# EFFORT TESTING IN CHILDREN ON

# NEUROPSYCHOLOGICAL

# AND SYMPTOM VALIDITY MEASURES

By

PHILIP LOUIS RAMBO

Bachelor of Science in Psychology
Texas A&M University
College Station, Texas
2002

Master of Science in Clinical Psychology
Oklahoma State University
Stillwater, Oklahoma
2008

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2011

EFFORT TESTING IN CHILDREN ON

NEUROPSYCHOLOGICAL

AND SYMPTOM VALIDITY MEASURES

Dissertation Approved:

Larry Mullins

Dissertation Adviser

John Chaney

Committee Member

Melanie Page

Committee Member

John Romans

Outside Committee Member

Sheryl A. Tucker

Dean of the Graduate College

TABLE OF CONTENTS

# LIST OF TABLES

TABLES

# CHAPTER I

## INTRODUCTION

Malingering in adult populations has received considerable scientific inquiry in the recent decades (Larrabee, 2003). As a result, measures of malingering are quickly becoming a standard component of adult neuropsychological evaluations (Iverson, 2003). Despite this expansion in effort testing of adults, paucity exists in the literature with regard to malingering in younger populations. While it is unclear why this area has received so little attention, the potential for malingering in children certainly warrants serious consideration, as children as young as four years of age have been found capable of utilizing deceptive strategies for material gain or to avoid punishment (Polack & Harris, 1999; Netwon, Reddy, & Bull, 2000; Ruffman, Olson, Ash, & Keenan, 1993; Sodian, Taylor, Harris, & Perner, 1991).

The extant literature on malingering in children suggests that children are capable of intentionally lowering their cognitive performances (Blaskewiz, Merten, & Kathman, 2008; Constaniou & McCaffrey, 2003); however, it appears that they may utilize different approaches to feigning impairment than adults. For example, it appears that children are more likely to exhibit an inconsistent malingered presentation, shifting in-and-out of underperforming, whereas adults are more likely to lower their entire cognitive performance or focus on a specific cognitive domain (e.g., memory) (Blaskewiz, Merten, & Kathman, 2008). As a result of this inconsistency, identifying children who are not providing a valid effort may prove more difficult than identifying adults.

One approach to identifying malingering is to include symptom validity measures during evaluations. Symptom validity measures are specifically designed to differentiate sub-optimal effort from true mild-moderate cognitive impairment (Larrabee, 2003). Recent efforts to cross-validate adult symptom validity measures in child populations have yielded mixed findings. More specifically, the Test of Memory Malingering (TOMM) shows promise (Blaskewiz, Merten, & Kathman, 2008; Nagle, Everhart, Durham, McCammon, & Walker, 2006) but other symptoms validity measures have not fared as well, particularly with younger children (Constaniou & McCaffrey,2003; Courtney, Dinkins, Allen, & Kuroski, 2003).

Another strategy to identify malingering in adults involves detecting patterns of performance on standard neuropsychological measures (Larrabee, 2003; Mittenberg, Azrin, Millsaps, & Heilbronner, 1993; Suhr & Boyer, 1999). While this approach has demonstrated promise in adult populations, it has received little attention in children (Blaskewitz, Merten, and Kathmann, 2008; Faust, Hart, and Guilmette, 1988).

The specific aim of the present study was to expand the literature on effort testing with children in three specific areas. First, the study further examined the validity of the TOMM in a child sample, as this measure has demonstrated promise with younger populations (Blaskewitz, Merten, and Kathmann, 2008; Constantinou & McCaffrey, 2003; Nagle, Everhart, Durham, McCammon, and Walker, 2006). Second, this study was the first to gather data on children's malingered performances on the Dot Counting Test. Third, this study was the first to measure malingered performances on a few standard neuropsychological measures thereby providing additional means of effort testing.

