group of non-ADHD individuals instructed to feign symptoms compared individuals with ADHD instructed to exaggerate their symptoms.

Numerous studies reported that, when results of symptom validity test failure were combined, positive predictive power increased. Jasinki et al. (2011) found that when two or more tests were failed, positive predictive power was 1.00 while Sollman et al. (2010) reported failure of three or more symptom validity tests produced 100% positive predictive power. Marshall et al. (2010) also reported that the probability of identifying malingering increased when three or more SVTs were failed.

Regarding embedded validity indices, Suhr et al. (2008) compared scores on embedded effort measures for a clinical sample of individuals that failed at least one effort index of the WMT to a clinical sample of ADHD individuals that gave good effort on the WMT. Individuals who failed the WMT failed few other embedded effort indices including: the Exaggeration Index of an expanded Auditory Verbal Learning Test (EIAVLTX; $n = 2$; Barrash, Suhr, & Manzel, 2004); the WAIS-III Digit Span score < 5 ($n = 5$; Iverson & Franzen, 1994), a Working Memory Index score < 70 ($n = 1$; Etherton, Bianchini, Ciota, Heinly, & Greve, 2006), the Auditory Verbal Learning Test recognition score < 10 ($n = 3$; Boone, Lu, & Wren, 2005), the Vocabulary-Digit Span score ≥ 2 ($n = 8$; Greve, Bianchini, Mathias, Houston, & Crouch, 2003). They concluded

38

that these embedded validity indices demonstrated unacceptably poor sensitivity to feigned ADHD.

Harrison and colleagues (2010) classified a subgroup of individuals (n = 17) as malingering based on Slick et al. (1999) criteria including failure of the WMT and failure of an additional SVT. They reported that WAIS-III embedded validity indices including the Digit Span (≤ 5; Iverson & Tulsky, 2003), Vocabulary – Digit Span (cutoff ≥ 5; Iverson & Tulsky, 2003), and Reliable Digit Span scores (cutoff ≤ 6) yielded poor sensitivity to malingered ADHD. When Receiver Operating Characteristic curves were examined, no measures produced strong classification accuracy; however, because the Digit Span and Reliable Digit Span scores incorrectly identified very few honest participants, failure of one or both of these subtests could indicate poor effort during psychoeducational evaluations.

Marshall et al. (2010) used the following standalone and embedded effort measures: Sentence Repetition Test (Strauss, Sherman, & Spreen, 2006) < 10 (Schroeder & Marshall, 2010), California Verbal Learning Test, Second Edition (Delis et al., 2000) Forced Choice Recognition Test (Root, Robbins, Chang, & Van Gorp, 2006), b Test (Boone, Lu, & Herzberg, 2002b; E-score, commission errors, total time, d error, omission errors), and the Reliable Digit Span score < 6 (Babikian et al., 2006). They also examined variables from the Conner's CPT (omission and commission t-scores) and the Test of Variables of Attention (Omission Errors, Commission Errors and RT Variability) that are considered indicative of suspect effort. With the exception of the Conner's Omission t-score cutoff < 20, most of the embedded measures examined produced excellent specificity (> 90%). However, few measures demonstrated acceptable sensitivity. There was little overlap in shared variance for most measures indicating that the SVTs are independent measures of effort, supporting the use of multiple SVTs. It should

be noted that in cases where there are numerous presentations for a disorder, low sensitivity is not uncommon, but high specificity is imperative. The results of Marshall et al. (2010) suggest that, while sensitivity was relatively low for most indices examined (e.g., instruments would not detect the vast majority of malingerers), individuals who fail multiple validity indices could confidently be classified as malingering.

This review of the literature suggests the most promising means of detecting malingered ADHD appear to be symptom validity tests. Specifically, many current studies found the WMT to be useful in detecting feigned ADHD symptoms. In addition, multiple studies used failure of the WMT as the criterion for suspect effort in their studies. However, it should be noted most symptom validity tests were originally designed to detect malingered cognitive symptoms. Most other embedded and standalone effort measures demonstrated excellent specificity, but few of the measures examined produced adequate sensitivity. While combining symptom validity tests may increase positive predictive power (Marshall et al., 2010; Sollman et al., 2010; Jasinski et al., 2011), this approach must be tempered by the fact that as the number of validity indices given increases so does probability of failing an index by chance. Clinicians should account for number of validity indices administered in practice in a manner similar to that of statisticians using Bonferroni correction to account for chance findings in multiple comparisons. In other words, clinicians should adjust the standard (i.e. failure of three effort indices) based on the number of validity indices administered rather than using it as a concrete rule.

Strategies Used to Simulate ADHD. Quinn (2003) asked ADHD Simulators what strategies they used to feign ADHD. Most participants reported using multiple strategies. The majority (61%) of participants reported using a strategy of general inattention. Forty-three percent and 17% reported that they used strategies of ignoring visual and auditory stimuli,

40

respectively. Fifty-seven percent intentionally made errors of commission, 35% reported that they made errors of omission, and 9% reported that they responded randomly. The ADHD Simulators also reported using behavioral strategies including: double-clicking the mouse in an attempt to demonstrate hyperactivity (35%), fidgeting (13%), and slowing down response time at the end of the test compared to the beginning (9%). This study suggests that college students asked to simulate ADHD chose to feign symptoms of inattention more often than symptoms of hyperactivity.

Harrison et al. (2007) found the most common strategies used by ADHD Simulators were: completing tasks slowly (31%) or responding quickly and carelessly while completing tasks (20%), attempting to act like someone that they knew who had ADHD (29%), "zoning out" /attending to distracting noises (26%), purposefully choosing incorrect answers, particularly on harder items (23%), and skipping items (23%). Other strategies reported were: being inattentive or disobeying instructions (14%), purposefully selecting items from the CAARS that matched DSM-IV criteria (11%), making their eyes lose focus or only focusing on the middle of the page (9%), reading questions repeatedly (6%), beginning tasks prior to being told (3%), and letting focus wane at the ends of tasks (3%). Frazier and colleagues (2008) presented students with a questionnaire in which they were asked to respond with either a yes or no to various strategies. The most common strategy endorsed was trying to show difficulty paying attention and attempted to appear less intelligent (90%). Other items commonly endorsed include: trying to miss difficult items (87%), responding inconsistently (87%) or responding slowly (77%), pretending to have trouble remembering things (74%), demonstrating difficulty reading or reading more slowly (71%), trying to miss easy items (71%), acting confused (68%) or acting

nervous (55%), pretending to have difficulty organizing information (45%), or difficulty understanding the tests (42%).

Marshall et al. (2010) performed ANCOVAs, covarying FSIQ, to determine if individuals who exhibited non-credible performance differed from credible clinical sample in responses to neuropsychological tests. They wanted to examine whether individuals feigning symptoms of ADHD employed specific strategies. Results indicated that individuals who exhibited noncredible performance attempted to feign specifically on tests that assess attention and processing speed. It also appeared that the manner in which individuals chose to distort their responses was associated with failure of particular measures. For example, individuals that failed two SVTs appeared to be strategic in choosing to fake on measures of memory and sustained attention. Individuals suspected of distorting responses on rating scales tended to feign only on measures of sustained attention, and individuals whose ratings of symptoms differed from psychometrists' ratings did not appear to attempt to distort their responses on any cognitive tests. This study offers unique insight into strategies that individuals use to simulate ADHD symptoms and warrants replications and extension to validate these findings.

Based on this limited literature, it appears that individuals asked to simulate ADHD chose multiple strategies to feign symptoms. Interestingly, Marshall et al. (2010) reported that, in their clinical sample suspected of poor effort, individuals chose specific strategies when completing neuropsychological and effort measures. The study conducted by Marshall et al. (2010) is the first and only study to date that sought to examine strategy use in a clinical sample, and this study highlights the need for more research in this area.

There are few neuropsychological, self-report, or symptom validity measures that offer adequate diagnostic statistics for malingered ADHD in college students. It appears current

neurocognitive validity tests may offer the best balance between sensitivity and specificity, but there is want for better measures. Suhr and colleagues (2011) recently proposed a new validity index for the CAARS that demonstrated adequate psychometric properties but requires further validation before being used clinically. There is a need for development and validation of new measures designed specifically for use in this population.

2. RATIONALE FOR THE PRESENT STUDY

The sheer number of college students seeking psychoeducational evaluations for ADHD symptoms has increased over the past decade (Harrison et al., 2007). There are multiple challenges to assessing adult ADHD including ongoing debate regarding the construct and assessment of adult ADHD. Further, it has become increasingly recognized that there are numerous external incentives to feign ADHD symptoms including the attainment of stimulant medications, which many college students use illicitly (White et al., 2006; Advokat et al., 2010; Rabiner et al., 2009; MacCabe et al., 2006), and academic accommodations that may further enhance performance in college level courses (Sullivan, 2007). Researchers have begun to acknowledge the need for investigation of malingered ADHD in this population. Multiple studies identified a subset of clinical data that was considered non-credible based on effort test failure (Harrison & Edwards, 2010; Suhr et al., 2008; Suhr et al., 2011; Sullivan et al., 2007) or Slick et al. (1999) criteria (Harrison, Rosenblum, & Currie, 2010; Marshall et al., 2010). Disturbingly, these authors found that current self-report measures of ADHD symptoms as well as many neurocognitive tests do not adequately distinguish between individuals feigning ADHD and college students that actually have it. Other studies have used analog ADHD malingerers and have found many current neuropsychological and self-report measures do not adequately distinguish between the simulated ADHD and clinical ADHD (Sollman et al., 2010; Booksh et al., 2010; Tucha et al., 2008). Despite the attention that malingered ADHD as begun to receive, little is known about the construct of malingered ADHD and few new indices have been successful in detecting college students who malinger ADHD.

Most current self-report questionnaires that assess ADHD symptoms are face-valid and easily faked. A recent study examined utility of the MMPI-2 validity scales for the detection of

feigned ADHD and reported that the MMPI-2 has a couple of promising scales among its previously established embedded indices designed to detect feigned psychiatric disorders (Young & Gross, 2011). The MMPI-RF scales, the Fp-r and RC1, also demonstrated potential utility (Harp et al, 2011). Sullivan and colleagues (2007) reported few individuals who failed the WMT failed any of the current PAI validity indices. However, it is suspected that the PAI contains a subset of items that will adequately differentiate individuals that are faking ADHD from individuals that actually meet diagnostic symptom criteria for the disorder.

The purpose of the present study is to empirically develop a scale from the items of the PAI that has both high sensitivity and specificity for detecting feigned ADHD. This study will use a group of college students asked to malinger ADHD symptoms (ADHD Simulators) and clinical data from students that have completed psychoeducational evaluations and were diagnosed with ADHD (Clinical ADHD). In the primary analyses, responses of ADHD Simulators will be compared to the Clinical ADHD group in order to identify items that adequately differentiate the two groups and develop a scale for the PAI to detect malingered ADHD in college students. Performance of ADHD Simulators on PAI and cognitive effort measures will also be compared to clinical groups including: individuals with psychopathology, comorbid ADHD-Psychopathology, individuals with no diagnosis, and individuals suspected of putting forth poor effort. In addition, a college control group and data collected from a community diagnosed sample of individuals prescribed stimulant medications for ADHD symptoms were used.

45

3. RESEARCH QUESTIONS AND HYPOTHESES

3.1. Question 1: Do college students asked to feign ADHD (ADHD Simulators) successfully manipulate self-report measures of psychopathology and childhood symptoms of ADHD?

Hypothesis 1: In keeping with findings that individuals asked to feign ADHD successfully obtain believable scores on self-report measures, it is hypothesized that ADHD Simulators in this study will obtain scores on the WURS, STAI-Y, and PAI that will be significantly higher compared to clinical groups and controls. However, it is expected ADHD Simulators' scores will fall within clinically believable ranges operationalized as t-scores < 80 or raw WURS scores < 55.

3.2. Question 2: Will the ADHD Simulators perform differently from clinical groups and the College Control group on already established PAI validity indices and embedded cognitive measures?

Hypothesis 2**:** As previous research has found that students suspected of feigning ADHD do not perform more poorly on other embedded indices of the PAI, it is expected that the ADHD simulators will not fail the Inconsistency, Infrequency, Negative Impression Management, Malingering Index, or Rogers Discriminant Function scales of the PAI more frequently than the clinical sample. However, based on previous research, it is expected that the Reliable Digit Span and Vocabulary – Digit Span will be more sensitive to malingered ADHD than the PAI validity scales.

3.3. Question 3: Is there a subset of PAI items that differentiate college students that were formally diagnosed with ADHD (Clinical ADHD) from ADHD Simulators?

Hypothesis 3: It is hypothesized ADHD Simulators will endorse significantly more PAI items related to psychopathology compared to the Clinical ADHD group. It is expected that items related to hyperactivity (Mania scale) and poor concentration/confusion (Schizophrenia – Thought Disorder Subscale) will be endorsed at significantly higher rates by individuals

46

attempting to feign ADHD because the items are face valid. Further, it is expected that the items

that differentiate simulated from clinical ADHD groups on a univariate level will be summed

into a scale that adequately differentiates the two groups. The scale will be labeled the newly

developed scale of the PAI.

3.4. Question 4: Will the newly developed scale demonstrate adequate sensitivity and specificity for malingered ADHD when applied to individuals that meet diagnostic criteria for other psychological disorders and comorbid ADHD and Psychopathology?

Hypothesis 4: It is expected the newly developed scale will differentiate simulated

ADHD from other clinical groups, including diagnoses of comorbid psychopathology and

ADHD. As mentioned previously, it is expected that items endorsed significantly more by

ADHD Simulators will be associated with Mania and Schizophrenia subscales, and few college

students seen for psychoeducational evaluation at this clinic are diagnosed with severe mental

illness. Therefore, it is hypothesized ADHD Simulators will select items suggestive of thinking

problems and hyperactivity more than college students diagnosed with psychopathology or

comorbid ADHD and psychopathology.

3.5. Question 5: How will ADHD Simulators compare to individuals from the Suspect Effort Group (i.e. identified among the clinical dataset by virtue of failing two or more WAIS-III or WMS-III embedded validity indices) on the newly developed scale?

Hypothesis 5: It is expected the Suspect Effort group's scale scores on the newly

developed scale will be comparable to the ADHD Simulators.

3.6. Question 6: What strategies will ADHD Simulators use to feign ADHD symptoms on self-report measures of ADHD, psychopathology, and neurocognitive effort measures?

Hypothesis 6: Previous literature suggests that the strategies selected are more consistent

with symptoms of inattention compared to symptoms of hyperactivity. Therefore, the author

expects that the ADHD Simulators in the current study will use more strategies associated with

inattention compared to hyperactivity.

47

4. METHOD

4.1. Participants

Participant Sample. A sample of 310 undergraduate student volunteers from a large university in the southeast United States was recruited by the University's online recruitment system to participate in this study for extra credit in Psychology courses. Of this sample, 73.5% were female, 80% were Caucasian, 7.4% were African American, 5.5% were Hispanic, 3.5% were Asian, and 3.5% identified as another ethnicity. The average age of the undergraduate student volunteers was 19.78 (SD ± 1.47; Range = 18-25). The student volunteers had an average of 13.10 years of education (SD ± 1.09; Range = 12-15) and the estimated Shipley-II FSIQ was 105.6 (SD ± 9.58; Range = 70-130).

Clinical data. This study also employed archival data from two university-affiliated psychology clinics. The majority of archival data (n = 401) was obtained from a psychology clinic affiliated with a large, southern University. This data was supplemented by archival data (n = 40) from a smaller, southern University. Data included in this study was collected from students that have completed psychoeducational evaluations from 2005-2012. In order to be included in this study, individuals had to have completed the Personality Assessment Inventory as part of the neuropsychological or psychoeducational evaluation. Of the clinical sample, 49.9% were female. The average age of individuals from the archival data was 23.23 (SD ± 6.73, Range = 16 - 56). The sample had 14.00 years of education (SD ± 1.9, Range = 12 - 22) and the mean FSIQ for individuals that had completed the WAIS-III (n = 381) was 103.72 (SD ± 12.52, Range = 70 - 151) and for individuals that completed the WAIS-IV (n = 45), FSIQ was 98.53 (SD ± 12.46, Range = 74 - 130).

Graduate students training to be clinical psychologists performed all of the assessments. Diagnoses were made with the supervision of a licensed psychologist and neuropsychologist. Archival data was grouped by type of diagnosis. In this sample, 24.7% were diagnosed with ADHD, 20.4% were diagnosed with psychopathology, 19.5% received no diagnosis, 14.7% were suspected of inadequate effort, 20.7% were diagnosed with comorbid ADHD and psychopathology.

4.2. Materials

Consent Form and Demographic Questionnaire for Study Participants. The analog participants were asked to complete a consent form and a demographic questionnaire. The consent form was the only document that contained identifying information, and was to ensure that students receive extra credit for their participation in this study. Consent forms were kept in a separate folder in a secure location. There was no documentation to associate participants' names with their subject numbers and subsequent test performance. All other information obtained in this study was anonymous. The demographic questionnaire queried information such as age, education, race, sex, and any history of or current psychological or neuropsychological conditions.

Shipley Institute of Living Scale, Second Edition (Shipley-II). The Shipley-II (Shipley, 2008) is divided into verbal and abstraction subtests in order to assess both crystallized knowledge and fluid reasoning abilities. The Verbal subtest consists of 40 vocabulary items. For each item, individuals are presented with a word and asked to choose a synonym for that word among three foils. Individuals are given one point for every correct answer and one additional point for every four items left unanswered. The Abstraction subtest consists of 25 sequence-completion problems that the individual is asked to solve. The items are initially relatively

49

simple but increase in complexity. One point is given to each correct answer. Participants are given 10 minutes to complete the Verbal portion and 10 minutes to complete the Abstraction portion of the test. Standard score for Vocabulary, Abstraction, and an FSIQ estimate are provided.

