THE EFFECTS OF INCENTIVIZING ADHD RELATED NON-CREDIBLE RESPONDING ON NEUROPSYCHOLOGICAL MEASURES AND PERFORMANCE VALIDITY TESTS

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Psychology

Jesse A. Siegel

Indiana University of Pennsylvania

August 2019

Indiana University of Pennsylvania
School of Graduate Studies and Research
Department of Psychology

We hereby approve the dissertation of

Jesse A. Siegel

Candidate for the degree of Doctor of Psychology

Laura Knight, Ph.D.
Assistant Professor of Psychology, Advisor

Margaret Reardon, Ph.D.
Associate Professor of Psychology

Anthony Perillo, Ph.D.
Assistant Professor of Psychology

ACCEPTED

Randy L. Martin, Ph.D.
Dean
School of Graduate Studies and Research

Title: The Effects of Incentivizing ADHD Related Non-Credible Responding on Neuropsychological Measures and Performance Validity Tests

Author: Jesse A. Siegel

Dissertation Chair: Dr. Laura Knight

Dissertation Committee Members: Dr. Margaret Reardon
Dr. Anthony Perillo

This study evaluated the effects of a contingent reward on college students' ability to feign Attention-Deficit/Hyperactivity Disorder (ADHD) on neuropsychological measures while avoiding detection on performance validity tests (PVT). It also sought to ascertain strategies used by college students to feign ADHD. This investigation extends previous research on malingering by including both incentivized and unincentivized experimental groups to establish the impact of incentivizing ADHD-related non-credible responding (NCR) on neuropsychological test performance.

Participants were randomly assigned to a control group or one of three experimental groups. Experimental groups included an unincentivized group and two incentivized groups, offered either $10 on the spot or entry into a $100 raffle, both of which were conditional upon participants' ability to successfully feign ADHD. Participants were administered five neuropsychological measures, three PVTs, and an ADHD self-report by an examiner blind to participant condition. Participants were asked an open-ended question about strategies used to feign ADHD. A total of 68 undergraduate students enrolled in an introductory psychology course participated in this study.

Results from this study showed that an incentive did improve participants' ability to feign ADHD compared to no incentive; however, most participants failed to successfully feign ADHD. Although participants were generally able to produce scores suggestive of ADHD on

neuropsychological measures and an ADHD self-report, most participants failed to avoid detection on PVTs. Participants in this study reported feigning strategies related to inattentive behavior, fidgety behavior, and giving slow responses, which is consistent with strategies reported in other studies, although methodological differences complicated efforts to compare results.

TABLE OF CONTENTS

Chapter Page

1 REVIEW OF LITERATURE 1

Introduction 1
Non-Credible Responding 5
Malingering 6
Effort 8
Inadvertent NCR 8
Performance Validity Tests 10
PVT Sensitivity and Specificity 12
Specific PVTs and SVIs 13
Test of memory malingering 13
21-Item test 15
Word memory test and medical symptom validity test 15
Embedded PVTs 16
The MMPI-2 17
Reliable digit span and advanced clinical solutions 18
Trail making test 18
Rey-Osterrieth complex figure task 19
Continuous performance tests 20
Issues Related to PVTs 21
Summary of PVTs 23
ADHD 23
ADHD in College Students 25
Neuropsychological Assessment of ADHD 27
ADHD Neuropsychological Battery 27
The tower test 27
The Stroop color-word test 28
Standard ADHD neuropsychological battery 28
Efficacy of Neuropsychological Measures in Diagnosing ADHD 30
Non-Credible Responding and ADHD 31
Strategies Used to Feign ADHD 33
Detecting NCR Within the ADHD Assessment 33
Limitations to Current Research 36
Present Study 38

2 METHODS 39

Participants and Recruitment 39
Measures 40

Chapter Page

Prescreen Measure 40
Demographics 40
Digit Span (Wechsler, 2008) 40
21-Item Test (Iverson et al., 1991) 41
Test of Memory Malingering (Tombaugh, 1997) 42
Trail Making Test (Partington & Leiter, 1949) 42
Continuous Performance Test (Rosvold et al., 1956) 43
Wender Utah Rating Scale (Ward, 1993) 44
Procedures 45
Hypotheses 48
Hypothesis 1 48
(i) 48
(ii) 48
(iii) 48
Hypothesis 2 48
(i) 49
(ii) 49
(iii) 49
Hypothesis 3 49
Hypothesis 4 50

3 RESULTS 51

Descriptive Analyses 51
Sensitivity and Specificity 53
Test of Memory Malingering 54
Reliable Digit Span 54
21-Item 55
Overall Sensitivity and Specificity 55
Results of Hypotheses 55
Neuropsychological Measures (Hypothesis 1) 55
Performance Validity Tests (Hypothesis 2) 60
Sophistication (Hypothesis 3) 64
Strategies (Hypothesis 4) 65

4 DISCUSSION 70

Overview of Findings 70
Strengths and Limitations 85
Future Directions 90
Conclusion 91

Chapter Page

REFERENCES 93

APPENDICES 110

Appendix A - Prescreening Measure: ASRS-V1.1 Symptom Checklist (Gray et al., 2014) 110
Appendix B- Demographic Questionnaire: Online Survey 111
Appendix C- Trail Making Test A and B (Partington & Leiter, 1949) 112
Appendix D- Continuous Performance Test (Rosvold et al., 1956) 115
Appendix E- Wender Utah Rating Scale (Ward, 1993) 117
Appendix F- Informed Consent Form 119
Appendix G- Prompt for Control Group 121
Appendix H- Prompt for Unincentivized Group 122
Appendix I- Prompt for Incentivized-Immediate Group 123
Appendix J- Prompt for Incentivized-Delayed Group 124
Appendix K- Attention Deficit Hyperactivity Disorder Diagnostic Criteria 125
Appendix L- Affirmation of Instructions 127
Appendix M- Strategy Record Form 128
Appendix N- Debriefing Form 129

LIST OF TABLES

Table		Page
1	Demographic Characteristics of Participants	51
2	Mean and Standard Deviation by Measure and Group	52
3	Tukey HSD of Control Group Versus NCR Groups on Neuropsychological Tests	58
4	Participants in Sophistication Groups	65
5	Participants Meeting Sophistication Criterion	65
6	Percentage of NCR Participants Reporting ADHD Symptoms (%)	68
7	NCR Participants Reporting Specific Errors (%)	69
8	Percent Reporting Specific Strategies in Harrison et al. (2007) Versus Current Study	69

LIST OF FIGURES

Figure Page

1 Digit span (forward & backward) and trail making test (A & B) by mean group t-score 59

2 CPT mean error rate by group 60

3 Mean TOMM score by group 62

4 Mean reliable digit span score by group 63

5 Mean 21-Item test score by group 63

CHAPTER 1

REVIEW OF LITERATURE

Introduction

Non-credible responding (NCR) occurs within the course of a neuropsychological assessment when an examinee disingenuously responds to questions or puts forth sub-optimal effort on a performance-based measure, with varying degrees of deliberateness, such that the scores derived from the measures do not contribute to an accurate depiction of the examinee's neuropsychological functioning (Suhr, 2015). It is an umbrella term that encompasses malingering (i.e., deliberate NCR) for the purpose of manipulating the outcome of litigation or receiving misappropriated accommodations or medication. NCR may also occur inadvertently because an examinee believes they are impaired (Suhr & Gunstad, 2005) or is distracted (Suhr & Wei, 2013). NCR is a barrier within the field of neuropsychological assessment because it masks the true neuropsychological functioning of an examinee (Larrabee, 2014).

Historically, NCR has been problematic when psychological assessments are utilized in forensic settings (Rogers & Granacher, 2011). Research suggests that it is a frequent occurrence, estimated to some extent in 22% of all psychological evaluations (Locke, 2008). NCR also presents problems in non-forensic settings, when examinees stand to gain misappropriated accommodations (e.g., disability services or funds) and/or medication (e.g., narcotic analgesics and/or psychostimulants). In the past 15 years, researchers have become aware that this problem extends into the neuropsychological assessment of Attention Deficit Hyperactivity Disorder (ADHD), where estimates show that 27 to 57% of college students engage in NCR during ADHD evaluations (Marshall et al., 2016; Sullivan et al., 2007).

ADHD is a neurological developmental disorder that emerges in childhood and is most frequently first diagnosed in elementary school-age children (APA, 2013). According to the American Psychiatric Association (APA), ADHD affects about 5% of children and 2.5% of adults. This disorder is characterized by inattentive and/or hyperactive symptoms. Hyperactive symptoms manifest differently in adults compared to children, and often include fidgetiness and restlessness. It is a behaviorally-based diagnosis (APA, 2013) rendered through observation, self-report, informant-report, and school records (Barkley, 1997).

Research suggests that ADHD diagnoses are on the rise. In the 1970s, the APA recognized that the symptoms of ADHD persist into adulthood. As a result, the number of adults meeting criteria for ADHD grew 3-fold between 1994 and 1997, and another 3-fold between 2002 and 2007, with most new diagnoses given to individuals ages 18-24 years (Montejano et al., 2011). College campuses across the USA have also seen an increase in the number of students with ADHD (Musso & Gouvier, 2014).

Neuropsychologically, ADHD is characterized by deficits in executive functions (EF), which refer to a collection of neuropsychological abilities that allow a person to appropriately respond to novel situations (Chan et al., 2008). Impairments in processing speed and attentional abilities are also common among individuals with ADHD (Adams et al., 1996). However, the severity of these neuropsychological deficits is inconsistent among individuals with ADHD, and may be indistinguishable from deficits due to other neurological impairments (Barkley & Murphy, 2010; Conant, 2014). While low scores on measures of EF are useful for quantifying the degree of the attentional impairment and imparting the strengths and weaknesses of an individual with ADHD, they are not specific to ADHD (Pettersson et al., 2018).

By law, students with ADHD are offered accommodations through university disabilities offices (Americans with Disabilities Act of 1990). These accommodations include, but are not limited to, extra test-taking time and priority registration. Students with ADHD may also be prescribed psychostimulant medication (Suhr & Wei, 2013; Weyandt et al., 2017). However, when these accommodations and medications are misappropriated through NCR to students without ADHD, they provide an unfair advantage.

Unsurprisingly, abuse and misuse of psychostimulant medications are on rise across college campuses, increasing from 6.2% in 2000 to 9.3% in 2011 (Rabiner, 2013). Medications may be obtained illicitly or directly from a prescriber. In a sample of 4,297 adults who admitted to illicitly using prescription psychostimulants, 20% admitted that they had engaged in NCR with a healthcare provider to gain access to these medications (Novak et al., 2007). To qualify for ADHD-related accommodations, students are often referred for neuropsychological evaluations, where they may engage in NCR. Likewise, NCR is estimated to occur within 25-50% of these assessments (Suhr et al., 2008; Sullivan et al., 2009)

Engagement in NCR requires a strategy, and several researchers have explored the methods that students use to feign ADHD in an effort to improve NCR detection methods (Frazier et al., 2008; Harrison et al., 2007; Quinn, 2003). One of these methods is the utilization of specialized measures referred to as Performance Validity Tests (PVTs; Larrabee, 2014). PVTs may be stand-alone, individually administered measures (Green, 2005) or embedded within a neuropsychological measure used as part of the evaluation (Schroeder et al., 2012). PVTs vary in their specificity, with the most effective ones correctly identifying NCR more than 85% of the time (Green, 2004; Suhr et al., 2008). The use of PVTs during an adult ADHD assessment has become a recommended standard practice (Conant, 2014).

The research regarding the utility of PVTs in ADHD evaluations has focused exclusively on simulated malingerers in a college sample. In most of these studies, all participants were rewarded, typically with research participation credit, regardless of the outcome (Booksh et al., 2010; Frazier et al., 2008). This a problem, because the ecological validity of this research may be compromised by the lack of incentive for students to successfully evade detection and feign ADHD. There are currently two studies that have offered extra incentives to students who can successfully evade detection (Fisher, 2007; Sollman et al., 2010). In both studies, however, all members of the simulated NCR groups were told that they would receive a reward if they remained undetected. Since this NCR group lacked a unincentivized comparison condition to control for the effects of this incentive, it is not known whether these incentives produced differences in participants' NCR style and test scores. Because there is considerable gain for successful feigning that is not present in most ADHD feigning studies, the current body of research potentially lacks real-life applicability to genuine NCR ADHD groups.

The current study sought to add an external incentive and use a comparable NCR group to determine whether there are differences in performance between these groups. By offering additional incentives to a subgroup of participants randomly assigned to engage in NCR and comparing this group's performance to that of an incentivized as usual NCR group (i.e., research participation credit for class), the current study examined differences in scores on PVTs (stand-alone and embedded) and commonly used neuropsychological measures in the assessment of ADHD. Differences between these groups may suggest that the ecological validity of future ADHD-related NCR studies can be improved by adding an external incentive to the NCR group.

This study utilized four groups: a control group, an extra-incentivized NCR group, and an incentivized-as-usual NCR group. Participants were selected from the same subject pool and

randomly assigned to a group. The examiner was blind to all conditions. All experimental groups were told to malinger ADHD; however, the extra-incentivized group was told, additionally, that they may receive a reward for evading detection on the PVTs. The control group was administered the same battery given to both experimental groups without any additional instructions.

In addition, after this experiment, students were surveyed regarding the strategies that they used to engage in NCR. The goals of the current research were to determine whether (1) an incentive produced disparate scores on neuropsychological measures between groups, (2) an incentive produced disparate scores on PVTs, (3) students' qualitative reports regarding their strategy differs between NCR groups. A difference between these groups may suggest that future studies should use incentivized groups to increase the ecological validity of their findings.

Non-Credible Responding

According to Suhr et al. (2008), non-credible responding (NCR) includes any behavior that may occur during the course of a psychological assessment that results in inaccurate scores. NCR may occur on self-report and/or performance-based measures. Morgan and Boone (2008) described NCR as an umbrella term encompassing any behavior that belies an examinee's true score. According to Suhr (2015), individual differences complicate the detection of NCR because methods of engaging in NCR vary by examinee; what may look like NCR from one examinee may reflect the true ability of another examinee. Suhr also pointed out that because NCR is not solely a result of malingering, inaccurate results may arise from the examinee's perception that he/she has a cognitive impairment or may result from an examinee's over-exaggeration of symptoms to draw attention to a bona fide neurological disorder. However, since there is no secondary gain sought, this behavior cannot be labeled malingering, and is thus called NCR.

Research suggests that NCR is a frequent occurrence, estimated in 22% of psychological evaluations (Locke, 2008).

NCR lies on a spectrum that ranges from unintentional to deliberate. On the unintentional side of the spectrum, an examinee may give erroneous or distorted responses due to false self-perception (which may be accidental or the result of a psychological disorder) or misunderstood directions (Larrabee, 2012; Lezak, 2012). In some cases, symptom exaggeration may be iatrogenic, which occurs when a medical professional produces the perception of a psychological disorder within a patient. Patients may then act in accordance with their perception of the symptoms (Broshek et al., 2015).

Variables such as the potential outcome of the assessment should be taken into consideration in determining whether NCR may be an issue. This is often an issue when the outcome of the assessment may be utilized to justify academic accommodations, the prescribing of medication, and/or add to evidence in a forensic context (Larrabee, 2014). NCR should be suspected whenever examinees stand to gain something external as a result of the evaluation.

Malingering

Intentional NCR is referred to as 'malingering.' Malingering is defined by the American Psychiatric Association (2013) as "…the intentional reporting of symptoms for personal gain (e.g., money, time off work)…" (p. 326). Lipman (1961) described 4 types of malingering behavior: *Invention* denotes the complete fabrication of neuropsychological issues. *Perseveration* denotes the continuation of symptoms, even though the initial symptoms have resolved. *Exaggeration* describes the embellishment of genuine symptoms that are present; this often appears when an examinee is 'crying for help,' and wants to garner the attention of a clinician. Lastly, *transference* refers to genuine, severe, and current symptoms that resulted from

something other than what the individual is reporting. Malingering is frequently suspected during forensic assessment, as litigants are incentivized to win their respective court cases; thus, measures to detect malingering are a routine part of forensic assessment (Larrabee, 2014).

The Malingering Criteria Checklist (MCC) is a commonly used measure for determining the likelihood of malingering (Slick, 1999). The first criterion is the presence of a discernable secondary gain. The term 'secondary gain' denotes the benefits that one may reap from successfully exaggerated physical or psychological symptoms (Davidhizar, 1994). The second criterion includes evidence from neuropsychological testing, such as inconsistencies between scores from two or more instruments that assess the same construct, or between test scores and observed behavior during testing, medical history, or reported injuries, and/or the derivation of statistically improbable scores. The third criterion pertains to evidence from the examinee's self-report, such as discrepancies between information provided by the examinee and information gathered through documented history, known patterns of functioning, behavioral observations, reports from reputable informants, and/or evidence of exaggerated or feigned dysfunction. The second and third criteria must not be fully accountable to a bona fide neurological condition.

Malingering may differ from feigning a psychological disorder. The Diagnostic and Statistical Manual, 5th edition (DSM-5), the handbook used by health care professionals to diagnose mental disorders, includes *Factitious Disorder,* which is characterized by the same behaviors that one may display when malingering, without any discernible reward for such behavior (APA, 2013). Spreen et al. (2006) stated that once factitious disorder is ruled out (by the absence of any obvious external gain), the presence of litigation increases the possibility of malingering.

Overall, malingering refers to an extreme example of NCR, whereby examinees deliberately mislead examiners by performing sub-optimally/dishonestly on assessment measures and self-reports for personal gain. Barring the existence of a factitious disorder diagnosis, malingering should be suspected in all medicolegal-related litigation wherein the examinee stands to gain (or retain, as in personal freedom) something.

Effort

PVTs measure effort, and an examinee putting forth adequate effort should, in theory, pass any given PVT (Bigler, 2012). However, measuring effort can be problematic. The term 'inadequate effort' is broad and unspecific (Lezak, 2012), and conveys that NCR behavior has occurred without indication of malice, deceit, and/or the presence of a secondary gain.

Low effort may indicate NCR, but also may occur naturally, resulting from boredom or disinterest in the assessment process (Mittenberg, 1993). According to Larrabee (2014), effort exists on a continuum and can vary substantially over the course of a lengthy assessment battery. An examinee's attention may be at its lowest point during the administration of an assessment of effort, which may give the false impression that overall effort was low. However, an examinee putting forth optimal effort is not likely to fail more than one of measure of effort in a neuropsychological battery. Thus, administering more than one assessment of effort is considered best practice to avoid mistaken attributions of low effort (Larrabee, 2012).

Inadvertent NCR

Examinees may engage in NCR without malintent. One example is the *diagnosis threat*, which occurs when an examinee with a genuine disability, primed to consider themselves disabled, performs sub-optimally on measures of cognitive functioning (Suhr & Gunstad, 2005). The diagnosis threat construct was born from research on stereotype threat, which demonstrates

that members of social groups for whom negative stereotypes are prevalent, primed to think of these stereotypes, subsequently perform poorly on cognitive tests when compared to unprimed members of the same group (Steele & Aronson, 1995).

To demonstrate the existence of the diagnosis threat, Suhr (2005) recruited participants with mild traumatic brain injury. All participants were given a neuropsychological battery, but half of the participants (the experimental group) were made aware immediately prior to test administration that their traumatic brain injury was the reason they had been selected to participate in the study. Results showed that the experimental group obtained poorer scores on measures of memory, attention, working memory, and psychomotor speed than the control group, suggesting that being reminded of their injury negatively impacted their test performance.

Another form of inadvertent NCR is *self-handicapping*, which occurs when examinees endorse physical or mental symptoms to rationalize poor test scores (Smith et al., 1983). In a recent study, Suhr and Wei (2013) discovered that priming students for failure led to an over-endorsement of ADHD symptoms. The authors recruited 85 undergraduate participants who were given a difficult memory task, followed by a self-report measure of ADHD symptoms. Half of the sample (i.e., the experimental group) was told that the difficult test was a measure of intelligence, while the other half (i.e., the control group) was given no further instruction. The authors found that the experimental group endorsed higher rates of symptomology related to ADHD. Participants in this group also endorsed higher rates of depressive symptoms, anxiety, and neuroticism after the test, when compared to the control group. The authors theorized that the examinees had felt threatened by the prospect of low scores on a test of intelligence and had subsequently engaged in self-handicapping.

Overall, NCR is a frequent occurrence within neuropsychological and forensic assessments, estimated to occur in 22% of all psychological evaluations (Locke, 2008). Malingering refers to deliberate attempts at NCR for the purposes of obtaining an external reward (e.g., money or medication (APA, 2013). Such forms of NCR are problematic within the field of neuropsychological and forensic assessments because they threaten the integrity of these assessments and their applicability to the real-world. If the purpose of forensic and/or neuropsychological assessments is to determine the existence and/or degree of an impairment, NCR acts as a barrier toward these efforts. Fortunately, PVTs offer a method by which individuals engaging in NCR may be identified.

Performance Validity Tests

Performance validity tests (PVTs) are measures meant to detect NCR (Larrabee, 2014). PVTs may be embedded within other measures or stand-alone. According to Larrabee (2014), PVT is an umbrella term for two categories of tests: Symptom Validity Inventories (SVIs) and Performance Validity Inventories (PVIs). 'Symptom' refers to information reported by an examinee, either in person or on questionnaires/self-report measures, whereas 'performance' refers to examinees' skills and/or abilities that are demonstrated on a clinician-administered measure. The stratification of PVTs into PVI and SVI reflects manners in which examiners may detect NCR. On a questionnaire, NCR may take the form of an examinee answering haphazardly or endorsing non-existent symptoms. On a performance-based measure, examinees engaging in NCR perform below their true ability level by putting forth sub-optimal effort (Morgan & Boone, 2008).

The validity of PVTs for detecting NCR is demonstrated in two ways: First, obviously unimpaired examinees, who demonstrate a reasonable degree of autonomy, should not score

below the level of an individual with a major neurocognitive disorder (Bigler, 2012; Lezak, 2012). When this does occur, it is indicative of NCR. Second, examinees should not score significantly below chance level on measures that require examinees to choose one of two answers. Referred to as a 'forced-choice measure,' a guessing examinee would obtain a score around 50%. As such, it is unlikely that an examinee responding truthfully would obtain a score that is significantly below chance level (Vickery et al., 2001). Scores on forced-choice measures with two option that are less than or equal to 50% indicate NCR.

The standardization of PVTs is conducted primarily on two populations: simulated malingerers and individuals suspected of NCR. Since it is impossible to recruit a sample of true malingerers, a simulated malingering population is made up of participants without neurological disorders who are instructed to feign neurological impairments. The latter population is made up of examinees who are suspected of NCR because of their involvement in litigation and suspiciously low scores obtained on neuropsychological measures. For both populations, researchers are interested in a response pattern profile that can then be compared to profiles of individuals with bona-fide neurological impairments. Cut-off scores are then established such that the individuals with bona-fide impairments can pass, but individuals in NCR groups cannot. Oftentimes, a passing rate of 90% for those with true impairments is used (Bigler, 2012). Theoretically, an individual who is conscious, lucid, sober, and with a reasonable degree of autonomy should not perform below the level of an individual with true neurological impairment.

Traditionally, PVTs have been used to ascertain the legitimacy of reported memory impairments within a forensic context (Rogers & Granacher, 2011; Weinborn et al., 2003). Thus, most PVTs appear to examinees as measures of memory ability – usually, a free-recall task

followed by a recognition task. Recognition tasks are simpler than free-recall tasks, although they often appear to be complicated tasks (Bigler, 2012). Many PVTs consist of several words or word-pairs that are read or shown to the examinee. After the presentation of these stimuli is a recognition trial in which examinees are asked to identify target words or word-pairs. The majority of individuals with moderate neurological impairment can accomplish these recognition tasks with a high degree of accuracy (Lezak, 2012), and thus, PVTs cannot reliably discriminate between individuals with and without neurological impairments, and are not valid assessments of memory (Rogers et al., 1993).

PVT Sensitivity and Specificity

According to Lezak (2012), the efficacy of a PVT is measured by its sensitivity and specificity. Sensitivity refers to the number of individuals with a target disorder or behavior who are correctly identified by the measure; in other words, the true-positive rate. Tests with high sensitivity are desired when the neuropsychologist wants to know if a general impairment or undesired behavior exists, and correct identification of true positives is the purpose of the assessment. Tests with high sensitivity (and low specificity) are often referred to as screeners, because they are effective at identifying people with impairments or specific behaviors but are ineffective at determining the degree and type of impairment, as distinguished from other similar impairments.

In contrast, specificity, according to Lezak (2012, p. 127), is "the proportion of people without the target disorder whose test scores fall within the normal range;" or in other words, the true-negative rate. A test's specificity refers to its ability to correctly detect individuals who do not have a particular disorder or impairment. For example, a measure designed to detect NCR is easily passable by an honest and healthy examinee. In theory, it should also be passable by

individuals with moderate neurocognitive impairment, since PVTs are normed on impaired populations to establish a cut-off score that 90% will pass. If a PVT could detect all honest responders in a group of malingerers, it would have 100% specificity.

Measures vary in the degrees of both sensitivity and specificity that are preferred depending on the neuropsychologist's goal. When the neuropsychologist is interested in assessing the possibility of NCR, a quick and sensitive screener may be desired to detect all possible cases of NCR, but this measure may yield a high false-positive rate. In such cases, neuropsychologists must add PVTs with high specificity to clarify the results, although these might take longer to administer.

Specific PVTs and SVIs

Test of memory malingering. The Test of Memory Malingering (TOMM) is a 50-item measure that is disguised as a visual memory task (Tombaugh, 1997). The TOMM contains two learning trials, during which examinees are shown a series of 50 drawn pictures for three seconds each, followed by recognition trials. The recognition trial entails a forced-choice between two pictures, only one of which was presented during the learning trial. An optional delayed recall trial may be administered 15 minutes later. NCR is suspected when examinees fail to select the pictures displayed during the learning trials more than five times during the recognition trial.

The TOMM was standardized on 405 community-dwelling individuals who were recruited by word of mouth and advertisements placed in the community. On average, this non-clinical sample correctly identified 94% of the items on the first trial, and 99% on the second trial. Before testing, participants were asked how they thought they would perform, and underestimated their abilities, predicting 63.6% correct on the first trial, followed by 86% for the second. This implies that the TOMM appears deceptively difficult to a nonclinical sample. The

second phase of development entailed testing a clinical sample of individuals with various cognitive impairments. Scores are reported as a percentage of correctly identified items: cognitive impairment (97.2), aphasia (98.6), traumatic brain injury (98.8), and dementia (91.4; (Tombaugh, 1996). Since most participants with known neurocognitive disorders obtained scores exceeding 90%, the cutoff score of <45 was established to identify NCR. The last phase of development included the administration of the TOMM to suspected malingerers involved in litigation and simulated malingerers. Overall data suggested that across clinical and non-clinical samples, obtained scores were rarely below 46; thus, a cut-off of <45 was retained (Tombaugh, 1997). With these norms, the TOMM demonstrated 100% sensitivity and 100% specificity.