CHAPTER II

# REVIEW OF LITERATURE

Malingering is defined as faking a condition in order to obtain some secondary gain (e.g., miss work, obtain financial compensation). While this area has not received significant scientific inquiry until fairly recently (Larrabee, 2007), feigning psychiatric and/or medical symptoms easily predates modern society, with examples dating as far back as the Old Testament when David "acted like a madman" and "allowed saliva to run down his beard" because he was afraid of King Achish (I Samuel 21:13, New International Version). Over the years, different systems for identifying those who were exaggerating or feigning a condition were developed which typically involved observing for overt behavioral inconsistencies, questioning the individual regarding their symptoms, or threatening/conducting various medical procedures (Geller, Erlen, Kaye, & Fisher, 1990). However, it was not until the early 20$^{th}$ century with the Compensation Acts of 1906 and the First World War that the identification of malingering began receiving serious consideration (Chesterman, Terbeck, & Vaughan, 2008).

More recently, identifying malingerers has received considerable attention in the area of neuropsychological assessment (Larrabee, 2007). This is primarily due to the nature of this form of evaluation, as neuropsychological testing relies almost entirely on the validity of the individual's performance, which is susceptible to a variety of factors including motivation, fatigue, other psychological contributors, and malingering. In fact one recent study found that effort accounts for over 50% of the variability in neuropsychological evaluations (Green, Rohling, Lees-Haley, & Allen, 2001). Given these findings, it is of utmost importance for clinicians to be accurate in recognizing when cognitive performances are inconsistent with actual neuropsychological functioning.

*Base Rates*

Malingering base rates are difficult to establish given that malingerers are rarely likely to admit that they are not putting forth an honest effort, thus the identification of these individuals relies on the clinician's own observations and interpretation of test performance. In addition, it must be supposed that a certain subset of malingerers, who are more trained and/or more sophisticated in their feigned clinical presentation, are likely never suspected of sub-optimal effort. This issue is further complicated by those individuals who have true cognitive deficits but still intentionally underperform on testing (symptom exaggeration). This would be a difficult base rate to establish but presumably would be quite high, particularly in litigious circumstances. Despite these limitations, some studies have been conducted that have provided some insight into the prevalence of this spurious behavior. In general, it has been found the type of case (e.g., criminal, mild traumatic brain-injury) plays a large role in the presence of this behavior. The highest rates of malingering have been found to be associated with personal injury cases involving mild traumatic head injury, with roughly 40% of this population meeting criteria for malingering (Larrabee, 2003; Mittenberg, Patton, Canyock, & Condit, 2002;

Carroll, Abrahamse, & Vaian, 1995). This inherently makes sense given that these are circumstances with a high probability of potential secondary gain. Personal injury cases are followed by criminal cases, where as many as 20% of this population have been estimated to be malingering (Adolf, Denney, & Houston, 2007). This is followed closely by disability evaluations, where malingering has been estimated to be as high as 15% (Chafetz & Abraham, 2005). When expanding malingering criteria to include probable malingering, rates of malingering increase for these populations, with slightly over half estimated as meeting criteria in disability evaluations (Chafetz,& Abraham, 2005), while estimates in criminal settings have been as high as one-third (Adolf, Denney, & Houston, 2007).

*Malingering Criteria*

When diagnosing malingering, it is important to determine how malingering is going to be defined. In the most recent edition of the Diagnostic and Statistical Manual of Mental Disorders (*DSM-IV-TR*) malingering is listed as an additional condition that may be a focus of clinical attention and is defined as the intentional production of false or grossly exaggerated physical or psychological symptoms which are motivated by external incentives. In addition, the DSM notes that malingering should be of particular concern when clients are referred by an attorney for clinical examination, when there is marked discrepancy between subjective suffering and objective data, and when there is a lack of cooperation during the evaluation. Finally, this set of criteria also includes the warning that there is an increased likelihood of malingering when Antisocial Personality Disorder is present (*DSM-IV-TR*).

While the *DSM-IV-TR* can be helpful in providing some general guidance in diagnosing malingering, it is insufficient in providing a thorough evaluative framework as it does not provide a definitive set of diagnostic criteria, instead only offering some guidelines and suggestions regarding this behavior. One problem with this approach is that by not providing explicit criteria for malingering the diagnosis relies heavily on the clinician's subjective interpretation of these guidelines. Slick, Sherman, and Iverson (1999) attempted to offer a more comprehensive set of criteria for maligning that was based on the discrepancy method (the discrepancy between the presentation of the client and what is typically expected with the given condition).