Personality Assessment Inventory (PAI). The PAI (Morey, 1991) is a self-report measure of psychopathology that is composed of 344 questions. It is computer scored. The computer scoring program provides information about 4 validity scales (Infrequency, Inconsistency, Positive Impression Management, and Negative Impression Management). Four additional validity scales (Cashel Discriminant Function, Defensiveness Index, Malingering Index, and Rogers Discriminant Function) may be calculated by hand. There are 11 clinical scales (subscales) designed to assess pathology: Somatization (conversion, health-related concerns, somatization), Anxiety (affective anxiety, physiological anxiety, cognitive anxiety), Anxiety-related Disorders (obsessive compulsive, phobias, traumatic stress), Depression (affective-depression, physiological-depression, cognitive-depression), Mania (activity level, grandiosity, irritability), Paranoia (hyper-vigilance, resentment, persecution), Schizophrenia (psychotic experiences, social detachment, thought disorder), Borderline Features (identity problems, affective instability, self-harm, negative relationships), Antisocial Features (antisocial behaviors, stimulus seeking, egocentricity), and Substance Abuse Scales (alcohol problems and drug problems). There are also five treatment scales: Aggression (aggressive attitude, verbal aggression, physical aggression), Suicidal Ideation, Stress, Nonsupport, and Treatment Rejection. Finally, the PAI has two interpersonal scales: Dominance and Warmth. In completing the test, individuals are asked to choose one answer from four choices (false, slightly true, mainly true, and very true), providing their own opinion of themselves. Scores are expressed as t-

50

scores. Scores at or above 70t fall in the 96th percentile and are interpreted as clinically significant because they are unlikely to occur in individuals without psychopathology. The reliability of the individual PAI scales is reported as ranging from .85-.94 for adults and .66-.90 for college students (Morey, 1991).

       State-Trait Anxiety Inventory, Form Y (STAI-Y). The STAI was originally published in 1970 and was designed to measure severity of anxiety symptoms (Spielberger, Gorsuch, & Vagg, 1970). The STAI differentiates acute (state) anxiety from chronic (trait) anxiety. The STAI was revised in 1983 (STAI-Y; Spielberger et al., 1983). The revised version was used for the current study. For the STAI-Y1, which measures state anxiety, individuals are asked to read 20 statements and "indicate how you feel right now, that is, at this moment." They respond along a four point Likert-type scale: not at all, somewhat, moderately so, very much so. The STAI-Y2 measures Trait anxiety, and asks individuals to respond to 20 statements by indicating "how you generally feel" on a four-point Likert-type scale. Responses for the STAI-Y2 include: almost never, sometimes, often, almost always. Many items are reverse scored. Scores range from 20-80 for both the state and trait forms. Raw scores are converted to age-normed t-scores. T-scores ranging from 60-64 are considered Borderline Elevated while scores over 65 are considered Significantly Elevated. The STAI-Y state reliability has been reported to range from .40-.54 and the trait reliability has been reported as .86 in two separate studies (Spielberger et al., 1970; Rule & Tarver, 1983).

       Wender Utah Rating Scale (WURS). The WURS (Ward, Wender, & Reimherr, 1993) is a self-report questionnaire that purports to assess childhood symptoms of ADHD. Thus, individuals are asked to provide retrospective ratings of their ADHD symptoms "as a child." There are 61 items scored on a five-point Likert-type scale. Some items are reverse scored, and

51

of the 61 items, only 25 items are included in the raw score that is interpreted. A raw score of ≥36 falls above the recommended cutoff for childhood ADHD symptoms. The authors reported that the WURS adequately discriminated individuals with ADHD from controls. Cronbach's alpha and test-retest reliabilities are above .85 (Weyandt et al., 1995). It should be noted that one study compared the WURS to the PAI and cognitive measures and found that it was more correlated with psychopathology and dysfunctional personality traits than with attentional problems (Hill et al., 2009).

Wechsler Adult Intelligence Scale, Third Edition (WAIS-III). The Digit Span and vocabulary subtests were used in the present study. The Digit Span subtest consists to two parts, Digit Span Forward and Digit Span Backward. Each part has multiple trials, consisting of two strings of numbers that are the same length. In the Forward portion, individuals are required to repeat each string of numbers in the same order without making any mistakes. The trials begin with a sequence of two numbers and subsequent trials increase the length of the string by one unit. Numbers are presented at a pace of one number per second. The subtest is discontinued when an individual incorrectly recalls both strings in a given trial. The Digit Span Backward task is similar to the Forward task except individuals are asked to repeat the string backward, in reverse order presentation. The subtest is discontinued after inaccurate responses to both strings at a given trial. Subscale scores are converted into scaled scores (Mean = 10; SD = 3).

Reliable Digit Span (RDS; Grieffenstien et al., 1994) is an effort index that can be calculated by hand from the WAIS-III Digit Span subtest. Scores are calculated by summing the longest digits forward and the longest digits backward for the last trial in which both responses are correctly answered. Scores below seven are indicative of poor effort.

WAIS-III Vocabulary subtest. In the Vocabulary subtest of the WAIS-III is a measure of expressive language and crystallized knowledge. Individuals are presented with words, one at a time, and asked give a brief verbal definition of the word. There are a total of 33 items scored on a three point Likert-type scale: incorrect (0), partially correct (1), correct (2). The subtest is discontinued after the individual obtains six consecutive, incorrect responses. Scores are reported as scaled scores (Mean = 10; SD = 3). The Vocabulary subtest is required to complete the Vocabulary-Digit Span (VDS) embedded effort index (Mittenberg et al, 1995). In this index, the Digit span score is subtracted from the Vocabulary score. A recommended cutoff score of greater than or equal to six has been proposed to detect poor effort.

Participant Effort-Rating Scale. Following the conclusion of testing, participants in the ADHD Simulator group were asked to complete an effort rating scale. Participants were asked to indicate on a Likert-type Scale of 1 to 5 (0 = not at all; 3 = somewhat; 5 = very much so) how much effort the put forth in performing as someone with ADHD and how successful they think they were at performing at that level without being caught. A final multiple-choice question assessed strategies that analog participants used to fake ADHD. This question will list strategies that Harrison et al (2007) reported in their study.

Consent Form for Clinical Samples. When individuals presented for a psychoeducational or neuropsychological evaluation, they were initially given a consent form. The consent form was read to them, and they were allowed to ask any questions. During the consent, they were asked whether their information may be used anonymously in research. They indicated their consent on the form my marking Y for "yes" (information may be used) or N for "no" (information may not be used for research). They were informed their participation is voluntary and declining participation would not have any negative impact on their testing.

4.3. Procedure

Prospective undergraduate sample. Participants were randomly assigned to the College Control condition or the ADHD Simulator condition. The ADHD Simulators were randomly divided into two groups. One group was used for development of the scale (n = 87) and a hold-out sample was used for cross-validation of the scale (n = 106). The ADHD Simulators were provided with the following scenario:

> Your roommate has been diagnosed with ADHD. She had trouble with classes, but then was given some medication for ADHD, and now does well. She even got a couple of A's recently, and has more time to socialize because studying is not as hard! During your midterms, you decided to try your roommate's medication, and ended up surprising yourself with how much easier things went. You may think that you have undiagnosed ADHD, so you "Google" the disorder to learn more about it. On the following pages are some of the things that you find.
>
> When you are done reviewing these materials, please use the colored paper to jot down symptoms that will help you remember how to fake on the tests you will be given. Tell the examiner when you are done.
>
> Please take the following tests as if you are trying to convince someone that you have ADHD. It is not necessary for you to try to act like you have ADHD; you only need to respond to the test items as if you do. Remember, you are trying to fake ADHD as a college student, so you must perform at least as well as someone who enrolled in a university would. Also, you want to respond in a way that would make it likely to get diagnosed with ADHD but you don't want to overplay the part to avoid being detected as a faker (modified from Sollman et al., 2010).

The College Control group received instructions to perform to the best of their ability (see instructions below). Data from individuals that completed the Control scenario was divided into two groups. One control group was used for during the initial scale development process (n = 45) and the other group (n = 43) was a hold-out sample used in cross-validation of the scale. The control scenario was balanced for length and type of task; however, participants in the control group were instructed to respond to the best of their ability:

54

Your roommate was having trouble with some of her classes. She went to a psychologist for testing. After the evaluation, she tells you that she has been diagnosed with a learning disability, which explains why she had trouble with classes. Specifically, she was diagnosed with a mathematics disorder. She expresses a lot of concern because she had difficulty understanding her diagnosis. You decided to do some research to help explain it to her. You "Google" the disorder to learn more about it. On the following pages are some of the things that you find.

When you are done reviewing these materials, please use the colored paper to jot down symptoms that will help you remember how to explain it to her. Tell the examiner when you are done.

Now, as part of this study, you will be asked to complete a few of the same tests that she took. Please take the following tests giving your best effort. It is necessary for you to try to as hard as you can in order to advance the research that we are conducting. Remember, you are trying to take these tests as a college student, smart enough to be accepted and enrolled at a large university. You are taking these tests because you are responsible and would like to get credit in you Psychology classes. Please sit quietly and fill out the following forms, answering as best as you can.

A group of individuals diagnosed with ADHD by a physician was also recruited for the prospective study in order to increase with number of individuals diagnosed with ADHD in the dataset. In order to be eligible to participate in the Prospective ADHD group, individuals had to bring proof of their diagnosis in the form of a diagnostic report, proof of prescription stimulant or non-stimulant medication approved for the treatment of ADHD with their name on it, etc. These individuals had no obvious external incentive to distort their responses. The instructions given to this group were not balanced for length nor were any additional materials provided. This was done in order to avoid response bias by priming them. They were provided with further rationale for why they were asked to present documentation of their disorder and asked to respond honestly. These individuals received the following scenario:

You have been diagnosed with ADHD. In order to participate in this study, you have to provide documentation of a formal ADHD

diagnosis. Unfortunately, some individuals attempt to fake ADHD.
We are conducting this experiment in order to be better able to detect
those who fake the disorder from those who actually have the
disorder. Your role is important because we will compare individuals
that fake to your data in order to be better able to detect faking. Please
take the following tests giving your best effort. We will use your data
as a person who has ADHD to compare to people that have been
asked to fake ADHD.

All test materials were administered to both groups using the different, group-specific

instructions. Trained undergraduate research assistants administered all measures. The consent

form, demographic questionnaire, Shipley-II, instructions, and self-report questionnaires were

group-administered. The maximum group size was eight participants per group. The WAIS-III

Vocabulary and Digit Span subtests were administered individually, in a separate room, by a

second trained undergraduate research assistant.

Each participant was provided with a copy of the consent form and asked to follow along

as it was read aloud. Afterwards, students were given a chance to have questions and concerns

addressed before their signatures were collected. Consent forms were stored in a separate folder,

in a separate location so that individuals' data could not be identified. After informed consent

was obtained, participants were assigned an identification number to ensure anonymity. All

individuals were asked to complete a demographic questionnaire, and the Shipley-II while giving

their best effort. After completing the Shipley-II, each group was asked to read and follow the

scenario that was provided to them with instructions for approaching the remaining tests.

As part of the scenario, students were presented with a pseudowebpage format adapted

from Sollman et al. (2010) that provided information about ADHD (materials to be provided in

meeting). This information was obtained by typing the keywords "Adult ADHD" into Google.

The top five webpages that were found were copied to create pseudowebpages. The ADHD

simulators were asked to take notes from the information provided on the pseudowebpages. Similarly, Control participants were asked to read a scenario that encouraged them to put forth their best effort. In order to balance the scenario and instructions with those received by the ADHD Simulators, the control participants were provided with pseudowebpages about mathematics disorder and asked to read and take notes (to be provided in meeting).

After the demographic information, Shipley, and scenarios and instructions were completed, all subjects were handed the PAI, STAI-Y, and WURS. The first subject (lowest subject number) was escorted to another examination room by a second trained research assistant who administered the WAIS-III Vocabulary and Digit Span subtests. When he/she returned and resumed filling out questionnaires, the next participant was asked to stop filling out questionnaires and go to the other examination room to complete the WAIS-III measures. This process continued until all subjects were tested individually. Participants were instructed in this procedure immediately before self-report questionnaires are administered. They were asked to be as quiet and least disruptive as possible. Administration time took approximately 2.5 hours per group. After completing the evaluation, participants were debriefed on the purposes and utility of the study. Analog ADHD simulators were asked to complete the Participant Effort Rating Scale. All protocols were scored by an undergraduate research assistant and checked by a Master's level examiner (MM).

Clinical Data. All data was anonymously entered into a database so that no identifying information was recorded. This study was approved by the university's Institutional Review Boards (LSU IRB # 3209; Appendix A).

During psychoeducational evaluations, individuals are asked to complete the following neuropsychological measures: WAIS, WMS, CPT, Trail Making Test A and Trail Making Test

B, and the Woodcock Johnson, Third Edition. They are also asked to complete self-report questionnaires including: the Beck Depression Inventory, PAI, STAI, and WURS. All measures were administered according to their standardized protocols. Test order was not strictly controlled during the course of these evaluations. Evaluations typically take six to eight hours to complete. Students that were prescribed stimulant medications had not taken their medications on the day of the evaluation.

Trained, undergraduate research assistants entered all data into a database at a later date. The majority of items were double-checked one time. Discrepancies were checked a third time by a master's level doctoral candidate (MM) and either resolved or rejected. For the purposes of illustrating data integrity, a sub-sample of the data (data from 53 individuals) was analyzed. The 53 individuals produced 29,921 data points. Of these, 343 (1.2%) errors were found upon the first check. Of these 343 discrepancies, 85 (24%) were overturned when checked by a maser's level clinician. It should be noted, most (80) of these discrepancies were associated with the STAI-Y form as students confused side Y-1 with side Y-2, incorrectly coding 40 items both times.

The data from individuals that exhibited adequate effort on embedded effort measures was categorized by type of diagnosis. The Clinical ADHD group was randomly separated into two groups. One group (n = 85) was used during the scale development phase of the study, and the other group (n = 24) was used during the cross-validation. The No Diagnosis group was also randomly divided into two groups to be used during the scale development (n = 45) and cross validation phases (n = 41), respectively. Other clinical groups included individuals diagnosed with psychopathology (Psychopathology; n = 123) and Comorbid ADHD-Psychopathology (n = 58).

58

During psychoeducational evaluations, effort is measured by six embedded effort indices of the WAIS-III and WMS-III. The WAIS-III embedded indices include the Reliable Digit Span, Voc-Digit Span, and the Mittenberg Discriminant Function. The WMS-III embedded indices include the Logical Memory Rarely Missed Items Index, Faces-I raw score, and Auditory Recognition Delayed Raw Score. As part of the study, individuals that failed at two or more effort indices were classified as the Suspect Effort group (n = 65) in order to determine whether individuals that exhibited suboptimal effort differ from those who do not fail any effort indices.

4.4. Statistical Analyses

A power analysis was run using G*Power 3.1.0 (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). Due to the large number of items in the Personality Assessment Inventory (344), a conservative statistical approach was used for item identification. The alpha level accepted as statistically significant was set to < .01 to minimize significance due to chance. The first power analysis was conducted to determine the power of detecting a medium effect size for t-tests. It revealed that power for detecting a medium effect (Cohen's d = .5) with an alpha level at < .05 would be .98 using a t-test with two equal groups of 100 people each. Power was also examined for DFA analysis proposed in Phases 2 and 3. When Multivariate Analysis of Variance (MANOVA; mathematically equivalent to a Discriminant Function) tests are used to detect a medium effect size ($f^2 = .20$) in a total sample of 200 and alpha = 0.05, the power is .84. SPSS 18.0 was used in all comparisons between and within groups.

Comparing the analog sample to the archival data sample: **Demographic Data.**

Univariate analyses of variance (ANOVA) were used to determine significant differences in age, and years of education between groups. In addition, the study participants' FSIQ standard scores were compared to the clinical samples' FSIQ standard scores using an ANOVA. Type of IQ test

given (Shipley-II, WAIS-III, WAIS-IV) was dummy coded and used as a covariate to examine whether differences are associated with psychometric properties of the tests. Pearson's Chi-square tests were used to determine differences in race/ethnicity and gender for the analog ADHD group and other groups. Subsequent Pearson's Chi-square tests examined differences between the subgroups of the prospective and archival data. Significant demographic variables were placed into a multiple regression analysis in order to determine effect on each of the PAI clinical and validity scales as well as the WURS, STAI-Y, and Cognitive Effort measures (RDS and VDS). Next, the ADHD Simulator group was compared specifically to the Clinical ADHD group using the methods described above.

   Diagnostic Statistics. Diagnostic statistics, also known as operating characteristics, are used to determine the accuracy of a specific test in determining whether are not individuals have specified disorders (Larrabee & Berry, 2007). Diagnostic statistics relevant to determining the utility of this scale include base rates, sensitivity, specificity, positive predictive power (PPP), and negative predictive power (NPP) (Gouvier, Hayes, & Smiroldo, 1998). Base rates are defined as the prevalence of a given condition in the population and are crucial for determining accuracy of diagnostic classification (Gouvier, 2001). Sensitivity reflects the probability that an individual who has a particular disorder will be diagnosed as such. It is calculated as the ratio of true positives divided by the sum of true positives and false negatives. Specificity refers to the probability that a person who does not have a specified disorder will obtain a negative test result. Specificity is calculated by dividing true negatives by the sum of true negatives and false positives (Larrabee & Berry, 2007). Positive predictive power is the probability that a disorder will be diagnosed given a positive test finding, and NPP is the probability of an individual

60

without the disorder being classified as such, given a negative test finding. Both PPP and NPP are influenced by the base rate of a given disorder and will change as the base rate changes.