In a study designed to replicate the results of Tombaugh (1997), Teichner and Wagner (2004) tested elderly participants deemed to be cognitively intact, cognitively impaired, or to have bona-fide dementia. The cognitively impaired group (*N*=21) was composed of participants who had primarily suffered strokes or had age-associated cognitive decline. These participants were deemed low-risk malingerers because none had any discernable external reason to engage in NCR. All participants in the cognitively intact group and 92.7% of participants in the cognitively impaired group obtained scores above 45. Participants in the dementia group (*N*=21*)*, however, obtained average scores below 40. The authors concluded that the TOMM is a strong indicator of effort in non-demented participants, but recommended caution when memory impairment is suspected (Teichner & Wagner, 2004).

The TOMM is a highly regarded PVT in forensic settings (Lezak, 2012). Weinborn et al., (2003) tested the TOMM on a forensic inpatient population divided into two groups: one group (control group) had been civilly committed, or had been deemed not guilty by reason of insanity (NGRI), the other group (at-risk group) were awaiting trial for murder, and were expected to

plead NGRI. All participants had been diagnosed with psychiatric disorders. Results indicate that the TOMM had moderate sensitivity, in that 60% of the at-risk group were identified. Further, a high 100% specificity rate indicates that there were no false-positive results. Overall, the TOMM is moderately efficient at identifying NCR within a psychiatric population, and highly efficient at producing low false-positive rates.

21-Item test. The 21-Item Test (Iverson et al., 1991) is a stand-alone PVT that takes roughly 5 minutes to administer. An examiner reads a list of 21 nouns, and the examinee repeats back as many words as possible during a free-recall trial, followed by a forced-choice trial in which the examiner presents word pairs and the examinee selects the word previously presented. The word pairs share semantic similarity, phonetic similarity, or are dissimilar (Iverson et al., 1994). Poor effort is suspected if an examinee fails to recognize words initially recalled on the free-recall trial, and/or if they score above a cut-off of nine incorrect items (Iverson, 1996). The 21-Item test demonstrates a 2.5% false positive rate in individuals with bona fide memory impairments, and 70% sensitivity in a community sample. Ultimately, the 21-Item Test is a quick PVT with high specificity, but poor sensitivity. It is best used a screener to determine if further PVTs are needed (Lezak, 2012).

Word memory test and medical symptom validity test. The Word Memory Test (WMT; Green, 1996) includes two learning trials in which examinees are shown semantically similar word-pairs. The recognition trial is a forced-choice test with two words, one familiar and one not. Failure rates below 90% are indicative of NCR. Gervais et al. (2004) found that, within a litigating sample of over 500 participants, the WMT demonstrated higher sensitivity than the TOMM. Green (2004) stated that multiple studies have found the WMT to be almost 100% sensitive to poor effort, and five separate independent simulator studies in three different

countries have found the MSVT to be 100% specific. A computerized version, Green's Medical Symptom Validity Test (MSVT), is also available and has demonstrated comparable sensitivity (Green, 2004). During the administration of the MSVT, examinees are twice shown a list of word-pairs that are semantically similar and are then asked, in several forced-choice recognition trials, to identify which of two words was on the original list. After a 10-minute delay, examinees are administered the recognition task again, followed by a prompt of one of the words of the word-pair, and then a free-recall trial. The MSVT is quicker to administer than the WMT, and it includes a genuine memory test, providing additional utility.

Both the MSVT and WMT are passable even by individuals with intellectual disabilities, and are typically only failed by individuals in the late stages of dementia (Green, 2005; Green et al., 2003). Although they measure the same construct, the MSVT displays only 10 word-pairs, rather than 20 as in the WMT, making the test easier and quicker to administer. Green (2004) found that of 24 participants who failed the TOMM, 22 participants also failed the MSVT; however, of 50 participants who failed the MSVT, only 22 participants also failed the TOMM. Thus, the MSVT is more specific than the TOMM at detecting NCR.

Embedded PVTs

In contrast to PVTs that stand-alone, effort may also be abstracted from measures of neuropsychological functioning that were not initially developed to assess NCR. On SVIs, embedded PVTs are a measure of uncommonly endorsed symptoms. When used to assess NCR on performance-based measures, these measures identify 'floor effects,' which occur when examinees perform sufficiently low such that the test or subtest cannot measure the intended construct (Rogers & Granacher, 2011). Floor effects are indicative of NCR because, like stand-alone PVTs, it is suspicious when functioning examinees obtain scores below those of examinees

with a moderate cognitive impairment. The research suggests that honest examinees who have the ability to live autonomously should not perform worse than examinees with moderate cognitive impairments who cannot live alone (Lezak, 2012; Nigg, 2005; Quinn, 2003; Willcutt et al., 2005). Like stand-alone PVTs, cut-off points are based upon research with suspected NCR groups, simulated NCR groups, and individuals with moderate cognitive impairments, and are often set around a 90% passing rate for individuals with impairment.

The MMPI-2. The Minnesota Multiphasic Personality Inventory – Second Edition (MMPI-2; Hathaway, 1989), a 567-item self-report measure of personality, was the first psychological test to assess an examinee's test-taking attitude for use in the interpretation of the results (Green, 2000). The validity scales on the MMPI-2 represent the first embedded PVT and the first SVI. The MMPI-2 Lie (L) scale, Infrequency (F) scale, and Defensiveness (K) scale contain items drawn from questions that contribute to the clinical scales. The L scale includes items that detect an examinee's unwillingness to admit to common faults and follies (Hathaway, 1989). The F scale is calculated by summing the number of items in the first two-thirds of the test endorsed by the examinee that were only endorsed by 10% or less of the standardization group. High F scores often appear in profiles produced by examinees who are trying to 'fake bad' or send a distress signal to the examiner. The F-back (Fb) scale, made up of items in the last portion of the test, operates just as the F scale does. The F-Psychopathology (Fp) scale was introduced in the reformatted format version of the MMPI-2 (MMPI-2-RF; Ben-Porath & Tellegen, 2008) and is made up of items that are not endorsed by a known psychopathologic population 90% of the time. The K scale is a measure of defensive responding and is comprised of items from clinical scales that distinguished the non-clinical sample in the standardization sample from psychiatric patients who produced normal profiles. Green (2000) stated that some

validity scales on the MMPI-2 are measures of internal consistency rather than veracity. For instance, the Variable Response Inconsistency Scale (VRIN) and True Response Inconsistency Scale (TRIN) measure an examinee's consistency in responding to questions with similar content throughout the measure. Low scores on either TRIN or VRIN suggest that the profile is invalid because the examinee responded inconsistently to pairs of questions matched on content and may have been responding haphazardly or without reading the questions.

Reliable digit span and advanced clinical solutions. The Advanced Clinical Solutions (ACS) is a battery of stand-alone PVTs and embedded PVTs abstracted from the Wechsler Adult Intelligence Scale - 4th edition (WAIS-IV; Wechsler, 2008) and Wechsler Memory Scale – 4th edition (WMS-IV; Wechsler, 2009). This battery of tests provides measures of effort, social perception, and premorbid intellectual functioning (Pearson, 2009). Reliable Digit Span (RDS) is a popular embedded PVT within the ACS that is abstracted from the WAIS-IV subtest Digit Span. The Digit Span subtest requires examinees to repeat strings of digits forward, backward, and in sequential order. RDS is calculated by summing the number of digits in the longest digit spans forward and backward for which the examinee completed both trials correctly. For example, if an examinee completed two perfect trials of three digits forward, and completed two perfect trials of five digits backward, RDS equals eight. Pearson (2009) established a cut-off of 6, which yields strong specificity (86.3%) and moderate sensitivity (61.4%) for detecting NCR (Jasinski et al., 2011). In a cross-validation study of RDS, Schroeder et al., (2012) obtained similar results, except for examinees with a history of stroke, intellectual disability, or memory impairment.

Trail making test. The Trail Making Test (TMT; Partington & Leiter, 1949) is a neuropsychological measure of processing speed and executive functioning. There are two

forms, A and B. TMT A, a measure of processing speed and visual scanning, requires examinees to rapidly connect encircled numbers, scattered on the page, in ascending order with a pencil. TMT B is a measure of cognitive switching ability that requires an examinee to rapidly connect encircled letters and numbers, scattered on a page, in alternating ascending order with a pencil. Time to completion of both forms comprises the examinee's final score.

As a PVT, NCR can be abstracted from unusually high error rates (connecting the encircled numbers/ letters out of order) and unusually slow completion times. In a review of TMT's effectiveness as a PVT, Iverson, Lange, Green, and Franzen (2002) compared the ratio of scores on form A to B (and B to A), the number of errors committed, and amount of time needed to complete within a sample of patients with mild to severe TBIs. Of this sample, 228 were involved in litigation and deemed at-risk of NCR, while 571 were not involved in litigation and comprised the control group. Data suggested that, within these groups, the TMT has moderate predictive ability for detecting feigning (88.3% for TMT A and 66.7% for TMT B for both mild and severe TBI combined). Sensitivity was low for TMT A and TMT B (14.5% and 11.6%, respectively); however, specificity was high (98.8% and 97.5%, respectively), suggesting that this measure identifies most examinees who are not engaging in NCR, but may miss examinees who are non-credible responders. The authors concluded that the TMT is an adequate screener for unsophisticated NCR, in that examinees identified by the TMT as engaging in NCR likely are correctly identified. However, since sensitivity is low, other PVTs must be used to reduce the number of missed cases.

Rey-Osterrieth complex figure task. The Rey-Osterrieth Complex Figure Test (ROCFT) was developed in 1941 by Andre Rey to investigate visual memory impairments and perceptual organizational skills (Lezak, 2012; Rey, 1941). It continues to be used by

neuropsychologists to assess visuoconstructional abilities, perceptual memory, and planning/organization (Deckersbach et al., 2000). The complex figure is first placed in front of the examinee, who is told to copy the figure using a pencil with an eraser. The examinee is not forewarned that a recall trial will take place. The complex figure and the copy are taken from the examinee, and the examinee engages in a distractor task, such as reciting the months of the year backward. Three minutes later, the examinee is instructed to reproduce the complex figure from memory in the immediate recall condition. A delayed recall condition follows 20 minutes later, followed by a recognition task (Spreen et al., 2006). Scores are calculated using criteria that rate the examinee's accuracy and placement of the various elements of the design.

Knight and Meyers (1995) found that compared to a sample of individuals with brain injuries, simulated feigners demonstrated poorer accuracy, slower copying ability, and poorer delayed and recognition ability. Most of the simulators reproduced recognition scores that fell two and half standard deviation below those obtained by normal, honest participants. Lu et al. (2003) found that the recognition task distinguished litigating examinees who were suspected of NCR from both a bona fide brain injury sample and from a sample not suspected of NCR. Further, they denoted that this group committed 'atypical' recognition errors of simple items.

Continuous performance tests. Continuous Performance Tests (CPT) are standardized computer programs that measure vigilance and sustained attention (Conner, 1995; Dobson-Patterson et al., 2016; Sandford & Turner, 1995). Stimuli appear on the monitor, some of which require the examinee to react (targets), and others that require no response (non-targets). Several scores are generated based on the examinee's responses that indicate impulsivity, inattention, sustained attention, variable attention, and waning attention. For instance, a fast reaction time with several commission errors is indicative of impulsive test-taking behavior. CPTs distinguish

normal controls from subjects with ADHD or schizophrenia with high specificity (Hervey et al., 2004; Lezak, 2012) and have strong test-retest reliability when examinees have not taken stimulant medication (Conners, 1994).

Quinn (2003) demonstrated that simulated malingerers consistently perform worse on CPTs than do normal controls and ADHD samples. She found that normal controls (M = 90.4, SD = 16.1) scored higher than an ADHD sample (M = 65.3, SD = 30.1), who scored less impaired than a NCR sample (M = 15.8, SD = 21.3). However, although there is a significant difference between the ADHD and NCR groups, cut-off scores have not been established. As the research currently stands, unusually low scores on a CPT could be indicative of severe neurological impairment or severely impaired attention and not NCR. Using a CPT as a PVT is problematic because it is difficult to determine *how* poor an NCR group would do, on average, compared to true ADHD groups. Quinn suggested that if other scores (on stand-alone PVTs or embedded PVTs) implicate NCR, poor performance on a CPT can be carefully used as further evidence of NCR.

Issues Related to PVTs

While the research supports the inclusion of PVTs in assessments in which NCR is a possibility, these measures have limitations (Larrabee, 2014). One issue concerns establishing proper cut-off points. As aforementioned, cut-offs are generally set such that 90% individuals with moderate cognitive impairment can pass. Oftentimes, this means functional examinees need only miss a few items to score below the cut-off (Bigler, 2012). This engenders a high probability of false-positive hits, especially for honest examinees with dysfunctional attention. A false-positive failure on a PVT may have dire consequences for an examinee, such as the denial of services or accommodations to which the examinee should be entitled.

Because scores on PVTs are considered measures of effort, examinees putting forth adequate effort are generally expected to achieve perfect or near-perfect scores; lower scores call into question the validity of all the results obtained during the assessment. Thus, scores achieved on PVTs should fall somewhere within a curtailed distribution. For example, missing just three out of 20 items (<90%) on the Medical Symptom Validity Test (Green, 2004) suggests probable suboptimal effort. Such a high floor indicates that an honest examinee need not stray too far from perfection to effectively fail this PVT.

Failure of one PVT does not always result in a spoiled assessment. In many cases when an examine fails a PVT, the clinical judgement of the examiner may be sufficient to determine whether the results obtained indicate NCR (Lezak, 2012). As stated above, if 90% of individuals with a cognitive impairment pass the MSVT, 10% are expected to fail, indicating a false positive result. However, because moderate cognitive impairment is a salient condition, a clinician is not likely to suspect NCR. For example, Brandt et al. (1985) found that more than half of individuals with Huntington's disease failed a PVT, none of whom were suspected of NCR. However, because symptoms of a neurocognitive disorder may be subtle, it is recommended best practice to use multiple PVTs to lower rates of false-positive failures (Schroeder et al., 2012; Sollman et al., 2010).

A more controversial issue among neuropsychologists concerns informed consent regarding PVT use. Suhr (2000) reported that examinees' awareness of the use of a PVT in an assessment produced a more sophisticated malingering style. In this study, nonclinical participants were separated into a control group, a naïve simulated NCR group, and a 'PVT aware' simulated NCR group. Findings indicated that the control group produced the most credible performance, followed by the group aware of the PVTs, with the naïve group

performing the worst. The author suggests that, for examinees engaging in NCR, awareness of the use of a PVTs may produce a more sophisticate style of NCR and thus decrease detectability.

Summary of PVTs

PVTs are effective methods of detecting NCR that exist in two forms: as stand-alone measures (e.g., TOMM and WMT) with the sole purpose of evaluating NCR, and as measures embedded within other measures (e.g., RDS and MMPI-2) intended to measure constructs other than NCR. Cut-off scores are created by comparing the scores of a normal control group, neurologically-impaired samples, and samples deemed 'at-risk' of malingering and/or simulated malingerers; however, they range in sensitivity and specificity. These scores reflect points at which normal controls and neurologically-impaired individuals are distinguishable from the 'malingering' groups. Scores from true malingering groups would be valuable to this body of research but are difficult to obtain because of the nature of malingering.

Historically, much of the data on at-risk malingerers originated from individuals involved in litigation, whose complaint was a memory impairment. This field is currently expanding, however, to include individuals, oftentimes college students, suspected of feigning ADHD. The nature of ADHD impairs college students' ability to compete with peers, often negatively affecting studying habits, organizational abilities, test taking abilities, and social abilities. Further, accommodation services intended to help these students gives an unfair advantage to students without ADHD. As a result, the use of PVTs within an ADHD assessment have become regarded as best practice toward ensuring the proper allocation of these services.

ADHD

ADHD is a neurological developmental disorder that is present in childhood and typically first diagnosed in elementary school-aged children (APA, 2013). The APA recognizes that

ADHD may persist from childhood into adulthood. Around 2.5% of children are diagnosed with ADHD, and of those individuals, 50 to 65% continue to have symptoms into adolescence and adulthood. ADHD falls into three diagnostic subtypes: inattentive, hyperactive, and combined presentation. To meet criteria for ADHD, symptoms must be present before age 12, must have been present for at least 6 months, and must impair day-to-day functioning. Inattention is characterized by failure to pay attention to detail; difficulty sustaining attention; inability to listen when spoken to directly; failure to follow through with projects, directions, schoolwork and/or instructions; difficulty organizing tasks and activities; distractibility; misplacing items; and forgetfulness. The presentation of hyperactivity varies from children to adults. In adults, it is characterized by fidgeting hands and feet, impulsivity, restlessness, interrupting others, difficulty engaging in leisure activities quietly, loquaciousness, and impatience. Individuals meeting criteria for both inattentive and hyperactive symptoms are diagnosed with combined presentation.

Prior to 1970, ADHD was considered a disorder exclusive to childhood, and as recognition of adult-ADHD grew, research proliferated, increasing three-fold between 1994 and 1997 (Hervey et al., 2004). As such, adults with ADHD not diagnosed in childhood began to seek evaluations. ADHD symptoms change over the lifespan, and with age, symptoms tend to become internalized; hyperactive symptoms diminish, making the disorder more subtle and harder to diagnose (Barkley et al., 2008; Hallowell, 1995). Adulthood evaluations differ from childhood ADHD evaluations because guardian reports are often not available, and hyperactive symptoms may have subsided. For adults, evaluations often include an interview with the client, and when available, an informant, as well as a standardized retrospective self-report, like the

Wender Utah Rating Scale (WURS; APA, 2013; Ward, 1993; Weiss & Murray, 2003).

ADHD in College Students

Individuals with ADHD are increasingly attending college. Out of 197 children diagnosed with ADHD in 1979-1980, 75% were attending college by 2002 (Barkley et al., 2002), and in the last two decades, neuropsychologists have seen a significantly large influx of college students seeking evaluations for ADHD (Musso & Gouvier, 2014). Upon entering college, these students face challenges related to test taking, organization, and peer relationships.

Without accommodations, college students with ADHD are at a disadvantage when compared to their peers. Norwalk et al. (2009) examined day-to-day functioning, perception of impairment, and academic achievements and ambitions in 321 college students with ADHD. They found that severity of impairment was negatively correlated with overall grade point average, career decision-making ability, self-efficacy, time-management abilities, study skills, and overall academic adjustment to college life, which tends to be more unstructured than high school (Diller, 2010). These impairments became more robust when these individuals had at least one comorbid psychological disorder. In addition, these students often struggle with interpersonal and romantic relationships (Bruner et al., 2015), and in general, tend to report elevated levels of distress (Gray et al., 2016).

College students with ADHD comprise 25% of students who receive disability services across college campuses in the United States (DuPaul et al., 2009), which some authors have argued is fewer than the true number of students with ADHD (Kooij et al., 2010). In fact, some authors estimate that only one-third of students with ADHD are receiving the benefits they deserve (Advokat et al., 2007).

To assist these students, the Americans With Disabilities Act of 1990 ("Americans with disabilities act of 1990," 1990) stipulates that students diagnosed with ADHD receive accommodations. These accommodations may include extra time on exams, the ability to take exams in a separate quiet room, peer support to enrich organization and planning abilities, and help controlling impulsive and self-destructive behavior (Suhr & Wei, 2013; Weyandt et al., 2017). Psychiatrically, ADHD is often treated with psychostimulant medications, including Ritalin (methylphenidate), Adderall (dextroamphetamine), and Focalin (dexmethylphenidate), that when used as prescribed are effective at bringing attention to normal levels of functioning, and tend to improve academic functioning (Norwalk et al., 2009). These medications may cause adverse side-effects, such as appetite suppression and insomnia.

These medications are currently listed as schedule II-controlled substances, which means they have an accepted medical value and high potential for abuse. A preliminary study at a small university found that 35.5% of the student body had used psychostimulants illicitly, and almost 20% reported using psychostimulants in combination with alcohol (Low & Gendaszek, 2002). According to Low & Gendaszek (2002), psychostimulant use is expected to become more widespread as prescriptions of these medications become easier to attain.

In a more recent study, Kiernan et al. (2016) sought to extend the work of Low and Gendaszek (2002) and examined the self-reported pattern of use of 141 college students which included college seniors (51%), juniors (31%), sophomores (16%), and freshmen (19%). They found that 30% of the overall sample admitted to using Adderall to help them study at least one time, with the majority stating that they would reuse Adderall in the future. The authors ultimately found small overall GPA differences (3.28 for non-users versus 3.16) between

students not using and using Adderall. They noted that cavalier attitudes regarding the use of Adderall for many students, with many equating it to a strong caffeinated beverage.

Neuropsychological Assessment of ADHD

The neuropsychological assessment of ADHD measures areas of impairment often detected in individuals with ADHD. According to Lezak (2012), a typical neuropsychological battery begins with a clinical interview, whereby an examiner inquires into an examinee's presenting complaint, medical/psychological history, family history, and academic history. In the assessment of ADHD, self-reports and informant reports are collected, which provide the necessary behavioral data needed to distinguish an attentional impairment from normal functioning (Barkley et al., 2002). Finally, neuropsychological data is collected to quantify the degree of the impairment, and impart strengths and weaknesses as areas for treatment (Pettersson et al., 2018). Two neuropsychological measures used in the assessment of ADHD, which have not been discussed, include the Tower Test (Culbertson & Zillmer, 2001) and Stroop Color-Word Test (Stroop, 1935).

ADHD Neuropsychological Battery

The tower test. Several analogous Tower Tests (Culbertson & Zillmer, 2001) share an identical conceptual structure in which examinees are given a board with pegs on which they must arrange various pieces in particular positions (Lezak, 2012). Each test has its own administration rules but shares the goal of recreating the target picture with the pieces within an allotted amount of time using the least number of moves possible. The tower test is a commonly used measure to assess planning abilities, which are often impaired in ADHD samples (Riccio et al., 2004).

The Stroop color-word test. The Stroop test, originally developed in 1935, is used to measure response inhibition and processing speed (Jensen & Rohwer, 1966; Stroop, 1935). Administration entails three trials over three conditions, each of which lasts 45 seconds. In trial one, examinees are asked to read words that denote colors as rapidly as possible. In the second trial, examinees are asked to identify colors (denoted by the first trial) as rapidly as possible. The third trial contains words that denote colors, written in colored ink. Examinees are asked to name the color of the ink rather than read the word. Examinees are scored based on how many items they can correctly identify. Trials 1 and 2 measure processing speed (Shanahan et al., 2006). Trial 3 measures response inhibition (Jensen & Rohwer, 1966). In a newer version of the Stroop (Golden & Freshwater, 2002), these three conditions yield four scores, which include normative comparisons of speed based on age and education level for all three conditions, plus a score that compares an examinee's speed on the first two conditions. The third score yields interference ability compared to peers. The fourth score is a pure measure of interference, since it does not take processing speed into account.

The first 2 trials of the Stroop task have been demonstrated to distinguish individuals with ADHD from age-matched controls (Shanahan et al., 2006). In regard to the third trial, a meta-analysis of 25 studies found no relationship between ADHD and normal controls; however, they did find that, on average, ADHD groups were 1.14 times slower to name words and colors (Schwartz & Verhaeghen, 2008). Both studies indicate that the Stroop task is a better measure of processing speed within an ADHD assessment than of response inhibition.

Standard ADHD neuropsychological battery. In addition to a clinical interview, Gallagher and Blader (2001) recommended using an abbreviated measure of overall cognitive functioning to set a baseline with which other scores will be compared, followed by EF

measures. Research has demonstrated that students are indistinguishable from their normal peer counterparts on intelligence measures (Weyandt et al., 2017). Other researchers suggest expanding this battery to include measures of focused attention, simple attention, sustained attention, and memory (Adams et al., 1996; Schoechlin & Engel, 2005).

Focused attention requires the recruitment and orchestration of several other domains, including verbal memory, visual memory, spatial memory, visual scanning, cognitive abilities, and perceptual abilities. As it pertains to ADHD, EF assessment instruments measure purposeful cognitive control, such as inhibition, detecting environmental mismatches, abstraction, set shifting, sustaining working memory, strategic encoding (often in 'chunks'), and planning (Chan et al., 2008). These assessments are often timed and require examinees to engage in multiple tasks simultaneously. Focused attention is measured using neuropsychological instruments and includes a processing speed component. Processing speed is measured using the Stroop task trials 1 and 2 and TMT A. Aspects of EF are measured using neuropsychological tests, such as TMT B (set-shifting), Tower Test (planning), Digit Span Backward (working memory), Stroop task trial 3 (inhibition), and ROCFT (planning and strategic encoding).

Simple attention entails the ability to focus on one or more stimuli without the need to manipulate mental information (Botvinick et al., 2001; Willcutt et al., 2005). Digit Span Forward is an adequate measure of simple attention. Sustained attention tasks require simple attention over a longer period of time. These tasks may entail attending to a stimulus that changes over time, and that requires an examinee's response when these changes occur. Sustained attention can be measured using a CPT.

The assessment of memory is split into visual and verbal components. Visual memory is measured immediately after the examinee is exposed to the stimulus (immediate recall) and

approximately 20 minutes later (delayed recall). Visual memory tasks include a visual stimulus and often require examinees to copy a drawing (e.g., ROCFT). Verbal memory tests require examinees to learn words over successive trials. An examinee's ability to retain the words is measured after approximately 20 minutes and can be measured using the Word-Pairs subtest on the WMS-IV (Wechsler, 2009).

Lastly, an assessment of ADHD requires several self/informant reports. Researchers have found that when participants answered honestly, self-reports had the highest specificity for diagnosing ADHD (Sollman et al., 2010). One such measure, the Wender Utah Rating Scales (WURS), is a retrospective self-report of ADHD symptoms. The WURS has been shown to identify 86% of actual ADHD cases, based on retrospective reporting of symptoms (Booksh et al., 2010).