The Slick criteria for malingering are as follows:

1. The presence of substantial external incentive (at least one)
2. Evidence from neuropsychological data (e.g., discrepancy between observed behavior and test data, discrepancy between client history and test data)
3. Evidence from self-report (e.g., self-report is inconsistent with known patterns of the disorder, self-report evidences exaggeration)
4. The behaviors meeting criteria for 2 & 3 are not fully accounted for by a psychiatric, neurological, or developmental condition and are intentionally aimed at gaining an external incentive.

An advantage to Slick et al.'s diagnostic approach is that it provides a more concrete diagnostic foundation for use in empirical studies. While research comparing malingerers from controls in the past has relied on more subjective methods of categorizing malingerers, the Slick criteria should lead to more precision and agreement in identifying this behavior. Ultimately, this should lead to a more accurate conceptualization of those who malinger.

*Methods of Detecting Malingering*

Currently, there are three primary methods that are utilized in the detection of malingering in cognitive testing. These include personality measures that utilize validity scales in determining the individual's approach to the measures, measures that are specifically designed to capture sub-optimal effort in specific cognitive areas (e.g., memory), and the identification of malingering through patterns of performance on standard neuropsychological measures. Because this review is concerned with malingering as it occurs during neuropsychological testing, only the symptom validity measures and patterns of performance will be addressed.

*Modern Malingering Measures*

One of the most popular methods of detecting malingering involves the use of a group of specialized tests that are aimed at identifying those who are intentionally underperforming cognitively (Larrabee, 2007). In general, these measures are designed to be sensitive to sub-optimal effort, while at the same time insensitive to true cognitive impairment. It also should be noted that when severe cognitive impairment is suspected it is important to consider the possibility of a false positive because these malingering measures presume that the examinee has no significant impairments (Larrabee, 2003). While a description of all the measures that fall in this category is beyond the scope of this review, a few of the more common symptom validity measures (Slick, Tan, Strauss, & Hultsch, 2004) are described below:

*Test of Memory Malingering (TOMM)*

The TOMM (Tombaugh, 1996) is one of the most studied and commonly used tests for assessing malingered effort during cognitive testing. It is a forced-choice recognition test that consists of an initial presentation of 50 items consisting of simple drawings of common items, followed by an immediate and a delayed recognition trial of the items. During the recognition trials, the subject is presented with 50 items (both new and old) and indicates whether they recognize the item, the administrator then indicates whether they are correct or incorrect (the first trial is actually considered a recognition task in which the subject is oriented to the correct items a second time, whereas the second trial measures for sub-optimal effort). During the second (final) recognition trial, the subject is not prompted regarding their correctness. Malingering is suspected on this measure if the individual scores below 45 correct (worse than chance) on the second trial (Tombaugh, 1996).

*Rey 15 Item Test*

The Rey 15 Item Test (or Rey 15) is a brief measure of malingering that is presented as a memory test. The Rey 15 consists of 3 columns and 5 rows containing 15 total items; however, the items are broken into 5 categorical groups containing 3 items in each set. In addition, 2 of the categorical groups are variants of the same set. Although there are multiple ways of administering this measure, the typical administration involves presenting the stimuli for 10 seconds after which the subject writes all of the items down that they can remember. Administration also sometimes involves an introduction to the measure that cues the participant that this test will be quite difficult, as it involves the memorization of several items over a short span of time. The typical cut-off for malingering on this measure is a score less than 9 out of 15, although this varies (Strauss, Sherman, & Spreen, 2006). While the Rey 15 is one of the more commonly used tests in assessing malingering (this is likely due to its brevity), it has consistently demonstrated low rates of specificity and sensitivity and thus should not be considered sufficient from the standpoint of assessing effort (Rey, 1958). Finally, in an effort to address the problems with the Rey, the Rey was recently revised to the Rey-II; however, one recent study which administered the Rey-II in conjunction with the TOMM found that the Rey-II still suffers from low sensitivity and specificity (Whitney, Hook, Steiner, Shepard, & Callaway, 2008).