Question 1: Do college students asked to feign ADHD (ADHD Simulators) successfully manipulate self-report measures of psychopathology and childhood symptoms of ADHD? A univariate multivariate analysis of covariance (MANCOVA) was used. Independent variable was group (five clinical groups: Clinical ADHD, No Diagnosis, Suspect Effort, Comorbid ADHD-Psychopathology, Psychopathology and three prospective groups: Prospective ADHD, Control, and ADHD Simulators). The dependent variables were WURS total raw score, and STAI-S and STAI-T T-scores. A second MANCOVA was performed with the 8 groups as independent variables and the t-scores of the 11 PAI clinical scales as dependent variables. Age, gender, and FSIQ were used as covariates in both analyses. Planned post-hoc analyses included a series of univariate ANCOVAs in order to compare ADHD Simulators to the other Clinical and Prospective groups. Age, gender, and FSIQ were used as covariates.

Question 2: Will the ADHD Simulators perform differently from clinical groups and the College Control group on already established PAI validity indices and embedded cognitive measures? First, a univariate MANCOVA was run with Group (Clinical ADHD, No Diagnosis, Psychopathology, Comorbid ADHD-Psychopathology, Suspect Effort, Prospective ADHD group, ADHD Simulators, and College Controls) as the independent variable and validity indices (RDS, V-DS, INF, INC, NIM, RDF, MAL) as the dependent variables to determine whether the groups differ in their scores on the validity indices. Age, gender, and FSIQ were used as covariates. Bonferroni correction of P < .01 was used to potentially account for large sample size and the possibility of obtaining significant results due to chance. Univariate

ANCOVAs were run as planned post-hoc analyses to compare ADHD Simulators to the other Clinical and Prospective groups. Age, gender, and FSIQ were used as covariates.

Next, failure of the PAI Validity indices was coded dichotomously (Pass or Fail) using the cutoff scores proposed by Morey (2003). The percentages of individuals in each group that failed PAI indices were examined using Chi-square analyses. Groups included in this analysis were the Clinical ADHD, No Diagnosis, Psychopathology, Comorbid ADHD-Psychopathology, and Suspect Effort groups as well as the Prospective ADHD group, the ADHD Simulators, and the College Controls.

Question 3: Is there a subset of PAI items that differentiate college students that were formally diagnosed with ADHD (Clinical ADHD) from ADHD Simulators? The PAI has 344 items, so in order to reduce significant findings due to chance, alpha was set to < .01. This study has adequate power to detect medium effect sizes with an alpha of .01. It was suspected that a large number of items may be significant. The purpose of this was to narrow the item pool of the PAI to 40 items for further analysis. This was done by identification and ranking of significance and effect size for all items. Univariate Analysis of Covariances (ANCOVAs) were run for each PAI item using ADHD Simulators and the Clinical ADHD group. Gender and age were covaried because the ADHD Simulator and Clinical ADHD groups differed on these items. Years of education were not covaried because it is highly correlated with age. Items were rank-ordered based on effect size (Partial eta-squared). Initially, the 40 items with highest effect sizes were retained for further consideration.

Next, comparisons were made between the College Control group and the Clinical No Diagnosis group in order to determine whether a significant number of previously identified items were confounded by condition. In order to do so, univariate ANCOVAs comparing

College Controls and Clinical No Diagnosis groups were run on the 40 items retained in previously. Age and years of education were employed as covariates. This phase of the study was concerned only with items that were highly confounded, so a Partial eta-squared of .8 or greater was used as criterion for exclusion. Items that differ significantly for these two groups removed and replaced by the items, next in line, that yielded the highest effect sizes for differences between ADHD simulators and Clinical ADHD.

In order to further narrow the item pool into a subset of items that, when weighted and summed, best differentiates the ADHD Simulators from the Clinical ADHD sample, a Discriminant Function Analysis (DFA) was run with ADHD Simulator and Clinical ADHD groups as dependent variables, and the 40 items identified in the preceding analyses as entry variables. Structural coefficients were used to determine what items best differentiate the two groups.

As the primary goal of this study was to develop a clinician-friendly scale that differentiates ADHD Simulators and individual who meet diagnostic criteria for ADHD, a unit weighting strategy was employed. The items identified by the DFA were summed into the newly developed scale of the PAI. The total score of the newly developed scale was entered into a logistic regression with the ADHD Simulators and the Clinical ADHD sample as independent variables. The logistic regression was run in order to determine whether the scale adequately predicted group membership. ROC curves were employed to determine a cutoff score that provides the best trade-off between sensitivity and specificity of ADHD Simulators and the Clinical ADHD group.

Question 4: Will the newly developed scale demonstrate adequate sensitivity and specificity for malingered ADHD when applied to individuals that meet diagnostic criteria

for other psychological disorders and comorbid ADHD and Psychopathology? The newly developed scale was cross-validated using a hold-out sample of ADHD Simulators, College Controls, Clinical ADHD, and No Diagnosis groups as well as the Clinical Psychopathology, and Comorbid ADHD-Psychopathology and the Prospective ADHD group. First, scores for the newly developed scale were calculated for all individuals. Next, sensitivities for ADHD Simulators and specificities for other clinical and control groups were calculated.

      Question 5: How will ADHD Simulators compare to individuals from the Suspect Effort Group (i.e. identified among the clinical dataset by virtue of failing two or more WAIS-III or WMS-III embedded validity indices) on the newly developed scale? Scores on the newly developed scale were calculated for the Suspect Effort Group. Scores were coded dichotomously as Pass or Fail in order to determine whether the Suspect Effort Group failed the newly developed scale as frequently as the ADHD Simulators. The dichotomous scores were compared the ADHD Simulator Group using Chi-square analysis. Sensitivity was calculated for the Suspect Effort group.

      Question 6: What strategies will ADHD Simulators use to feign ADHD symptoms on self-report measures of ADHD, psychopathology, and neurocognitive effort measures? The ADHD Simulators were provided with a list of strategies commonly used by individuals asked to feign ADHD. Percentages of individuals that endorsed using each strategy were calculated and reported.

5. RESULTS

5.1. Demographic Information

The demographic variables from the prospective data (ADHD Simulators, College Controls, Prospective ADHD) were compared to archival data (Clinical ADHD, No Diagnosis, Psychopathology, Comorbid ADHD-Psychopathology, and Suspect Effort). Significant differences were noted in age and gender. The prospective group was significantly younger than the clinical group (F $(1, 750)$ = 78.73, $p < 0.001$; η^2 = .095). The prospective group obtained higher FSIQ scores compared to the clinical group; however it should be noted that while the difference reached statistical significance, the effect size was small (F $(1, 726)$ = 2.51, $p < 0.05$; η^2 = .02) with the differences in FSIQ of 6.77 points between the control group and the psychopathology group. Type of IQ test administered was a significant covariate (F $(1, 726)$ = 8.67, $p < 0.01$; η^2 = .01). The clinical group had significantly more years of education compared to the prospective group (F $(1, 743)$ = 52.004, $p = 0.001$; η^2 = .065). In regards to gender, there were a significantly greater proportion of females to males in the prospective group compared to the archival data (χ^2 (1) = 22.74; $p < .001$). There was a trend toward significance for race (χ^2 (1) = 14.7; $p < .04$); however, it did not reach a priori determined significance level of $p < .01$. Means and Standard Deviations can be found in Table 3.

A series of multiple regression analyses were run in order to further examine the effects of age, gender, and FSIQ on the PAI clinical and validity scales. The ADHD Simulator group was left out of the present analyses as their responses to the variables in question are intentionally distorted. Older age was associated with higher WURS ($\beta = 0.13$), STAI- State ($\beta = 0.20$), and STAI-Trait ($\beta = 0.23$) scores. There was a significant effect of age ($\beta = 0.19$) and

Table 3. Mean (SD) demographic variables.

Group	n	% male	% Caucasian	Age	Ed	FSIQ
Clinical	*441*	*49.7*	*81.4*	*23.34 (6.83)*	*13.96 (1.80)*	*103.3 (12.6)*
ADHD	109	56.9	81.7	21.82 (4.05)	13.85 (1.68)	106.2 (13.5)
No Diagnosis	86	54.7	74.4	22.84 (7.01)	13.73 (2.01)	102.4 (11.1)
PSYC	123	39.0	87.8	24.83 (8.59)	13.90 (1.88)	101.1 (12.2)
Comorbid ADHD & PSYC	58	50.0	81.0	23.21 (6.15)	14.40 (1.64)	103.5 (11.3)
Suspect Effort	65	50.8	78.5	23.91 (6.84)	14.13 (1.68)	103.1 (13.8)
Prospective	*310*	*26.5*	*80.0*	*19.76 (1.46)*	*13.09 (1.09)*	*105.9 (9.6)*
ADHD Simulators	193	25.9	77.2	19.77 (1.47)	13.12 (1.11)	105.9 (9.5)
Control	88	28.4	81.8	19.50 (1.15)	12.92 (1.00)	107.0 (9.4)
Prospective ADHD	29	24.1	93.1	20.48 (1.94)	13.41 (1.60)	103.1 (9.5)

FSIQ (β = -0.11) on the Somatization scale with older individuals and individuals with lower FSIQ scores obtaining higher scores. The Anxiety scale was significantly affected by FSIQ (β = -0.11) and gender (β = 0.17) with females and individuals with lower FSIQ scores obtaining higher scores. The PAI Anxiety Related Disorders scale was affected by age (older; β = 0.13), FSIQ (lower; β = -0.13), and gender (female; 0.20). The older individuals (β = 0.11) and lower FSIQ (β =- 0.12) were associated with higher Depression scale scores. Lower FSIQ was associated with high Paranoia scale scores (β = -0.13), and older age was associated with higher Schizophrenia scale scores. Being male was related to higher scores on the Antisocial Features (β = -0.28), Alcohol (β = -0.12), and Drug scales (β = -0.19). There were no significant effects on the Mania or Borderline Features scales.

In regards to the validity indices, lower FSIQ was associated with the Negative Impression Management (β = -0.11), Inconsistency (β = -0.21), Infrequency (β = -0.26), and Malingering Indices (β = -0.16). In addition, younger age was associated with higher scores on the Infrequency scale (β = -0.11). Being female was associated with lower scores on the Positive Impression Management scale (β = -0.14). The higher scores on the Rogers Discriminant Function Index were associated with younger age (β = -0.21), lower FSIQ (β = -0.16), and being male (β = -0.12).

Univariate ANOVAs were performed to determine whether the ADHD Simulators differed significantly from the Clinical ADHD group as the primary analyses involve these two groups. There were no significant differences in FSIQ (F (1, 297) = 0.046, p = ns; η^2 = .001) or race (χ^2 (5) = 6.75; p = 0.24) between the two groups. The Clinical and Simulated ADHD groups differed by 0.79 IQ points. The clinical ADHD group was significantly older (F (1, 300) = 37.18, p < 0.001; η^2 = .110) than the ADHD Simulators. Also, the Clinical ADHD group completed

13.81 years of education while the ADHD Simulators completed an average of 13.12 years of education. This difference was significant (F (1, 300) = 18.35, p < 0.001; η^2 = .058). In addition, there were a significantly greater proportion of males to females (75% male) in the Clinical ADHD group (χ^2 (1) = 28.64; p < 0.001) while the ADHD Simulator group had a higher percentage of females (65%).

5.2. Hypothesis 1: It is expected that the ADHD Simulators will obtain significantly higher scores on the PAI clinical scales, the WURS, and the STAI-Y compared to clinical groups and controls but that their scores will not exceed a t-score of 80 (raw score > 55 on WURS).

A MANCOVA was performed with the group (Clinical ADHD, No Diagnosis, Suspect Effort, Psychopathology Comorbid ADHD-Psychopathology, and Prospective ADHD, Control, and ADHD Simulators) as independent variable and the WURS Total Score, STAI-S t-score, and STAI-T t-score as dependent variables. Means and standard deviations are presented in Figure 1. Age, gender, and FSIQ were used as covariates. Using Wilks' criterion, the combined dependent variables were significantly related to the combined covariates (F (21, 619) = 22.01, p < 0.001; η^2 = .20); however, upon examining individual covariates, only age was significant (F (3, 637) = 5.32, p < 0.001; η^2 = .03). Test of between subjects indicated significant group differences for the WURS Total score (F (7, 629) = 67.76, p < 0.001; η^2 = .43), STAI-S T-score (F (7, 629) = 31.14, p < 0.001; η^2 = .26), and STAI-T T-score (F (7, 629) = 40.64, p < 0.001; η^2 = .31). The association between age and STAI-S was not significant, but the associations between age and WURS Total score (η^2 = .01) and STAI-T were small (η^2 = .02). Univariate ANCOVAs were used to determine whether the ADHD Simulator group differed from the other Clinical and Prospective groups. Age, gender, and FSIQ were used as covariates. In regards to the WURS Total score, the ANCOVA revealed that the ADHD Simulator group obtained significantly higher scores compared to
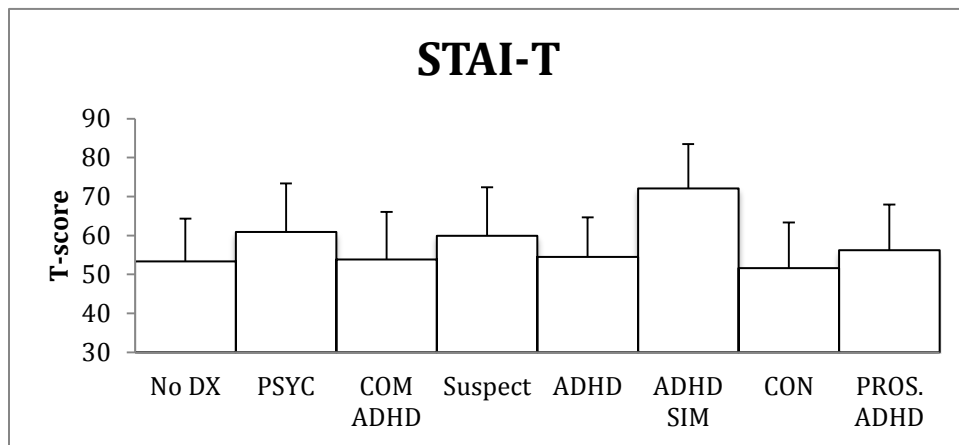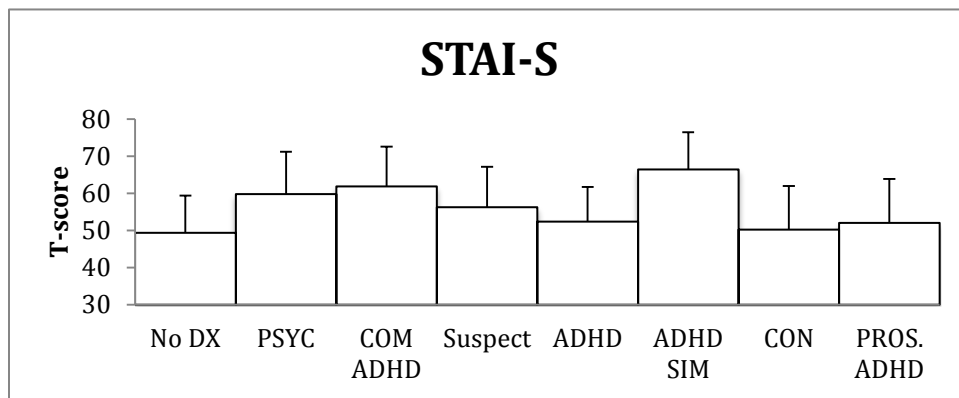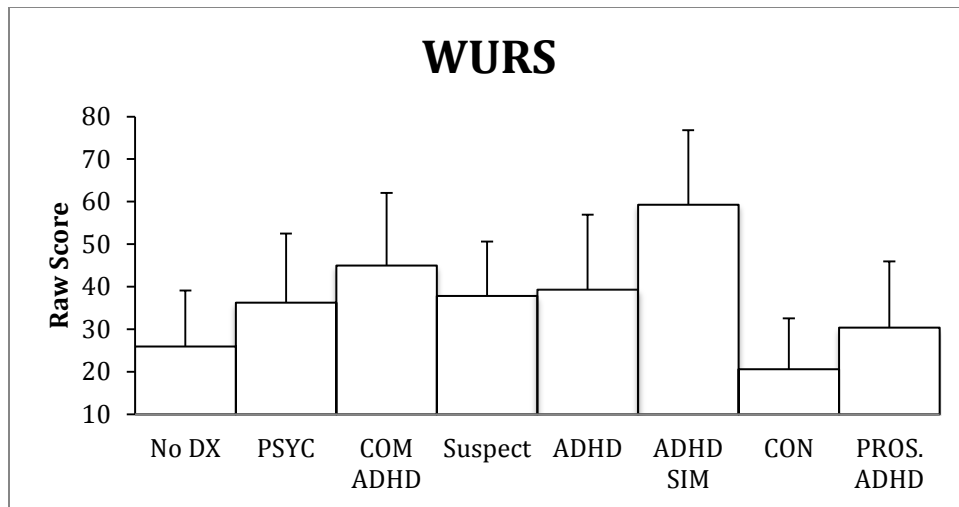
68

Figure 1. ADHD Simulators obtained higher scores on the WURS Raw score and STAI-State and STAI-Trait T-scores compared to all other groups: No Diagnosis (No DX), Psychopathology (PSYC), Comorbid ADHD-Psychopathology (COM ADHD), Suspect Effort (Suspect), Clinical ADHD (ADHD), ADHD Simulators (ADHD SIM), Controls (CON) and Prospective ADHD group (PROS. ADHD).

the Clinical ADHD (F (1, 287) = 81.21, $p < 0.001$; η^2 = .22), No Diagnosis (F (1, 247) = 164.35, $p < 0.001$; η^2 = .40), Psychopathology (F (1, 286) = 106.80, $p < 0.001$; η^2 = .27), Comorbid ADHD-Psychopathology (F (1, 245) = 26.41, $p < 0.001$; η^2 = .10), Suspect Effort (F (1, 252) = 69.57, $p < 0.001$; η^2 = .22), College Control (F (1, 278) = 337.36, $p < 0.001$; η^2 = .55), and Prospective ADHD (F (1, 219) = 67.30, $p < 0.001$; η^2 = .24) groups. The ADHD Simulator Group obtained a mean raw score of 58.85 (SD = 17.54), which was only slightly higher than hypothesized. On the STAI-S, the ADHD Simulator group obtained significantly higher scores (66.45 ± 11.56) than the Clinical ADHD (F (1, 275) = 88.00, $p < 0.001$; η^2 = .24), No Diagnosis (F (1, 250) = 108.63, $p < 0.001$; η^2 = .30), Psychopathology (F (1, 289) = 36.95, $p < 0.001$; η^2 = .11), Comorbid ADHD-Psychopathology (F (1, 239) = 8.92, $p < 0.01$; η^2 = .04), Suspect Effort (F (1, 251) = 29.96, $p < 0.001$; η^2 = .11), College Control (F (1, 272) = 117.49, $p < 0.001$; η^2 = .30), and Prospective ADHD (F (1, 216) = 37.69, $p < 0.001$; η^2 = .15) groups. The ADHD Simulators obtained significantly higher STAI-T scores (71.96 ± 11.38) compared to the Clinical ADHD (F (1, 272) = 130.79, $p < 0.001$; η^2 = .33), No Diagnosis (F (1, 247) = 147.38, $p < 0.001$; η^2 = .37), Psychopathology (F (1, 286) = 62.70, $p < 0.001$; η^2 = .18), Comorbid ADHD-Psychopathology (F (1, 235) = 24.32, $p < 0.001$; η^2 = .09), Suspect Effort (F (1, 247) = 49.92, $p < 0.001$; η^2 = .17), College Control (F (1, 270) = 178.12, $p < 0.001$; η^2 = .40), and Prospective ADHD (F (1, 213) = 45.27, $p < 0.001$; η^2 = .18) groups.