Efficacy of Neuropsychological Measures in Diagnosing ADHD

Research demonstrates that neuropsychological assessments can elucidate the symptoms of ADHD. A meta-analysis of 24 studies including 867 individuals with ADHD and 806 controls concluded that there are robust neuropsychological differences between individuals with and without ADHD. Tasks that require focused attention, sustained attention, verbal memory, and abstract problem-solving discriminate ADHD participants from controls with a moderate effect size between 0.5 and 0.6. Simple attention and memory deficits can distinguish participants with ADHD from controls, with effect sizes of -0.38, and -0.56, respectively (Schoechlin & Engel, 2005). In another study, the TMT B, Digit Span Backward, and the CPT-II discriminated ADHD participants from healthy participants 53%-66% of the time (Pettersson et al., 2018). However, after controlling for age and IQ, these groups became indistinguishable.

While there are differences between individuals with ADHD and healthy individuals on various EF tasks, these results are not always robust (Conant, 2014), and the variability among individuals with ADHD makes determining a specific ADHD-profile futile. This is further compounded by the occasional variable performance by normal controls, who may occasionally obtain low scores on EF tasks because of individual variability rather than a neuropsychological dysfunction (Dobson-Patterson et al., 2016). Because test scores vary between individuals with ADHD, and scores vary within a normal population, a distinguishable ADHD profile cannot be determined. That is, the neuropsychological deficits associated with ADHD are unspecific (Barkley et al., 2010).

However, there is still utility for neuropsychological measures within ADHD assessments. Conant (2014) highlights two: First, several diagnoses may resemble ADHD, which may include behavior associated with bipolar disorder or strokes. Neuropsychological assessment can assist in ruling out other diagnoses to clarify the etiology of the attentional impairment. Second, once ADHD has been diagnosed, neuropsychological testing can elucidate the examinee's strengths and weaknesses. These in turn may inform treatment. Gallagher and Blader (2001) reviewed the practical guidelines for using neuropsychological measures within an ADHD assessment, identifying steps neuropsychologists should undertake. Consistent with Harrison et al. (2007), the first step is to determine whether an impairment exists. The second is to determine the extent of this impairment. Third, determine if the attentional impairment can be attributed to another disorder or condition.

Non-Credible Responding and ADHD

Misappropriated accommodations for college students with ADHD give an unfair advantage to students without ADHD (Conant, 2014). Students without ADHD may fake the symptoms of

ADHD to obtain extra time on exams, a quiet room apart from peers to take an exam, and/or a prescription for psychostimulant medication. Montejano et al. (2011) attributed a three-fold increase in ADHD evaluations of college students between 2002 and 2007 to students without ADHD seeking accommodations and psychostimulant medication. Sullivan et al. (2007) reviewed the WMT scores of LD/ADHD assessment cases in a college clinic and found signs of NCR in 25-50% of the assessments. In a similar study, Suhr et al. (2008) found that 31% of 26 examinees in ADHD assessments failed the WMT. This suggests that while many students may genuinely have undiagnosed ADHD (Advokat et al., 2007), many are also malingering to gain accommodations or medication.

The use of psychostimulants within the college population is on the rise. Rabiner (2013) found that between 2000 and 2011, rates of psychostimulant use across college campuses rose from 6.6% to 9.3%. Of students who admitted to using nonmedical psychostimulants, 32% admitted to a sole use, while 34% admitted to using 3-4 times per month. In a sample of 4297 adults between the ages of 18 and 49, Novak et al. (2007) found that 4.3% of 18- to 25-year-olds admitted to using psychostimulants non-medically in the past year, 33% of whom stated they used them to increase productivity.

Misuse of psychostimulant medications may occur for recreational purposes, to improve attention and concentration, and/or to lose weight (DeSantis & Hane, 2010; Julien et al., 2004; Rabiner, 2013). Most respondents who endorsed the use of psychostimulants admitted to having never been formally prescribed these medications, and further admitted to getting them from family members or friends. Also, of those who admitted to using psychostimulants illicitly, 20% stated that they had, at one point, feigned the symptoms of ADHD to a healthcare provider whom they knew would not ask “too many questions” and would give them their desired prescription.

Given the potentially high number of college students who engage in NCR to obtain ADHD-related accommodations and medication, and the ease with which many students have been found to malinger ADHD, it follows that neuropsychologists must assess for NCR within ADHD evaluations, especially in college sample.

Strategies Used to Feign ADHD

Several researchers have investigated the strategies that college students report using in simulation studies. For instance, Quinn (2003) administered post-hoc surveys to participants in the feigning group, and found that they reported engaging in more than one strategy. She found that 61% tried to be inattentive, 43% ignored visual stimuli, 35% made deliberate errors by omission, 17% ignored auditory stimuli, 13% feigned fidgetiness, and 9% responded slowly Harrison et al. (2007) also reviewed malingering strategies and found that 31% reported slow responding, 29% attempted to appear 'zoned out,' 26% made commission errors, 23% made omission errors, and 23% responded impulsively. Frazier et al. (2008) gave participants a series of yes/no questions regarding self-reported strategies of feigning ADHD. They found that 77% reported responding slowly, 87% reported inconsistent responding, and 90% tried to appear less intelligent. In addition, 90% reported trying to demonstrate "difficulty with paying attention," (p. 505).

Detecting NCR Within the ADHD Assessment

Erdodi and Roth (2016) found that because the diagnosis of ADHD relies on self-report, engaging in NCR can be simple. Most individuals with ADHD perform worse than healthy individuals on tests of EF and attention, but again, these patterns of results are not specific to ADHD. The only way to limit NCR, according to the authors, is the addition of PVTs to ADHD

assessments. In addition, neuropsychologists should be on the lookout for variable, staggered, and/or unusually low scores among the testing data (Booksh et al., 2010).

Subsequent simulation studies since Quinn (2003) have demonstrated that feigners do not face much adversity in their endeavors to obtain an ADHD diagnosis. In one study, Booksh et al. (2010) split 110 college students into normal controls and simulated ADHD malingerers. Archival data of 56 students diagnosed with ADHD through the author's university clinic were also used. These students had been diagnosed via structured interview and neuropsychological measures. All participants in the study were given the same battery as the clinical ADHD group, and in addition the WMT. All participants were told prior to testing that a PVT would be used. The researchers had an independent psychologist review the masked data and sort participants into the control, simulated ADHD, or clinical ADHD groups. Overall, they found that the psychologist mismatched 44% of the simulated groups into the clinical ADHD group. Inclusion of the WMT data improved decision-making by 2%, which the authors stated was low because participants were aware that a PVT would be utilized. The authors concluded that college students can successfully simulate ADHD, and in addition, produce scores that are generally more impaired than individuals with bona fide ADHD. Further, participants should not be forewarned that a PVT will be used.

In a similar study, Sollman et al. (2010) used the MSVT, TOMM and the Miller Forensic Assessment of Symptoms (MFAST; Miller, 2001), which is a measure of uncommon symptomatology. In this study, students instructed to feign ADHD were told to spend some time using Google to study ADHD symptoms before attempting the assessment. The authors ultimately concluded that all PVTs demonstrated "a moderate level of sensitivity" (p. 333) in detecting NCR. The authors were not concerned with specificity, although they stated that the

TOMM fared the best. Instead, they were interested in how specificity was increased by the administration of multiple PVTs. They found that, on average, one PVT rendered a specificity of 0.633, two rendered specificity at 0.931, and three or more yielded 1.00. They concluded that all ADHD neuropsychological evaluations should feature several PVTs, especially those that appear to require sustained attention.

In a recent landmark study, Marshall et al. (2016) examined 428 archived ADHD assessment cases that all included a clinical interview, a comprehensive neuropsychological battery, and several PVTs, both stand-alone and embedded. Of these completed assessments, the authors identified 115 participants who had put forth 'suspected effort,' as determined by multiple failed PVTs and the Slick et al. (1999) criteria, and who ultimately had not been diagnosed with ADHD. The authors then had psychologists blinded to the above information diagnose these 'suspected effort' participants on the basis of one of three categories of test information: only interview data; interview data and self-reports; and interview data, self-reports, and neuropsychological test results. Based solely on the interview notes alone, 71% of patients would have been given a diagnosis of ADHD. Based on an interview and self-reports, this number dropped to 65%. With an added CPT, this number dropped to 62%. In addition, the authors noted that, on average, the suspected effort group performed significantly worse on most of neuropsychological measures administered. These measures included select subtests from the WAIS-IV, Delis-Kaplan Executive Functioning Systems (D-KEFS), and WMS-IV, and the ROCFT. Thus, the authors suggest the use of PVTs in all ADHD evaluations.

In sum, PVTs should be an integral part of an adult ADHD assessment, given the potential for malingering, and research has consistently demonstrated that college students are capable of successfully feigning ADHD (Booksh et al., 2010; Marshall et al., 2016; Quinn, 2003;

Sollman et al., 2010). Research also demonstrates that NCR occurs with some regularity (Suhr, 2000; Sullivan et al., 2007). Data suggest that NCR can be detected by PVTs, the assessment of which become more effective when several are utilized (Harrison et al., 2007). Subsequent research has focused on the methods used by examinees when engaging in NCR and has generally found that participants act in an inattentive manner, respond slower, ignore stimuli, and made deliberate errors on neuropsychological measures (Harrison et al., 2007; Quinn, 2003). Research suggests that PVTs are of variable efficacy when examinees are aware they are being used (Suhr, 2000).

Limitations to Current Research

As it stands, all but two ADHD NCR studies have utilized unincentivized groups. That is, participants are instructed to feign ADHD, and receive research participation credit regardless of their feigning ability. In these studies, undergraduate college students are recruited from subject pools and randomly assigned into samples of simulated malingers and normal controls. Simulated malingers are then instructed to feign ADHD on neuropsychological measures, PVTs, and/or symptom-based measures. Differences in these scores have then been used, traditionally, to identify effective means to discriminate simulated malingers from controls. However, since true malingers are motivated by an incentive, the ecological validity of these unincentivized groups are questionable. Thus, one potential limitation of these simulation studies is the uncertainty surrounding participants' intrinsic motivation to feign ADHD.

For example, in the initial ADHD simulation study, Quinn (2003) recruited individuals for the ADHD sample from the Office of Disabilities and did not offer these participants a reward or research credit. The control and NCR groups were comprised of undergraduate students recruited from a psychology class, all of whom received research credit for their

participation, regardless of their performance. Similarly, both Frazier et al. (2008) and Harrison et al. (2007) also offered all participants research credit, regardless of the ability to successfully feign ADHD. Others have provided more enticing rewards; for example, Booksh et al. (2010) offered participants entry into a raffle to win a $50 gift certificate to a local restaurant. However, this incentive was offered to *all* participants, regardless of their success at feigning ADHD. That is, participants did not stand to lose these offers if unable to successfully feign ADHD. Thus, the degree to which these groups represent true malingering groups is questionable.

As noted above, two studies have offered incentives beyond some form of credit to improve the ecological validity of these NCR samples. Sollman et al. (2010) recruited their sample from the Office of Disabilities and from a psychology class. All participants were offered two research credits. However, individuals in the feigning group were offered an additional $45 if they could successfully feign ADHD (e.g., pass PVTs). This offer was made to every member of the feigning group. Ultimately, they found hit rates between .877 and .542 (out of 30 participants) depending on the PVT. In a dissertation, Fisher (2007) offered participants the chance to win $25 if ADHD could be successfully feigned on 3 self-reported measures and found that 77-93% were successful.

However, these studies did not use a comparison group of participants who were not incentivized to engage in NCR, which potentially limits the ability to determine the impact of an incentive on performance. Although both studies included an external incentive, the lack of an unincentivized comparison group means that the extent to which these added incentives produced differences in participants' scores is unknown, and thus, their resemblance to a true malingering sample is unclear. Increasing the ecological validity of the ADHD NCR groups has

implications for future research; however, it is not known if adding an external incentive produces differences in participants scores on neuropsychological measures and PVTs.

Present Study

The current study was conducted to examine the ecological validity of utilizing simulated Non-Credible Responding (NCR) samples to examine performance on measures routinely administered in ADHD evaluations within a college setting. The study aimed to address several gaps in the literature concerning the effects of incentivizing ADHD-related behavior compared to an unincentivized group and identifying strategies participants used to do so. Participants were randomly assigned to one of four groups: two groups of participants incentivized to successfully feign ADHD, a group of participants unincentivized to successfully feign ADHD, and control group, given no further instructions other than to put forth adequate effort. As such, this study had four primary goals: The first goal was to examine differences in neuropsychological test scores between groups. The second goal of this study was to examine differences specific to Performance Validity Test (PVT) scores between groups. The third goal of the study was to examine the level of sophistication used by NCR groups to successfully feign ADHD. As such, these terms were operationally defined and tested. The fourth goal of the current study was to qualitatively examine the strategies used by NCR participants to successfully feign ADHD and evade detection by PVTs. It was hypothesized that all NCR groups would produce scores in the borderline impaired range on all neuropsychological measures, below the cut-off on all PVTs, and relay strategies similar to past research. It was also hypothesized that both incentivized groups would produce scores reflecting significantly less impairment than the unincentivized group in the direction of the control group, since they were predicted to behave in a "sophisticated" manner compared to the unincentivized group.

CHAPTER 2

METHODS

Participants and Recruitment

Participants were undergraduate students enrolled at Indiana University of Pennsylvania and recruited from the Psychology Department subject pool, which consisted of PSYC 101 students. Sona software operated by the subject pool was used to allow eligible participants to register to participate in the study. To be eligible for participation, students had to between the ages of 18 and 22 years and produce a neurotypical profile on a six-question ADHD symptom prescreening measure (Appendix A).

Since this study required students to feign ADHD, students reporting current symptoms of ADHD were excluded from participation. In addition, since healthy students were sought for participation, students with a reported history of a Specific Learning Disorder (LD) and/or students receiving academic accommodations were excluded. Lastly, exclusion criteria included experience of a traumatic brain injury within the past three months due to the unpredictable nature of the healing process. The rationale for exclusion criteria was to maintain the integrity of the data and to obtain a sample able to perform in the normal range on neuropsychological measures. Upon arrival to the testing session, all participants were asked to complete a demographic form, which also asked if they had ever been diagnosed with ADHD or LD. Participants who had produced a neurotypical profile on the initial prescreen but indicated that they had, at some point in their life, been diagnosed with ADHD or LD were given credit and excused. All participants received research participation credit via the aforementioned Sona software.

Measures

Prescreen Measure

As required for enrollment in the Sona system, participants completed a prescreen survey to determine eligibility for participation in research studies. For the current study, the prescreen survey included part A of the Adult ADHD Self-Report Scale (ASRS-v1.1; Schweitzer *et al*., 2001; See Appendix A), which included questions about current attentional functioning and symptoms of ADHD. The ASRS-v1.1 is a demonstrably strong predictor of ADHD diagnosis in a college sample (Gray et al., 2014) and has a sensitivity of 100% and specificity of 71% in a primary-care sample (Hines et al., 2012). Students who endorsed current symptoms of ADHD or age outside of the 18- to 22-year range were excluded from this study.

Demographics

Participants completed a brief demographic questionnaire online via Google Forms in which they reported their age, gender, ethnicity, native language, handedness, academic class, history of head injury, history of ADHD diagnosis and/or Specific Learning Disorder, and whether they received academic accommodations for any reason (Appendix B). This questionnaire was completed during the testing session.

Digit Span (Wechsler, 2008)

Participants completed the first two sections of the WAIS-IV subtest Digit Span, which required them to repeat strings of digits forward and backward. This measure served two purposes: to measure neuropsychological functioning and to identify NCR. Digit Span Forward (DsF) is a neuropsychological measure of simple attention, which is often impaired in ADHD samples (Rosenthal et al., 2006). Digit Span Backward (DsB) is a measure of working memory, which is an executive function (EF) that is also commonly impaired in individuals with ADHD

(Chan et al., 2008). In addition, DsF and DsB are used to calculate the embedded PVT, Reliable Digit Span (RSD; Pearson 2009). Norms for Digit Span Forward and Digit Span Backward from the WAIS-IV manual, which are based on a total of 2200 community-dwelling individuals with 220 participants for the age groups used in this study (i.e., 18:0-19:11 and 20:0-24:11; (Wechsler, 2008) were used to identify cut-off scores.

RDS was originally formulated by Greiffenstein (1994) and is calculated by summing the number of digits that comprise the last item for which both the forward and backward trials were completed correctly. RDS had demonstrably strong specificity (86.3%) and moderate sensitivity (61.4%) for detecting NCR when a cut-off score of 6 was used (Jasinski et al., 2011). Reliable Digit Span was later licensed by Pearson (2009) and re-normed on several specific populations, including a simulated ADHD group; however, a cut-off score of 6 was maintained. As such, participants obtaining scores below 6 were considered 'unsophisticated' in their feigning response style and were categorized as having failed this PVT. Participants obtaining scores at 6 or above were considered 'sophisticated' in their NCR style.

21-Item Test (Iverson et al., 1991)

The 21-Item Test is a brief stand-alone PVT that takes 5 minutes to administer. To administer this measure, an examiner reads a list of 21 nouns, which is followed by a free-recall trial wherein participants are asked to repeat back as many words as possible. This trial is then followed by a forced-choice trial wherein participants are required to pick a target word from two given words (a word pair). The word pairs are either similar semantically or phonetically or are dissimilar. The authors stated that poor effort is indicated when an examinee makes more than 9 errors (Iverson et al., 1994). According to one meta-analysis, a cut-off score of 9 on the 21-Item Test fields high specificity (100%), but low sensitivity (21.96%), and a hit-rate of

60.65% (Vickery et al., 2001). Thus, this measure is most effective at detecting unsophisticated NCR. For the purposes of this study, participants obtaining scores above 9 were considered 'unsophisticated' in their response style and to have failed this PVT. Participants instructed to feign ADHD who obtained scores equal to or below 9 were considered 'sophisticated' in their response style. The 21-Item test may be obtained directly from the authors.

Test of Memory Malingering (Tombaugh, 1997)

The Test of Memory Malingering (TOMM) is a stand-alone PVT that takes about 15 minutes to complete. It is a 50-item measure that is disguised as a visual memory task. The TOMM contains two learning trials, during which participants are shown a series of 50 drawn pictures for three seconds each, followed by recognition trials that entail a forced choice between two pictures, only one of which was presented during the learning trial. According to the author, NCR is indicated when examinees make more than five errors, resulting in a cut-off score of 45. With this cut-off score, the TOMM demonstrated 100% sensitivity and 100% specificity by the author of the measure (Tombaugh, 1997), although studies demonstrate that these rates vary (Batt et al., 2008). In the current study, participants who obtained scores below a cut-off of 45 were considered 'unsophisticated,' and to have failed this PVT. Participants asked to feign ADHD who obtained scores at or above 45 were considered 'sophisticated.' The TOMM is purchasable through the publisher, Pearson.

Trail Making Test (Partington & Leiter, 1949)

The Trail Making Test (TMT) is a neuropsychological measure of processing speed and executive functioning, both of which may be impaired in individuals with ADHD (Barkley, 1997). The TMT has two trials (A and B). In TMT A, participants rapidly connect scattered encircled numbers with a pencil in ascending order. In TMT B, participants rapidly alternate

between connecting scattered encircled letters and numbers with a pencil in ascending order. Individuals with ADHD often have longer completion times on TMT B than their healthy counterparts (Nigg et al., 2005) as a result of EF deficits. Traditional Halstead-Reitan norms (Heaton et al., 2004) could not be used to establish a score for the current study, since the majority of participants were predicted to fall below age 20 years. The norms selected for this study (Tombaugh, 2004) were based on 911 Canadian community-based participants ages 18-89 years old (M = 58.5, SD = 21.7). The normative group used to derive scores for the current study had 155 individuals between 18-24 with a mean education level of 12 years (Appendix C).

Continuous Performance Test (Rosvold et al., 1956)

This version of the CPT, operated by Millisecond Software LLC, is a computerized version of the original CPT designed by Rosvold et al. in 1956. The original CPT utilized an apparatus that "consisted essentially of a revolving drum on which two series, each of 31 letters, were mounted side by side" (Rosvold et al., 1956, p. 345) and required examinees to sustain attention by responding (by pushing a button) to target stimuli whilst ignoring non-target stimuli. Like other versions of the CPT, this test measures attention and response control (Spreen et al., 2006).

Administration time is 13 minutes and consists of two trials of 6.5 minutes each. Instructions inform examinees of target and nontarget stimuli. During the first trial, examinees are instructed to press the spacebar every time the letter "X" appears on the screen, while the second trial requires examinees to press the spacebar only when the letter "X" is followed the letter "A." Both trials consist of two practice "blocks" and 10 experimental blocks, each of which consists of 31 letters, eight of which are targets. Letters are presented to participants for 600 milliseconds, with a 200-millisecond blank screen between presentations. After trial one, a 30-second break is presented before trial two begins. Four scores are ultimately derived from this

measure, consisting of omission and commission rates from both trials. A high number of omission errors generally indicates inattention while a high number of commission errors indicates deficient response control (Spreen et al., 2006). Since data regarding this version of the CPT and NCR do not exist, for the purposes of this study, cut-off points for both commission and omissions errors were established based on the control group scores. Scores were determined to indicate NCR if they differed from the control group's mean score by more than two standard deviations (Appendix D).

Wender Utah Rating Scale (Ward, 1993)

The WURS is a retrospective self-report questionnaire assessing childhood symptoms of ADHD in the domains of behavioral functioning, academic functioning, and history of ailments. Unlike other retrospective inventories, the WURS does not include a validity index to deter NCR. Norms for the WURS were selected based on the author's recommendation (Ward, 1993) of two cut-off scores: A score of 36 yields 96% sensitivity and specificity, and a score of 46 yields 86% sensitivity and 99% specificity. However, the author cautioned that a score of 46 should be utilized because a cut-off score of 36 does not adequately distinguish unipolar depression from ADHD. Thus, a score of 46 was used for this study, with participants obtaining scores of 46 or higher determined to have an "ADHD-profile."

Questions on the WURS reflect commonly known symptoms of ADHD (e.g., frequent daydreaming) and lesser known comorbid symptoms (e.g., low self-esteem and/or frequent anxiety). Participants producing scores that indicated the presence of ADHD symptoms (e.g., impulsivity, persistent attentional problems, motor hyperactivity) during childhood were considered to have engaged in successful NCR (Appendix E).

Procedures

The study took place in an assessment room in the basement of Uhler Hall at Indiana University of Pennsylvania. Five examiners were involved in the collection of data, including the principle investigator (PI), a graduate student in the clinical Psy.D. program, and three undergraduate psychology honors students. All assistants completed the Collaborative Institution Training Initiative (CITI) and were trained in the administration of the battery by the PI. Research assistants were observed administering each measure by the PI before they administered the battery to participants without supervision. The PI and research assistants were in frequent contact during data collection to ensure adherence to standardized procedures. The PI administered 36 batteries, while the four research assistants administered 32.

After signing a consent form (Appendix F), participants were asked to complete a brief demographics questionnaire that was reviewed by the examiner before continuing (See Appendix B). If the participant indicated a childhood history of ADHD or other exclusionary criteria, they were excused from the testing session and received credit for their participation. Next, participants were instructed to pick an envelope with instructions out of a shoebox. These instructions randomly assigned participants to one of four conditions: an unincentivized control group, an unincentivized NCR group, an incentivized-delayed NCR group, and an incentivized-immediate NCR group. The incentivized-immediate group was added halfway through data collection, after a cursory analysis revealed that an incentive did not appear to be affecting performance. It was theorized that the process of winning the reward (faking ADHD followed by winning a raffle) may have seemed unachievable or was not sufficiently motivating to participants; thus, an immediate reward condition was added. Participants in this group were offered $10 on the spot for passing the same reward criteria as the incentivized-delayed group.

Members of the control group received instructions to perform their best on all measures (Appendix G), while members of all three experimental groups were given three items: instructions/vignette (Appendices H-J), the DSM-5 diagnostic criteria for ADHD (Appendix K), and an affirmation of instructions (Appendix L). The examiner was blind to participants' group assignment.

The instructions/vignette for the experimental groups depicted a college student interested in dishonestly obtaining prescription psychostimulant medication. All participants in the experimental groups were instructed to "convince the examiner that [they had] ADHD;" however, participants in both incentivized groups were additionally offered rewards for successfully doing so. Instructions for the incentivized-immediate group informed participants that, "if you are able to successfully convince the examiner that you have ADHD, you will be given $10 at the end of this session" (See Appendix I), whereas instructions for the incentivized-delayed group stated, "If you are able to successfully convince the examiner that you have ADHD, you will be entered in a $100 Visa gift card raffle." (See Appendix J). After selecting an envelope, participants were given time alone to review the instructions, and were instructed to notify the examiner (who was sitting outside the examination room) when they were ready to begin.

The DSM-5 diagnostic criteria for ADHD (APA, 2013) given to participants included lists of the inattentive symptoms of ADHD and several hyperactive symptoms that specifically pertain to ADHD in adulthood (e.g., fidgeting, impulsivity; See Appendix K). Although participants were told that judgment regarding their ability to feign ADHD would be based solely on test scores, the lists of symptoms of ADHD was provided to assist participants in feigning ADHD.

An “affirmation of instructions” was added two weeks into data collection to ensure that participants were aware of the instructions after one participant reported difficulty with the task because she did not want the examiner to think she was “crazy” and thus, did not follow directions (this participant was removed from the data). This affirmation of instructions required examinees to copy two sentences verbatim, which affirmed that participants were aware that they were pretending to be a character (and not behaving as they normally would), and that the evaluation of ADHD would be based only on scores, and not on behavioral observations (See Appendix L).

Measures were administered in a counterbalanced order to control for fatigue, in which even-numbered participants completed two trials of the TOMM, followed by the CPT, 21-Item Test, TMT A and B, Digit Span Forward and Backward, and WURS, whereas odd-numbered participants completed these measures in the opposite order. At the end of the testing session, participants in the experimental groups were given a form to record the strategies they used to produce impaired ADHD-like scores (Appendix M). Afterwards, all participants were given a debriefing form (Appendix N), thanked for their participation, and asked not to reveal the procedure of the experiment to others. With respect to NCR participants who passed the reward criteria, participants in the incentivized-immediate group were given $10, while participants in the incentivized-delayed group had their emails added to a raffle list (that was not attached to their name or data). Additionally, participants were asked for informal feedback about their experience as participants in the study.

Hypotheses

Hypothesis 1

The first hypothesis concerned participants' ability to feign ADHD on neuropsychological tests, including the CPT, Digit Span, and Trail Making Test, and a measure of ADHD symptoms, the WURS. There were three parts to this hypothesis, in line with work by Quinn (2003), Suhr (2000), Iverson (2002) and Jasinski et al. (2011):

(i). The control group was hypothesized to produce normal scores on all six neuropsychological measures, whereas the other three groups were expected to produce impaired scores on all six neuropsychological measures (Quinn, 2003; Iverson, 2002; Jasinski et al., 2002). The NCR groups are hypothesized to produce scores that are significantly lower than the control group.