*Word Memory Test*

The Word Memory Test is a computerized measure of malingering that involves the presentation of semantically linked word pairs. Similar to the TOMM, the malingering portion of the Word Memory Test is a forced choice task that requires the participant to choose between one of the originally presented words and a new word that is in the same semantic category. This forced choice portion occurs twice, immediately following the initial presentation of the stimuli and after a 30 minute delay. During each of these trials the participant is prompted regarding the correctness of their response. One advantage of the Word Memory Test over other malingering measures is that it also includes a series of memory tests that are administered following the recognition tasks. These include a multiple choice task (in which the participant is given a word and chooses from eight other words to identify its pair), paired associates (in which the person receives the first word and recalls the pair), and a free recall task that requires the participant to recall as many words as they can. These memory tests are administered from least difficult (multiple choice) to most difficult (free recall) to facilitate memorization of the pairings (Green, Allen, & Astner, 1996).

*Computerized Assessment of Response Bias (CARB)*

The CARB is a computerized measure of malingering that was formulated on the Hiscock and Hiscock procedure (1989). Administration involves the participant viewing 5 digits, followed by a 2 second delay, and then having to choose between two 5 digit numbers, one of which is the original. The CARB is similar to other digit tasks of malingering, however, it is computer administered which offers a couple of unique advantages. First, by being computer administered, the CARB allows for stricter control with regard to standard administration. This ensures that the CARB should be essentially the same task regardless of administrator as the likelihood of administrative error is virtually removed. Second, the computer administration also allows for the respondent's latency to be measured. This is important as delayed responding has been found to be associated with malingered performances (Allen, Conder, Green, & Cox, 1997; Green & Iverson, 2001).

*Patterns of Neuropsychological Performance*

While measures of malingering are useful in providing relatively straightforward methods of assessing for feigned effort, they also suffer from a few potential weaknesses. First, it is possible that more sophisticated individuals are capable of detecting the purpose of these measures. This is mainly due to the possibility that malingering measures (e.g., Rey 15) may appear relatively easy when compared to other measures that constitute a typical cognitive battery. For this reason, it has been recommended that malingering measures be used at the beginning of the evaluation, when expectations regarding the difficulty of these tests is likely shrouded in ambiguity. In addition, as these have been used more and more frequently, awareness of these measures has risen outside of the psychological assessment community (Mittenberg, Patton, Canyock, & Condit, 2002). As a result, the likelihood that examinees have been provided with general information regarding the presence of symptom validity measures or have been strategically coached on how to perform on specific measures has risen (Youngjohn, 1995). In order to circumvent this possibility, abnormal patterns on standard measures of neuropsychological functioning have been identified (Larrabee, 2003).

Several standard neuropsychological measures have been identified as offering some degree of utility in identifying sub-optimal levels of effort. These include digit span on the Wechsler scales (scaled scores that are 5 or less, total digit forward that is 4 or less, etc.) (Griffenstein, Baker, & Gola, 1994; Larrabee 2003), the Wisconsin Card Sorting Task (higher occurrence of perseverative errors and failures to maintain set) (Suhr & Boyer, 1999; Bernard, McGrath, & Houston, 1996), and poorer performances on gross motor tasks relative to fine motor skill (Griffenstein, Baker, & Gola,1996; Heaton, Smith, Lehman, & Vogt, 1978; Mittenberg, Rotholic, Russell, & Heilbronner, 1996), to name a few. Overall, the use of neuropsychological measures in assessing malingering has distinct advantages over relying solely on the standard symptom validity measures in that they (1) allow for the continual assessment of malingering throughout the evaluation

(Meyers & Volbrecht, 2003), (2) are more resistant than symptom validity measures to being identified as malingering tests because they are actual measures of cognitive performance (Youngjohn, 1995), (3) allow for assessment across a broad array of cognitive conditions, whereas symptom validity measures typically address one cognitive domain (e.g., recognition memory), and (4) are more time efficient given that they do not require additional effort testing measures be added to the battery administered (Meyers & Volbrecht, 2003). Overall, when addressing malingering in the context of a neuropsychological evaluation, it is recommended to not over rely on either of these methods and, rather, consider both symptom validity measures and performance patterns on neuropsychological tests when establishing the validity of a cognitive performance (Slick, 1999).