Regarding PAI scales, the MANCOVA was significant (F (77, 652) = 7.39, $p < 0.001$; η^2 = .10). A relatively large association was found with gender (η^2 = .14) and lesser associations were found with age (η^2 = .05) and FSIQ (η^2 = .06). Examination of between-subjects effects revealed gender was significantly associated with the Anxiety, Anxiety Related Disorders, Antisocial Behaviors, and Drug subscales at the $p < 0.01$ level. Age was significantly associated

70

with the Somatization and Anxiety Related Disorders subscales and FSIQ was significantly

associated with the Anxiety Related Disorders and Paranoid subscales. The ANCOVAS

examining group performances on the clinical scales of the PAI were significant ($p < 0.001$). The

ADHD Simulators did not obtain mean t-score above 80 on any of the PAI clinical scales. Effect

sizes from the between subjects MANCOVA are presented in Table 4. Figure 2 illustrates

groups' score on the PAI Clinical Scales.

Table 4. ANCOVA statistics for comparison of groups on PAI clinical Scales.

	F-ratio	p-value	Partial eta^2
SOM	36.84	<0.001	0.263
ANX	43.97	<0.001	0.298
ARD	29.68	<0.001	0.223
DEP	40.82	<0.001	0.283
MAN	21.90	<0.001	0.175
PAR	31.23	<0.001	0.232
SCZ	41.10	<0.001	0.284
BOR	42.44	<0.001	0.291
ANT	46.58	<0.001	0.311
ALC	23.57	<0.001	0.186
DRUG	28.69	<0.001	0.217

Note: Somatization (SOM); Anxiety (ANX); Anxiety Related Disorders (ARD); Depression
(DEP); Mania (MAN); Paranoia (PAR); Schizophrenia (SCZ); Borderline Features (BOR);
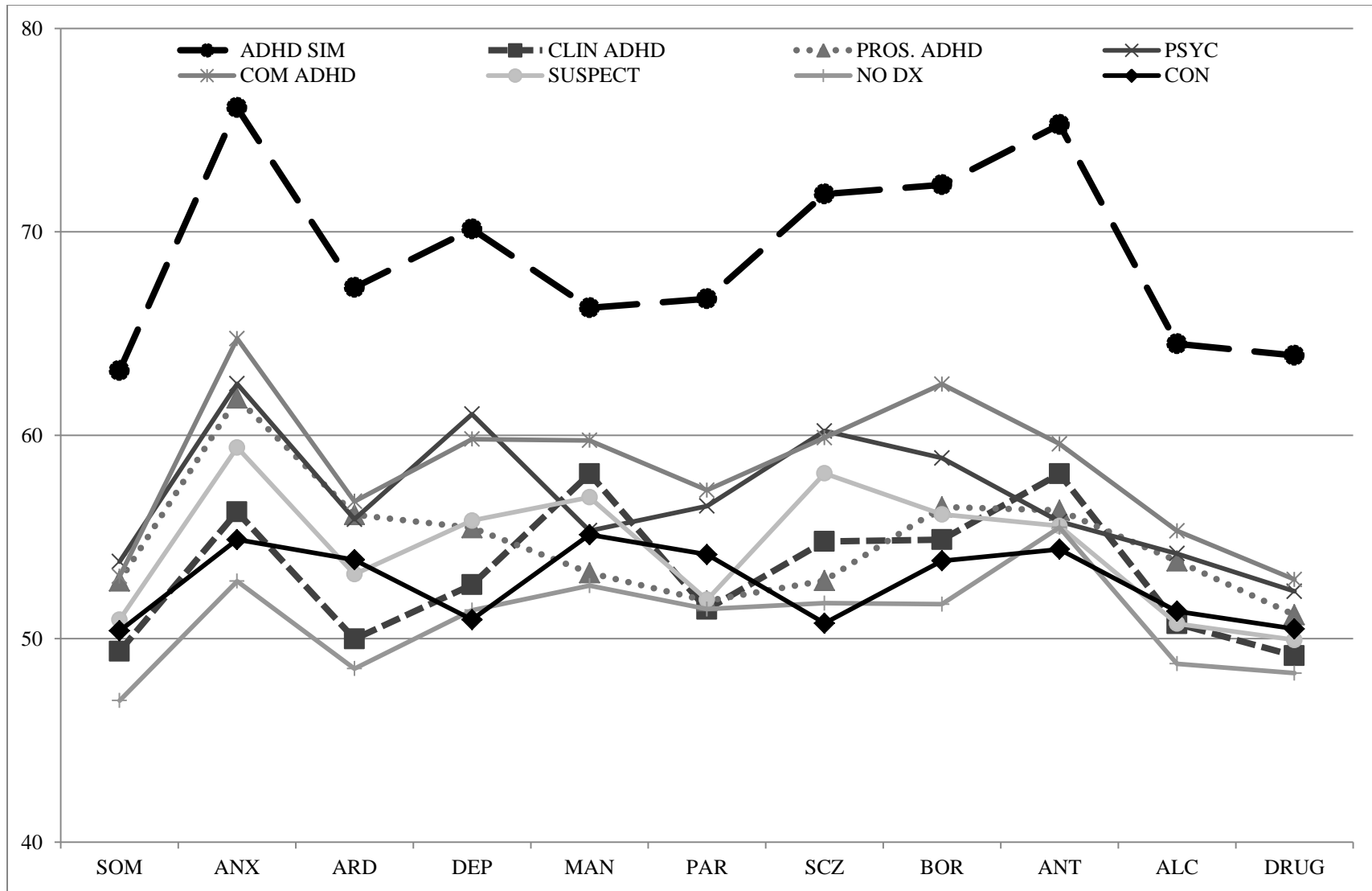Antisocial Features (ANT); Alcohol (ALC); Drug (DRG).

71

Figure 2. Comparison of the five clinical groups (Clinical ADHD, No Diagnosis, Psychopathology, Comorbid ADHD-Psychopathology, and Suspect Effort) and the three Prospective groups (ADHD Simulators, College Controls, and Prospective ADHD) on the clinical scales of the PAI.

Examination of Univariate ANCOVA analyses comparing the ADHD Simulators to all other groups revealed the ADHD Simulators obtained significantly higher scores than the other prospective and clinical groups on all of the clinical scales (Somatization, Anxiety, Anxiety Related Disorders, Depression, Mania, Paranoid, Schizophrenia, Borderline Behaviors, and Antisocial Behaviors). When the ADHD Simulator group was compared to the Clinical ADHD, No Diagnosis, Suspect Effort, Control, and Prospective ADHD groups, effect sizes for differences in scores were generally large $\eta^2 > 0.10$ with the largest effect size being $\eta^2 = .41$ for differences between the ADHD Simulators and the No Diagnosis group on the Somatization scale. While all comparisons between ADHD Simulators and the Comorbid ADHD-Psychopathology groups were significant, effect sizes were smaller with the largest effect size being $\eta^2 = .13$ for differences on the Schizophrenia scale, but differences on all other scales were $\eta^2 < 0.10$ or less. F-ratios, p-values, and effect sizes are included in Table 5.

5.3. Hypothesis 2: It is expected that ADHD simulators will obtain scores on cognitive effort measures that are similar to the Suspect Effort group but significantly poorer than other groups. It is also expected ADHD Simulator's scores on self-report validity indices will be comparable to the other clinical groups.

Percentages of individuals who failed each effort index are presented in Figure 3. A MANCOVA examining differences in PAI validity indices revealed significant associations between the combined covariates and the combined dependent variables when using Wilks' criterion (F (42, 687) = 8.45, $p < 0.001$; $\eta^2 = 0.08$). Age ($\eta^2 = 0.06$), gender ($\eta^2 = 0.05$), and FSIQ ($\eta^2 = .07$) were all modestly associated with the dependent variables. Between-subjects tests indicated significant main effects for each validity index (Table 6).

A series of Univariate ANCOVAs were used as planned orthogonal contrasts. They revealed that ADHD Simulators' scores on the RDS were significantly higher than the Suspect Effort Group but comparable to all other clinical and the prospective groups, and scores on the

73

Table 5. ANCOVA Analyses for Post-Hoc Comparisons of PAI clinical scale scores for ADHD Simulators compared to other Clinical and Prospective Groups.

	Clinical ADHD		No Diagnosis		Psychopathology		Comorbid ADHD-Psychopathology		Suspect Effort		College Controls		Prospective ADHD	
	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2
SOM	87.39	0.23	110.52	0.29	61.30	0.17	25.06	0.09	40.63	0.14	74.78	0.21	16.44	0.07
ANX	145.74	0.33	188.73	0.41	62.02	0.17	26.52	0.10	62.96	0.20	135.78	0.33	27.20	0.11
ARD	98.18	0.25	101.43	0.27	47.10	0.13	19.44	0.07	45.06	0.15	60.16	0.18	14.34	0.06
DEP	116.94	0.28	129.93	0.32	25.97	0.08	24.48	0.09	47.92	0.16	120.01	0.30	32.07	0.13
MAN	35.93	0.11	85.16	0.24	59.43	0.16	15.35	0.06	85.16	0.24	58.89	0.18	33.79	0.13
PAR	94.97	0.24	84.82	0.24	39.86	0.12	19.85	0.07	67.13	0.21	57.28	0.17	35.43	0.14
SCZ	90.11	0.23	126.39	0.32	40.93	0.12	21.98	0.08	31.33	0.11	126.08	0.31	41.48	0.16
BOR	113.95	0.28	161.22	0.37	63.56	0.17	22.89	0.09	67.91	0.21	124.83	0.31	39.62	0.15
ANT	78.89	0.21	101.55	0.27	100.18	0.23	36.40	0.13	57.89	0.18	113.50	0.29	37.68	0.15
ALC	61.60	0.17	66.46	0.20	31.70	0.09	12.61	0.05	29.05	0.10	43.00	0.13	14.99	0.06
DRG	62.33	0.17	57.39	0.17	40.42	0.12	18.48	0.07	28.66	0.10	35.69	0.11	15.81	0.07
	DF = 1,296 $p < .001$		DF = 1, 272 $p < .001$		DF = 1, 306 $p < .001$		DF = 1, 247 $p < .001$		DF = 1, 255 $p < .001$		DF = 1, 277 $p < .001$		DF = 1, 219 $p < .001$	

Note: Somatization (SOM); Anxiety (ANX); Anxiety Related Disorders (ARD); Depression (DEP); Mania (MAN); Paranoia (PAR); Schizophrenia (SCZ); Borderline Features (BOR); Antisocial Features (ANT); Alcohol (ALC); Drug (DRG).

74
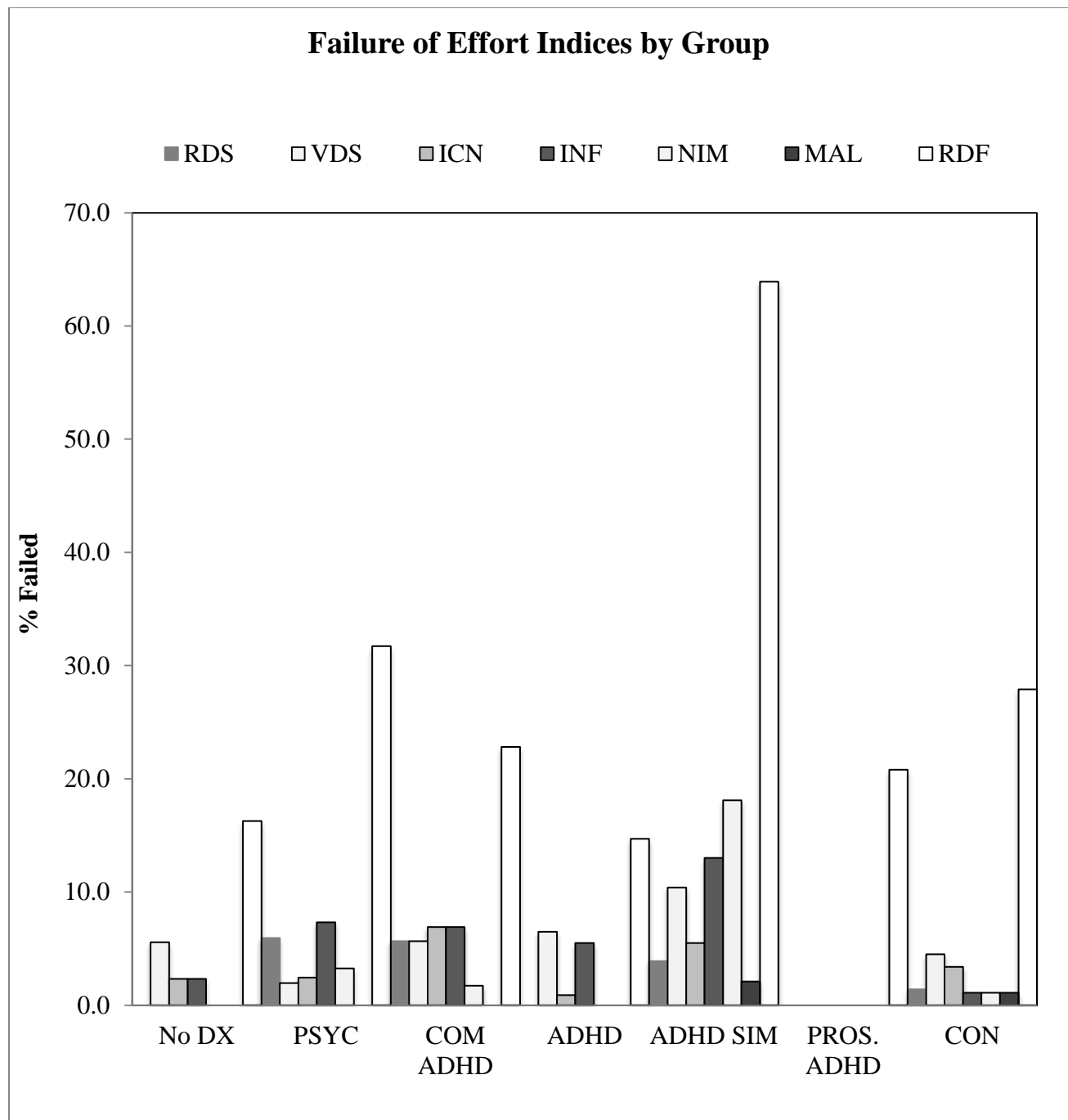
Figure 3. Percentages of individuals who failed effort indices in each clinical and control group. ADHD Simulators obtained significantly higher scores compared to all groups except the Comorbid ADHD-Psychopathology group on the Inconsistency (ICN) and Infrequency (INF) Scales. ADHD Simulators obtained higher scores than the clinical and prospective groups on the Negative Impression Management (NIM), Rogers Discriminant Function (RDF), and Malingering Index (MAL). ADHD Simulator's scores were comparable to the clinical and prospective groups (except Suspect Effort) on the Reliable Digit Span (RDS) and Vocabulary-Digit Span (VDS).

Table 6. ANCOVA results comparing group performances on validity indices.

	F*	P-value	Partial Eta Squared
RDS	14.01	<0.001	0.14
VDS	14.17	<0.001	0.14
ICN	10.67	<0.001	0.09
INF	7.35	<0.001	0.07
NIM	33.17	<0.001	0.24
PIM	20.80	<0.001	0.17
MAL	11.03	<0.001	0.10
RDF	26.16	<0.001	0.20

*Degrees of freedom = 1, 7

Note: RDS = Reliable Digit Span; VDS = Vocabulary – Digit Span; ICN – Inconsistency Scale; INF = Infrequency Scale; NIM = Negative Impression Management; PIM = Positive Impression Management; MAL = Malingering Index; RDF = Rogers Discriminant Function.