(ii). The unincentivized NCR group was hypothesized to produce scores that were significantly lower than both incentivized groups on all six neuropsychological measures (Suhr, 2000).

(iii). The control group was hypothesized to produce a mean WURS total score in the normal range, whilst all three NCR groups were hypothesized produce mean total scores consistent with an ADHD-profile (above 46).

Hypothesis 2

The second hypothesis concerned participants' ability to feign ADHD on PVTs, which included the TOMM, RDS, and the 21-Item Test. As such, this hypothesis had three levels, in line with work by Tombaugh (1997), Iverson et al. (1991), Pearson (2009), Schroeder et al. (2012), and Suhr (2000).

(i). Participants in the control group were expected to perform significantly less impaired than all NCR groups on all PVTs.

(ii). The control group was hypothesized to produce normal scores, within the optimal effort range on the TOMM (≥45), RDS (≥6), and 21-Item (<9), whilst the NCR groups were hypothesized to produce scores within the malingering range (Iverson et al., 1991; Pearson, 2009; Schroeder et al., 2012; Tombaugh, 1997).

(iii). Participants in both incentivized groups were expected to perform significantly less impaired than students in the unincentivized group on all PVTs.

Hypothesis 3

The third hypothesis of this study concerned NCR groups' sophistication in simulating ADHD. Research has suggested that participants in malingering studies behave in a more careful manner when made aware of the risks associated with detection (Suhr, 2000). To this end, these participants have been found to produce scores that were more sophisticated, or less clearly indicative of malingering, than participants unaware of risks associated with detection. Thus, participants in both incentivized NCR groups were hypothesized produce scores that were more "sophisticated," or less clearly noncredible, than members of the unincentivized group.

Specifically, participants in the NCR groups were categorized as engaging in "unsophisticated" NCR if they failed two or more PVTs, scored below a T-score of 35 on five or more neuropsychological measures, or produced a normal (non-ADHD) profile on the WURS (scoring below 46). Participants were considered to have engaged in "sophisticated" NCR if they failed no more than one PVT, scored below a T-score of 35 on no more than five neuropsychological measures and failed the WURS. All other combinations were deemed "inconsistent," which was also considered "unsophisticated NCR."

Hypothesis 4

The fourth hypothesis concerned all three NCR groups' strategies to feign ADHD. To this end, participants were given a pen and paper and an open-ended prompt to record the strategies they used to feign ADHD. Reported strategies were predicted to be similar to those found by Harrison et al. (2007), Frazier et al. (2008), and Quinn (2003), which included working slowly, deliberately making mistakes, missing items, and becoming distracted.

CHAPTER 3

RESULTS

Descriptive Analyses

A total of 76 undergraduate students from the Psychology Department subject pool participated. Seven participants' data were excluded from analyses because they reported a diagnostic history of ADHD and/or Specific Learning Disorder. One additional participant was excluded after she revealed that she did not follow the examiner's instructions. Data from the 68 remaining participants were analyzed for this study. As per the demographic questionnaire, the mean participant age was 19.06 years (*SD* = 0.86). The sample primarily identified as female (63.2%), white or Caucasian (85.3%), native English speaking (94.1%), right-handed (89.7%), and Freshmen (75%). Other demographic data are listed in Table 1. Means and standard deviations for all four groups for each measure are reported in Table 2. A one-way ANOVA was conducted to determine whether significant differences on neuropsychological measures and PVTs were based on battery order (balance/counter balanced), examiner, sex, ethnicity, academic class, language, and/or handedness. An ANCOVA was used to compare potential associations between age (a continuous variable) and scores. No relationships between performance and demographic factors, order, or examiner were found for any measure.

Table 1
Demographic Characteristics of Participants

		Group			
	N (%)	*Control n (%)*	*Unincentivized n (%)*	*Incentivized Delayed n (%)*	*Incentivized Immediate n (%)*
TOTAL	68	18 (26.5)	18 (26.5)	16 (23.5)	16 (23.5)
Sex					
Female	43(63.2)	10	13	12	8
Male	8(36.8)	8	5	4	8
Academic Class					
Freshman	51 (75)	13	13	12	13
Sophomore	12 (17.6)	2	3	4	3
Junior	5 (7.4)	3	2	0	0
Handedness					
Right	61 (89.7)	16	16	14	15
Left	6 (8.8)	1	2	2	1
Ambidextrous	1 (1.5)	1	0	0	0
Ethnicity					
Caucasian	58 (75)	13	16	15	14
African American	6 (8.8)	2	2	1	1
Asian Origin	3 (4.4)	3	0	0	0
Hispanic	1 (1.5)	0	0	0	1
Native Language					
English	64 (94.1)	16	17	16	15
Mandarin	3 (4.4)	3	0	0	0
Spanish	1 (1.5)	0	0	0	1

Table 2
Mean and Standard Deviation by Measure and Group

	Control (*n = 18)* M(SD)	Unincentivized *(n = 18)* M(SD)	In-de (*n = 16)* M(SD)	In-me (*n = 16)* M(SD)
TOMM (cutoff ≤ 45)	49.78 (.73)	32.39 (13.87) **	41.69 (7.61)	35.81 (9.33) **
21-Item (cutoff ≥ 9)	3.28 (2.24)	8.33 (4.13) **	5.63 (2.96)	7.25 (1.84) **
RDS (cutoff ≤ 6)	9.50 (2.04)	3.28 (7.88) **	7.88 (1.71)	7.69 (1.84)
DsF	49.61 (11.51)	36.56 (13.93) **	42.81 (9.28)	39.56 (9.48)
DsB	49.61 (10.10)	37.94 (12.32) **	41.56 (6.82)	41.13 (7.15)
TMT A	44.96 (10.82)	35.54 (39.05)	39.04 (12.37)	33.15 (15.10) *
TMT B	41.46 (11.90)	33.22 (12.66)	36.96 (15.45)	33.51 (13.59)
CPT Omissions X	3.53 (3.47)	39.61 (25.61) **	19.75 (18.60)	35.13 (19.12) **
CPT Omissions AX	3.53 (4.71)	26.00 (18.23) **	12.94 (12.17)	21.50 (12.77) **
CPT Commissions X	1.41 (1.29)	13.28 (12.88) **	11.19 (12.70) *	11.00 (8.29) *
CPT Commissions AX	3.47 (2.50)	17.06 (19.16)	19.13 (21.44) *	16.31 (15.12)
WURS	25 (13.10)	50.67 (16.65) **	49.19 (23.90) **	50.19 (20.16) **

In-de = Incentivized-Delayed, In-me = Incentivized-Immediate; TOMM=Test of Memory Malingering; RDS=Reliable Digit Span; DsF= Digit Span Forward; DsB = Digit Span Backward; TMT = Trail Making Test; CPT= Continuous Performance Test; WURS= Wender Utah Rating Scale; *P<0.05 difference from control group, **P<0.01 difference from control group

Sensitivity and Specificity

Sensitivity analyses of the PVTs was predicated on participants deliberately engaging in NCR. Thus, the researcher examined the data for participants who may not have followed directions and removed these scores from this analysis. This included participants who produced scores that were indistinguishable from the control group, passed all PVTs, produced a neurotypical profile on the WURS, and who produced scores in the borderline impaired range on two or fewer neuropsychological measures. The borderline impaired range included scores on Dsf, Dsb, TMT A, and TMT B that were t<35, and errors on the CPT (all four trials) that fell 1.5 standard deviations above the control group: Omission X > 8.74, Omission AX > 10.60, Commission X > 3.12, and Commission AX > 7.22. These scores were used for the CPT rather than t-scores because the control group produced low means and standard deviations, and as a

result, many NCR group t-score means fell within the negative range. Thus, CPT scores were calculated by summing the number of errors per trial.

The production of scores in the borderline impaired range on two neuropsychological measures was permitted because 25% of participants in the control group produced scores in this range on two neuropsychological measures. This is not an unusual finding, as normal individuals often produce low scores on isolated measures in a battery due to individual characteristics (Schretlen et al., 2003). Overall, six examinees were identified - one from the incentivized-immediate group, three from the incentivized-delayed group, and two from the unincentivized group - who produced scores that were similar to the control group, and their data were omitted from this analysis. Ultimately, this left 44 participants who were deemed to have engaged in NCR and who were included in the sensitivity/specificity analyses.

Test of Memory Malingering

This study found that nine out of 44 participants in the NCR groups passed the TOMM, while all participants in the control group passed. Thus, the TOMM demonstrated 80% sensitivity and 100% specificity at detecting NCR with this sample. When NCR groups were stratified, the TOMM demonstrated 69% sensitivity (and 100% specificity) detecting nine of 13 participants in the incentivized-delayed group, 81% sensitivity detecting 13 out of 16 participants in the unincentivized group, and 87% sensitivity detecting 13 of 15 of the incentivized-immediate group.

Reliable Digit Span

In the current study, the RDS correctly detected 11 of 44 NCR participants as engaging in NCR, and falsely identified one of 18 control participants, yielding 25% sensitivity and 94.44% specificity. When NCR groups were stratified, RDS correctly identified two of 13 members of

the incentivized-delayed group as engaging in NCR (15% sensitivity), one out of 13 participants of the incentivized-immediate group (7% sensitivity), and seven of 16 participants of the unincentivized group (44% sensitivity).

21-Item

Results from the current study found that 11 of 44 NCR participants were correctly detected as engaging in NCR, with one member of the control group incorrectly identified, leading to an overall sensitivity of 25% with 94.44% specificity. The 21-Item detected two of 13 members of the incentivized-delayed group (sensitivity of 15%), two of 15 members of the incentivized-immediate group (sensitivity of 13%), and seven of 16 members of the unincentivized group (sensitivity of 44%).

Overall Sensitivity and Specificity

Results from the current study show that the sensitivity in detecting NCR varied by PVT, although the TOMM was far more efficient than the 21-Item or RDS. Participants within the unincentivized group were detected more often than both incentivized groups, which was most apparent on the 21-Item and RDS, for which sensitivity was 44% on both measures, compared to sensitivity of 15% for both incentivized groups. With respect to the unincentivized group, of the five participants who failed only two PVTs, all failed the TOMM, two failed RDS, and three failed the 21-Item. These data suggest that the 21-Item and RDS are only sensitive to blatant and unsophisticated forms of NCR, while the TOMM is best suited for all cases.

Results of Hypotheses

Neuropsychological Measures (Hypothesis 1)

In the first part of Hypothesis 1, it was predicted that participants in the control group would perform in the normal range while participants in the NCR group would produce scores

that were in the borderline impaired range. Additionally, it was predicted that participants in the control group would produce scores reflecting significantly less than all three NCR groups on all neuropsychological measures. As predicted, the control group produced normal scores on all neuropsychological measures; thus, this group consisted of 'normal' participants as intended by the researcher. Surprisingly, the majority of NCR participants also produced scores in the normal range, although variability across measures was noted. All NCR groups produced borderline impaired scores (or below) on all four trials of the CPT. Additionally, both the unincentivized group and incentivized-immediate group produced borderline impaired scores on TMT B, while only the incentivized-immediate produced borderline impaired scores on TMT A. It should be noted that all NCR groups' scores that did not fall within the borderline impaired range on Dsf, Dsb, TMT A, and TMT B, all fell within the low average range.

To compare each NCR groups' performance on neuropsychological measures to the control group, eight one-way between-subjects ANOVAs were conducted. Significant differences ($p < 0.05$) were found between groups on DsF, DsB, TMT A (Figure 1), and all four trials of the CPT (Omission trials A and AX, Commission trials A and AX; see Table 2 and Figure 2). Post-hoc comparisons using Tukey HSD indicated that the control group was significantly different from at least one NCR group on each measure; however, the control group did not differ from each NCR group. For example, the control group was significantly different from the unincentivized group on DsF, but not from either incentivized group. Also, the control group was significantly different from the incentivized-immediate group on TMT A, but not the unincentivized group or the incentivized-delayed group (see Table 3). Overall, these data show that the control group produced scores reflecting significantly less impairment than the unincentivized group on five out of eight measures, four out of eight compared to the

incentivized-immediate group, and two out of eight compared to the immediate-delayed group, meaning that this part of Hypothesis 1 was partially supported.

It was also hypothesized that students in both incentivized groups would produce scores reflecting significantly less impairment than students in the unincentivized group on all neuropsychological measures. Performances by all three NCR groups on all neuropsychological measures were compared using eight one-way between-subjects ANOVAs. Significant differences were found between groups on both CPT Omission trials: trial X [(F 2, 47) = 3.858, $p < 0.028$, $\eta^2 = 0.14$] and trial AX [(F 2, 47) = 3.400, $p < 0.042$, $\eta^2 = 0.13$]. Post hoc comparisons using Tukey HSD showed that the unincentivized group differed from the incentivized-delayed group on both Omission trials ($p < 0.05$) in the hypothesized direction. No other significant differences were found between NCR groups on any other neuropsychological measure. Overall, although significant differences were found between the incentivized-delayed and unincentivized groups on two CPT scores, no differences were found on the other six measures, so this part of Hypothesis 1 was not supported.

The final part of the first hypothesis was that students in the control group would produce a normal profile on the WURS, while students in all three NCR groups were expected to produce an ADHD profile. To address this, group means and standard deviations on WURS scores were derived for all four groups (see Table 2). Results demonstrate, as predicted, that the control group mean score fell within the normal range, while all three NCR group mean scores fell within the ADHD range. A one-way between subjects ANOVA [(F 3, 64) = 7.961, $p < 0.004$, $\eta^2 = 0.27$] and subsequent Tukey HSD ($p < 0.002$) showed that the mean score of the control group was significantly lower than the mean scores for all three NCR groups (see Table 2). With respect to individual participants, data show that with the exception of one participant,

the control group produced a normal profile on the WURS, whilst most NCR participants did not. Of the NCR groups, 11 out of 17 (64.7%) of the unincentivized group exceeded the cut-off for ADHD, 10 out of 15 (66.7%) of the incentivized-immediate group exceeded the cut-off, and 9 out of 15 (60%) of the incentivized-immediate group scored above the cut-off. In sum, this last portion of Hypothesis 1 was supported, as the control group produced a neurotypical mean score on the WURS that was significantly lower than mean scores for all three NCR groups, which fell within the ADHD range. However, it should be noted that the range of scores was greater for the NCR groups, and a large minority of participants produced normal profiles.

Table 3

Tukey HSD of Control Group Versus NCR Groups on Neuropsychological Tests

			Mean		95% Confidence Interval	
Measures	Group	Group	Difference	Sig.	Lower Bound	Upper Bound
Digit Span - Forward	Control	Unincentivized	13.056*	.005	3.12	22.99
Digit Span - Backward	Control	Unincentivized	11.667*	.003	3.31	20.02
Trail Making Test - A	Control	In-im	11.80653*	.045	.1945	23.4185
CPT Omissions X	Control	Unincentivized	-36.222*	.000	-52.57	-19.87
		In-im	-31.736*	.000	-48.59	-14.88
CPT Commissions X	Control	Unincentivized	-11.778*	.004	-20.49	-3.07
		In-de	-9.688*	.030	-18.67	-.71
		In-im	-9.500*	.034	-18.48	-.52
CPT Omissions AX	Control	Unincentivized	-22.500*	.000	-33.80	-11.20
		In-im	-18.000*	.001	-29.65	-6.35
CPT Commissions AX	Control	In-de	-15.458*	.032	-29.93	-.99

Note: In-de = Incentivized-Delayed, In-im = Incentivized-Immediate

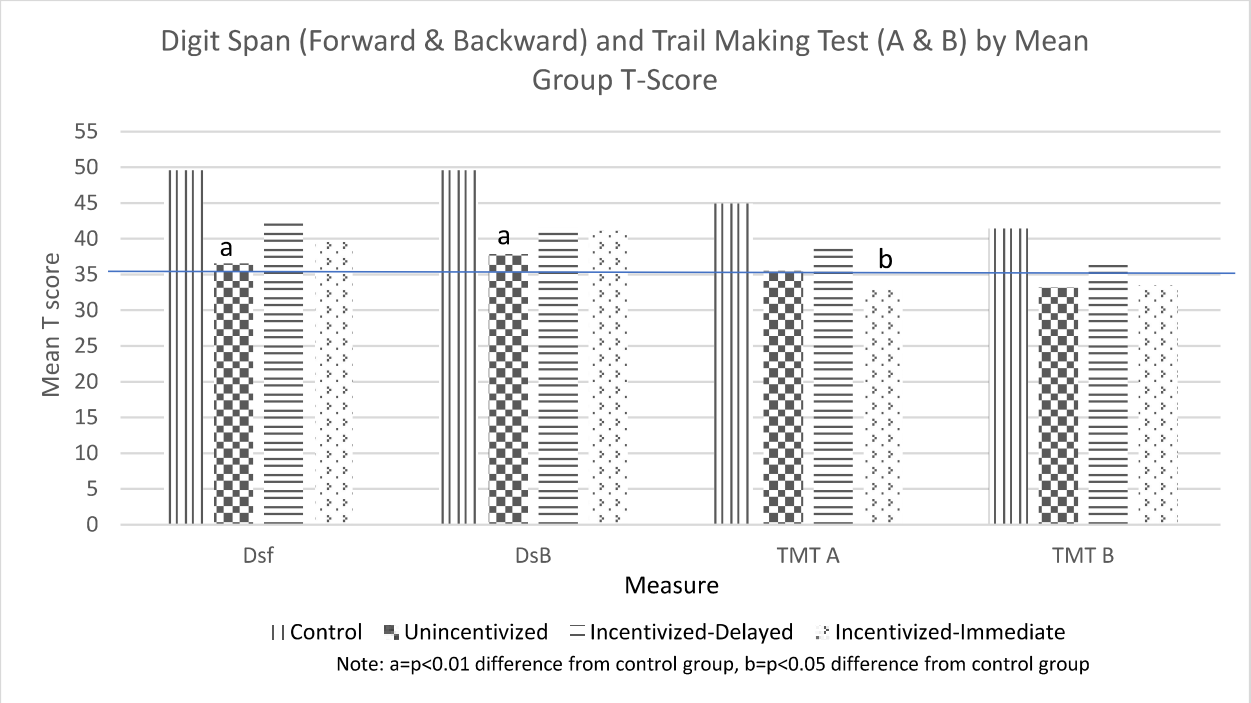
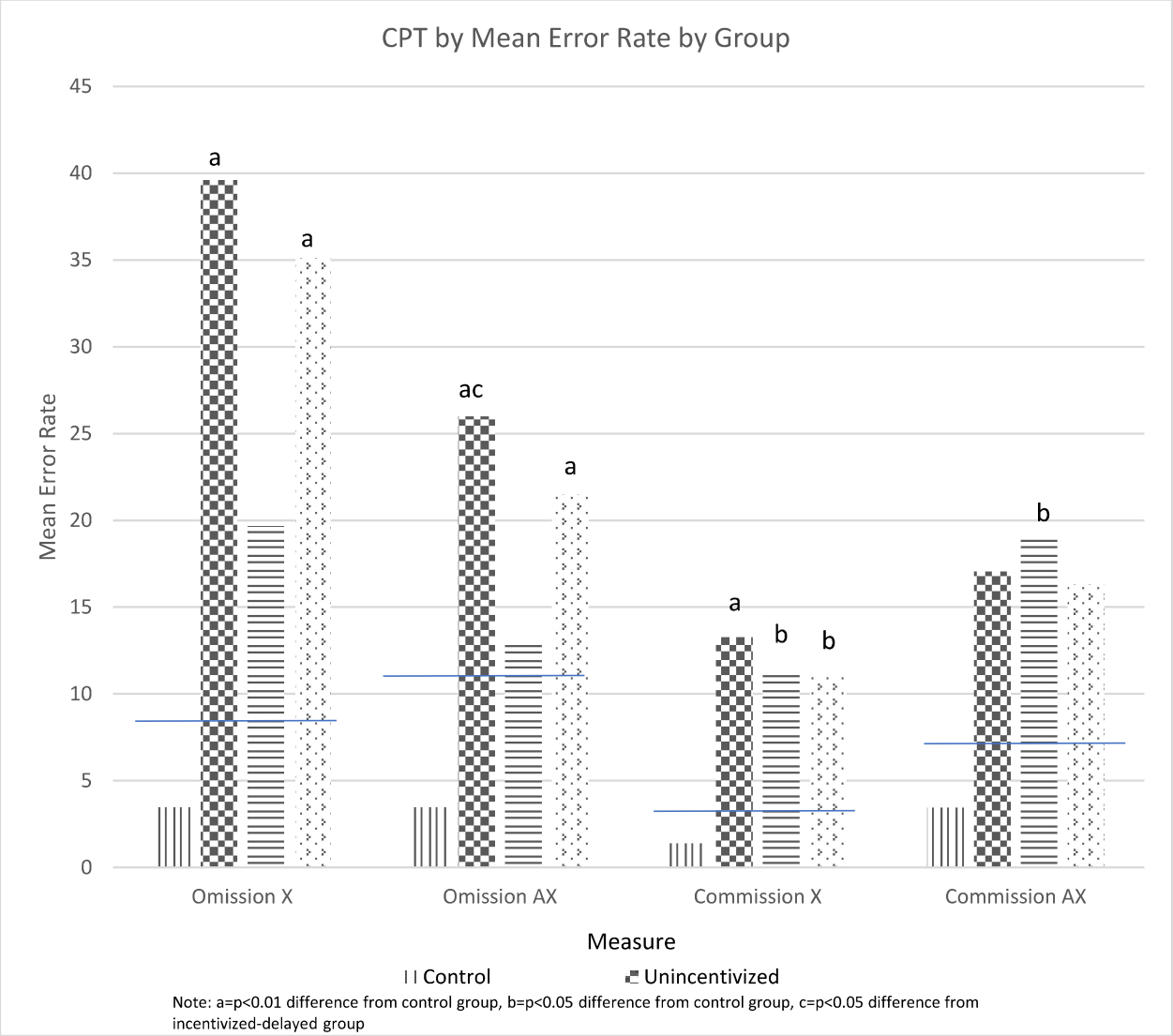


Figure 1. Digit span (forward & backward) and trail making test (A & B) by mean group t-score.

Figure 2. CPT mean error rate by group.



Performance Validity Tests (Hypothesis 2)

Hypothesis 2 had three parts. It was first hypothesized that participants in the control group would produce scores reflecting significantly less impairment than all NCR groups on all three PVTs. To address this, three one-way between-subjects ANOVAs were conducted to compare performance on PVTs by the control group with each NCR group. Results showed that there were significant between-group differences for the TOMM [(F 3, 64) = 12.087, $p < 0.000$, $\eta^2 = 0.36$], 21-Item [(F 3, 64) = 9.854, $p < 0.000$, $\eta^2 = 0.32$] and RDS [(F 3, 64) = 5.096, $p < 0.003$, η^2

= 0.19]. Post-hoc Tukey HSD analyses indicated that the control group mean scores were significantly different from the unincentivized group on all three PVTs (Figures 3 - 5), and from the incentivized-immediate group on the TOMM and 21-Item test in the hypothesized direction, but not the RDS. Contrary to expectations, no significant differences were found between the control group and the incentivized-delayed group on any PVTs. Thus, this part of Hypothesis 2 was partially supported.

It was also hypothesized that the mean score of the control group would fall in the normal range, above established cut-off scores for NCR, whilst mean scores of all NCR groups would fall below. To address this, group means and standard deviations for all four groups on all three PVTs were compared. PVT cut-off scores for NCR (see Table 2) were predetermined based on normative data as discussed in the prior chapter. As such, the control group produced mean scores within the normal range on all three measures, consistent with predictions. No member of the control group failed the TOMM, and in fact, only three participants did not obtain perfect scores of 50. One control participant failed the 21-Item and another failed RDS. These participants' scores were not outliers on any other tests, so they were retained for analyses. All three NCR groups produced mean scores below the TOMM cut-off, consistent with the hypothesis (Figure 3). Contrary to expectations, however, only the unincentivized group produced a mean score below the cut-off on RDS (Figure 4). Further, no NCR group produced a mean score that fell above the cut-off on the 21-Item Test (Figure 5). Overall, this hypothesis was partially supported, in that the control group did produce mean scores above the cut-off and all NCR groups produced mean TOMM scores below the cut-off; however, the three NCR groups did not meet expectations on the RDS or 21-item Test.

It was further hypothesized that participants in both incentivized groups would produce scores reflecting significantly less impairment than participants in the unincentivized group on all PVTs. Three between-subjects one-way ANOVAs were conducted to compare all three NCR groups across all three PVTs, as mentioned above. Post-hoc analyses showed that the incentivized-delayed group was significantly different from the unincentivized group on both the TOMM ($p < 0.05$) and the 21-Item ($p < 0.05$). No other significant differences were found. These data suggest that the incentivized-delayed group came closer to producing a "normal" profile than the unincentivized group, although this pattern was not prominent or pronounced, and did not extend to the incentivized-immediate group. Overall, these findings partially support the hypothesis, in that one of the incentivized groups performed significantly less impaired than the unincentivized group on two PVTs.

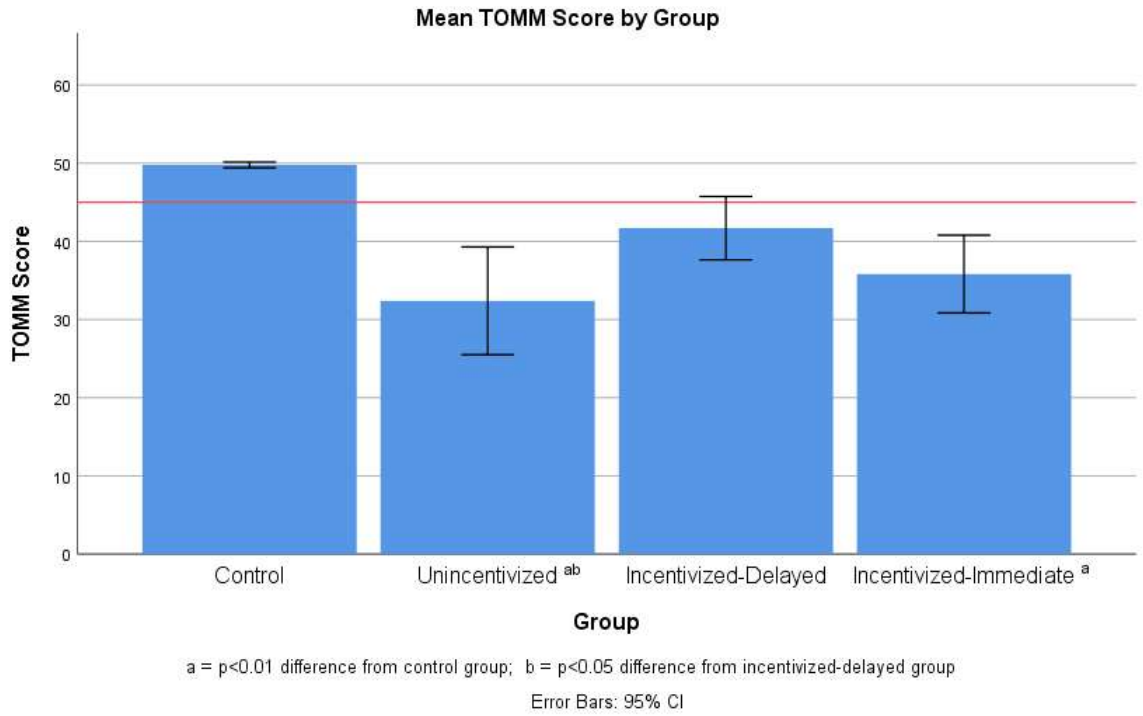


Figure 3. Mean TOMM score by group.