*Types of Designs*

Typically, studies addressing malingering utilize one of two study designs: simulation or known-group designs, with the former being the most common design. With simulation designs, the participants are instructed or "coached" to underperform on cognitive measures (in studies dealing with personality, the participants typically attempt to fake a mental disorder). As the name implies, the participants in these studies are presumably simulating malingering and are thus not malingering in the purest sense. As a result, the validity of the "malingering" that occurs in these studies is clearly a weakness of this design as it is difficult (if not impossible) to know for sure that these individuals are malingering in a fashion consistent with how malingering actually occurs in clinical settings. One attempt to address this concern has been the use of small rewards for successful malingering. In these studies the participant is typically informed that they will receive some financial incentive for lowering their performance, but not to an obvious degree. While this would seem helpful in eliciting performances that more likely represent actual malingering, the research has not always supported the effectiveness of providing these types of rewards (Bernard, 1990).

Another disadvantage to simulation designs is the simulation-malinger paradox (Rogers & Cavanaugh, 1983), which involves asking someone to comply with malingering and then drawing comparisons with individuals who do not comply when asked to, which is clearly a contradiction. Despite these weaknesses, one advantage to this design is that it offers a high degree of control and flexibility with regard to the overall study design. In addition, this design also offers the advantage of being able to easily compare between groups of simulated malingerers (e.g., coached vs. un-coached), clinical/neurologic populations, and/or controls (Rogers, Harrell, & Liff, 1993).

In contrast, known-group designs examine malingering in individuals who are presumed to be actively malingering (or whose data was collected when they were malingering). One obvious advantage to this approach is that it offers greater external validity given that these individuals have been identified as malingerers and thus a part of the population of interest. However, one problem with this design is that there is not always agreement regarding who fits the criteria for malingering. As mentioned before, the *DSM-IV-TR* does provide some general guidelines in determining malingered performance; however, there is clearly a certain degree of subjectivity when using these guidelines in real world settings. As a result, it is possible that some of the individuals in these studies do not represent malingering in its strictest since (Rogers, Harrell, & Liff, 1993).

*Children and Malingering*

Malingering in adults has been one of the more extensively researched areas in recent years (Larrabee, 2003). Despite the growing interest in adults' deceptive capacity, little effort has been made in determining children's ability to fake cognitive deficits. This is concerning given that past research has demonstrated children as young as pre-school ages have been identified as engaging in deceptive behaviors, particularly when they are benefited by avoiding punishment or through some material gain (Netwon, Reddy, & Bull, 2000; Polack & Harris, 1999).

Before examining the literature on malingering in children, the issue of why a child would malinger needs to be addressed. One reason that a child might engage in malingering involves one of the basic psychological mechanisms: avoidance. Avoidance is defined as mentally or physically avoiding something that causes some degree of distress (Merriam-Webster's, 2002). In terms of malingering in children, the classic avoidance scenario involves the child who feigns an illness in order to stay home from school. While this presumably common behavior seems benign on the surface, due to the secondary gain involved in avoiding the "distressful" situation, when a child engages in this type of behavior they are malingering.

Litigious circumstances provide the most obvious scenario in which a child would be motivated to feign cognitive deficits. Here, secondary gain would be through more traditional avenues of direct financial compensation or, perhaps, material compensation from some older entity. Given that personal injury claims, on the whole, appear to be on the rise (National Center for State Courts, 2008), the need to establish diagnostic malingering guidelines specific to younger populations clearly warrants further scientific inquiry.

One final area that has been identified more recently involving malingering and younger individuals involves the feigning of attention deficit hyperactivity disorder (ADHD) in order to obtain medications and/or favorable academic accommodations. In recent decades, the increase in young individuals seeking evaluations for this disorder has increased substantially (Harrison, Edwards, & Parker, 2007; Nichols, Harrison, McCloskey, & Weintraub, 2002). This issue is further complicated by diagnostic approaches that rely entirely on interview and self-report measures, thus making successful deception exceptionally easy (Harrison, Edwards, & Parker, 2007).

Those who have investigated malingering in younger populations have typically taken two approaches in attempting to address this neglected area, either through cross-validating adult measures of malingering on samples of children or through actually measuring the capacity of children to intentionally lower their performance on measures of cognition and/or symptom validity measures.