VDS were significantly lower than the Suspect Effort Group but comparable to other groups. When the Inconsistency and Infrequency scales were examined, the ADHD Simulators obtained significantly higher scores compared to all groups except the Comorbid ADHD-Psychopathology group. ADHD Simulators obtained higher scores than the clinical and prospective groups on the Negative Impression Management, Rogers Discriminant Function, and Malingering Index. Also, the ADHD Simulators obtained lower scores on the Positive Impression Management Scale compared to all of the other groups. The ADHD Means and standard deviations are presented in Table 7. Results of the ANCOVA are presented in Table 8.

Chi-square analyses were run to examine whether all groups differed in percentage of individuals who failed effort indices. There were significant differences on the Reliable Digit Span (RDS: $\chi 2$ (7) = 64.17; $p < 0.001$) and the Vocabulary-Digit Span (VDS: χ^2 (7) = 153.00; $p < 0.001$). Examination of individual groups indicated that significantly more individuals in the Suspect Effort Group failed the RDS and VDS compared to all other groups, including the ADHD Simulators. The ADHD Simulators did not fail cognitive effort measures at a higher rate than other clinical or control groups. The groups did not differ significantly in percentages of

76

individuals who failed the Inconsistency scale (χ^2 (7) = 10.80; p = ns) or the Malingering Index (χ^2 (7) = 9.20; p = ns) of the PAI. There were significant group differences on rates of failure of the Infrequency Scale (χ^2 (7) = 22.05; $p < 0.01$), Negative Impression Management Scale (χ^2 (7) = 82.68; $P < 0.001$), the Positive Impression Management Scale (χ^2 (7) = 31.63; $p < 0.001$), and the Rogers Discriminant Function (χ^2 (7) = 132.32; $p < 0.001$) scale.

Table 7. Means (SD) for PAI and Cognitive Validity Indices.

Group	RDS Mean (SD)	VDS Mean (SD)	ICN Mean (SD)	INF Mean (SD)	NIM Mean (SD)	PIM Mean (SD)	MAL Mean (SD)	RDF Mean (SD)
Clinical								
ADHD	10.13 (2.01)	1.91 (2.55)	48.54 (8.51)	54.54 (10.17)	50.05 (8.41)	44.36 (10.30)	0.82 (0.94)	-0.99 (1.04)
No Diagnosis	9.93 (1.78)	0.83 (2.95)	50.25 (9.42)	53.49 (9.01)	48.61 (6.67)	49.1 (11.10)	0.72 (0.79)	-0.95 (0.92)
PSYC	9.19 (2.81)	1.28 (2.56)	52.24 (9.31)	54.10 (11.98)	55.57 (11.95)	42.34 (12.46)	0.92 (0.97)	-0.64 (1.23)
Comorbid ADHD & PSYC	9.61 (2.79)	1.56 (2.96)	54.02 (9.93)	55.18 (10.97)	55.12 (11.95)	41.13 (11.86)	0.82 (0.77)	-0.54 (1.05)
Suspect Effort	7.12 (3.10)	5.00 (3.19)	49.28 (7.72)	52.70 (9.85)	52.92 (9.45)	45.66 (11.18)	0.67 (0.82)	-1.10 (1.08)
Prospective								
ADHD Simulators	9.74 (1.98)	2.02 (2.95)	55.56 (8.69)	59.65 (11.75)	68.88 (21.46)	33.61 (10.37)	1.54 (1.29)	0.46 (1.30)
Control	9.89 (1.62)	1.69 (2.99)	51.04 (9.93)	51.47 (8.29)	51.32 (10.38)	45.58 (10.60)	0.67 (0.89)	-0.58 (1.10)
Prospective ADHD	10.18 (2.23)	0.09 (2.95)	51.45 (9.31)	52.82 (11.24)	48.45 (5.03)	44.48 (11.86)	0.73 (0.65)	-1.06 (0.88)

Note: RDS = Reliable Digit Span; VDS = Vocabulary − Digit Span; ICN − Inconsistency Scale; INF = Infrequency Scale; NIM = Negative Impression Management; PIM = Positive Impression Management; MAL = Malingering Index; RDF = Rogers Discriminant Function.

77

Table 8. ANCOVA Analyses for Post-Hoc Comparisons of Validity Indices for ADHD Simulators compared to other groups.

	Clinical ADHD		No Diagnosis		Psychopathology		Comorbid ADHD-Psychopathology		Suspect Effort		College Controls		Prospective ADHD	
	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2	F-ratio	Partial eta^2
RDS	0.59	0.002	0.27	0.001	0.13	0.001	0.01	0.000	45.08	0.174	0.41	0.002	0.69	0.004
VDS	1.91	0.007	0.92	0.004	0.17	0.001	0.02	0.000	64.17	0.203	0.25	0.001	2.04	0.010
INF	11.26	0.037	18.60	0.064	11.20	0.036	3.88	0.016	8.39	0.032	53.46	0.163	8.26	0.037
INC	34.63	0.110	30.37	0.101	13.39	0.042	1.66	0.007	28.08	0.100	16.97	0.058	9.89	0.044
NIM	66.31	0.184	63.75	0.191	35.22	0.104	17.01	0.065	27.56	0.098	30.89	0.101	22.98	0.096
RDF	73.15	0.200	26.88	0.091	42.09	0.122	18.19	0.069	47.28	0.158	43.29	0.137	27.21	0.112
MAL	16.16	0.052	57.70	0.177	11.89	0.038	11.12	0.044	12.47	0.047	29.25	0.096	12.37	0.054
PIM	50.09	0.146	103.20	0.277	30.73	0.092	16.66	0.064	39.49	0.135	77.42	0.220	26.21	0.108

Note: RDS = Reliable Digit Span; VDS = Vocabulary – Digit Span; ICN – Inconsistency Scale; INF = Infrequency Scale; NIM = Negative Impression Management; PIM = Positive Impression Management; MAL = Malingering Index; RDF = Rogers Discriminant Function.

A series of Chi-square analyses were run as planned analyses to compare the percentages of individuals who scored above the cutoffs in the ADHD Simulator group to the other clinical and prospective groups for the Infrequency, Negative Impression Management, Positive Impression Management, and Rogers Discriminant Function scales. Significance level was set at $p < 0.01$ in order reduce the likelihood of finding significance due to chance. For the Infrequency Scale, there were no significant differences between the ADHD Simulator and Suspect Effort Groups ($\chi^2 (1) = 3.94$; $p = ns$), the Psychopathology group ($\chi^2 (1) = 2.86$; $p = ns$), and the Comorbid ADHD-Psychopathology ($\chi^2 (1) = 1.60$; $p = ns$). There were trends toward significance when ADHD Simulators were compared to the Clinical ADHD group ($\chi^2 (1) = 4.20$; $p = 0.04$) and the Prospective ADHD group ($\chi^2 (1) = 4.23$; $p = 0.04$). There were significant differences between the ADHD Simulators and the No Diagnosis ($\chi^2 (1) = 7.69$; $p < 0.01$) and College Control ($\chi^2 (1) = 10.05$; $p < 0.01$) groups.

When percentages of ADHD Simulators who failed the Negative Impression Management Scale were compared to other groups, significant differences were found for all groups: Clinical ADHD ($\chi^2 (1) = 22.36$; $P < 0.001$), No Diagnosis ($\chi^2 (1) = 17.83$; $p < 0.001$), Psychopathology ($\chi^2 (1) = 15.38$; $p < 0.001$), Comorbid ADHD-Psychopathology ($\chi^2 (1) = 9.78$; $p < 0.01$), Suspect Effort ($\chi^2 (1) = 13.34$; $p < 0.001$), Prospective ADHD ($\chi^2 (1) = 6.24$; $p < 0.01$), and College Control ($\chi^2 (1) = 15.36$; $p < 0.001$). For the Positive Impression Management scale, group differences reached statistical significance for the Clinical ADHD ($\chi^2 (1) = 9.00$; $p < 0.01$), No Diagnosis ($\chi^2 (1) = 25.70$; $p < 0.001$), Psychopathology ($\chi^2 (1) = 15.38$; $p < 0.001$), Comorbid ADHD-Psychopathology ($\chi2 (1) = 10.10$; $p < 0.01$), Suspect Effort ($\chi^2 (1) = 6.36$; $p < 0.01$). Finally, there were significant differences between percentages of ADHD Simulators who failed the Rogers Discriminant Function all other clinical and groups: Clinical ADHD ($\chi^2 (1) = 67.16$; p

< 0.001), Psychopathology (χ^2 (1) = 30.99; $p < 0.001$), Comorbid ADHD-Psychopathology (χ^2 (1) = 29.85; $p < 0.001$), No Diagnosis (χ^2 (1) = 29.85; $p < 0.001$), Suspect Effort (χ^2 (1) = 53.74; $p < 0.001$), Prospective ADHD (χ^2 (1) = 22.44; $p < 0.001$), and College Control (χ^2 (1) = 30.78; $p < 0.001$) groups.

5.4. Hypothesis 3: It is hypothesized that ADHD Simulators endorse significantly more items related to psychopathology compared to individuals that meet diagnostic criteria for ADHD. Also, it is expected that these items will be summed into a scale that adequately differentiates the two groups.

When ADHD simulators were compared to the clinical ADHD sample, 268 of the 344 PAI items were significant at $p < 0.001$. Partial eta-squared (η^2) was examined to determine the 40 items with the highest effect sizes. Effect sizes were large for the 40 items, ranging from 0.18 to 0.34. When the college control group was compared to the clinical No Diagnosis Group, 5 of the 40 items were significant at the $p < 0.01$ level and had $\eta^2 > 0.08$. The next five items with the highest effect sizes for differentiating ADHD Simulators from the Clinical ADHD group. One additional item yielded significant differences between controls and the No Diagnosis group and was replaced. These five items were added to the 35 items previously identified.

The 40 items that best differentiated ADHD Simulators from the Clinical ADHD sample were placed into a discriminant function analysis with the ADHD Simulators and Clinical ADHD groups as the grouping variables. Seven cases were dropped from analysis due to missing data leaving 82 ADHD Simulators and 83 individuals in the Clinical ADHD group. Mahalanobis d indicated there were no multivariate outliers. Examination of Box's M indicated significant heterogeneity of covariance ($p < 0.001$). In order to determine the effect of heterogeneity of covariance, a second discriminant function was run using separate group covariances. Box's M was no longer significant and the classification statistics were similar compared to the first discriminant function. Therefore, it was determined that analyses could proceed using within

80

group covariances for identification of items as using separate group covariance can lead to over-fitting.

The discriminant function was significant ($\chi2$ (39) = 168.52; $p < 0.001$) and accounted for 100% of the variance between the two groups. The discriminant function demonstrated 91.5% classification accuracy misidentifying six Clinical ADHD individuals as Simulators and eight ADHD simulators as having clinically significant ADHD. Examination of the Structure Matrix Coefficients indicated that all 40 items loaded between .132 and .507. Fourteen items had loadings > 0.40. The fourteen items with Structure Matrix Loadings > 0.40 were retained. They were summed into the newly developed scale of the PAI, and it should be noted that Items 235 and 301 are reverse scored per PAI scoring procedure:

SUM = PAIItem4 + PAIItem14 + PAIItem54 + PAIItem73 + PAIItem84 +

PAIItem105 + PAIItem113 + PAIItem233 + PAIItem235 + PAIItem274 +

PAIItem275 + PAIItem276 + PAIItem301 - PAIItem344.

The scale demonstrated very good reliability in the current sample (Cronbach's alpha = 0.88). The total scale score (sum of the 14 items) was placed into a logistic regression with group (Clinical ADHD vs. ADHD Simulators) as the dependent variable. The logistic regression indicated that, when the fourteen items were summed into the new scale, it was a significant predictor of malingered ADHD in these two groups [χ^2 (1, n = 168) = 124.76, $p < .001$, odds ratio = 1.32, CIs (1.22, 1.43)]. The Hosmer and Lemeshow Test of goodness-of-fit was not significant ($p = 0.73$) indicating a good fit for the model. When the scale was placed into an ROC curve area under the curve was .935, which is considered outstanding.

Next, various cutoff scores were examined for the newly developed scale using hold-out samples of the Clinical ADHD (n = 24) and ADHD Simulators (n = 106) as well as the No

Diagnosis (n = 41), College controls (n = 43), and Prospective ADHD (n = 29) groups (see Table 9). Examination of the coordinates of the ROC curve suggests a cutoff scores of > 14 yielded 90.4% specificity and 76.5% sensitivity in the development sample of ADHD Simulators and Clinical ADHD groups. However, the cutoff score of > 14 yielded specificities below 90% for the College controls and the Prospective ADHD group. Specificity of the scale improved as cutoff scores were raised, and a cutoff score of > 16 yielded adequate specificities for the Clinical ADHD groups as well as the No Diagnosis and Control groups; however, a cutoff score of > 22 yielded Specificity above 90% for the Prospective ADHD Group. Positive Predictive Power (PPP) and Negative Predictive Power (NPP) were calculated at different hypothesized base rates of malingering for the ADHD Simulators, the Clinical ADHD group, the Clinical No Diagnosis group, and Prospective ADHD group and are available in Table 10.

5.5. Hypothesis 4: **It is expected that the scale that is developed will differentiate simulated ADHD from other clinical groups, including diagnoses of comorbid psychopathology and ADHD.**

The various cutoff scores were examined for the Psychopathology and Comorbid ADHD-Psychopathology groups (See Tables 9 and 10). A cutoff score of > 14 did not yield adequate (> 0.90) specificity for either group. Specificities of 86% were found for both groups at a cutoff score of 20, and a cutoff score of > 22 yielded 91% and 93% specificity for the Psychopathology and Comorbid ADHD-Psychopathology groups, respectively. A univariate analysis of variance was conducted and indicated significant differences between groups (F (7, 726) = 60.78, $p <$ 0.001; ŋ2 = 0.37). Scheffe's post-hoc analyses revealed the analog group obtained significantly higher scores compared to all other groups, but the Psychopathology and Comorbid ADHD-Psychopathology group's scores were significantly higher compared to the Clinical ADHD and No Diagnosis groups.

Table 9. Examination of Sensitivity and Specificity of various cutoff scores for newly developed in the current sample.

	Mean (SD)	Cutoff > 14	Cutoff > 16	Cutoff > 18	Cutoff > 20	Cutoff > 22
Specificity		**81**	**86**	**90**	**92**	**95**
Clinical ADHD for Development	5.6 (5.8)	90	95	96	98	100
Clinical ADHD Hold-out Sample	7.0 (6.4)	91	96	96	96	100
No Diagnosis	4.5 (5.0)	95	98	99	100	100
PSYC	10.7 (8.5)	68	74	82	86	91
Comorbid ADHD & PSYC	11.7 (7.7)	67	72	77	86	93
Control	6.8 (7.8)	85	91	92	92	93
Prospective ADHD	9.1 (8.9)	75	82	86	89	93
Sensitivity		*71*	*64*	*59*	*51*	*44*
ADHD Simulators for Development	21.9 (8.5)	*77*	*69*	*65*	*55*	*48*
ADHD Simulators Hold-out Sample	19.8 (10.2)	*66*	*60*	*54*	*48*	*40*

Table 10. Positive Predictive Power and Negative Predictive Power for various cutoff-scores for the newly developed scale at different base-rates.

	PPP BR=0.14	PPP BR = 0.25	PPP BR= 0.50	NPP BR = 0.14	NPP BR = 0.25	NPP BR= 0.50
Hypothesis 3[a]						
Cutoff > 14	51	68	86	62	49	26
Cutoff > 16	62	77	91	69	55	30
Cutoff > 18	65	79	92	71	58	32
Cutoff > 20	66	80	92	73	60	35
Cutoff > 22	78	88	96	75	61	34
Hypothesis 4[b]						
	PPP BR=0.14	PPP BR = 0.25	PPP BR= 0.50	NPP BR = 0.14	NPP BR = 0.25	NPP BR= 0.50
Cutoff > 14	38	55	79	51	41	22
Cutoff > 16	43	61	82	59	48	26
Cutoff > 18	48	65	85	64	52	29
Cutoff > 20	52	69	87	68	56	32
Cutoff > 22	66	80	92	72	59	33

[a]Hypothesis three examines Positive Predictive Power (PPP) and Negative Predictive Power (NPP) for the Clinical ADHD, No Diagnosis, and College Controls, Prospective ADHD, and ADHD Simulator Groups.

[b]Hypothesis four examines PPP and NPP for all Clinical and Prospective groups with the exception of the Suspect effort group as that is examined in Hypothesis 5.

5.6. Hypothesis 5: The ADHD Simulators' scores on the newly developed scale will be comparable to the Suspect Effort Group's scores.

The ADHD Simulator group was compared to the Clinical Suspect Effort group using

Chi-square analysis with the dichotomously coded (pass or fail) cutoff score of > 14, and the

Chi-Square was significant (χ^2 (1, n = 165) = 60.09, $p < .001$). The mean newly developed scale

score for the Suspect effort group was 7.85 (SD = 7.02), and few individuals (n = 10) in the

Suspect Effort group score above the cutoff score of > 14 (15% Sensitivity), and sensitivity declined as the higher cutoff scores were examined. As noted under Hypothesis 1 in the results section, the Suspect Effort group did not fail the PAI validity scales at higher rates than other clinical groups. The suspect effort group did not obtain scores that rose above the cutoff at rates that were higher than the other clinical groups. In fact, the Suspect Effort group's performance on the newly developed scale mirrored performance of the Clinical ADHD hold-out sample, as there were no clinically significant differences between the two groups using a cutoff score of > 14 (χ^2 (1) = 0.65, p = .724).

5.7 Hypothesis 6: It is expected that the ADHD Simulators in the current study will use more strategies associated with inattention compared to hyperactivity.