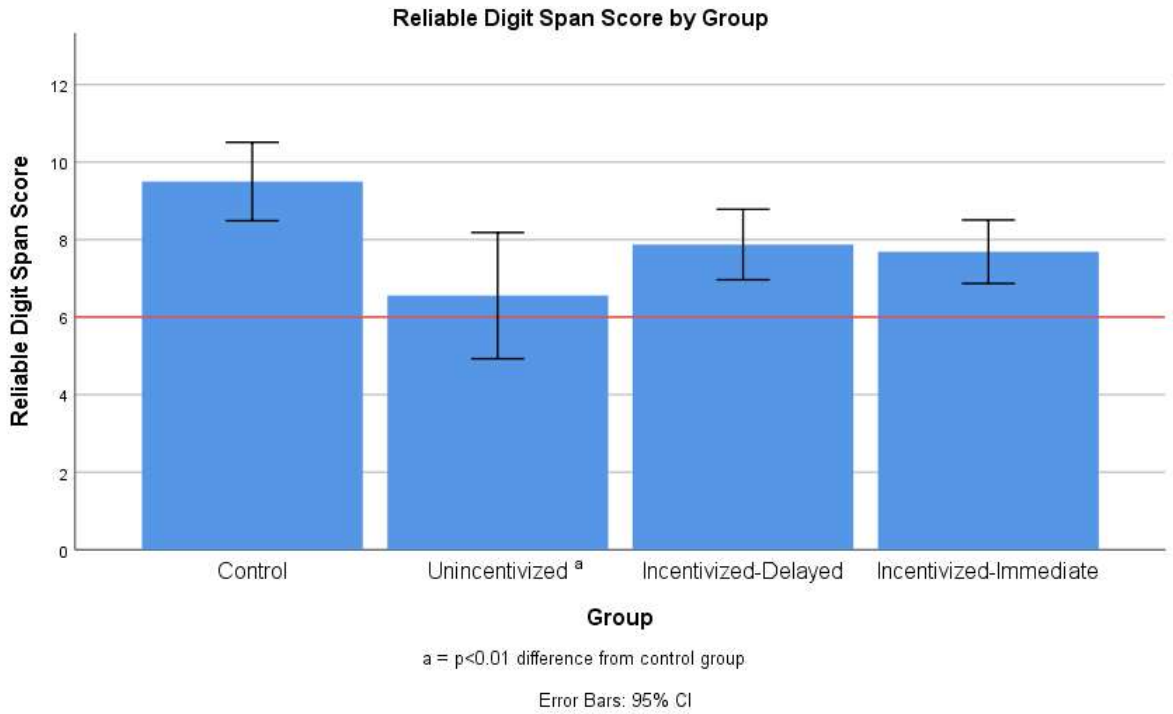


Figure 4. Mean reliable digit span score by group.

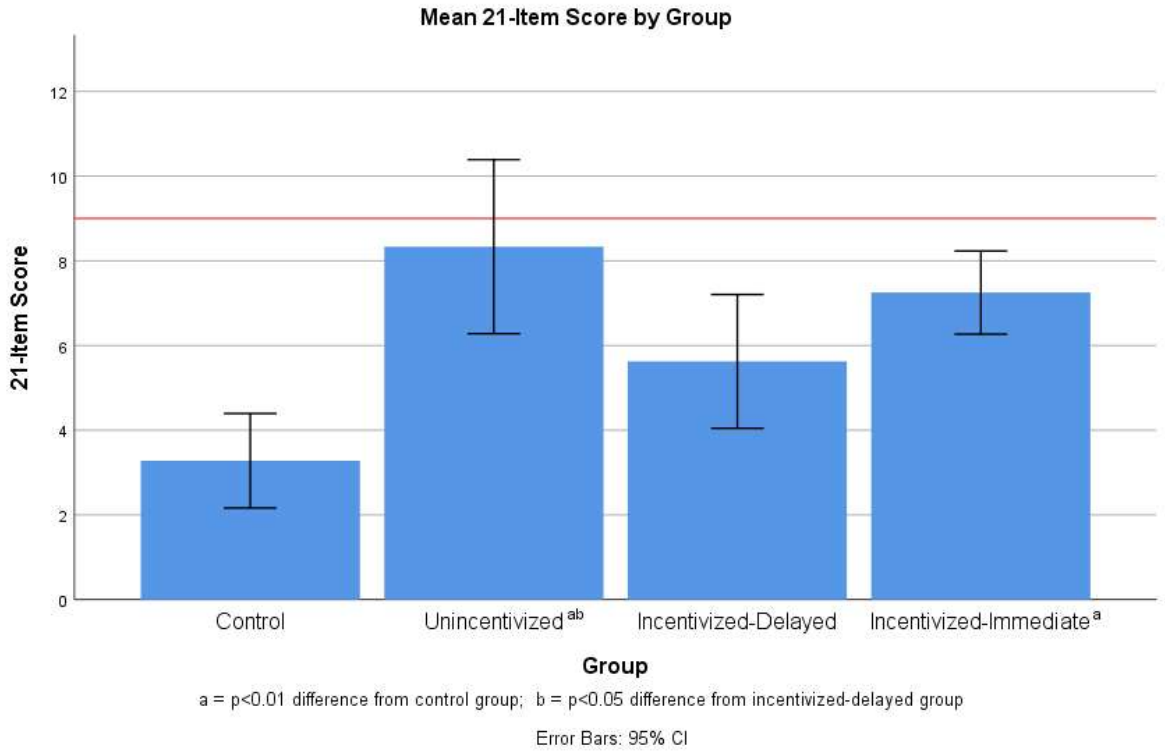


Figure 5. Mean 21-Item test score by group.

Sophistication (Hypothesis 3)

It was additionally hypothesized that members of the unincentivized NCR group would engage in "unsophisticated" NCR, while members of both incentivized NCR groups would engage in "sophisticated" NCR as a result of the incentive offered for performance. For this study, the criteria for "sophisticated" were met when a participant produced an ADHD profile on the WURS, failed no more than one PVT, and produced borderline impaired scores (T<35) on no more than five neuropsychological measures. One failed PVT was permitted in the sophisticated group because a) there is a nonzero false-positive rate for all PVTs (Larrabee, 2012), and b) the specificity for identifying NCR is substantially higher when more than one PVT is utilized [i.e., 100% according to Sollman et al. (2010)]. The rationale to categorize participants who produced five or more impaired scores as "unsophisticated" was that this pattern would not typically be observed in an ADHD population, and would likely reflect a more serious neurological impairment (Woods et al., 2002). Another classification ("no diagnosis") was given to participants who did not produce any results indicative of a clinical impairment in any category (i.e., failed ≤1 PVT, passed WURS, impaired on ≤5 neuropsychological measures).

Overall, data suggest that the unincentivized group did not differ from the incentivized groups in their NCR sophistication (Table 4). Using a 3x3 chi-square test of independence, group differences were not statistically significant $\chi^2(2, N = 68) = 0.16$, $p = 0.93$. All three groups produced similar ADHD-profiles on the WURS and produced low scores (T<35) on fewer than five neuropsychological measures (Table 5); however, more members of the unincentivized group failed more than one PVT (50%) than in both the incentivized-delayed (25%) and incentivized-immediate groups (19%). Thus, while there did appear to be differences in some areas, the null hypothesis failed to be rejected.

Table 4
Participants in Sophistication Groups

Group	Sophisticated n (%)	Unsophisticated n (%)	No Diagnosis n (%)
Incentivized - Immediate	4 (25.2)	7 (43.75)	5 (31.25)
Incentivized - Delay	5 (31.25)	6 (37.50)	5 (34.25)
Unincentivized	3 (16.67)	11 (61.11)	4 (22.22)

Table 5
Participants Meeting Sophistication Criterion

	ADHD Profile on WURS n(%)	Failed <2 PVTs n(%)	Produced Low Scores (T<35) on <5 Neuropsychological Measures n(%)
Incentivized -Delayed (N=16)	10 (62.5)	12 (75)	13 (81)
Incentivized -Immediate (N=16)	9 (56)	13 (81)	11 (69)
Unincentivized (N=18)	11 (61)	9 (50)	13 (72)
Total (N=50)	30 (60)	34 (68)	37 (74)

Strategies (Hypothesis 4).

Finally, it was hypothesized that participants in the unincentivized NCR group would report similar methods of NCR as the incentivized NCR groups. Methods endorsed were expected to be similar to those found by Quinn (2003), Frazier et al. (2008), and Harrison et al. (2007), which included responding slowly, quickly, and/or inconsistently, making deliberate errors (commission/omission), zoning out, and acting "hyperactive." To address this, students were asked in an open-ended fashion to report everything they did to produce scores that resembled ADHD. Responses were coded by the PI into three general categories: endorsement of general symptoms specific to DSM-5 ADHD (Table 6); purposely giving wrong answers or withholding a correct answer (Table 7); or specific measure-related mistakes (e.g., the response "I didn't press the spacebar when I saw the X," was coded as a CPT omission error; Table 7). Data were examined both by individual experimental group and as a combined group to permit comparison with other studies. It should be noted that participant responses were coded into any of the three general categories into which they fit. For example, a response of "clicking the spacebar too

much" was coded as both an impulsive response and a specific error on the CPT. However, measure-specific responses were only coded once into any symptom-specific category (e.g., a CPT commission error could not be coded as both impulsivity and inattention).

With respect to coding strategies by ADHD symptoms, responses were categorized by either attention deficit symptoms or hyperactive symptoms and were based upon the DSM-5 ADHD criteria (APA, 2013). Attention deficit symptoms included behaviors that suggested the participant was distracted, inattentive, avoidant, and/or forgetful. A strategy was categorized as "distracted" if an examinee indicated that something had captured their attention away from the task or that they had been looking around the room. The criteria for "inattention" included participant responses that indicated difficulty sustaining attention, zoning out, missing instructions or making careless mistakes (like on TMT). "Avoidant" included behaving in a bored or frustrated manner (e.g., "sighing"). "Forgetful" included responses such as "forgot some letters, numbers, pictures."

Hyperactive ADHD symptoms included fidgetiness, restlessness, and impulsivity. "Fidgetiness" entailed any responses that suggested minor physical movements, like tapping hands or feet or playing with hair or clothing. "Restlessness" included leg shaking and more overt physical behaviors, such as shifting in chair, and/or standing up. "Impulsivity" entailed interrupting the examiner, answering quickly, and "asking off the wall questions." An additional symptom category "slow" was added based on participant responses and previous studies. Although this is not a symptom of ADHD, participant responses such as "took my time drawing lines," and "answer slowly" were coded as such.

Overall, the NCR combined group reported that they had demonstrated inattention (70%), had acted distracted (56%), were forgetful (16%), and were avoidant of attentional tasks (8%;

Table 7). A total of 58% stated they were fidgety, 42% reported that they were restless, and 34% reported being impulsive in their responding patterns. An additional 28% reported that they had responded slowly (a non-ADHD behavior). In terms of error type, 34% said they purposefully committed commission errors, while 8% engaged in omission errors.

The current study found results that varied in comparison to previous research. Quinn (2003) found that 61% of participants engaged in "general" inattentive symptoms, including ignoring visual (43%) and auditory (17%) stimuli. Regarding hyperactive symptoms, they found that 30% reported "double clicking the mouse" (during a CPT), while 13% engaged in "general" fidgety behavior. Nine percent behaved in a slow manner. Additionally, 57% reported commission errors, while 35% engaged in omission errors. Quinn's (2003) results are consistent with the current study with respect to the reported behaviors, but inconsistent in the rates at which participants reported engaging in these behaviors. For example, 58% of all NCR participants in the current study reported "fidgety" behavior, compared to 13% in Quinn (2003).

Harrison et al. (2007) did not categorize data based on ADHD criteria, and instead on specific behaviors. Comparison with the current study can be found in Table 8. For example, Harrison et al. included categories such as "fidgeting" (40%) "acting bored," (3%), "zoning out" (26%), "completing tasks slowly" (26%), and "completing tasks quickly and carelessly" (20%), "skipping items "(23%) and "deliberately choosing incorrect answers" (23%; p.584). These results were similar to the current study, such as acting bored (4%), responding in a slow manner (28%). However, other results differed, such as fidgeting (58%), responding quickly (10%), making deliberate mistakes (34%) and purposely skipping items (8%). Overall, like the comparison with Quinn (2003), similar strategies were used, but rates differed.

Frazier et al. (2008) employed a different method of collecting participants' strategies, and instead gave examinees a series of yes/no questions, (e.g., "Did you respond slowly?") As such, they found that 90% reported engaging in inattentive behaviors, 74% were forgetful, 77% were slow, and 87% made deliberate errors (both commission and omission errors), which also differed markedly from this study's data.

Overall, this hypothesis was supported, in that general strategies employed by participants in the current study were similar to prior research; however, the percentage of participants engaging in these respective strategies varied by method of data collection. For example, while all studies found that acting in an inattentive manner was a common strategy, endorsement varied from 26% to 90% across all four studies. Also, with the exception of Frazier et al. (2008), studies found that participants frequently committed commission errors and engaged in fidgety behavior.

Additionally, the current study found that behaving in a hyperactive manner was more common than in other studies; however, this was more the case with the unincentivized group. Results also showed that between 28% and 77% of participants across three of the four studies found that behaving in a slow manner (e.g., answering slowly) was a common strategy, even though this is not a symptom of ADHD.

Table 6
Percentage of NCR Participants Reporting ADHD Symptoms (%)

		Unincentivized	Incentivized (Delayed)	Incentivized (Immediate)	Total
Hyperactivity	Fidgetiness	72.22	56.25	43.75	58.00
	Restlessness	44.44	37.50	43.75	42.00
	Impulsivity	22.22	56.25	25.00	34.00
Attention Deficit	Distracted	50.00	50.00	68.75	56.00
	Inattention	66.67	68.75	75.00	70.00
	Dislikes/Avoids Attention tasks	0.00	6.25	18.75	8.00
	Forgetful	0.00	25.00	25.00	16.00
Non-ADHD	Slow	33.33	25.00	25.00	28.00

Table 7

NCR Participants Reporting Specific Errors (%)

	WURS	CPT	Digit Span	TMT	TOMM	21-Item	Total Commission Errors	Total Omission Errors
Unincentivized	27.78	0.00	22.22	0.00	22.22	11.11	27.78	11.11
Incentivized (Delayed)	31.25	12.50	18.75	6.25	6.25	0.00	31.25	12.50
Incentivized (Immediate)	12.50	6.25	25.00	0.00	6.25	0.00	43.75	0.00
Total	24.00	6.00	22.00	2.00	12.00	4.00	34.00	8.00

Table 8

Percent Reporting Specific Strategies in Harrison et al. (2007) Versus Current Study

	Current study	Harrison et al. (2007)
Fidgeting	58	40
Acting bored	4	3
Zoning out	6	26
Completing tasks slowly	28	26
Completing tasks quickly and carelessly	10	20
Skipping items	8	23
Deliberately choosing incorrect answers	34	23

CHAPTER 4

DISCUSSION

Overview of Findings

This study evaluated the effects of a contingent reward on college students' ability to feign Attention-Deficit/Hyperactivity Disorder (ADHD) on neuropsychological measures while avoiding detection on performance validity tests (PVT). It also included a retrospective ADHD self-report. The rationale for this study was initially based upon on Suhr (2000), who found that motivation and performance can be altered by making participants aware of PVT utilization. As such, this study sought to extend her findings by investigating changes that financial incentives have on participant behavior beyond instructions to feign a particular response style. While it is well established that individuals alter behavior upon becoming aware of observation (i.e., the Hawthorne Effect; Landsberger, 1958), Suhr (2000) was the first to demonstrate that assessment-related NCR is altered based on knowledge of observation. In this study, participants were randomly assigned to one of four conditions: a control group, an unincentivized group, or one of two incentivized groups. Financial incentives for the incentivized groups were either $10 at the conclusion of the testing session (for the incentivized-immediate group) or the opportunity to win $100 at the conclusion of the experiment via a raffle (for the incentivized-delayed group).

In previous studies of simulated malingerers, researchers relied upon participants to put forth their best effort to feign impairments (Frazier et al., 2008; Harrison, Edwards, & Parker, 2007; Quinn, 2003). In the first study that utilized simulated malingerers during an ADHD evaluation, Quinn (2003) concluded that future researchers should address issues related to the difficulty of assembling a genuine malingering sample by offering an external reward, since the motivation of undergraduates may be dissimilar to a true malingering sample. Similarly,

Harrison et al. (2007) concluded that future researchers may want to incentivize ADHD-related behavior because these participants do not have a stake in the outcome of the researcher's findings. As such, this study attempted to provide such a reward to examine the impact on participant responding.

Since individuals who engage in malingering, by definition, stand to gain (or by extension lose) an external reward, a contingent reward was included in this study of the effects of incentivizing simulated malingering behavior. Within the body of research on NCR and malingering, three studies have offered participants an external reward, but did not include an unincentivized group for comparison purposes (Booksh et al., 2010; Fisher, 2007; Sollman et al., 2010); thus, it is unknown whether or how these incentives altered behavior beyond standard instructions. Additionally, keeping with past studies (Frazier et al., 2008; Harrison et al., 2007; Quinn, 2003), this study examined the subjective strategies utilized by participants to produce ADHD profiles. If a pattern of common strategies is evident in the reports of the NCR participants, the observation of such behavior provides clinicians with a model to compare against, or specific behaviors (in addition to a failed PVT) to watch for when NCR is suspected in an evaluation.

The first hypothesis of the current study concerned differences in neuropsychological test and self-report scores between the groups. It was first predicted that performance of participants in the control group would fall in the normal range, while performance of participants in the NCR groups would fall in the borderline impaired range on all neuropsychological measures (Dsf, Dsb, TMT A & B, CPT [all trials]). This part of the hypothesis was partially supported, as the control group mean scores did fall in the normal range, but NCR group mean scores varied by measure and specific NCR group, with some scores falling within or below the borderline

impaired range, but most falling within the low average range. These results show that the NCR groups performed similarly to the control group. This was unforeseen because it is in direct contrast with several previous studies (Booksh et al., 2010; Quinn, 2003; Sollman et al., 2010), all of whom reported grossly impaired scores by their respective ADHD NCR groups. However, the current study's NCR group scored comparably with these authors' clinical ADHD samples. Thus, these results show that the current NCR group, contrary to expectations based on previous research, were able to produce scores that resembled a true ADHD sample, and did not produce grossly impaired scores.

The first part of Hypothesis 1 also predicted that the control group would produce scores on neuropsychological measures that were significantly different than NCR groups. Results show that the control group was significantly different, in the hypothesized direction, from the unincentivized group on the majority of neuropsychological measures (DSf, DSb, CPT-all four trials). Both incentivized groups produced scores that were significantly different from the control group: four for the incentivized-immediate group (TMT A and CPT-three trials) and two for the incentivized-delayed group (CPT-two trials). Overall, there were significant differences between the control group and NCR groups on all measures except Trails B.

The second part of the first hypothesis predicted that participants in both incentivized groups would produce scores reflecting significantly less impairment than participants in the unincentivized group on all neuropsychological measures. This was predicated on the assumption that participants in the incentivized groups would put forth effort to avoid detection while performing in a way that was suggestive of ADHD, while participants in the unincentivized group, lacking an incentive, would be more easily detected as feigning. As such, both incentivized groups produced scores that were closer (but not significantly) to the control

group than to the unincentivized group. However, the unincentivized group differed only from the incentivized-delayed group, and on only CPT omission rate (both trials); thus, this portion of Hypothesis 1 was not supported. These results indicate that an incentive did not alter behavior between the NCR groups to a significant degree on neuropsychological tests**.** It is unclear why differences were found for only the CPT, and notably, no members of the unincentivized group reported making deliberate omissions on the CPT; however, this difference may be due to chance, as 24 analyses with a p-value of 0.05 (eight neuropsychological measures across three NCR groups) were conducted to test this part of the hypothesis, inflating the risk of Type 1 error.

The final portion of the first hypothesis correctly predicted that participants in the control group would produce a neurotypical profile on the WURS, while participants in all three NCR groups would produce an ADHD profile. The WURS has demonstrably strong sensitivity (86%) to identify ADHD and strong specificity (99%) to identify normal controls (Ward, 1993). It is also susceptive to NCR, as demonstrated by Jachimowicz and Geiselman (2004), who found that 65% of an NCR sample were able to produce an ADHD profile despite not having ADHD. The current study found results that were fairly commensurate, in that 60% of the NCR sample produced a profile consistent with ADHD (specifically, 61% of the unincentivized group, 56% of the incentivized-immediate group, and 62.5% of the incentivized-delayed group). Additionally, Jachimowicz and Geiselman (2004) found that of the four ADHD self-report measures studied (Conners Adult ADHD Rating Scales, Brown Adult ADHD Scale, ADHD Rating Scale, and WURS), participants were least successful at producing a profile consistent with ADHD on the WURS, suggesting that it is more difficult to produce an ADHD profile on the WURS than other ADHD self-report measures, which may explain why 40% of the NCR participants in the current study failed produce such a profile on the WURS. Overall, Jachimowicz and Geiselman (2004)

suggested that it is best practice to examine scores on self-report measures with the understanding that they are susceptible to NCR; this idea is echoed by other researchers as well (Harrison et al., 2007; Suhr et al., 2008).

Results from the first hypothesis indicate, contrary to expectations, that NCR participants were efficient at producing scores that resemble ADHD on neuropsychological measures. It was initially predicted that NCR group scores would be similar to Booksh et al. (2010) NCR group, and fall mostly below the borderline impaired range. However, most NCR group mean scores fell above the borderline impaired range, albeit below the average range (within the low average range). This is in line with previous research that suggests that individuals with ADHD often produce scores on neuropsychological measures that are lower than normal individuals (Schoechlin & Engel, 2005); however, these differences seldomly represent widespread impairments (Conant, 2014). Thus, these groups succeeded in producing scores that may be expected from an individual with ADHD. This was the case even with the unincentivized group, though group means were lower, but still in the true ADHD range. Still, several scores between NCR groups (mainly the unincentivized group) and the control group were significantly different. The unincentivized group produced scores that were consistently lower, although not significantly so, than both incentivized groups. Overall, results show that an incentive does appear to alter NCR behavior in the hypothesized direction, but not to the extent that it produces significant differences between groups, as initially anticipated. Further, neuropsychological measures were only efficient at detecting egregious forms of NCR, but that is unsurprising, given that is not their intended purpose.

This finding is consistent with previous research showing that the Trail Making Test (A & B) may identify overt malingering (Iverson, 2002), but is not recommended as a PVT since it

is likely to miss most non-credible respondents (Powell et al., 2011). In the current study, only the incentivized-immediate group produced significantly different scores from the control group on TMT A, showing that it did, in fact, miss the majority of NCR participants. Similar guidelines have been recommended for Digit Span forward and backward (Whitney et al., 2013). Results from this study are commensurate such that only the unincentivized group produced scores that were significantly lower than the control group on Digit Span (forward and backward); thus, Digit Span also failed to detect most NCR participants. These results suggest that the neuropsychological measures used in the current study were efficient at detecting egregious instances of NCR. However, in a clinical sample, low scores on neuropsychological measures may reflect true impairments. Therefore, low scores on neuropsychological measures should be used in conjunction with failed PVTs to identify NCR and should never be used to indicate NCR on their own.

The second hypothesis of this study concerned differences in PVT performance between groups. Within the first part of this hypothesis, it was expected that the control group would produce scores that were significantly different from all three NCR groups. As opposed to neuropsychological measures that are used to detect impairments across multiple cognitive domains, PVTs are used to assess suboptimal effort; therefore, it was expected that the control group would obtain scores above the cut-off on all three PVTs, indicating normal or neurotypical performance. This hypothesis has support in the literature, as shown by a recent meta-analysis of 11 simulated-ADHD malingering studies that found evidence of statistically significant differences between normal participants and NCR participants (Wallace et al., 2019), suggesting PVTs are strong indicators of NCR in ADHD evaluations.

As predicted in the current study, the control group produced scores that fell in the normal range. The unincentivized group produced scores showing NCR on all PVTs, and all scores were statistically different from the control group in the hypothesized direction; however, a different pattern of results emerged for the incentivized groups. The incentivized-immediate group produced scores that were significantly different from the control group on the 21-Item and the TOMM in the predicted direction, but not on the RDS. The incentivized-delayed group, however, did not differ significantly from the control group on any PVT. Hence, this portion of Hypothesis 2 was partially supported, as it was expected that all NCR groups would differ significantly from the control group on all PVTs.

The second portion of Hypothesis 2 predicted that participants in the control group would produce scores that were within the normal range (i.e., above cut-offs), while participants in the NCR groups would score in the NCR range. This portion of the hypothesis was partially demonstrated, as all participants in the control group passed the TOMM, while one failed the 21-Item, and another failed RDS. All three NCR groups produced mean scores that were below the cut-off of the TOMM, but scores were mixed with respect to the 21-Item and RDS. Only the unincentivized group produced a mean score that fell with the NCR range on RDS, and no group produced a mean score within the NCR range on the 21-Item.

The final portion of Hypothesis 2 predicted that both incentivized groups would produce scores that were significantly closer to the normal range than the unincentivized group on all three PVTs. As such, the incentivized-delayed group was significantly different from the unincentivized group on the TOMM and 21-Item in the hypothesized direction. No differences were found, however, between the incentivized-immediate and unincentivized groups. Thus, this portion of Hypothesis 2 was partially supported.