The ADHD Simulators were provided with a list of common strategies for feigning ADHD that are reported in the literature. Percentages of individuals who responded to each item are reported in Table 11. They were asked to circle all of the strategies that they used to fake ADHD during the study. The number of strategies used by students varied greatly (Range = 0 - 13) but students typically employed five different strategies (mean = 4.94; SD = 2.76). Of the five responses more frequently endorsed by the ADHD Simulators, two were associated with inattention and one was associated with hyperactivity/impulsivity. The two most commonly endorsed responses were associated with deliberate attempts to malinger that did not specify feigning symptoms of inattention or hyperactivity.

Table 11. Percentages of ADHD Simulators who endorsed each strategy.

Question	% Endorsed
Attempting to select responses that match what I know about the criteria	55.7
Thinking of Someone that has ADHD and trying to act like that	50.5
Letting my mind wander	42.7
Completing tasks quickly and carelessly	40.6
Zoning out	40.1
Fidgeting	36.5
Re-reading questions	36.5
Completing tasks slowly	29.7
Letting my focus wane toward the end of a task	29.3
Acting Bored	27.6
Choosing incorrect answers	25.5
Focusing on extraneous things	24.5
Skipping items	16.1
Not attending to the researcher's instructions	8.9
Beginning before being told	8.9
Un-focusing my eyes to make items harder to read	7.8
Interrupting the Researcher	6.3
Ignoring sides of the page and focusing on the center	4.7
Disobeying verbal instructions	3.7

6. DISCUSSION

Malingered ADHD in college students has received growing interest over the past several years. College students may wish to feign ADHD in order to obtain academic accommodations (Sullivan et al., 2007) or stimulant medications (White et al., 2006; Advokat et al., 2010; MacCabe et al., 2006; Rabiner et al 2009). The current literature suggests that college students are able to simulate ADHD on self-report questionnaires as well as on many neurocognitive tests. At this point, it appears that malingering indices designed to detect feigned neurocognitive impairments demonstrate the best sensitivity and specificity for malingered ADHD, though their utility is limited because they may only catch the most unsophisticated malingerers in this population (Musso & Gouvier, 2012). Only one embedded validity index, the CAARS Infrequent Index, has been derived for self-report questionnaires specifically for detecting ADHD in college students (Suhr et al., 2011), and while it demonstrated promise in discriminating individuals believed to be responding honestly from those believed to be exaggerating deficits, it requires further validation. The purpose of the current study was to develop an embedded validity index within the Personality Assessment Inventory that provided adequate sensitivity and specificity for detecting malingered ADHD in college students.

6.1 Hypothesis 1: It is expected that the ADHD Simulators will obtain significantly higher scores on the PAI clinical scales, the WURS, and the STAI-Y compared to clinical groups and controls but that their scores will not exceed a t-score of 80 (raw score > 55 on WURS).

This hypothesis was supported. The ADHD Simulators scored significantly higher on the WURS, STAI-Y State and Trait, and all of the PAI clinical scales. However, it should be noted that while the average scores for all of these scales were significantly higher compared to other groups, they fell within clinically believable ranges defined as group mean T-scores < 80 for the STAI-S and STAI-T. Scores (mean of 58.85) were slightly higher than expected for the WURS.

In addition, the two PAI scales with the highest mean scores were the Anxiety (76) and Antisocial Features (75), which, while elevated, are not exceedingly elevated as to attract the attention of clinicians.

The PAI clinical scale profile of the ADHD Simulator group is consistent with what clinicians often refer to and interpret as a "cry for help." The cry for help profile is described in the MMPI literature by Dalhstrom and colleagues as "pleading for special attention" and they conceptualize the cry for help profile as a patient's dramatization of symptoms as a means of ensuring that he will receive the attention that he feels he needs. (Dahlstrom, Welsh, & Dalstrom, 1972). Little empirical research has been conducted on the cry for help profile, but it is generally interpreted as an individual with actual problems overstating his or her problems "in a flag-waving fashion" as a means of signaling for help (Marks, Seeman, & Haller, 1974). However, the current data suggests that, at least in the college student population, this cry for help profile may actually reflect attempts to malinger ADHD.

6.2 Hypothesis 2: It is expected that ADHD simulators will obtain scores on cognitive effort measures that are similar to the Suspect Effort group but significantly poorer than other groups. It is also expected ADHD Simulator's scores on self-report validity indices will be comparable to the other clinical groups.

This hypothesis was not supported. In fact, few individuals in the ADHD Simulator group failed the Reliable Digit Span (3.9%) or the Vocabulary-Digit Span (10.4%). Few ADHD Simulators failed the Malingering Index (2.1%), the Inconsistency Index (5.5%), and the Infrequency Index (13.0%) of the PAI. Higher percentages of ADHD Simulators failed the Negative Impression Management Index (18.1%) and the Rogers Discriminant Function (63.9%) compared to all other groups. It should be noted that individuals who were clinically diagnosed with a psychiatric disorder or ADHD, as well as individuals with no diagnosis and controls failed

88

the Rogers Discriminant Function more frequently compared to other scales suggesting the Rogers Discriminant Function lacks specificity in psychoeducational evaluations.

The current findings are consistent with Sullivan et al. (2007) who reported that few college students suspected of malingered ADHD failed the validity indices of the PAI. The finding that few individuals failed the Infrequency Index of the PAI is not consistent with Young and Gross (2011) who reported that individuals suspected of malingered ADHD are more likely to endorse infrequent symptoms of psychopathology (Fp scale of the MMPI-2). However, it should be noted that the Infrequency scale of the PAI was designed to detect carelessness and idiosyncratic response styles and is not specific to infrequently endorsed items of psychopathology (Morey, 2003). In addition, the findings of the present study suggest that individuals malingering ADHD would not be detected using the neurocognitive measures (Reliable Digit Span or Vocabulary-Digit Span) nor would they be detected using the validity scales of the PAI.

6.3 Hypothesis 3: It is hypothesized that ADHD Simulators endorse significantly more items related to psychopathology compared to individuals that meet diagnostic criteria for ADHD. Also, it is expected that these items will be summed into a scale that adequately differentiates the two groups.

This hypothesis was supported. The ADHD Simulators obtained significantly different scores on 268 items of the PAI. Effect sizes were examined to determine the 40 most significant items that differentiated the two groups, and effect sizes of all 40 items were large (.18-.34). However, an unexpected finding was that half of the items (6) were derived from the anxiety scale of the PAI. One item related to the Mania-Irritability subscale and two items related to the Borderline Affective Instability scale were significant. The other four items were derived from the Anxiety Related Disorders scale, the Suicide scale, the Positive Impression Management scale, and the Depression scale. No items assessing Hyperactivity or Thought Disorder were

included in the 40 most significant items. This finding is surprising because the college students in the current sample were not diagnosed with severe mental illness such as Bipolar disorder or Schizophrenia. This finding is not inconsistent with Musso et al. (2011) who examined college students' responses on the Thought Disorder Subscale of the PAI and reported that college students with ADHD obtained elevated scores on this subscale. Musso et al. (2011) suggested that the Thought Disorder Subscale of the PAI may measure self-reported symptoms of ADHD. In a subsequent unpublished manuscript, no items from the PAI Thought Disorder scale emerged as significant predictors of clinical ADHD suggesting individuals with psychopathology and individuals with comorbid ADHD and psychopathology may also endorse more items on this scale. Further research is needed to determine why the Thought Disorder Subscale is elevated in college students.

The 40 most significant items were placed into a discriminant function analysis, and fourteen items emerged as having structure matrix coefficient loadings > 0.4. These items were summed into a scale. The diagnostic statistics were calculated for the scale. When malingering is in question, it is better to fail to identify malingerers than to falsely classify clinical samples as malingering (Type I error). Therefore, a range of cutoff scores was examined to determine the best sensitivity and specificity for the clinical sample. Ultimately two cutoff scores emerged, a cutoff of > 16 proved useful for distinguishing malingered ADHD from clinical ADHD, and a cutoff of > 22 that distinguish other groups with psychopathology or comorbid ADHD and psychopathology. Therefore, it is recommended that, when an individual does not complain of comorbid psychiatric symptoms during an ADHD evaluation, a cutoff score of > 16 should be used as it is highly specific to malingered ADHD. However, when comorbid psychopathology is in question, an individual should not be suspected of malingering unless his score on the newly

90

developed scale exceeds 22. This use of alternative cutoff scores for various populations is not uncommon in neuropsychological testing. For example, the WURS proposes a cutoff score of $\geq$ 36 for individuals presenting with ADHD symptoms, but a cutoff of $\geq$ 46 when an individual complains of comorbid depressive symptoms. Similarly, the Dot Counting Test (Boone et al., 2002), a stand-alone symptom validity test, proposes a range of cutoff scores based on clinical conditions.

It should be noted that the specificity of the scale was somewhat lower in individuals that were recruited for this study who had been diagnosed with ADHD by a health care provider (Prospective ADHD group) compared to the Clinical ADHD group that consisted of individuals who had been diagnosed through neuropsychological evaluation. It is unclear why more individuals in the prospective ADHD group would have failed because they had no study-related external incentive to malinger, and on most of the self-report questionnaires and validity indices, the Prospective ADHD group responded more similarly to controls compared to individuals diagnosed with ADHD. This finding brings into question the validity of individuals diagnosed with ADHD without formal neuropsychological assessment, as a majority of individuals in the Prospective ADHD group did not endorse significant childhood symptoms of ADHD on the WURS. Elevations on the PAI consistent with the Clinical Psychopathology and Comorbid ADHD-Psychopathology groups suggest these individuals may have more symptoms of psychopathology that are misdiagnosed as ADHD by physicians. As mentioned in the Introduction, symptoms of inattention and difficulty concentrating are included in diagnostic criteria of depressive and anxiety disorders increasing the potential for misdiagnosis. Further research is needed to explore whether healthcare providers adhere to DSM-IV-TR criteria when diagnosing Adult ADHD and administering prescription stimulant medications.

6.4 Hypothesis 4: **It is expected that the scale that is developed will differentiate simulated ADHD from other clinical groups, including diagnoses of comorbid psychopathology and ADHD.**

This hypothesis was supported. A cutoff score of >22 yielded excellent specificity in all groups including the Psychopathology and Comorbid ADHD-Psychopathology groups. While sensitivities reported for the newly developed scale of the PAI are higher than all of the current PAI validity indices for detecting malingered ADHD, only 44% of individuals feigning ADHD would be identified at a cutoff score of > 22. Sensitivity reported for the newly developed scale is comparable to what has been reported in the MMPI literature. However, the two studies that have examined MMPI cutoff scores did not examine diagnostic statistics for detecting malingered ADHD in a sample with comorbidity. Harp and colleagues reported sensitivities ranging from .045 to .18 for the MMPI-RF cutoff scores suggested in the manual F-r, Fp-r, and Fs scales, and sensitivities improved when with their proposed cutoffs, ranging from .36 to .64. Young and Gross (2011) reported sensitivities ranging from .16 to .59 for various cutoff scores of the MMPI-2 validity indices. No studies to date have examined the use of MMPI-2 scales in detecting malingered ADHD in a general clinical setting; however, it is suspected poorer sensitivities and specificities would be found for MMPI-2 and MMPI-2-RF validity indices compared to studies that employed only individuals who meet diagnostic criteria for ADHD and ADHD simulators (Harp et al., 2011; Young & Gross, 2011).

Walters et al. (2009) examined the taxometric structure of malingering and reported that malingering is a dimensional construct rather than a dichotomous construct making the use of one cutoff score as indicative or not indicative of malingering less useful. Therefore, this study reported diagnostic statistics for a range of cutoff scores. While cutoff scores are suggested for both the presence and absence of clinical ADHD and comorbid ADHD and psychopathology, it

92

is recommended that clinicians consider the sensitivity and specificity of a given score in light of additional sources of information.

6.5 Hypothesis 5: The ADHD Simulators' scores on the newly developed scale will be comparable to the Suspect Effort Group's scores.

One interesting finding of this study is that individuals classified as having suspect effort based on failure of two or more cognitive embedded effort indices did not feign on the PAI. None of the individuals in the Clinical Suspect Effort group failed the ICN, NIM, or MAL indices of the PAI. Only 4.6% failed the INF, and while 10.9% failed the RDF, the measures proved insensitive in the college students. In fact, a similar number of individuals in the Suspect Effort group failed PAI validity when compared to other clinical and control groups including College Control and Clinical No Diagnosis groups. There are several hypotheses for why the Clinical Suspect Effort group did not fail the embedded indices of the self-report measure. One hypothesis is that cognitive and self-report effort indices measure distinct constructs, and individuals that chose to perform poorly on cognitive measures may not feign on measures of psychopathology. This hypothesis is consistent with (Nelson et al. 2007) who reported malingered psychopathology and malingered neurocognitive dysfunction are distinct constructs. Marshall (2010) also found that individuals that feigned in their study chose very specific cognitive and/or self-report strategies. In their study, individuals that failed two neurocognitive effort measures were strategic in feigning on measures of memory and sustained attention. Those suspected of feigning on rating scales also tended to feign on cognitive measures of sustained attention, and those whose ratings of symptoms differed from psychometrists' ratings did not appear to attempt to distort their responses on any cognitive tests. It appears the findings of this study are consistent with these findings that individuals select specific measures of feigning.

93

Also, it this study, feigned psychopathology appears to be a distinct construct from suspect effort on embedded cognitive measures.

Alternatively, the discrepancy found between the Clinical Suspect Effort group's failures on neurocognitive effort indices may reflect lack of specificity for college students. Pella et al. (2009) reported that some embedded cognitive effort indices lacked adequate specificity in a sample of college students. In particular, the Mittenberg Discriminant Function and Vocabulary-Digit Span subtests demonstrated poor specificity in college students.

6.6 Hypothesis 6: It is expected that the ADHD Simulators in the current study will use more strategies associated with inattention compared to hyperactivity.

The college students were surveyed about strategies used to feign ADHD. Most students employed multiple strategies. Over 50% of the students surveyed reported that they attempted to select responses that match what they knew about the criteria and/or attempted to think of someone they knew who had ADHD and tried to act like them. Other cognitive strategies used by over one-third of the sample included zoning out and letting one's mind wander. Behavioral strategies employed by over one-third of the sample included: fidgeting, completing tasks quickly and carelessly, and/or completing tasks slowly. These findings are consistent with other researchers who report that ADHD simulators employ various strategies when asked to simulate ADHD, and they typically employ strategies that are associated with inattention rather than hyperactivity (Quinn, 2003; Frazier et al., 2008; Harrison et al., 2010; Marshall, 2010).

6.7 Limitations

This study has several limitations. Most pertinent, as with all analog research, the generalizability of findings is an issue. Individuals asked to feign ADHD for class credit do not have as much at stake as individuals that are completing these measures for an external incentive such as medication or academic accommodations. Therefore, it is possible that responses of the

94

ADHD Simulator group are exaggerated compared to the way that college students with an external incentive may perform. Also, as noted previously, individuals identified as having suspect effort did not perform similarly to the ADHD Simulators in the current study. These differences may indicate that cognitive effort tests and validity indices of self-report instruments measure distinct constructs. Some studies use the WMT as a gold-standard criterion for malingered ADHD; however, while the WMT initially appears intimidating, it may be may only detect less sophisticated malingerers because the task is ultimately relatively easy.

Another limitation of the current study is that individuals in the current study receive information in the form of psuedowebpages, but they were only provided with five minutes to read the material. It is likely that individuals presenting to the clinic with the intent of malingering or exaggerating responses would take more time to prepare. They would be more likely to spend a reasonable amount of time familiarizing themselves with diagnostic criteria. In addition, this study utilized a significantly greater proportion of females than the clinical data, which could further impede generalizability. All analyses were run using gender as a covariate to attempt to correct for this. Further research should continue to investigate the utility of this scale in populations with comparable numbers of males and females.

In addition, there is no gold-standard instrument for measuring malingered ADHD, and the literature indicates few current measures are adequate at detecting malingered ADHD. The lack of a gold standard undermines the diagnostic accuracy of clinicians by increasing the chance that many individuals who feigned ADHD during the clinical evaluation were undetected. This was taken into account by adjusting the criterion (clinical diagnosis) for a hypothetical error rate (Please see Appendix B). For the purposes of the current study, the hypothetical error rate chosen was 25%. The diagnostic statistics of the newly developed scale were recalculated adjusting for

95

this error rate. This is a novel idea in the malingering literature as the implications of Type II errors are often dire in forensic settings (where much of the literature has focused). However, this study acknowledges that, in the absence of a gold standard, diagnostic accuracy was likely not 100% and this should be accounted for in the diagnostic statistics. When this adjustment was made, sensitivity and specificity improved for all cutoff scores; however, the improvements were negligible.

6.8 Conclusion

This study makes several important contributions to the literature. First, this study illustrated that college students asked to simulate ADHD are successfully able to manipulate the Wender Utah Rating Scale and remain undetected by the Personality Assessment Inventory's embedded validity indices. More importantly, this study developed the first embedded effort index for a broad, objective personality measure that is designed to detect malingered ADHD. In order to do so, this study compared ADHD Simulators' responses to individual PAI items to a Clinical ADHD group. Fourteen items were identified and summed into a scale, the newly developed scale of the PAI. A cutoff score > 16 was yielded excellent specificity for the Clinical ADHD and No Diagnosis groups. This cutoff score is should be used when there are no complaints or evidence of comorbid psychiatric symptoms. However, a cutoff score of > 22 is recommended for students who present with complaints of comorbid psychopathology, in order to avoid misdiagnosing individuals with complaints of ADHD and psychiatric comorbidities as malingering.

This scale demonstrated utility for detecting malingered ADHD, but should not be used in isolation. As with all malingering measures, there is a risk of false-positive errors in clinical samples when using only one measure. However, the current scale offers better sensitivity and

specificity than most other self-report measures and has also been validated in a sample of college students diagnosed with psychopathology. Numerous indices have been investigated, but few have been examined in more than one study and fewer, still, have been validated for use in in the presence of psychopathology (Musso and Gouvier, 2012). At present, much more research is needed to validate newly developed scales such as the newly developed scale and CAARS Infrequency Index. Future research should also include groups with comorbid ADHD and psychopathology. This study demonstrated that cutoff scores that prove useful in differentiating malingerers from clinical samples with ADHD did not offer adequate specificity in groups with comorbid ADHD and psychiatric symptoms. In addition, future research should continue to pursue the development of measures that offer better sensitivity and specificity for malingered ADHD, as it appears to be an elusive construct.

REFERENCES

Advokat, C. D., Guidry, D., & Martino, L. (2010). Licit and illicit use of medicaitons for attention-deficit hyperactivity disorder in undergraduate college students. *Journal of American College Health, 56*(6), 601-606.

American Psychiatric Association (APA). (2000). *Diagnostic and Statistical Manual ofMental Disorders: DSM-IV-TR*. Washington, DC.

American Psychiatric Assosciation. (2012, May 20). Retrieved August 10, 2011, from American Psychiatric Association: DSM-5 Development: http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=383#

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

Babikian, T., Boone, K.B., Lu, P., Dean, A., Wen, J., Nitch, S., et al. (2005). Sensitivity and specificity of the finger tapping scores for the detection of suspect effort. *The Clinical Neuropsychologist, 19,* 105-120.

Baity, M. R., Siefert, C. J., Chambers, A., & Blais, M. A. (2007). Deceptiveness on the PAI: A study of naive faking with psychiatric inpatients. *Journal of Personality Assessment, 88*(1), 16-24.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disrderinto young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology, 111*, 279-289.

Barkley, R. A., & Murphy, K. R. (2006). *Attention deficit hyperactivity disorder: A clinical workbook* (3rd Edition ed.). New York: Guilford Publications.

Barkley, R. A., & Murphy, K. R. (2010). Impairent in occupational functioning and adult ADHD: The predictive utility of executive function (EF) ratings versus EF tests. *Archives of Clinical Neuropsychology, 25*, 157-173.

Barkley, R. A., Murphy, K. R. & Fischer, M. (2008). *ADHD in adults; What the science says.* New York: The Guilford Press.

Barrash, J., Suhr, J., & Manzel, K. (2004). Detecting poor effort and malingering with an expanded version of the Auditory Verbal Learning Test (AVLTX): Validation with clinical samples. *Journal of Clinical and Experimental Neuropsychology*, *26*, 125–140.

Ben-Porath, Y. S., & Tellegen, A. (2008). *Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF): Manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press.

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

Bigler, E. D. (2012). Response to Larrabee. *Journal of the International Neuropsychological Society, 18(4),* 640-642.

Blase, S. L., Gilbert, A. N., Anastopoulos, A. D., Costello, E. J., Hoyle, R. H., Swartzwelder, H. S., & Rabiner, D. L. (2009). Self-Reported ADHD and Adjustment in College Cross-sectional and Longitudinal Findings. *Journal of Attention Disorders*, *13*(3), 297-309

Booksh, R. L., Pella, R. D., Singh, A. N., & Gouvier, W. D. (2010). Ability of college students to simulate ADHD on objective measures of attention. *Jouranl of Attention Disorders, 13*(4), 325-338.

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

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

Boone, K. B., Lu, P., & Wen, J. (2005). Comparison of various RAVLT scores in the detection of noncredible memory performance. *Archives of Clinical Neuropsychology*, *20*, 301–319.

Bracken, B. A., & Boatwright, B. S. (2005). *Examiner's manual: Clinical Assessment of Attention Deficit – Child and Adult*. Odessa, FL: Psychological Assessment Resources.

Brown, T. E. (1996). Brown Attention Deficit Disorder Scales for Adolescents and Adults. San Antonio, TX: The Psychological Corporation.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny Reading Test manual for scoring and interpretation*. Chicago, IL: Riverside.

Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., Reynolds, C. R., & Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity: NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology, 20(4),* 419-426.

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

Cashel, M. L., Rogers, R., Sewell, K., & Martin-Cannici, C. (1995). The Personality Assessment Inventory and the detection of defensiveness. *Assessment, 2*, 333-324.

Conners, C. K. (1995). *Conners' continuous performance test*. Ottawa, Ontario, Canada: MultiHealth Systems.

Conners, C. K., Erhardt, D., & Sparrow, E. (1999). *CAARS: Conners'Adult ADHD Rating Scale*. North Tonawanda, NY: Multi-Health Systems.

Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. (1972). *An MMPI handbook* (Vol. 1, pp. 155-174). Minneapolis: University of Minnesota Press.

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001). *Delis-Kaplan Executive Function System.* San Antonio, TX: The Psychological Corporation.

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

Diller, L. (2010). ADHD in the college student: Is anyone else worried? *Journal of Attention Disorders, 14*(1), 3-6.

DuPaul, G. J., Guevermont, D. C., & Barkley, R. A. (1991). Attention-deficit hyperactivity disorder in adolescence: Critical assessment parameters. *Clinical Psychological Review, 11*, 233-245.

DuPaul, G. J., Thomas, T. J., Anastopoulos, A. D., & Reid, R. (1998). ADHD Rating Scale-IV: Checklists, norms, and clinical interpretation. New York, NY: Guilford Press.

DuPaul, G. J., Schaughency, E. A., Weyandt, L. L., Tripp, G., Kiesner, J., Ota, K., & Stanish, H. (2001). Self-Report of ADHD Symptoms in University Students Cross-Gender and Cross-National Prevalence. *Journal of Learning Disabilities, 34(4),* 370-379.

Edens, J. F., Watkins-clay, M. M., & Poythress, N. G. (2007). Detection of malingering in psychiatric unit and general populaiton prison inmates: A comparison of the PAI, SIMS, and SIRS. *Journal of Personality Assessment, 88*(1), 33-42.

Faraone, S. V., Biederman, J., Spencer, T., Wilens, T., Seidman, L. J., Mick, E., et al. (2000). Attention-Deficit/Hyperactivity Disorder in adults: An overview. *Biological Psychiatry, 48*, 9-20.

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

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

Fedele, D. A., Hartung, C. M., Canu, W. H., & Wilkowski, B. M. (2010). Potential symptoms of ADHD for emerging adults. *Journal of Psychopathology and Behavioral Assessment*, *32*(3), 385-396.

Fisher, A. B., Watkins, M. W. (2008). ADHD Rating Scales' Susceptibility to faking in a College Student Sample. *Journal of Postsecondary Education and Disability, 2*, 81-92.

Frazier, T. W., Frazier, A. R., Busch, R. M., Kerwood, M. A., & Demaree, H. A. (2008). Dection of simulated ADHD and reading disorder using symptom validty measures. *Archives of Clinical Neuropsychology, 23*, 501-509.

Frederick, R. I. (2000). A personal floor effect strategy to evaluate the validity of performance on memory tests. *Journal of Clinical and Experimental Neuropsychology*, *22*(6), 720-730.

Frederick, R. I. (2003). *VIP (Validity Indicator Profile) manual.* Minneapolis: MCS Pearson Inc.

Garnier, L. M., Arria, A. M., Caldeira, K. M., Vincent, K. B., O'Grady, K. E., & Wish, E. D. (2010). Sharing and selling of prescription medications in a college students sample. *Journal of Clinical Psychiatry, 71(3),* 262-269.

Garnier-Dykstra, L. M., Caldeira, K. M., Vincent, K. B., O'Grady, K. E., & Arria, A. M. (2012). Nonmedical use of prescription stimulants during college: Four-year trends in exposure opportunity, use, motives, and sources. *Journal of American College Health, 60(3),* 226-234.

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2007). Development and validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment*, 14, 196–208.

Glutting, J. J., Sheslow, D., & Adams, W. (2002). CARE: College ADHD response evaluation manual. Wilmington, DE: Wide Range.

Golden, C. J. (1978). *The Stroop Color and Word Test: A manual for clinical and experimental uses*. Chicago, IL: Stoelting.

Golden, C. J., & Freshwater, S. M. (2002). *The Stroop Color and Word Test: A manual for clinical and experimental uses*. Wood Dale, IL: Stoelting.

Gouvier, D., & Hayes, J. S. Y Smiroldo, B. B. (1998). The significance of base rates, test sensitivity, test specificity, and subject's knowledge of symptoms in assessing TBI sequelae and malingering. *Detection of malingering during head injury litigation*, 55-79.

Gouvier, W. D. (2001). Are you sure you're really telling the truth?. *NeuroRehabilitation-An Interdisciplinary Journal*, *16*(4), 215-220.

Green, P. (2003). *Green's word memory test for Microsoft Windows*. Edmonton, Alberta, Canada: Green's Publishing Company.

Green, P. (2006). *NV-MSVT for Windows, Version 1.3*. Edmonton, Alberta, Canada: Green's Publishing Company.

Greene, R. L. (1997). Assessment of malingering and defensiveness on multiscale inventories. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 169-207). New York: Guildord.

Greenberg, L., Kindschi, C., Dupuy, T., & Corman, C. (1996). *Test of Variables of Attention: Clinical guide*. Los Alamitos, CA: Universal Attention Disorders.

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

Harrison, A. G., & Edwards, M. J. (2010). Symptom exaggeration in post-secondary students: Preliminary base rates in a Canadian Sample. *Applied Neuropsychology, 17,* 135-143.

Harrison, A. G., Edwards, M. J., & Parker, K. C. (2007). Identifying students faking ADHD: Preliminary findings and strategies for detection. *Archives of Clinical Neuropsychology, 22*, 577-588.

Harrison, A. G., Rosenblum, Y. & Currie, S. (2010). Examining unusual digit span performane in a population of postsecondary students assessed for academic difficulties. *Assessment, 17,* 283-293. DOI: 10.1177/1073191109348590.

Harp, J. P., Jasinski, L. J., Shandera-Ochsner, A. L., Mason, L. H., & Berry, D. T. R. (2011). Dectetion of malingered ADHD using the MMPI-2-R.F. Psychological Injury and Law, 4, 32-43.

Haws, S. W., & Boccaccini, M. T. (2009). Detection of overreporting of psychopathology on the Personality Assessment Inventory: A meta-analytic review. *Psychological Assessment, 21*(1), 112-124.

Hayes, J. S. (1998). *A validation study of the Assessmnet of Dissiumuaiton Scale using criminally and civilly committed participants*. Unpublished doctoral dissertation, Louisiana State Univeristy, Baton Rouge, Louisiana.

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.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., & Millis, S. R. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. The Clinical Neuropsychologist, 23, 1093–1129. doi:10.1080/13854040903155063

Heiligenstein, E., Conyers, L. M., Berns, A., & Smith, M. A. (1998). Preliminary Normative Data of DSM-IV Attention Deficit Hyperactivity Disorder in College Students. *Journal of Americn College Health, 46*(4), 185-188.

Henry, G. K., Heilbronner, R. L., Mittenberg, W., Enders, C., & Stanczak, S. R. (2008). Comparison of the Lees-Haley Fake Bad Scale, Henry-Heilbronner Index, and Restructured Clinical Scale 1 in identifying noncredible symptom reporting. *The Clinical Neuropsychologist, 22*, 919-929.

Henry, G. K., Heilbronner, R., Mittenberg, W., & Enders, C. (2006). the Henry-Heilbronner Index: A 15-item empirically derived MMPI-2 subscale for identifying probable malingering in personal injury litigants and disability claimants. *The Clinical Neuropsychologist, 20*, 786-797.

Hervey, A. S., Epstein, J. N., & Curry, J. F. (2004). The neuropsychology of adults with attention deficit hyperactivity disorder: A meta-analytic review. *Neuropsychology, 18*, 485-503.

Hill, B. D., Pella, R. D., Singh, A. N., Jones, G. N., & Gouvier, W. D. (2009). The Wender Utah Rating Scale Adult ADHD Diagnostic Tool or Personality Index?. *Journal of Attention Disorders*, *13*(1), 87-94.

Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice methodfor the detection of malingering. *Journal of Clinical and ExperimentalNeuropsychology, 11,* 967–974.

Hopwood, C., Orlando, M., & Clark, T. S. (2010). The dection of malingered pain-related disability with the personality assessment inventory. *Rehabilitation Psychology, 55*(3), 307-310.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D. G., Edwards, C. L., & Smith, G. T. (1998). Development and initial validation of a newprocedure for evaluating adequacy of effort given during neuropsychologicaltesting: The Letter Memory Test. *Psychological Assessment, 10,*128–149.

Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. Applied Neuropsychology, 13(2), 77-90.

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.

Jachimowicz, G. &. Geiselman, R. E. (2004). Comparision of Ease of Falsification of Attention Defict Hyperactivity Disorder Diagnosis Using Standard beahioral rating scales. *Cogntive Science Online, 2*, 6-20.

Jasinski, L. J., Harp, J. P., Berry, D. T. R., Shandera-Ochsner, A. L., Mason, L. H., & Ranseen, J. D. (2011). Using sympotm validity tests to detect malingered ADHD in college students. *The Clinical Neuropsychologist, 25(8),* 1415-1428.

Kessler, R. C., Adler, L., Barkley, R., Biederman, J., Conners, C. K., Demler, O., ... & Zaslavsky, A. M. (2006). The prevalence and correlates of adult ADHD in the United States: results from the National Comorbidity Survey Replication. *The American journal of psychiatry*, *163*(4), 716-723.

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

Kucharski, L. T., Toomey, J. P., Fila, K., & Duncan, S. (2007). Detection of malingering of psychiatric disorder with the Personality Assessment Inventory: An investigation fo criminal defendants. *Jouranl of Personality Assessment, 88*(1), 25-32.

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

Larrabee, G. J. (2007). *Assessment of malingered neuropsychological deficits*. Oxford University Press, USA.

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

Larrabee, G. J. (2012b). Response to Bigler. *Journal of the International Neuropsychological Society, 18(4),* 630-631.

Latham, P. S. (1996). *Documentation and the law for professionals concerned wiht ADD/LD and those they serve.* Washington, D. C.: Latham and Latham.

Leark, R. A., Dixon, D., Hoffman, T., & Huynh, D. (2002). Fake bad test response bias effects on the test of variables of attention. *Archives of Clinical Neuropsychology*, 17, 335-342.

Lees-Haley, P. R. (1991). A fake bad scale for the MMPI-2 for personal injury claimants. *Psychological Reports, 68*, 203-210.

Mannuzza, S., Klein, R. G., Klein, D. F., Bessler, A., & Shrout, P. (2002). Accuracy of adult recall of childhood attention deficit hyperactivity disorder. *American Journal of Psychiatry, 159*, 1882-1888.

Marchetta, N. D., Hurks, P. P., Jolles, J., & Krabbendam, L. (2008). Interference control, working memory, concept shifting and verbal fluency in adult with Attention-deficit/hyperactivity disorder (ADHD). *Neuropsychology, 22*(1), 74-84.

Marks, P. A., Seeman, W., & Haller, D. L. (1974). *The actuarial use of the MMPI with adolescents and adults*. Williams & Wilkins.

Marshall, P. S. (2010). Effectiveness of symptom validity measures in identifying cognitive and behavioral symptom exaggeration in adult attention deficit hyperactivity disorder. *The Clinical Neuropsychologist, 24*(7), 1204-1237.

Martin, R. C., Hayes, J. S., & Gouvier, W. D. (1996). Differential vulnerability between postconcussion self-report and objective malingering tests in identifying simulated mild head injury. *Journal of Clinical andExperimental Neuropsychology, 18(2),* 265-275.

McCabe, S. E., Teter, C. J., & Boyd, C. J. (2006). Medical use, illicit use and diversion of prescription stimulant medication. *Journal of Psychoactive Drugs, 38*(1), 43-56.

McKee, T. E. (2008). Comparison of a norm-based versus criterion-based approach to measuring ADHD symptomatology in college students. *Journal of Attention Disorders, 11, I677-688.*

Merten, T., Bossink, L., & Schmand, B. (2007). On the limits of effort testing: Symptom validity tests and severity of neurocognitive symptoms in nonlitigant patients. *Journal of Clinical and Experimental Neuropsychology, 29(3),* 308-318.

Meyers, J. E., & Volbrecht, M. E. (2003). A validation of multiple malingering detection methods in a large clinical sample. *Archives of Clinical Neuropsychology, 18*, 261-276.

Miller, H. A. (2001). *MFAST: Miller Forensic Assessment of Symptoms Test professional manual.* Odessa, FL: Psychological Assessment Resources, Inc.

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

Mogge, N. L., Lepage, J. S., Bell, T., & Ragatz, L. (2010). the negative distortion scale: a new PAI validity scale. *Journal of Forensic Psychiatry & Psychology, 21*(1), 77-90.

Morey, L. C. (1991). *Personality Assessment Inventory: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Morey, L. C. (1993). Defensiveness and malingering indices for the PAI. *Paper presented at the meeting of the American Psychological Association.* Toronto, Canada.

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

Morey, L. C., & Lanier, V. W. (1998). Operating characteristics of six response distortion indicators for the Personality Assessment Inventory. *Assessment, 5*(3), 203-214.

Murphy, K. & Barkley, R. A. (1996a). Attention deficit hyperactivity disorder in adults: Comorbidities and adaptive impairments. *Comprehensive Psychiatry, 37,* 393-401.

Murphy, K., & Barkley, R. A. (1996b). Updated adult norms for the ADHD Behavior Checklist for adults. *The ADHD Report*, 12–16.

Musso, M. W., Barker, A. A., Jones, G. N., Roid, G. H., & Gouvier, W. D. (2011). Development and validation of the Stanford binet-5 rarely missed items-nonverbal index for the detection of malingered mental retardation. *Archives of clinical neuropsychology*, *26*(8), 756-767.