Findings from Hypothesis 2 show that an added incentive does appear to improve effort on PVTs, as the incentivized-delayed group produced PVT scores that were not significantly different from the control group. However, despite statistically significant differences in mean scores, participants in all NCR groups were overwhelmingly detected by PVTs. Thus, these results are not clinically significant, suggesting that even with an incentive, most individuals engaging in NCR are still detected by PVTs.

Clearly, not all PVTs are created equal. Results with the TOMM showed that all three NCR groups produced mean scores identifying NCR and all control participants produced scores within the normal range. This was predicted given the relatively high sensitivity of the TOMM according to the author (Tombaugh, 1996) and a meta-analysis (Batt et al., 2008). As previously stated, only two of 18 control participants did not achieve a perfect score on the TOMM; however, all produced scores above the cut-off. In fact, according to the author, similar scores were produced by individuals with cognitive impairment, aphasia, and TBI, but not dementia (Tombaugh, 1996). The author of the TOMM did not include an ADHD sample, but current literature suggests that individuals with ADHD are overwhelmingly able to score above the cut-off on the TOMM (Sollman et al., 2010). Current results suggest, as does past research, that the TOMM is an effective measure at detecting NCR within ADHD evaluations. The RDS and the 21-Item, however, were less successful at detecting NCR, consistent with the lower sensitivity reported for these measures (Schroeder et al., 2012; Vickery et al., 2001) Additionally, one control participant failed the 21-Item and another separate control participant (neither of whom was considered to be an outlier according to performance on other measures) failed RDS, which is consistent with the higher false-positive rate compared to the TOMM.

Prior to calculating both sensitivity and specificity of the PVTs in this study, the researchers identified six examinees whose performance indicated that they did not engage in NCR as instructed. NCR participants who did not fail any PVTs, produced a normal profile on the WURS, and produced T-scores below 35 on two or fewer neuropsychological measures were not included in the sensitivity and specificity analysis because they were considered to have performed normally and thus, were not engaging in NCR. Since determining sensitivity requires the accurate identification of true-positives, false-positives, false-negatives, and true-negatives, these participants' data were eliminated to provide a clearly NCR sample for analyses. The rationale for allowing participants judged as performing "normally" to fail two measures was that one quarter of control participants in this study produced scores at or below the borderline impaired range on two neuropsychological measures, due to random participant characteristics. This is not an unusual finding, as Schretlen et al. (2003) found in a study of 197 normal adults that an individual's scores in a neuropsychological battery vary an average of 1.6 standard deviations across measures (e.g., T=45 on measure A and T=29 on measure B), with some scores varying as much as 6.1 standard deviations between measures. Hence, NCR participants in this study performing in a consistent manner with control participants were eliminated from sensitivity and specificity analyses to ensure a sample of clearly noncredible responders.

As such, one unanticipated finding concerned the relatively low sensitivity of the 21-Item and RDS. The authors reported sensitivities of 70% (Iverson et al., 1994) and 86.3% (Pearson, 2009), respectively, in community samples. However, one meta-analysis demonstrated significant variability in sensitivity for the 21-Item, ranging from 2.5% to 65% across six studies using college students as simulated malingerers (Vickery et al., 2001). A meta-analysis of 21 studies using the RDS, five of which utilized simulated malingerers, found sensitivity to range

between 27% and 71%, with a weighted mean sensitivity of 38% across 165 participants (Schroeder et al., 2012). The current study found a 25% sensitivity rate for detecting NCR for both measures, consistent with the lower estimates from these meta-analyses. These results suggest that these two measures are less useful than the TOMM at detecting NCR in college students feigning ADHD. However, since they are able to detect NCR in cases of unsophisticated NCR and take less than five minutes to administer, they are useful to include in evaluation, as recommended by Lezak (2012).

A literature review by Batt et al. (2008) of the TOMM's sensitivity/specificity demonstrated that its effectiveness varies, sometimes dramatically, by study and by population. Given the consequences of both false-positives (an examinee is incorrectly believed to have engaged in NCR) and false-negatives (undetected NCR), the integrity of this measure is vital in assessing suboptimal effort. Three studies discussed in this analysis that utilized simulated malingering students found that sensitivity ranged from 1.00 to 0.96; however, this range was expanded to 0.34 to 1.00 when several other samples were added (litigating and forensic). Larrabee (2012) offered one reason for this variability, stating that PVT distributions are both positively-skewed and leptokurtotic, because anything under a 90% success rate (although this varies by PVT) tends to reflect poor effort; that is, minor changes in motivation, which may result from true NCR or fatigue, may have drastic results for overall scores. Thus, since PVTs are impacted by minor changes in motivation, it is perhaps unsurprising (given the variability in participant's reported strategies) that the current study not only found sensitivity rates that differed from the author's, but that varied between groups.

Another unanticipated finding in this study concerned the CPT's ability to distinguish control participants from NCR participants more effectively than both RDS and the 21-Item.

This CPT is a computerized version of Rosvold *et al* (1956) mechanical CPT and has not been established as a clinical instrument. It presents stimuli at a constant rate for several minutes, so it is a relatively simple and boring task. Scoring information for this measure was not available, so for the purposes of this study, the sum total of commission errors and sum total of omission errors were calculated. The control group produced a low mean and standard deviation, such that around 10 errors (depending on the CPT trial) resulted in a score that fell one and a half standard deviations below the control group mean and was considered to be in the borderline impaired range. Scores on both total commission and omission errors distinguished control from NCR participants, correctly identifying all members of the control group and 33 out of 44 NCR participants, yielding 100% specificity. With respect to detecting NCR, sensitivity was 0.62 for the incentivized-delayed group, 0.80 for the incentivized-immediate group, and 0.812 for the and unincentivized group. These results suggest that this CPT may have utility as a PVT in ADHD evaluations and warrants further research in this capacity.

The third hypothesis of this study concerned the sophistication with which participants engaged in NCR. It was anticipated, based on previous research (Suhr, 2000), that incentivized participants, aware of the conditional nature of the reward, would direct more energy toward their overall performance, behaving in a more meticulous and sophisticated manner. As such, it was expected that members of the unincentivized group would produce scores that represented clear signs of NCR, while both incentivized groups would produce scores that were less obviously noncredible. Although participants within the unincentivized group failed more than one PVT at a higher frequency than participants in both incentivized groups, there were no statistical differences between these groups; thus, this hypothesis was not supported. This may be

due in part to weak statistical power given the small size of the groups (i.e., 18 in the unincentivized group and 16 in both incentivized groups).

Nevertheless, most participants were not successful in feigning ADHD. Forty-two percent failed one PVT (all of whom failed the TOMM), 24% failed two (varying equally between RDS and the 21-Item), and 8% failed all three. Of the remaining 26% (13 participants) who did not fail a PVT, six participants also produced normal profiles on the WURS and neuropsychological measures, suggesting that they did not engage in NCR (and were eliminated from only the sensitivity/specificity analyses, as discussed above). Further, out of all 50 participants in the NCR groups, only five passed the stringent reward criteria during data collection (four in the incentivized-delayed group and one in the incentivized-immediate group). Additionally, although a reward was not offered to the unincentivized group for successful NCR, it is worth noting that no participants from that group passed according to these criteria.

It is important to note that the criteria for ADHD used in this study is insufficient for a genuine clinical evaluation, in which symptom checklists or diagnostic interviews of both the patient and collateral sources are needed. Moreover, neuropsychological tests are not diagnostic in the evaluation of ADHD; in fact, a primary issue with using neuropsychological measures in the evaluation of ADHD is that there is no specific ADHD profile (Barkley & Murphy, 2010). Although individuals with ADHD often perform in the impaired range across neuropsychological domains, including verbal ability, executive functioning, simple/sustained attention, verbal/visual memory, and problem solving (Schoechlin & Engel, 2005), test results are varied such that two people with ADHD may present heterogeneous neuropsychological profiles. Thus, the diagnosis of ADHD is based upon self- and collateral reports of current and childhood behavioral functioning and impairment on measures such as the WURS, CAARS,

Barkley Adult ADHD Rating Scale, or others (Barkley et al., 2010). Neuropsychological measures help support a diagnosis and identify areas of strength and weakness, but are not diagnostic instruments. Accordingly, since some impairments are common, the criteria for successfully feigning ADHD in this study with respect to producing scores in the borderline impaired range on neuropsychological measures were not specific; in other words, it was not predetermined which measures had to fall in the borderline impaired range.

The final hypothesis concerned strategies used by participants to engage in NCR. It was expected that participants in the unincentivized group would report similar methods of NCR as the incentivized NCR groups. Additionally, since other studies have employed similar methods of collecting strategy data (Frazier et al., 2008; Harrison & Edwards, 2010; Harrison et al., 2007; Quinn, 2003), it was predicted that participants in this study would report engaging in similar strategies to past studies.

Results ultimately supported the hypothesis, in that strategies were similar across studies; however, the rates with which participants reported engaging in these strategies varied. The data were coded and organized by ADHD symptom category, per the DSM-5 (APA, 2013) criteria: criterion A1 attention deficit symptoms (distracted, inattention, avoidance, forgetful), criterion A2 hyperactive symptoms (fidgetiness, restlessness, and impulsivity), and non-ADHD symptoms (which only included slowness). As such, for the hyperactive symptoms, the unincentivized group reported more "fidgety" behavior than both incentivized groups. Participants endorsing fidgety behavior remarked that they had tapped their writing utensils or fingers on the table, played with their jewelry/clothing/hair, and/or shook their legs. The incentivized-delayed group reported more behavior related to impulsivity, which included behavior such as hitting the space bar at inappropriate times (during the CPT), interrupting the

examiner, and/or rushing through the tasks. Other reported behaviors such as shifting positions, rocking back and forth in chair, and/or cleaning glasses were categorized as restlessness (subsumed under hyperactivity).

Common attention deficit-related behaviors included staring off into the distance, looking around the room, asking the examiner to repeat questions (because participant was not listening), and thinking about "other things" during testing. Additionally, participants reporting engagement in "slow" behavior remarked that they had taken their time, worked slowly, and hesitated before answering. The three participants who were most successful in feigning ADHD reported using strategies such as pronounced fidgetiness ("cracking ankles, playing with pen") and impulsivity ("answering too fast"). Other strategies reported included being distracted, incorrectly "marking symptoms of childhood questions," (on the WURS) and "talking inappropriately during experiment," (e.g., attempting to engage the examiner in several conversations about psychology).

Although the general strategies utilized by participants in this study were similar to those endorsed in past studies, rates differed as a function of assessment method. For example, about four times as many participants admitted to acting fidgety in this study compared to Quinn (2003), even though the word "fidgety" was used in both studies. One reason for this difference may be the interpretation or categorization of participants' responses. Quinn (2003) stated that participants reported fidgety behavior, but it is not clear how responses were coded or categorized. However, if only NCR participants in this study who used the specific term "fidgety" were counted as demonstrating fidgety behavior, the percentage would be halved, but still twice as high as Quinn (2003). It is unclear why there is disparity between these data.

Harrison et al. (2007) collected open-ended data and included highly specific terms such as "beginning tasks before being told to go ahead," and "completing tasks quickly and carelessly," both of which would have been classified as "impulsivity" in this study, and both of which may have been indicated by the same participant. Thus, these rates were compared with similar specific strategies reported by participants in this study. These results demonstrated, again, that similar strategies were used, but rates differed.

One interesting finding from both this study and past studies are the rates in which participants admitted to engaging in slow behavior, which is not a documented DSM symptom of ADHD. Since this was a commonly reported strategy, it may be a behavior for which clinicians should be on the lookout; however, future studies will want to examine what this looks like, as informal discussion with examiners suggested that no participants were specifically observed behaving in an obviously slow manner. Additionally, although NCR participants across studies were reported to engage in fidgety behavior, no other studies reported the specific "fidgety" behaviors endorsed (e.g., tapping a pencil). Also, it is unknown to what extent individuals with bona fide ADHD tend to fidget during the structured setting of an assessment, or how malingered fidgetiness may differ from genuine fidgetiness.

Overall, results regarding the effects of incentivizing ADHD-related NCR were mixed but suggest that external incentives may influence ADHD-related NCR; however, these results were not as consistent or as prominent as hypothesized. While the incentivized-delayed group produced scores on PVTs that were closer to the control group, they were still detectable by PVTs as NCR. Additionally, this group produced scores on neuropsychological measures that fell within the range of a true ADHD sample, according to previous authors (Booksh et al., 2010; Quinn, 2003; Sollman et al., 2010). In other words, the incentivized-delayed group produced

scores that were closer to "successful malingering" than the incentivized-immediate or unincentivized groups, but still detectable as NCR

Strengths and Limitations

A strength of the current study was the consistent methodology with prior simulated-malingering studies. For example, Quinn (2003), Harrison et al. (2007), Frazier et al. (2008), Booksh et al. (2010), Sollman et al. (2010), and Fisher (2007) all utilized a simulated-malingering paradigm to compare the performance of participants instructed to emulate ADHD symptoms on PVTs, SVTs, self-report measures, and neuropsychological measures. Like the current study, Fisher (2007) and Sollman et al. (2010) incentivized participants for engaging in NCR, but did not include an unincentivized group; ergo, it was unknown if the added incentive altered behavior beyond what would be expected from instructions. Thus, the current study extended the literature to examine the effects of incentives on ADHD-related NCR with the addition of a comparison unincentivized group.

Another strength of the current study was the addition of two incentivized groups that were offered different rewards. This sought to extend the findings of Sollman et al. (2010), who offered participants in the feigning group $45 to produce ADHD scores and Fisher (2007), who offered participants the chance to win $25, but concluded that future studies should use a larger reward. The incentivized "immediate" and "delayed" groups were intended to examine differences between groups as a function of the time between the receipt of the reward (i.e., at the end of the testing session versus at the end of the study); however, no hypothesis were made regarding differences these rewards might produce. The incentivized-immediate group was added to this study one week into data collection after it appeared that participants in the incentivized-delayed group were responding normally, rather than attempting NCR. It was

speculated that the two-step process of winning $100, by first successfully faking ADHD and then winning a raffle, may have seemed unachievable to participants and was not providing the motivation intended to influence performance. The incentivized-immediate group served to address these potential barriers by making the reward appear achievable and immediate. However, as the results suggest, the delayed group were significantly more similar to the control group than the immediate group, which may be a result of chance, but may also be related to the reward size. According to a meta-analysis that examined the effects of external rewards on motivation, Deci et al. (2001) noted that intrinsic motivation can be increased for a low-interest task when a sufficient reward is offered. Thus, it is possible that $10 was not a sufficient amount of money to increase intrinsic motivation.

Another strength of this study was the categorization of ADHD strategies reported by participants. In this study, qualitative responses were categorized based on all ADHD symptoms, types of errors, and errors on specific measures. These categories are also discrete and data-driven, since they represent the diagnostic criteria for ADHD in the DSM-5 (APA, 2013). This study sought to extend and consolidate the work of past researchers (Frazier et al., 2008; Harrison et al., 2007; Quinn, 2003), who reported differing specific strategies. By stratifying participant responses via symptom category, enumerable specific strategies could be combined in an informative, comprehensive, and repeatable manner. Since all participant strategies reported in this study easily fit into the DSM-5 ADHD A1 and A2 criterions, this method of organizing qualitative data was useful.

Despite the strength of the research design, there were several limitations. Notably, the primary limitation was the sample size of the study. A total of 68 participants were divided between four groups (16,16,18,18). Data from eight participants could not be used. Despite

efforts to screen potential participants for ADHD, six participants reported a history of ADHD and had to be excluded; two other participants were excluded for failing to follow directions. Thus, unless overwhelming differences were present between groups, the study lacked sufficient power to detect differences. The differences in means between participants categorized as "sophisticated" or "unsophisticated" in their response style approached significance for some measures, but were not statistically different; thus, it is unknown whether these differences were meaningful and might have reached significance with a larger sample, or whether they were due to error.

Another present and unexpected issue was participant effects, including poor motivation and/or confusion, as well as discomfort with the task. With respect to discomfort, several participants informally admitted, post-experimentally, that the odd/awkward nature of the experiment, which included the instruction to essentially "lie" to a stranger, was uncomfortable and/or difficult for them. This was surprising since it had not been mentioned in the existing body research regarding simulated NCR. In terms of participant motivation, it appeared that participants were not following instructions early in the data collection. This was addressed by the addition of an affirmation of instructions to ensure that participants were aware that they would not personally be judged based on their performance after several participants stated that they had experienced varying degrees of discomfort having to act in a dishonest fashion. One participant expressed worry that the examiners would think she was "crazy" if she had engaged in NCR. In another instance, one participant admitted, after the testing session, that she did not understand instructions and had not followed instructions. Data from these participants were excluded but revealed the need to ensure that participants read the instructions. Overall, six NCR participants produced scores that were congruent to control group participants, suggesting that

they did not understand either what was being asked of them or how to do it. Data from these participants were dropped from the sensitivity/specificity analyses due to concern that normal responders in an NCR sample would bias these results. However, the majority of NCR participants failed at least one PVT, suggesting that the current data do reflect proper task-engagement.

Nevertheless, future studies may want to ensure that participants are fully informed of instructions prior to the testing session and have the opportunity to ask questions about instructions, so participants are fully prepared to engage in the task. Ideally, this would be handled by an additional research assistant who could address concerns related to discomfort and confusion about the instructions so that all participants are comfortable attempting NCR. To keep the examiner blind to participant condition, the current study did not have a mechanism for answering participant questions that may have arisen after group assignment was made.

A further limitation was potential of experimenter effects (Cook et al., 2002). It was intended that examiners be blind to participant condition to reduce potential bias; however, examiners were generally aware when participants were assigned to an NCR group based on the observable change in behavior between reviewing informed consent and administering tests. Several participants immediately began tapping a writing utensil, impulsively glancing around the room, and/or acting in a hyperactive/distractible manner in a way that they had not displayed prior. Additionally, a failed TOMM, an extremely rare occurrence in a normal population, particularly in combination with otherwise normal scores, or multiple failed PVTs and borderline impaired scores (or lower) on all neuropsychological measures, was indicative that the participant was assigned to an NCR group. Although the examiner did not know to which NCR group the participant was assigned, it was clear when participants were simulating NCR or

ADHD. Since it was suspected that there would be differences between the NCR groups and control group, it is possible that awareness of NCR group assignment may have altered experimenter behavior in the form of unintended, disapproving body language, or other behaviors such as hurrying slow participants along. Results are also limited by the involvement of the principle investigator (PI) in the data collection and analyses. The PI conducted the majority of testing sessions (36 out of 68) and also coded/categorized the qualitative data. To guard against any experimenter effects, future studies may want to consider using test administrators who are blind to hypotheses, as well as experimental conditions, to increase the precision of the study (Cook et al., 2002).

The final limitation concerned the lack of established normative data for the CPT used in this study. Given that this CPT was not intended for a clinical application and is a computerized version of Rosvold et al (1956) mechanical CPT, normative data did not exist, nor were there norms for a clinical ADHD sample for comparisons. Similar to other CPTs in clinical use, normative data for the IVA-CPT (utilized in Quinn, 2003) was based upon both a larger sample of healthy individuals and individuals with ADHD. These measures provide additional scores (beyond just commission/omission error rates) that are vital in determining the quality of a participant's performance; for instance, variability is derived from the error rate between intervals and can distinguish ADHD from normal functioning (Advokat et al., 2007). Further, these measures present target stimuli in inconsistent intervals, allowing for the derivation of an accurate depiction of reaction time. These measures would have provided useful information to this study for comparison purposes since known ADHD participants frequently produce lower scores than normal participants (Willcutt et al., 2005). Thus, future studies may want to utilize CPTs with establish ADHD normative data for comparison purposes.

Future Directions

There are several areas that future researchers may consider addressing. Future studies may want to consider participants' initial intrinsic motivation to engage in the task. As per Deci et al. (2001), participants feel less intrinsic motivation to put forth effort (such as would be required for escaping detection of NCR) when they do not have an initial interest in the task, especially when an insufficient reward is offered. The current study utilized college students in an introductory psychology class, all of whom must collect research participation credit as part of a curriculum. While they do get to select the studies in which they participate, it is possible that intrinsic motivation to follow instructions is low, since credit toward completing the assignment is not based on effort. In a real-world setting, malingerers have a high degree of intrinsic motivation, which researchers should attempt to replicate. As such, future studies may want to consider using measures to assess participant motivation. This may be accomplished by adding a questionnaire as part of the debriefing process to assess participants' motivation regarding the respective reward and the extent to which they followed directions (to win the reward).

It was previously mentioned that several participants stated that they experienced a certain level of discomfort or anxiety upon being asked to engage in blatantly dishonest behavior. As such, future studies may want to examine the potential moderating impact of participant anxiety or social desirability on the ability to evade PVT detection. The Social desirability bias describes situations whereby participants respond in a manner inconsistent with their own views to appease researchers (King & Bruner, 2009). Thus, given the imposing nature of the task in this study and further potential for participants to have difficulty assuming the role of someone engaging in NCR, a measure of social desirability (e.g., The Social Desirability

Scale) would be beneficial for gaining clarity into this potential moderator. Additionally, to examine the effects of anxiety, future studies might consider using a brief questionnaire to examine the degree to which participants were anxious, distracted and/or had difficulty engaging in the task because of discomfort. These questions may also be devoted to specific aspects of the testing setting, (e.g., "How uncomfortable did you feel making deliberate errors in front of the examiner versus the computer?"). Both may be assessed as part of the debriefing process.

Lastly, future studies may want to consider examining the utility of the CPT used in this study (Rosvold et al., 1956) as a PVT. Initially, data would have to be collected from a true ADHD sample to ensure that individuals with ADHD are able to score similarly to a normal sample. However, since this measure was 75% effective at distinguishing control participants from NCR participants, this warrants future research.

Conclusion

In this study, college students were tasked with simulating ADHD on a typical ADHD battery. Participants in two of the groups were incentivized with a monetary reward to increase the ecological validity of the simulated-malingerer paradigm. The reward's effect on participant ability to evade detection on performance validity tests and to produce a realistic ADHD profile were the primary constructs of analysis.

Overall, results were mixed. The general differences between the control group and the NCR groups on PVT performance suggests that, consistent with past studies, PVTs are effective in detecting NCR. In this study, all but five NCR participants were detected. Participants in the incentivized-delayed group produced scores that were significantly closer to the control group than both the incentivized-immediate and unincentivized groups. Further, both incentivized groups produced scores that were statistically closer to the control group than the unincentivized

group, suggesting that an incentive did alter behavior. Nonetheless, since the majority of participants failed to successfully feign ADHD, these results were not clinically significant.

Despite variability between the three NCR groups and the control group, when compared on a construct of "sophistication" of NCR, no differences between the three NCR groups were detected. This shows that while these groups varied when compared to the control group, they did not vary significantly when compared to each other. Additionally, this study found, as have past studies, that many simulated malingerers engage in fidgety behavior and also "slow" behavior, which is not a symptom of ADHD. Thus, this should be met with caution when observed during real-life ADHD evaluations.

Although this study has notable limitations, including a small sample size, the finding that an added incentive does alter behavior within an ADHD NCR paradigm is novel. Given that arranging a true NCR sample is not possible, incentivizing NCR-related behavior is a model that future studies may want to consider using to increase the ecological validity of similar simulated NCR studies. Additionally, given variability within current research regarding strategies used by participants to feign ADHD, future studies may also want to consider collecting open-ended qualitative data by ADHD criterion to add to the current literature, since symptoms of ADHD are already established by APA and are discreet. These data would also be easily repeatable in future studies.

References

Adams, R. L., Parsons, O. A., Culbertson, J. L., & Nixon, S. J. E. (1996). *Neuropsychology for clinical practice: Etiology, assessment, and treatment of common neurological disorders*. Washington D.C.: American Psychological Association. doi: 10.1037/10198-000

Advokat, C., Martino, L., Hill, B., & Gouvier, W. (2007). Continuous performance test (cpt) of college students with ADHD, psychiatric disorders, cognitive deficits, or no diagnosis. *Journal of Attention Disorders, 10*(3), 253-256. doi:10.1177/1087054706292106

Americans With Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 328 (1990).

APA. (2013). *Diagnostic and statistical manual of mental disorders: DSM 5*. Washington D.C.: American Psychiatric Association.

Barkley, R. A., Murphy, K. R., & Fischer, M. (2010). *ADHD in adults: What the science says*. New York: Guilford Press.

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

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65. doi:10.1037/0033-2909.121.1.65

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology, 111*(2), 279. doi:10.1037/0021-843x.111.2.279

Barkley, R. A., & Murphy, K. R. (2010). Impairment in occupational functioning and adult ADHD: The predictive utility of executive function (EF) ratings versus ef tests. *Archives of Clinical Neuropsychology, 25*(3), 157-173. doi:10.1093/arclin/acq014

Batt, K., Shores, E. A., & Chekaluk, E. (2008). The effect of distraction on the word memory test and test of memory malingering performance in patients with a severe brain injury. *Journal of the International Neuropsychological Society, 14*(6), 1074-1080. doi:10.1017/s135561770808137x

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF: Manual for administration, scoring and interpretation*. University of Minnesota Press. doi:10.1037/t15121-000

Bigler, E. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society, 18*(04), 632-642. doi:10.1017/s1355617712000252

Booksh, R., Pella, R. D., Singh, A. N., & Drew Gouvier, W. (2010). Ability of college students to simulate ADHD on objective measures of attention. *Journal of Attention Disorders*, *13*(4), 325-338. doi:10.1177/1087054708329927

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624. doi:10.1037/0033-295x.108.3.624

Brandt, J., Rubinsky, E., & Lassen, G. (1985). Uncovering malingered amnesia. *Annals of the New York Academy of Sciences*, *444*(1), 502-503. doi:10.1111/j.1749-6632.1985.tb37625.x

Broshek, D. K., De Marco, A. P., & Freeman, J. R. (2015). A review of post-concussion syndrome and psychological factors associated with concussion. *Brain injury, 29*(2), 228-237. doi:10.3109/02699052.2014.974674

Bruner, M. R., Kuryluk, A. D., & Whitton, S. W. (2015). Attention-deficit/hyperactivity disorder symptom levels and romantic relationship quality in college students. *Journal of American College Health, 63*(2), 98-108. doi:10.1080/07448481.2014.975717

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology, 23*(3), 201-216. doi:10.1016/j.acn.2007.08.010

Conant, L. (2014). Neuropsychological assessment of developmental disorders: Learning disabilities, attention-deficit/hyperactivity disorder, autism spectrum disorder. In Parsons & Hammeke (Eds.), *Clinical neuropsychology: A pocket handbook for assessment, third edition*. Washington D.C.: American Psychological Association. doi:10.1037/14339-023

Conner, C. (1995). *Conner's continuous performance test.* North Tonawanda (NY): Multi-Health Systems.

Conners, C. (1994). *The continuous performance test (CPT): Use as a diagnostic tool and measure of treatment outcome.* Paper presented at the Annual Meeting of the American Psychological Association, Los Angelas, CA

Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*: Houghton Mifflin Boston, MA. doi:10.1086/345281

Culbertson, W. C., & Zillmer, E. (2001). *Tower of london-drexel university (toldx)*: Multi-Health Systems. doi:10.1080/87565641.2011.646448

Davidhizar, R. (1994). The pursuit of illness for secondary gain. *The Health Care Manager, 13*(1), 10-15. doi:10.1097/00126450-200019020-00005

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research, 71*(1), 1-27. doi:10.3102/00346543071001001

Deckersbach, T., Savage, C. R., Henin, A., Mataix-Cols, D., Otto, M. W., Wilhelm, S., . . . Jenike, M. A. (2000). Reliability and validity of a scoring system for measuring organizational approach in the complex figure test. *Journal of Clinical and Experimental Neuropsychology, 22*(5), 640-648. doi:10.1076/1380-3395(200010)22:5;1-9;ft640

DeSantis, A. D., & Hane, A. C. (2010). "Adderall is definitely not a drug": Justifications for the illegal use of ADHD stimulants. *Substance Use & Misuse, 45*(1-2), 31-46. doi:10.3109/10826080902858334

Diller, L. (2010). ADHD in the college student: Is anyone else worried? *Journal of Attention Disorders, 14*(1), 3-6. doi:10.1177/1087054710361585

Dobson-Patterson, R., O'Gorman, J. G., Chan, R. C., & Shum, D. H. (2016). ADHD subtypes and neuropsychological performance in an adult sample. *Research in Developmental Disabilities, 55*, 55-63. doi:10.1016/j.ridd.2016.03.013

DuPaul, G. J., Weyandt, L. L., O'Dell, S. M., & Varejao, M. (2009). College students with ADHD: Current status and future directions. *Journal of Attention Disorders*, *13*(3), 234-250. doi:10.1177/1087054709340650

Erdodi, L. A., & Roth, R. M. (2016). Feigned adult ADHD. *National Academy of Neuropsychology Bulletin*, *30*(1), 16-17. doi:10.1007/s12207-017-9286-6

Fisher, A. G. (2007). ADHD rating scales' susceptibility to faking in a college student sample.

Dissertation Abstracts International, 68, 620.

Frazier, T. W., Frazier, A. R., Busch, R. M., Kerwood, M. A., & Demaree, H. A. (2008). Detection of simulated ADHD and reading disorder using symptom validity measures. *Archives of Clinical Neuropsychology, 23*(5), 501-509. doi:10.1016/j.acn.2008.04.001

Gallagher, R., & Blader, J. (2001). The diagnosis and neuropsychological assessment of adult attention deficit/hyperactivity disorder. *Annals of the New York Academy of Sciences, 931*(1), 148-171. doi:10.1111/j.1749-6632.2001.tb05778.x

Gervais, R. O., Rohling, M. L., Green, P., & Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. *Archives of Clinical Neuropsychology, 19*(4), 475-487. doi:10.1016/s0887-6177(03)00111-2

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

Gray, S., Woltering, S., Mawjee, K., & Tannock, R. (2014). The adult ADHD self-report scale (ASRS): Utility in college students with attention-deficit/hyperactivity disorder. *PeerJ, 2*, (324). doi:10.7717/peerj.324

Gray, S. A., Fettes, P., Woltering, S., Mawjee, K., & Tannock, R. (2016). Symptom manifestation and impairments in college students with ADHD. *Journal of Learning Disabilities, 49*(6), 616-630. doi:10.1177/0022219415576523

Green, P. (2004). *Medical Symptom Validity Test (MSVT) for microsoft windows: User's manual*. Paul Green Publication Incorperated.

Green, P. (2005). *Green's Word Memory Test for Microsoft Windows: User's manual*. Green's Publications Incorporated.

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

Green, P., Lees-Haley, P. R., & Allen, L. M. (2003). The word memory test and the validity of neuropsychological test scores. *Journal of Forensic Neuropsychology, 2*(3-4), 97-124. doi:10.1300/J151v02n03_05

Green, R. (2000). *The MMPI-2 an interpretive manual* (Vol. 2). Massachusetts: Allyn & Bacon.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 6*(3), 218. doi:10.1037/1040-3590.6.3.218

Hallowell, E. M. (1995). Psychotherapy of adult attention deficit disorder. A comprehensive guide to attention deficit disorder in adults: *Research, Diagnosis, and Treatment*, 146-167.

Harrison, A. G., & Edwards, M. J. (2010). Symptom exaggeration in post-secondary students: Preliminary base rates in a Canadian sample. *Applied Neuropsychology*, *17*(2), 135-143. doi:10.1080/09084281003715642

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*(5), 577-588. doi:10.1016/j.acn.2007.03.008

Hathaway, S. R., McKinley, J. C., & MMPI Restandardization Committee. (1989). *MMPI-2: Minnesota Multiphasic Personality Inventory-2: manual for administration and scoring*. University of Minnesota Press.

Heaton, R. K., Miller, S. W., Taylor, M. J., & Grant, I. (2004). Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological

norms for African American and Caucasian adults. *Lutz, FL: Psychological Assessment Resources*.

Hervey, A. S., Epstein, J. N., & Curry, J. F. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology, 18*(3), 485. doi:10.1037/0894-4105.18.3.485

Hines, J. L., King, T. S., & Curry, W. J. (2012). The adult ADHD self-report scale for screening for adult attention deficit–hyperactivity disorder (ADHD). *Journal of the American Board of Family Medicine*, *25*(6), 847-853. doi:10.3122/jabfm.2012.06.120065

Iverson, G. L., Franzen, M. D., & McCracken, L. M. (1991). Evaluation of an objective assessment technique for the detection of malingered memory deficits. *Law and Human Behavior, 15*(6), 667. doi:10.1007/bf01065859

Iverson, G. L., Franzen, M. D., & McCracken, L. M. (1994). Application of a forced-choice memory procedure designed to detect experimental malingering. *Archives of Clinical Neuropsychology, 9*(5), 437-450. doi:10.1093/arclin/9.5.437

Iverson, G. L., & Franzen, M. D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology*, *18*(1), 38-51. doi:10.1080/01688639608408260

Iverson, G. L., Lange, R. T., Green, P., & Franzen, M. D. (2002). Detecting exaggeration and malingering with the trail making test. *The Clinical Neuropsychologist, 16*(3), 398-406. doi:10.1076/clin.16.3.398.13861

Jachimowicz, G., & Geiselman, R. E. (2004). Comparison of ease of falsification of attention deficit hyperactivity disorder diagnosis using standard behavioral rating scales. *Cognitive Science Online, 2*(1), 6-20. doi:10.1037/a0018857

Jasinski, L. J., Berry, D. T., Shandera, A. L., & Clark, J. A. (2011). Use of the Wechsler Adult Intelligence Scale Digit Span subtest for malingering detection: A meta-analytic review. *Journal of Clinical and Experimental Neuropsychology*, *33*(3), 300-314. doi:10.1080/13803395.2010.516743

Jensen, A. R., & Rohwer, W. D. (1966). The stroop color-word test: A review. *Acta Psychologica, 25*, 36-93. doi:10.1016/0001-6918(66)90004-7

Julien, R. M. (2013). *A primer of drug action: A concise nontechnical guide to the actions, uses, and side effects of psychoactive drugs, revised and updated*. Holt Paperbacks. doi:10.1093/ajhp/59.11.1130a

Kiernan, J., Reid, C., & Zavos, P. (2016). Pulling an all-nighter: Current trends of college students' use of adderall. *MOJ Womens Health, 3*(1), 167-170. doi:10.15406/mojwh.2016.03.00057

King, M. F., & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology & Marketing*, *17*(2), 79-103. doi:10.1002/(SICI)1520-6793(200002)

Knight, J., & Meyers, J. (1995). *Comparison of malingered and brain-injured productions on the rey–osterrieth complex figure test.* Paper presented at the Poster session presented at the 23rd annual meeting of the International Neuropsychological Society, Seattle, WA. doi:10.1017/s1355617700001843

Kooij, S. J., Bejerot, S., Blackwell, A., Caci, H., Casas-Brugué, M., Carpentier, P. J., . . . Fitzgerald, M. (2010). European consensus statement on diagnosis and treatment of adult ADHD: The european network adult ADHD. *BMC Psychiatry, 10*(1), 67. doi:10.1186/1471-244x-10-67

Landsberger, H. A. (1958). *Hawthorne revisited, cornell studies in industrial and labor relations*. Vol. IX: Ithaca, NY: Cornell University Press.

Larrabee, G. J. (2012). Dialogue performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society, 18*, 1-7. doi:10.1017/S1355617712000240

Larrabee, G. J. (2014). Assessment of performance and symptom validity and the diagnosis of malingering. In Parsons, M.W., & Hammeke, T.A. (Eds.), *Clinical neuropsychology: A pocket handbook for assessment, third edition* (3 ed., pp. 734). Washinton D.C.: American Psychological Association. doi:10.1037/14339-006

Lezak, M. D. (2012). *Neuropsychological assessment* (5th ed.). Oxford ; New York: Oxford University Press.

Lipman, F. D. (1961). Malingering in personal injury cases. *Temple Law Review*, *35*, 141.

Locke, D., Smigielski, J., Powell, M., & Stevens, S. (2008). Effort issues in post-acute outpatient acquired brain injury rehabilitation seekers. *Neurorehabilitation, 23*(3), 273-281.

Low, K. G., & Gendaszek, A. (2002). Illicit use of psychostimulants among college students: A preliminary study. *Psychology, Health & Medicine, 7*(3), 283-287. doi:10.1080/13548500220139386

Lu, P. H., Boone, K. B., Cozolino, L., & Mitchell, C. (2003). Effectiveness of the rey-osterrieth complex figure test and the meyers and meyers recognition trial in the detection of suspect effort. *The Clinical Neuropsychologist, 17*(3), 426-440. doi:10.1076/clin.17.3.426.18083

Marshall, P. S., Hoelzle, J. B., Heyerdahl, D., & Nelson, N. W. (2016). The impact of failing to identify suspect effort in patients undergoing adult attention-deficit/hyperactivity disorder (ADHD) assessment. *Psychological Assessment*, *28*(10), 1290. doi:10.1037/pas0000247

Miller, H. A. (2001). *Miller forensic assessment of symptoms test: M-fast; professional manual*: Psychological Assessment Resources. doi:10.1037/e307362004-005

Mittenberg, W., Azrin, R., Millsaps, C., & Heilbronner, R. (1993). Identification of malingered head injury on the wechsler memory scale—revised. *Psychological Assessment, 5*(1), 34. doi:10.1037/1040-3590.5.1.34

Montejano, L., Sasané, R., Hodgkins, P., Russo, L., & Huse, D. (2011). Adult ADHD: Prevalence of diagnosis in a us population with employer health insurance. *Current Medical Research and Opinion, 27*(2), 5-11. doi:10.1185/03007995.2011.603302

Morgan, J. E., & Boone, K. B. (2008). Assessment of feigned cognitive impairment: A neuropsychological perspective. *Journal of the International Neuropsychological Society, 14*(4), 669. doi:10.1017/S1355617708080892

Musso, M. W., & Gouvier, W. D. (2014). "Why is this so hard?" A review of detection of malingered ADHD in college students. *Journal of Attention Disorders*, *18*(3), 186-201. doi:10.1177/1087054712441970

Nigg, J. T. (2005). Neuropsychologic theory and findings in attention-deficit/hyperactivity disorder: The state of the field and salient challenges for the coming decade. *Biological Psychiatry, 57*(11), 1424-1435. doi:10.1016/j.biopsych.2004.11.011

Nigg, J. T., Stavro, G., Ettenhofer, M., Hambrick, D. Z., Miller, T., & Henderson, J. M. (2005). Executive functions and ADHD in adults: Evidence for selective effects on ADHD

symptom domains. *Journal of Abnormal Psychology, 114*(4), 706. doi:10.1037/0021-843x.114.3.706

Norwalk, K., Norvilitis, J. M., & MacLean, M. G. (2009). ADHD symptomatology and its relationship to factors associated with college adjustment. *Journal of Attention Disorders, 13*(3), 251-258. doi:10.1177/1087054708320441

Novak, S. P., Kroutil, L. A., Williams, R. L., & Van Brunt, D. L. (2007). The nonmedical use of prescription ADHD medications: Results from a national internet panel. *Substance Abuse Treatment, Prevention, and Policy, 2*(1), 1. doi:10.1186/1747-597x-2-32

O'Bryant, S. E., Engel, L. R., Kleiner, J. S., Vasterling, J. J., & Black, F. W. (2007). Test of memory malingering (tomm) trial 1 as a screening measure for insufficient effort. *The Clinical Neuropsychologist, 21*(3), 511-521. doi:10.1080/13854040600611368

Partington, J., & Leiter, R. (1949). *Halstead-reitan trail making test*. Arizona: Reitan Neuropsychology Laboratory. doi:10.1007/springerreference_183009

Pearson, N. C. S. (2009). *Advanced clinical solutions for WAIS-IV and WMS-IV: Administration and scoring manual.* San Antonio: The Psychological Corporation. doi:10.1016/b978-0-12-386934-0.00004-3

Pettersson, R., Söderström, S., & Nilsson, K. W. (2018). Diagnosing ADHD in adults: an examination of the discriminative validity of neuropsychological tests and diagnostic assessment instruments. *Journal of Attention Disorders*, *22*(11), 1019-1031.. doi:10.1177/1087054715618788

Powell, M. R., Locke, D. E., Smigielski, J. S., & McCrea, M. (2011). Estimating the diagnostic value of the trail making test for suboptimal effort in acquired brain injury rehabilitation

patients. *The Clinical Neuropsychologist, 25*(1), 108-118. doi:10.1080/13854046.2010.532912

Quinn, C. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology, 18*(4), 379-395. doi:10.1016/s0887-6177(02)00150-6

Rabiner, D. L. (2013). Stimulant prescription cautions: Addressing misuse, diversion and malingering. *Current Psychiatry Reports, 15*(7), 1-8. doi:10.1007/s11920-013-0375-2

Rey, A. (1941). The psychological examination in cases of traumatic encepholopathy. Problems. *Archives de Psychologie, 28,* 215-285.

Riccio, C. A., Wolfe, M. E., Romine, C., Davis, B., & Sullivan, J. R. (2004). The tower of london and neuropsychological assessment of ADHD in adults. *Archives of Clinical Neuropsychology, 19*(5), 661-671. doi:10.1016/j.acn.2003.09.001

Rogers, & Granacher, J. (2011). *Conceptualization and assessment of malingering*. New Jersey: John Wiley & Sons.

Rogers, R., Harrell, E. H., & Liff, C. D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review, 13*(3), 255-274. doi:10.1016/0272-7358(93)90023-f

Rosenthal, E. N., Riccio, C. A., Gsanger, K. M., & Jarratt, K. P. (2006). Digit span components as predictors of attention problems and executive functioning in children. *Archives of Clinical Neuropsychology, 21*(2), 131-139. doi:10.1016/j.acn.2005.08.004

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome Jr, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20*(5), 343. doi:10.1037/h0043220

Sandford, J., & Turner, A. (1995). *Manual for the integrated visual and auditory continuous performance test.* Richmond, VA: Brain Train. doi:10.1037/t05189-000

Schoechlin, C., & Engel, R. R. (2005). Neuropsychological performance in adult attention-deficit hyperactivity disorder: Meta-analysis of empirical data. *Archives of Clinical Neuropsychology, 20*(6), 727-744. doi:10.1016/j.acn.2005.04.005

Schretlen, D. J., Munro, C. A., Anthony, J. C., & Pearlson, G. D. (2003). Examining the range of normal intraindividual variability in neuropsychological test performance. *Journal of the International Neuropsychological Society, 9*(6), 864-870. doi:10.1017/s1355617703960061

Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable digit span: A systematic review and cross-validation study. *Assessment, 19*(1), 21-30. doi:10.1177/1073191111428764

Schwartz, K., & Verhaeghen, P. (2008). ADHD and stroop interference from age 9 to age 41 years: A meta-analysis of developmental effects. *Psychological Medicine, 38*(11), 1607-1616. doi:10.1017/s003329170700267x

Schweitzer, J. B., Cummins, T. K., & Kant, C. A. (2001). Attention-deficit/hyperactivity disorder. *Medical Clinics of North America, 85*(3), 757-777. doi:10.1016/s0025-7125(05)70339-4

Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G., . . . DeFries, J. C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology, 34*(5), 584. doi:10.1007/s10802-006-9037-8

Slick, D. J., Sherman, E. M., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13*(4), 545-561. doi:10.1076/1385-4046(199911)13:04;1-y;ft545

Smith, T. W., Snyder, C., & Perkins, S. C. (1983). The self-serving function of hypochondriacal complaints: Physical symptoms as self-handicapping strategies. *Journal of Personality and Social Psychology, 44*(4), 787. doi:10.1037/0022-3514.44.4.787

Sollman, M. J., Ranseen, J. D., & Berry, D. T. (2010). Detection of feigned ADHD in college students. *Psychological Assessment*, *22*(2), 325. doi:10.1037/a0018857

Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. American Chemical Society. doi:10.1080/09084280701280502

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of african americans. *Journal of Personality and Social Psychology, 69*(5), 797. doi:10.1037/0022-3514.69.5.797

Stroop, J. (1935). Stroop color word test. *Journal of Experimental Physiology, 18*, 643-662. doi:10.1037/h0054651

Suhr, & Gunstad, J. (2005). Further exploration of the effect of “diagnosis threat” on cognitive performance in individuals with mild head injury. *Journal of the International Neuropsychological Society, 11*(01), 23-29. doi:10.1017/s1355617705050010

Suhr, & Wei. (2013). Symptoms as an excuse: Attention deficit/hyperactivity disorder symptom reporting as an excuse for cognitive test performance in the context of evaluative threat. J*ournal of Social and Clinical Psychology, 32*(7), 753. doi:10.1521/jscp.2013.32.7.753

Suhr, J. (2015). *Psychological assessment: A problem-solving approach*. New York, NY: Guilford Publications.

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*(5), 521-530. doi:10.1016/j.acn.2008.05.003

Suhr, J. A., & Gunstad, J. (2000). The effects of coaching on the sensitivity and specificity of malingering measures. *Archives of Clinical Neuropsychology, 15*(5), 415-424. doi:10.1093/arclin/15.5.415

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. doi:10.1080/09084280701509083

Sullivan, J. R., Riccio, C. A., & Castillo, C. L. (2009). Concurrent validity of the tower tasks as measures of executive function in adults: A meta-analysis. *Applied Neuropsychology, 16*(1), 62-75. doi:10.1080/09084280802644243

Teichner, G. (2004). The test of memory malingering (TOMM): Normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology, 19*(3), 455-464. doi:10.1016/s0887-6177(03)00078-7

Teichner, G., & Wagner, M. T. (2004). The test of memory malingering (TOMM): Normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology, 19*(3), 455-464. doi:10.1016/s0887-6177(03)00078-7

Tombaugh, T. N. (1996). *Test of memory malingering*. North Tonawanda, NY.

Tombaugh, T. N. (1997). The test of memory malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*(3), 260. doi:10.1037/1040-3590.9.3.260

Tombaugh, T. N. (2004). Trail making test a and b: Normative data stratified by age and education. *Archives of Clinical Neuropsychology, 19*(2), 203-214. doi:10.1016/s0887-6177(03)00039-8

Vickery, C. D., Berry, D. T., Inman, T. H., Harris, M. J., & Orey, S. A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology, 16*(1), 45-73. doi:10.1016/s0887-6177(99)00058-x

Wallace, E. R., Garcia-Willingham, N. E., Walls, B. D., Bosch, C. M., Balthrop, K. C., & Berry, D. T. (2019). A meta-analysis of malingering detection measures for attention-deficit/hyperactivity disorder. *Psychological Assessment, 31*(2), 265. doi:10.1037/pas0000659

Ward, M. F. (1993). The Wender Utah Rating Scale: an aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *150*, 885-885. doi:10.1176/ajp.150.6.885

Wechsler, D. (2008). *Wechsler adult intelligence scale–fourth edition (WAIS-IV).* San Antonio, TX: Pearson. doi:10.1037/t15169-000

Wechsler, D. (2009). *Wechsler memory scale–fourth edition (WMS-IV):* San Antonio, TX: Pearson. doi:10.1037/t15175-000

Weinborn, M., Orr, T., Woods, S. P., Conover, E., & Feix, J. (2003). A validation of the Test of Memory Malingering in a forensic psychiatric setting. *Journal of Clinical and Experimental Neuropsychology*, *25*(7), 979-990. doi:10.1076/jcen.25.7.979.16481

Weyandt, L. L., Oster, D. R., Gudmundsdottir, B. G., DuPaul, G. J., & Anastopoulos, A. D. (2017). Neuropsychological functioning in college students with and without ADHD. *Neuropsychology, 31*(2), 160. doi:10.1037/neu0000326

Whitney, K. A., Shepard, P. H., & Davis, J. J. (2013). WAIS-IV Digit Span variables: Are they valuable for use in predicting TOMM and MSVT failure?. *Applied Neuropsychology: Adult*, *20*(2), 83-94. doi:10.1080/09084282.2012.670167

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry, 57*(11), 1336-1346. doi:10.1016/j.biopsych.2005.02.006

Woods, S. P., Lovejoy, D. W., & Ball, J. (2002). Neuropsychological characteristics of adults with ADHD: A comprehensive review of initial studies. *The Clinical Neuropsychologist, 16*(1), 12-34. doi:10.1076/clin.16.1.12.8336

Appendix A

Prescreening Measure: ASRS-V1.1 Symptom Checklist (Gray et al., 2014)

Adult ADHD Self-Report Scale (ASRS-v1.1) Symptom Checklist

Patient Name		Today's Date	

Please answer the questions below, rating yourself on each of the criteria shown using the scale on the right side of the page. As you answer each question, place an X in the box that best describes how you have felt and conducted yourself over the past 6 months. Please give this completed checklist to your healthcare professional to discuss during today's appointment.	Never	Rarely	Sometimes	Often	Very Often
1. How often do you have trouble wrapping up the final details of a project, once the challenging parts have been done?					
2. How often do you have difficulty getting things in order when you have to do a task that requires organization?					
3. How often do you have problems remembering appointments or obligations?					
4. When you have a task that requires a lot of thought, how often do you avoid or delay getting started?					
5. How often do you fidget or squirm with your hands or feet when you have to sit down for a long time?					
6. How often do you feel overly active and compelled to do things, like you were driven by a motor?					