Musso, M. W. & Gouvier, W. D. (2012). "Why Is This So Hard?": A Review of Detection of Malingered ADHD in College Students. *Journal of Attention Disorders*.

Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Nueropsychological Society, 13*, 440-449.

Norwalk, K., Norvilitis, J. M., & MacLean, M. G. (2009). ADHD symptomatology and its relationship to factors associated with college adjustment. *Journal of Attention Disorder, 13*(3), 251-258.

Pella, R. D., Hill, B. D., Shelton, J. T., Elliott, E., & Gouvier, W. D. (2012). Evaluation of embedded malingering indices in a non-litigating clinical sample using control, clinical, and derived groups. *Archives of Clinical Neuropsychology, 27(1),* 45-57.

Price, J. R. (1995). Identificaiton of malingering and symtpom exaggeration. Workshop presetned at the 15th annual conference of the National Academy of Neuropyschology. November 4, 1995.

Quinn, C. A. (2003). Detection of malingerin gin assessment of adult ADHD. *Archives of Clinical Neuropsychology, 18*, 379-395.

Rabiner, D. L., Anastopoulos, A. D., Costello, E. J., McCabe, S. E., & Swartzwelder, H. S. (2009). The misuse and diversion of prescribed ADHD medications by college students. *Journal of Attention Disorders, 13*(3), 144-153.

Rapport, L. J., VanVoorhis, A., Tzelepis, A., & Friedman, S. R. (2001). Executive Functioning in Adult Attention-Deficit Hyperacitiviy Disorder. *The Clinical Neuropsychologist, 15*(4), 479-491.

Reitan. R. M. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology, 19*, 393-394.

Reitan. R. M. & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery: \ Theory and interpretation.* Tucson, AZ: Neuropsychology Press.

Resnick, R. J. (2005). Attention deficit hyperactivity disorder in teens and adults: They don't all outgrow it. *Journal of Clinical Psychology, 61*(5), 529-533.

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

Riccio, C. A., Wolfe, M., Davis, B., Romine, C., George, C., & Lee, D. (2005). Attention Deficit Hyperactivity Disorder: Manifestation in adulthood. *Archives of Clinical Neuropsychology, 20*, 249-269.

Rogers, R. (1990). Models of Feigned Mental Illness. *Progression Psychology: Research and Practice, 21(3),* 182-188.

Rogers, R. (Ed.). (1997). *Clinical assessmnet of malingering and deception* (2nd ed.). New York: Guildford.

Rogers, R. (2008). *Clinical Assessment of Malingering and Deception* (3rd ed.). New York, New York: Guilford.

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

Rogers, R., Sewell, K. W., Morey, L. C., & Ustad, K. L. (1996). Detection of feigned mentl disorders on the Personality Assessment Inventory: A discriminant analysis. *Journal of Personality Assessment, 67*, 629-640.

Root, J., Robbins, R., Chang, L., & van Gorp, W. (2006). Detection of inadequate effort on the California Verbal Learning Test (2nd ed.): Forced choice recognition and critical item analysis. *Journal of the International Neuropsychology Society, 12,* 688-696.

107

Rosselli, M., Ardila, A., Santisi, M., Bojan, D., Wolowitz, B., & Delaney, R. (2000). ADHD symptoms and neuropsychological test scores in a non-clinical college population. *Archives of Clinical Neuropsychology, 15*, 776.

Rule, W.R., & Traver, M.D. (1983). Test-retest reliabilities of State-Trait Anxiety Inventoryin a stressful social analogue situation. *Journal of Personality Assessment.* (47,3), 276-277.

Sandford, J. A., & Turner, A. (1995). *Intermediate visual and auditory Continuous Performance Test interpretation manual.* Richmond, VA: Braintrain.

Schipper, L. J., Berry, D. T. R., Coen, E., & Clark, J. A. (2008). Cross-validation of a manual from of the Letter Memory Test using a known groups methodology. *Clinical Neuropsychologist, 22,* 345–349.

Schoechlin, C., & Engel, R. R. (2005). Neuropsychological performance in adult attention-deficit hyperactivity disorder: Meta-analysis of empirical data. *Archives of Clinical Neuropsychology, 20*, 727-744.

Schroeder, R., & Marshall, P. (2010). Validation of the sentence repetition test as a measure of suspect effort. *The Clinical Neuropsychologist, 24,* 326-343.

Shipley, W. C. (2004). *Shipley Institue of Living Scale, Second Edition*. Ottawa, Ontario, Canada: MultiHealth Systems.

Slick, D., Hopp, G., Strauss, E., & Thompson, G. B. (1997). *VSVT: Victoria Symptom Validity Test, Version 1.0, Professional Manual.* Odessa, FL: Psychological Assessment Resources, Inc.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13(4),* 545-561.

Sobanski, E., Brüggemann, D., Alm, B., Kern, S., Deschner, M., Schubert, T., ... & Rietschel, M. (2007). Psychiatric comorbidity and functional impairment in a clinically referred sample of adults with attention-deficit/hyperactivity disorder (ADHD). *European archives of psychiatry and clinical neuroscience*, *257*(7), 371-377.

Sollman, M. J., Ranseen, J. D., & Berry, D. T. R. (2010). Detection of feigned ADHD in college students. *Psychological Assessment, 22*(2), 325-335.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). The state-trait anxiety inventory. *Palo Alto, Calif: Consulting Psychologists Press Inc*.

Spielberger, C.D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G.A. (1983).*Manual for the State-Trait Anxiety Inventory (Form Y).*Palo Alto, CA: Consulting Psychologist Press.

Steffan, J. S., Clopton, J. R., & Morgan, R. D. (2003). An MMPI-2 scale to detect malingered depression (Md scale). *Assessment*, *10*(4), 382-392.

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

Suhr, J. A., Buelow, M., & Riddle, T. (2011). Development of an Infrequency Index for the CAARS. *Jouranl of Psycholoeducational Assessment, 29*(2), 160-170.

Suhr, J., Hammers, D., Dobbins-Buckland, K., Zimak, E., & Hughes, C. (2008). The relationship of malingering test failure to self-reported symptoms and neuropsychological findings in adults referred for ADHD evaluation. *Archives of Clinical Neuropsychology, 23*, 521-530.

Suhr, J. A., Sullivan, B. K., & Rodriguez, J. L. (2011). The relationship of noncredible performance to Continuous Performance Test Scores in adults referred for Attention-Deficits/Hyperactivity Disorder evaluation. *Archives of Clinical Neurospychology, 26*(2), 1-7.

Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in Attention-Deficit Hyperactivity Disorder and Learning Disorder Assessments. *Applied Neuropsychology, 14(3),* 189-207.

Teter, C. J., McCabe, S. E., Cranford, J. A., Boyd, C. J., & Guthrie, S. K. (2005). Prevalence and motives for illicit use of prescription stimulants in an undergraduate student sample. *Journal of American College Health, 53*(6), 253-262.

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

Triolo, S. J., & Murphy, K. R. (1996). *Manual for scoring and interpretation.* New York: Mazel.

Tucha, L., Sontag, T. A., Walitza, S., & Lange, K. W. (2009). Detection of malingered attention deficit hyperactivity disorder. *ADHD Atten Deficit Hyperactivity Disorder, 1*, 47-53.

Tucha, L., Tucha, O., Laufkotter, R., Walitza , S., Klein, H. E., & Lange, K. W. (2008). Neuropsychological assessment of attention in adults iwth different subtypes of attention-deficit/hyperactivity disorder. *Journal of Neural Transmission, 115*, 269-278.

Walters, G. D., Rogers, R., Berry, D. T., Miller, H. A., Duncan, S. A., McCusker, P. J., ... & Granacher Jr, R. P. (2008). Malingering as a categorical or dimensional construct: The latent structure of feigned psychopathology as measured by the SIRS and MMPI-2. *Psychological assessment*, *20*(3), 238.

Ward, M., Wender, P., & Reimherr, F. (1993). The Wender Utah Rating Scale: An aid in retrospective diagnosis of children with ADHD. *American Jouranl of Psychiatry, 160*, 245-256.

Wasserstein, J. (2005). Diagnostic issues for adolescents and adults with ADHD. *Journal of Clinical Psychology, 61*(5), 535-547.

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

Wechsler, D. (1997b). *WMS-III Administration and Scoring Manual.* San Antonio, Texas: The Psychological Corporation.

Weyandt, L. L., & DuPaul, G. J. (2006). ADHD in college students. *Journal of Attention Disorders, 10*, 9-19.

Weyandt, L. L., & DuPaul, G. J. (2008). ADHD in college students: Developmental findings. *Developmental Disabilities, 14*, 311-319.

Weyandt, L. L., Janusis, G., Wilson, K. G., Verdi, G., Paquin, G., Lopes, J., ... & Dussault, C. (2009). Nonmedical Prescription Stimulant Use Among a Sample of College Students Relationship With Psychological Variables. *Journal of Attention Disorders*, *13*(3), 284-296.

Weyandt, L. L., Linterman, I., & Rice, J. A. (1995). Reported Prevalence of Attentional Difficulties in a Gernal Sample of College Students. *Journal of Psychopathology and Behavioral Assessment, 17*(3), 293-304.

White, B. P., Becker-Blease, K. A., & Grace-Bishop, K. (2006). Stimulant medication use, misures, and abuse in an undergraduate and graduate student sample. *Journal of American College Health, 54*(5), 261-268.

White, T., & Stern, R. (2003). *Neuropsychological Assessment Battery.* Lutz, FL: Psychological Assessment Resources, Inc.

Whiteside, D., Clinton, C., Diamonti, C., Stroemel, J., White, C., Zimberoff, A., & Waters, D. (2010). Relationship between suboptimal cognitive effort and the clinical scales of the Personality Assessment Inventory. *The Clinical Neuropsychologist*, *24*(2), 315-325.

Whiteside, D. M., Dunbar-Mayer, P., & Water, D. P. (2009). Relationship between TOMM performance and PAI validity scales in a mixed clinical sample. *The Clinical Neuropsychologist, 23*, 523-533.

Wilens, T. E., Biederman, J., Faraone, S. V., Martelon, M., Westerberg, D., & Spencer, T. J. (2009). Presenting ADHD symptoms, subtypes, and comorbid disorders in clinically referred adults with ADHD. *The Journal of clinical psychiatry*, *70*(11), 1557.

Williams, E. D., Reimherr, F. W., Marchant, B. K., Strong, R. E., Halls, C., Soni, P., ... & Robison, R. J. (2010). Personality disorder in ADHD Part 1: Assessment of personality disorder in adult ADHD using data from a clinical trial of OROS methylphenidate. *Annals Clinical Psychiatry*, *22*, 84-93.

Wilkinson, G. S. (1993). *The Wide Range Achievement Test-3*. San Antonio, TX: The Psychological Corporation.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). The Woodcock-Johnson III. Itasca, IL: Riverside.

Woods, S. P., Lovejoy, D. W., Stutts, M. L., Ball, J. D., & Fals-Stewart, W. (2002). Comparative efficiency of a discrepancy analysis for the classification of Attention-Deficit Hyperactivity Disorder in Adults. *Archives of Clinical Neuropsychology, 17*, 351-369.

Young, C. J., & Gross, A. M. (2011, April). Detection of Response Bias and Noncredible Performance in Adult Attention-Deficit/Hyperactivity Disorder. *Archives of Clinical Neuropsychology, 26*(2), 165-172.

Youngjohn, J. R., Wershba, R., Stevenson, M., Sturgeon, J., & Thomas, M. L. (2011). Independent validation of the MMPI-2-RF Somatic/Cognitive and Validity scales in TBI litigants tested for effort. *The Clinical Neuropsychologist, 25*(3), 463-476.

APPENDIX A: IRB APPROVAL FORM

ACTION ON PROTOCOL APPROVAL REQUEST

Institutional Review Board
Dr. Robert Mathews, Chair
131 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.6792
irb@lsu.edu | lsu.edu/irb

LSU

TO: Drew Gouvier
Psychology

FROM: Robert C. Mathews
Chair, Institutional Review Board

DATE: September 22, 2011
RE: **IRB#** 3209

TITLE: Detecting Malingered ADHD in College Students Using the Personality Assessment Inventory

New Protocol/Modification/Continuation: New Protocol

Review type: Full ___ Expedited _X_   **Review date:** 9/19/2011

Risk Factor: Minimal _X_ Uncertain _____ Greater Than Minimal_____
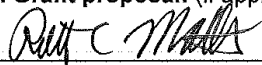
Approved_____X_____ **Disapproved**_____

Approval Date: 9/22/2011   **Approval Expiration Date:** 9/21/2012

Re-review frequency: (annual unless otherwise stated)

Number of subjects approved: 400

Protocol Matches Scope of Work in Grant proposal: (if applicable)_____

By: Robert C. Mathews, Chairman _____

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. SPECIAL NOTE:
All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb

APPENDIX B
DIAGNOSTIC STATISTICS ADJUSTED FOR CRITERION VALIDITY

Method

Typically, sensitivity and specificity are calculated using a 2X2 table (denoted in Table 12), but this study performed additional calculations to adjust for imperfect diagnostic accuracy (depicted in dark grey boxes in Table 13). Formulas are labeled accordingly. For the purposes of this study, diagnostic accuracy was calculated by assuming that 25% of the individuals identified as malingerers were correctly identified by the newly developed newly developed scale, but were not detected by clinicians. The method used in depicted in Table 3. The specificity of the newly developed scale was imperfect meaning that not of the 25% of malingerers in the database would have been identified by the scale. Therefore, specificity of the measure was multiplied by the theoretical number malingerers to calculate "accurately identified" clinical malingerers (E in the table below). Similarly, the number of individuals correctly identified was multiplied by 1- the specificity to account for the percentage of individuals in the database that would not have been detected by the newly developed newly developed scale (F in Table 3). It should be noted that this is theoretical and not to be used in clinical decision making without further scientific investigation.

Table 13. Formulas for calculating sensitivity and specificity adjusting for criterion validity.

	ADHD Simulators	Honest Clinical Group	Clinical Malingerers[a]	Adjusted Formulas
Malingering	a	c	E=(c*Error rate)*SP	SN=(a+E)/(a+b+E+F)
Not Malingering	b	d	F=(c*Error rate)*(1-SP)	SP = (d+F)/(c+d+E+F)
Customary Formulas	SN = A/A+B	SP=D/C+D		

[a]newly proposed method accounting for diagnostic accuracy in calculating sensitivity and specificity.

Results

The diagnostic statistics presented in Hypotheses three, four, and five were re-analyzed, assuming that 25% of the individuals identified by the newly developed scale were actually malingering ADHD at the time of the evaluation and were correctly identified by the scale. Sensitivity and specificity for the groups adjusting for this error rate are presented in Table 14, and PPP and NPP adjusting for the error rate are presented in Table 15. The diagnostic statistics changed very little with slight improved sensitivity for some groups; however, for the most part, changes were within one percentage point.

Table 14. Sensitivities and Specificities of group at varying cutoff scores assuming the clinical diagnostic classification error rate was 25% in each group.

	Cutoff > 14	Cutoff > 16	Cutoff > 18	Cutoff > 20	Cutoff > 22
Specificity	**82**	**86**	**90**	**93**	**96**
Clinical ADHD for Development	91	95	96	98	100
Clinical ADHD Hold-out Sample	92	96	96	96	100
No Diagnosis	95	98	99	100	100
Psychopathology	68	75	82	86	91
Comorbid ADHD & PSYC	67	73	78	86	93
Control	85	91	92	92	93
Prospective ADHD	76	83	86	89	93
Sensitivity	*71*	*66*	*59*	*55*	*46*
ADHD Simulators for Development	*77*	*70*	*65*	*55*	*52*
ADHD Simulators Hold-out Sample	*67*	*60*	*54*	*48*	*44*

Table 15. Positive and Negative Predictive Power of various cutoff scores for the newly developed scale adjusting for a 25% diagnostic error rate.

	PPP = 0.14	**PPP = 0.25**	**PPP= 0.50**	**NPP = 0.14**	**NPP = 0.25**	**NPP= 0.50**
Cutoff > 14	39	57	80	52	41	22
Cutoff > 16	45	62	83	59	47	26
Cutoff > 18	49	67	86	64	52	29
Cutoff > 20	53	70	87	68	56	32
Cutoff > 22	67	80	93	72	59	33

VITA

Mandi Wilkes Musso was reared in Simmesport, Louisiana by her father, Marcus Wilkes, and mother Wanda Wilkes. Mandi performed well in school and participated in multiple extracurricular activities. At the age of 13, Mandi decided that she wanted to be a Clinical Psychologist, a career path that she pursued when she began taking classes at the University of Louisiana at Lafayette in 2002.

While at the University of Louisiana at Lafayette, Mandi was an active member of Psi Chi and was elected President in 2005. She also worked as a research assistant at the New Iberia Research Center in New Iberia, Louisiana, completing an Honors Thesis on self-injurious behavior in Rhesus Macaques. She graduated Summa Cum Laude with an Honors Bachelors of Science degree in psychology and a minor in biology in 2006. She also received the Hait Lewis Award for Academic Excellence. That same year, she was married to M. Dustin Musso.

After graduating in 2006, Mandi worked full time as a research associate at the New Iberia Research Center while continuing to take classes in biology and chemistry. Mandi was accepted into Dr. Wm. Drew Gouvier's Neuropsychology Lab, and began taking graduate courses at Louisiana State University in 2008. She successfully defended her master's thesis which focused on developing an embedded effort index for the Stanford Binet-5 to detect feigned intellectual disability. Her general examination involved a case with HIV dementia, tertiary neurosyphillis, and cerebrovascular disease. She is currently completing her internship at Warren Alpert Medical School of Brown University in Providence, Rhode Island. Her primary research interests include neuropsychology and malingering assessment.