Part A

Appendix B

Demographic Questionnaire: Online Survey

1. **Participant # (filled in by examiner)**

2. **What is your age?**
18
19
20
21
22

3. **What is your sex?**
Female
Male
Other:

4. **What is your ethnicity?**
White or Caucasian
Black or African American
Hispanic
Asian Origin
Middle Eastern or Arabic
Native American
Native Hawaiian or Pacific Islander
Other:

5. **What is your native language?**

6. **What is your handedness?**
Right
Left
Ambidextrous

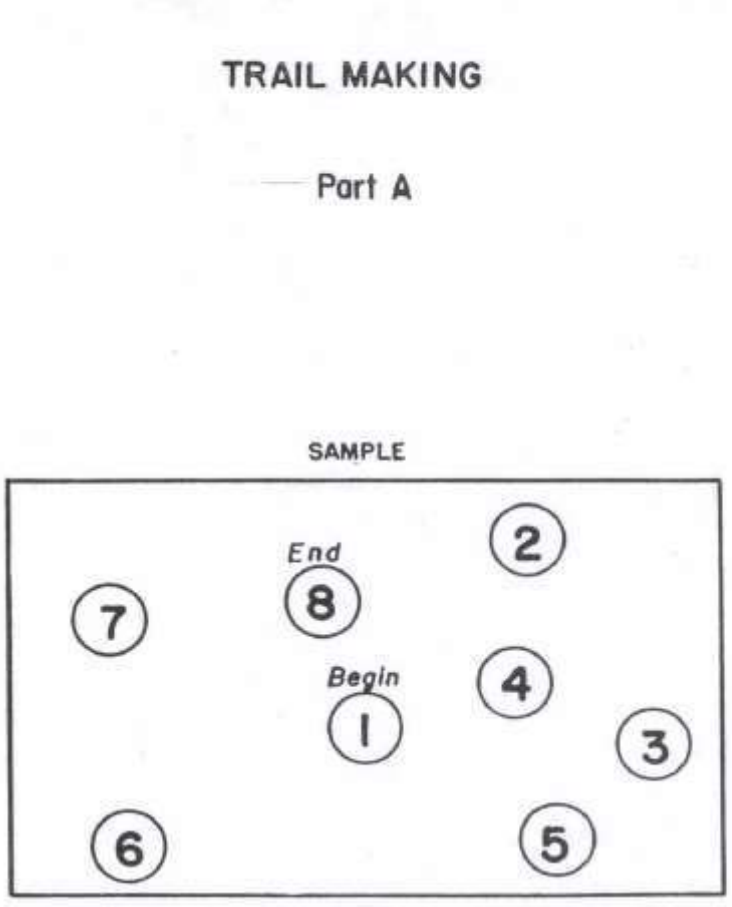
7. **What is your academic class?**
Freshman
Sophomore
Junior
Senior

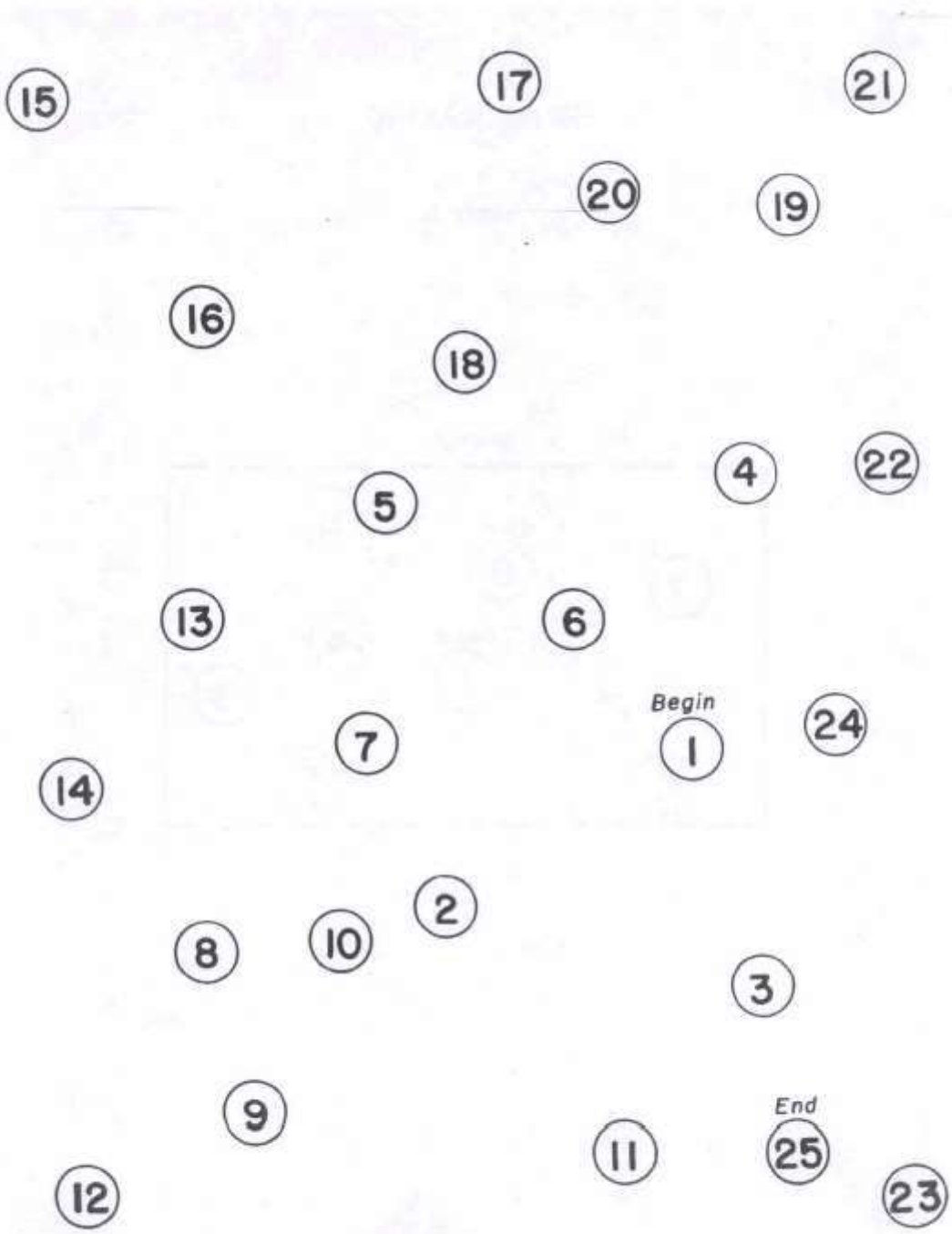
8. **Have you ever been formally diagnosed with any of the following? (SELECT ALL THAT APPLY)**.
Attention Deficit Hyperactivity Disorder
Learning Disorder
Traumatic Brain Injury (In the last 3 months)
Epilepsy or Seizures

9. **Do you receive academic accommodations through the Office of Disability? (Not including financial aid)**
Yes
No

Appendix C

Trail Making Test A and B (Partington & Leiter, 1949)

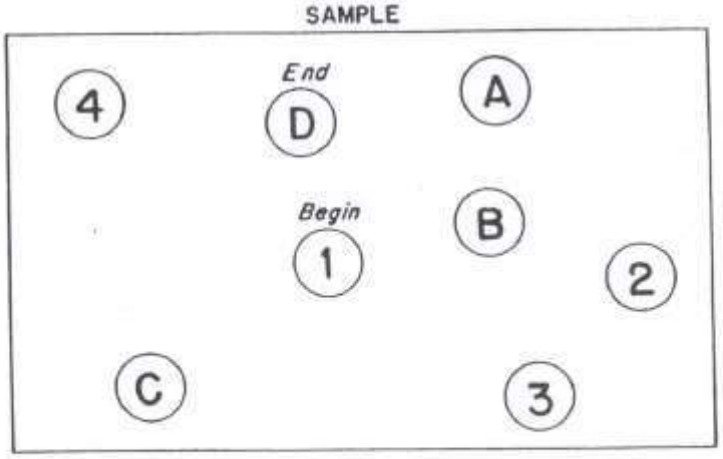


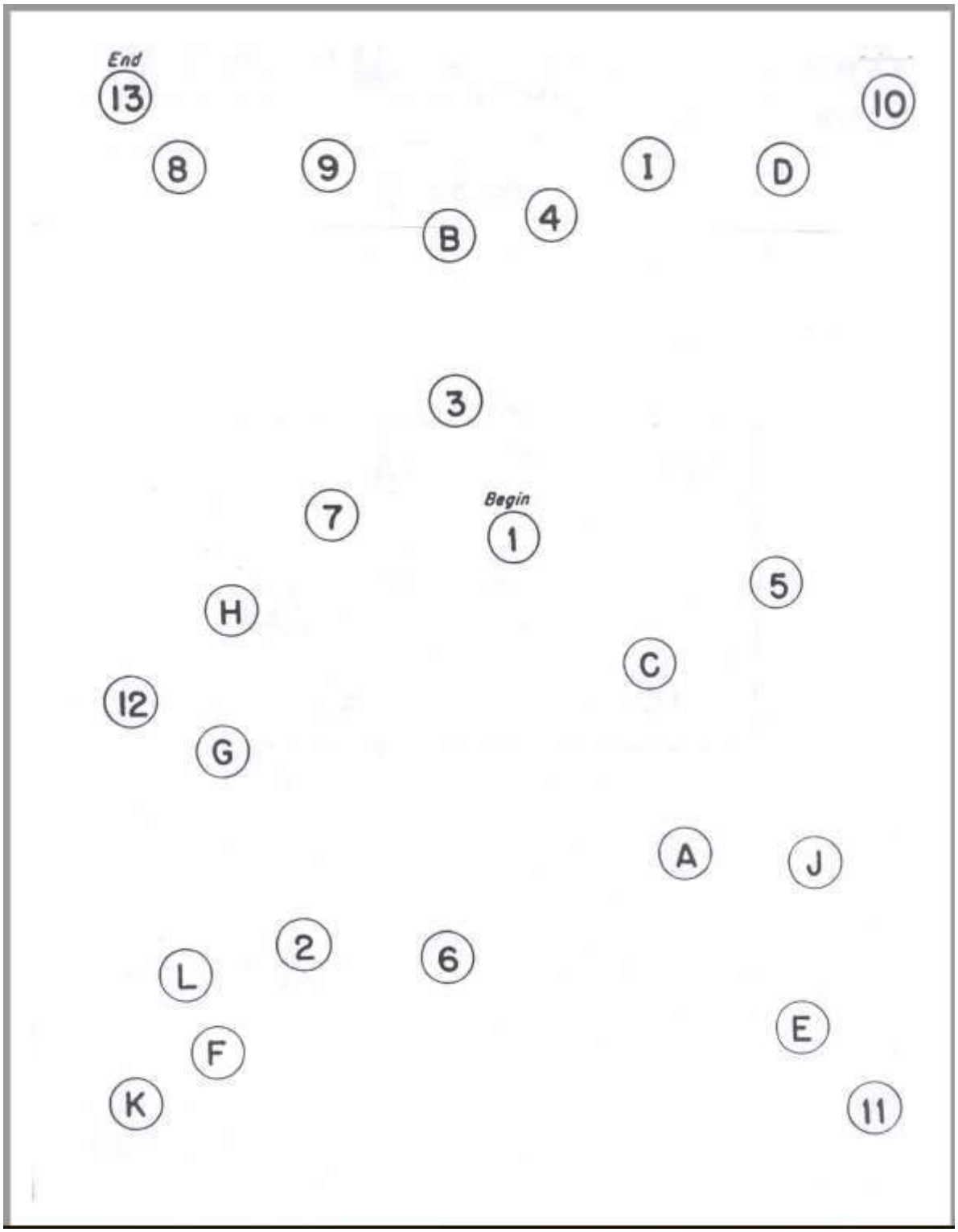

15
17
21
20
19
16
18
4
22
5
13
6
Begin
24
7
1
14
2
8
10
3
9
End
11
25
12
23

TRAIL MAKING

Part B

SAMPLE




End
13
10
8
9
1
D
4
B
3
7
Begin
1
5
H
C
12
G
A
J
2
6
L
E
F
K
11

Appendix D

Continuous Performance Test (Rosvold et al., 1956)

CONTINUOUS PERFORMANCE TEST (CPT)

SCRIPT INFO

last updated: 01-04-2018 by K. Borchert (katjab@millisecond.com) for Millisecond Software, LLC

BACKGROUND INFO

Purpose

This script implements the Continuous Performance Test as described

in:

Rosvold, H.E., Mirsky, A., Sarason, M., Bransome, E.D., Jr., and Beck, L.H. A Continuous Performance Test of brain damage. Journal of Consulting Psychology, 20, 343 (1956).

Task

Participants get presented a sequence of letters (one-by-one).

Task1: press the Spacebar whenever the letter is an X

Task2: press the Spacebar whenever the letter is an X that follows an A..

Appendix E

Wender Utah Rating Scale (Ward, 1993)

AS A CHILD I WAS (OR HAD):	Not at all or very slightly	Mildly	Moder-ately	Quite a bit	Very Much
1. Active, restless, always on the go					
2. Afraid of things					
3. Concentration problems, easily distracted					
4. Anxious, worrying					
5. Nervous, fidgety					
6. Inattentive, daydreaming					
7. Hot or short tempered, low boiling point					
8. Shy, sensitive					
9. Temper outbursts, tantrums					
10. Trouble stick-to-it-ivenessing, not following through, failing to finish things started					
11. Stubborn, strong willed					
12. Sad or blue, depressed, unhappy					
13. Uncautious, dare-devilish, involved in pranks					
14. Not getting a kick out of things, dissatisfied with life					
15. Disobedient with parents, rebellious, sassy					
16. Low opinion of myself					
17. Irritable					
18. Outgoing, friendly, enjoy company of people					
19. Sloppy, disorganized					
20. Moody, have ups and downs					
21. Feel angry					
22. Have friends, popular					
23. Well organized, tidy, neat					
24. Acting without thinking, impulsive					
25. Tend to be immature					
26. Feel guilty, regretful					
27. Lose control of myself					
28. Tend to be or act irrational					
29. Unpopular with other children, didn't keep friends for long, didn't get along with other children					
30. Poorly coordinated, did not participate in sports					
31. Afraid of losing control of self					
32. Well coordinated, picked first in games					
33. (for women only) Tomboyish					
34. Ran away from home					
35. Get in fights					
36. Teased other children					
37. Leader, bossy					
38. Difficulty getting awake					
39. Follower, lead around too much					

40. Trouble seeing things from someone else's point of view					
41. Trouble with authorities, trouble with school, visits to principal's office					
42. Trouble with the police, booked, convicted					
MEDICAL PROBLEMS AS A CHILD:					
43. Headaches					
44. Stomachaches					
45. Constipation					
46. Diarrhea					
47. Food allergies					
48. Other allergies					
49. Bedwetting					
AS A CHILD IN SCHOOL:					
50. Overall a good student, fast					
51. Overall a poor student, slow learner					
52. Slow reader					
53. Slow in *learning* to read					
54. Trouble reversing letters					
55. Problems with spelling					
56. Trouble with mathematics or numbers					
57. Bad handwriting					
58. Though I could read pretty well, I never really enjoyed reading					
59. Did not achieve up to potential					
60. Repeated grades (which grades?) _____					
61. Suspended or expelled (which grades?) _____					

Appendix F

Informed Consent Form

Informed Consent Form

You are invited to participate in this research study. The following information is provided in order to help you to make an informed decision whether or not to participate. If you have any questions, please do not hesitate to ask. You are eligible to participate because you are a student in the General Psychology course at Indiana University of Pennsylvania and you met criteria for inclusion in the study. The Department of Psychology at Indiana University of Pennsylvania supports the practice of protection of human participants in research. If you agree to participate, please be aware that you are free to withdraw at any point throughout the duration of the experiment without any penalty.

In this study we will ask you to complete several paper-pencil tasks, two questionnaires, and a computerized task. If for any reason during this study you do not feel comfortable, you may leave the laboratory and receive credit for the time you participated, and your information will be discarded.

The risks of participation in this study are minimal. It is possible that some participants may feel uncomfortable answering some of the interview or survey questions. If this occurs, you are free to skip any questions or end your participation in the study at any time. Participants who feel distress as a result of study participation may discuss this with the research assistants or principal investigator, a licensed clinical psychologist. Your participation in this study will require approximately 90-120 minutes for which you will be awarded two (2) research participation credits for PSYC 101. Participation or nonparticipation will not affect the evaluation of your performance in PSYC 101 and your instructor will not have access to your questionnaire responses or test data. Alternatively, if you do not wish to participate in research, you may complete "Read and Review" assignments, when available, through the SONA research participation website. When this study is complete you will be provided with the results of the experiment as a whole, if you request them, and you will be free to ask any questions.

If you choose to participate, all information will be confidential and will have no bearing on your academic standing or services you may receive from the University. Your decision will not affect your grade in PSYC 101. All information collected from you will be identified by number, not by name, and will not be used to identify particular individuals in need of psychological treatment. The information you provide us will be considered only in combination with that of other participants. The information obtained in this study may be published in scientific journals or presented at scientific meetings, but your identity will be kept confidential.

If you are willing to participate in this study, please sign below. When you complete your participation in this study, you will be given an information sheet that will provide you with information about counseling and other services that are available to you as a student at IUP.

If you have any further questions concerning this study, please feel free to contact us through phone or email: Jesse Siegel at j.a.siegel@iup.edu or Dr. Laura Knight at laura.knight@iup.edu. Please indicate with your signature on the space below that you understand your rights and agree to participate in the experiment.

Your participation is solicited, yet strictly voluntary. All information will be kept confidential and your name will not be associated with any research findings.

This project has been approved by the Indiana University of Pennsylvania Institutional Review Board for the Protection of Human Subjects (Phone: 724/357-7730).

I have read and understand the information on the form and I consent to participate in this study. I understand that my responses are completely confidential and that I have the right to withdraw at any time. I have received an unsigned copy of this informed consent form to keep in my possession.

Name:

Signature:

Date:

I certify that I have explained to the above individual the nature and purpose, the potential benefits, and possible risks associated with participating in this research study, have answered any questions that have been raised, and have witnessed the above signature.

Date: Investigator's signature:

Appendix G

Prompt for Control Group

INSTRUCTIONS: Please try your best on all tasks.

Appendix H

Prompt for Unincentivized Group

During this experiment, you will be given several tests often used in the evaluation of Attention Deficit Hyperactivity Disorder (ADHD). Please read the following vignette for specific instructions:

You are a student at Indiana University of Pennsylvania. Last month, on a Saturday, you went to the library with your roommate to study for an exam. Upon arriving at an empty table, your roommate offered you a pill to help you study and told you that it was Adderall. You took this pill and found that suddenly you had the ability to concentrate without once getting distracted for several hours. You were so impressed with the results, the next day you asked your roommate for another pill to help you with a project. Again, you were able to sit and concentrate for several hours and felt very productive.

The following Wednesday, you asked your roommate where to get this medication. Your roommate informed you that they had been diagnosed with ADHD. In addition to a prescription for Adderall, they were also allowed extra time on exams, and could take these exams in a separate room, away from the distraction of other students. You then asked your roommate how to go about getting diagnosed with ADHD. They told you that first they went to a psychiatrist, and then to a neuropsychologist. They had to take a lot of weird tests, and afterwards, they were diagnosed with ADHD.

You then decided that you too wanted a diagnosis of ADHD. You booked an appointment with a psychiatrist last week. During the appointment, they told you that it doesn't sound like you have ADHD but referred you for testing anyway. Eager for this diagnosis, you have decided to pretend that you have ADHD during this evaluation. Prior to the appointment, you looked at the criteria for ADHD. You understand that what you will be doing is dishonest, but you see this as means to an end.

INSTRUCTIONS: Your goal is to successfully convince the examiner that you have ADHD. **You will not be evaluated based on your acting ability**, but more-so on your ability to fake ADHD on the various tasks you will take today. This will be reflected in the scores that you produce.

You will now be given 10 minutes to review the diagnostic criteria for ADHD, after which the experiment will start.

Appendix I

Prompt for Incentivized-Immediate Group

During this experiment, you will be given several tests often used in the evaluation of Attention Deficit Hyperactivity Disorder (ADHD). Please read the following vignette for specific instructions:

You are a student at Indiana University of Pennsylvania. Last month, on a Saturday, you went to the library with your roommate to study for an exam. Upon arriving at an empty table, your roommate offered you a pill to help you study and told you that it was Adderall. You took this pill and found that suddenly you had the ability to concentrate without once getting distracted for several hours. You were so impressed with the results, the next day you asked your roommate for another pill to help you with a project. Again, you were able to sit and concentrate for several hours and felt very productive.

The following Wednesday, you asked your roommate where to get this medication. Your roommate informed you that they had been diagnosed with ADHD. In addition to a prescription for Adderall, they were also allowed extra time on exams, and could take these exams in a separate room, away from the distraction of other students. You then asked your roommate how to go about getting diagnosed with ADHD. They told you that first they went to a psychiatrist, and then to a neuropsychologist. They had to take a lot of weird tests, and afterwards, they were diagnosed with ADHD.

You then decided that you too wanted a diagnosis of ADHD. You booked an appointment with a psychiatrist last week. During the appointment, they told you that it doesn't sound like you have ADHD but referred you for testing anyway. Eager for this diagnosis, you have decided to pretend that you have ADHD during this evaluation. Prior to the appointment, you looked at the criteria for ADHD. You understand that what you will be doing is dishonest, but you see this as means to an end.

INSTRUCTIONS: Your goal is to successfully convince the examiner that you have ADHD. **You will not be evaluated based on your acting ability**, but more-so on your ability to fake ADHD on the various tasks you will take today. This will be reflected in the scores that you produce.

You will now be given 10 minutes to review the diagnostic criteria for ADHD, after which the experiment will start.

********If you are able to successfully convince the examiner that you have ADHD, you will be given $10 at the end of this session. Good luck!!!********

Appendix J

Prompt for Incentivized-Delayed Group

During this experiment, you will be given several tests often used in the evaluation of Attention Deficit Hyperactivity Disorder (ADHD). Please read the following vignette for specific instructions:

You are a student at Indiana University of Pennsylvania. Last month, on a Saturday, you went to the library with your roommate to study for an exam. Upon arriving at an empty table, your roommate offered you a pill to help you study and told you that it was Adderall. You took this pill and found that suddenly you had the ability to concentrate without once getting distracted for several hours. You were so impressed with the results, the next day you asked your roommate for another pill to help you with a project. Again, you were able to sit and concentrate for several hours and felt very productive.

The following Wednesday, you asked your roommate where to get this medication. Your roommate informed you that they had been diagnosed with ADHD. In addition to a prescription for Adderall, they were also allowed extra time on exams, and could take these exams in a separate room, away from the distraction of other students. You then asked your roommate how to go about getting diagnosed with ADHD. They told you that first they went to a psychiatrist, and then to a neuropsychologist. They had to take a lot of weird tests, and afterwards, they were diagnosed with ADHD.

You then decided that you too wanted a diagnosis of ADHD. You booked an appointment with a psychiatrist last week. During the appointment, they told you that it doesn't sound like you have ADHD but referred you for testing anyway. Eager for this diagnosis, you have decided to pretend that you have ADHD during this evaluation. Prior to the appointment, you looked at the criteria for ADHD. You understand that what you will be doing is dishonest, but you see this as means to an end.

INSTRUCTIONS: Your goal is to successfully convince the examiner that you have ADHD. **You will not be evaluated based on your acting ability**, but more-so on your ability to fake ADHD on the various tasks you will take today. This will be reflected in the scores that you produce.

You will now be given 10 minutes to review the diagnostic criteria for ADHD, after which the experiment will start.

******If you are able to successfully convince the examiner that you have ADHD, you will be entered in a $100 Visa gift card raffle. Good luck!!!******

For entry into this drawing, please indicate an email address where you may be contacted:
_____@iup.edu

Appendix K

Attention Deficit Hyperactivity Disorder Diagnostic Criteria

Attention-Deficit/Hyperactivity Disorder

1. A persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development, as characterized by (1) and/or (2):
    A. Inattention
        a. Often fails to give close attention to details or makes careless mistakes in schoolwork, at work, or during other activities (e.g., overlooks or misses details, work is inaccurate).
        b. Often has difficulty sustaining attention in tasks or play activities (e.g., has difficulty remaining focused during lectures, conversations, or lengthy reading).
        c. Often does not seem to listen when spoken to directly (e.g., mind seems elsewhere, even in the absence of any obvious distraction).
        d. Often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (e.g., starts tasks but quickly loses focus and is easily sidetracked).
        e. Often has difficulty organizing tasks and activities (e.g., difficulty managing sequential tasks; difficulty keeping materials and belongings in order; messy, disorganized work; has poor time management; fails to meet deadlines).
        f. Often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (e.g., schoolwork or homework; for older adolescents and adults, preparing reports, completing forms, reviewing lengthy papers).
        g. Often loses things necessary for tasks or activities (e.g., school materials, pencils, books, tools, wallets, keys, paperwork, eyeglasses, mobile telephones).
        h. Is often easily distracted by extraneous stimuli (for older adolescents and adults, may include unrelated thoughts).
        i. Is often forgetful in daily activities (e.g., returning calls, paying bills, keeping appointments).
    B. Hyperactivity and impulsivity:
        a. Often fidgets with or taps hands or feet
        b. Is often “on the go,” acting as if “driven by a motor” (e.g., is unable to be or uncomfortable being still for extended time, as in restaurants, meetings; may be experienced by others as being restless or difficult to keep up with).
        c. Often talks excessively.
        d. Often blurts out an answer before a question has been completed (e.g., completes people’s sentences; cannot wait for turn in conversation).
        e. Often interrupts or intrudes on others (e.g., butts into conversations, games, or activities; may start using other people’s things without asking or receiving permission; for adolescents and adults, may intrude into or take over what others are doing).

2. There is clear evidence that the symptoms interfere with, or reduce the quality of, social, academic, or occupational functioning.

Appendix L

Affirmation of Instructions

Now that you have read the instructions **CAREFULLY,** please copy the following sentences in the space provided:

"I understand that, from this point forward, I am playing the character described in the story until the testing session is completed. From this point forward, everything I do will be based on this character, not based on me."

"I understand that my ability to fake ADHD will be evaluated ONLY by my ability to produce test scores like a person with ADHD."

Appendix M

Strategy Record Form

For this experiment, you were asked to fake ADHD to the best of your ability. What are some things that you did in order to make your test scores convincing?

1_____

2_____

3_____

4_____

5_____

6_____

7_____

8_____

9_____

10_____

11_____

12_____

13_____

12_____

15_____

16_____

17_____

18_____

19_____

20_____

Appendix N

Debriefing Form

Thank you for your participation! This study aims to understand college students' behavior within Attention Deficit Hyperactivity Disorder (ADHD) evaluations, and factors that may change this behavior. Your participation will help us understand how changes in motivation may alter test taking behavior and performance. If you have any questions about this research, please feel free to contact Jesse Siegel at j.a.siegel@iup.edu and/or Dr. Laura Knight at laura.knight@iup.edu

The following resources are available at IUP to assist students:

Students who would like an ADHD-related evaluation are encouraged to contact:

Center for Applied Psychology
Uhler Hall, 210
(724) 357 – 6228
http://www.iup.edu/psychology/centers/default.aspx
"The CAP currently houses several clinics offering psychotherapeutic and evaluation services, staffed by IUP faculty members who are Pennsylvania licensed psychologists and by doctoral students in advanced training."

Students who would like to undergo behavioral interventions for ADHD-related symptoms are encouraged to contact:

Center for Applied Psychology
Uhler Hall, 210
(724) 357 – 6228
http://www.iup.edu/psychology/centers/default.aspx
"The CAP currently houses several clinics offering psychotherapeutic and evaluation services, staffed by IUP faculty members who are Pennsylvania licensed psychologists and by doctoral students in advanced training."

The Counseling Center
Suites on Maple East, G – 31
(724) 357 – 2621
www.iup.edu/counselingcenter
"Faculty members and staff at the center work collaboratively with students to foster the self-knowledge and skills necessary to succeed personally, academically, and professionally. The Counseling Center allows students the opportunity to integrate their personal goals with their academic goals."

Center for Health and Well-Being
Suites on Maple East, Ground Floor
(724) 357 – 3550
www.iup.edu/healthcenter
"The Health Service is the campus health-care provider for IUP students. Part of the Center for Health and Well-Being, it contains offices providing health services for students (formerly the

Pechan Health Center), counseling services, health education services, and alcohol, tobacco, and other drug services."

Students who would like assistance with study habits, and other academic needs are encouraged to contact:

Center for Student Success
Pratt Hall, 107
(724) 357 - 3936
http://www.iup.edu/success/default.aspx
"Pratt Hall is designated as the Center for Student Success, and our students and staff are learning to depend on the resources within Pratt to enhance their IUP learning experience." Within the Center for Student Success is the Veterans' Student Liaison, Peer Mentor Program, PATH: Project Assignment Technology Help, and College Prep 101 for Latinos.

Supplemental Instruction (SI)
Department of Developmental Studies
Pratt Hall, 202
(724) 357 - 2729
http://www.iup.edu/page.aspx?id=40085
"SI provides small-group study/review sessions for sections of courses with difficult content or high levels of failure and withdrawal rates."

Students who would like assistance in planning for their future, and organizing related materials are encouraged to contact:

Career Development Center
Pratt Hall, 302
(724) 357 – 2235
http://www.iup.edu/career/default.aspx
"The center functions as a comprehensive career planning and placement service within the university in order to meet the needs of IUP students and alumni by helping students develop self-awareness, enabling students to clarify and evaluate their career and educational goals, providing students with direction and information on the job market and educational opportunities, helping students develop a methodology to reach their educational goals, and assisting students in the career decision making process."

If you are interested in learning more about ADHD in college students, please consult the following sources:

Diller, L. (2010). ADHD in the college student: is anyone else worried? Journal of attention disorders, 14(1), 3-6.

Gray, S. A., Fettes, P., Woltering, S., et al. (2016). Symptom manifestation and impairments in college students with ADHD. Journal of learning disabilities, 49(6), 616-630.

Low, K. G., & Gendaszek, A. (2002). Illicit use of psychostimulants among college students: a preliminary study. Psychology, Health & Medicine, 7(3), 283-287.

Weyandt, L. L., Oster, D. R., Gudmundsdottir, B. G., et al. (2017). Neuropsychological functioning in college students with and without ADHD. Neuropsychology, 31(2), 160.