### *Cross-Validation Literature*

In 2003, Constantinou and McCaffrey conducted a study that examined children's performance on both the Test of Memory Malingering (TOMM) and on the Rey 15 item test. The authors listed four main goals for this project: 1) to collect a normative data for children on the TOMM, 2) to demonstrate that children's performance on the TOMM is comparable to adults, 3) to demonstrate that the TOMM is unaffected by age, education, and cultural background, and 4) to compare children's performance on the TOMM with the Rey-15-item test. This study utilized two samples of children, one in the United States and one in Cyprus. All of the children in this study (61 in Cyprus and 67 in the United States) were first administered the TOMM, followed by the Rey 15-item-test.

Overall, Constantinou and McCaffrey found support for the use of the TOMM as a valid measure of children's effort, with every child receiving a score above the cut-off by at least the retention trial ($3^{rd}$ trial). Two children from Cyprus did not score above cut-off on the second trial of this measure; however, these children were also the only children in this study who were accompanied by their parents and, as the authors point out, the presence of a third party can be detrimental to neuropsychological performance. No significant differences were found between TOMM performance and the children's age, education, and cultural background.

Performances on the Rey 15 item test were not as promising, as this test correlated highly with both age and education. Also, several of the younger children did not achieve the cut-off score for this measure, thus incorrectly suggesting less than reasonable effort by these children. As a result, the authors concluded that the Rey 15-item-test is not an appropriate measure of malingering in the very young. It was noted, however, that as the children's age increased their performance on the Rey-15 increased as well, suggesting that perhaps this measure is appropriate for children above the age of 9, although they recommended further investigation to fully clarify this possibility.

While Constanitnou and McCaffrey indicated that the TOMM appears to be a valid measure of effort in healthy children, it was still not known whether the TOMM would also be a valid measure in a clinically referred sample of children. Donders (2005) attempted to address this issue through an investigation that examined the validity of the TOMM on a large clinically referred pediatric sample. In this study, 100 children with documented clinical diagnoses were administered the TOMM and the California Verbal Learning Test, Children's Version (CVLT-C) as part of a flexible neuropsychological battery (the CVLT-C was included in order to offer a memory comparison for the TOMM performances). The working diagnoses varied, although roughly half were being evaluated following a traumatic brain injury.

Overall, only three of the 100 children did not reach the minimal cut-off for the TOMM. The author indicated that of these three, two were accurately identified as sub-optimally performing across the cognitive battery, while the last child was likely a false positive. Donders also found a slightly significant correlation between age and TOMM performance, even though 21 out of the 23 children in the youngest age group (6 – 8 years) still performed above cut-off for this measure. There was no correlation between the CVLT-C and the TOMM, which suggests that the TOMM is still assessing effort in children. Overall, the author concluded that this investigation provides further evidence for the potential of the TOMM as a measure of effort in children and that the normal "adult" cut-off of 45 appears to be appropriate with children as well (as 97% of participants met or exceeded this cut-off). This study also expanded the past literature by demonstrating the TOMM's utility in a pediatric sample.

Courtney, Dinkins, Allen, and Kuroski (2003) conducted a study that examined the effect of age on performance on two computerized effort tests, the Computerized Assessment of Response Bias (CARB) and the Word Memory Test (WMT). In this investigation, 111 children were administered these two measures in conjunction as part of a comprehensive neuropsychological evaluation. The children were between the ages of 6 and 17, with the majority being referred for attention deficit/hyperactivity disorder or learning disability evaluations. In line with the authors' hypotheses, the findings suggest that children who were below 11 years of age were significantly more likely to produce performances on both of these measures that would suggest the possibility of reduced effort, even though it is presumed that all children were putting forth normal effort. In addition, the authors also found that reading ability may have contributed to the poorer performance on the WMT, as children with lower reading levels had more difficulty achieving normal performances on this test. As a result, the authors advise against using the CARB or WMT measures with younger children, particularly when the children are suspected of developmental disabilities and/or have suspected reading difficulties.

Green and Flaro (2003) also conducted a study that examined the performance of children on the WMT (this study was an extension of Flaro and Green, 2000, which used similar methodology but included the CARB). The participants in this study were 135 children between the ages of 7 and 18 who were referred for a neuropsychological evaluation for a wide range of diagnostic purposes. The children were divided into two categories, a psychiatric group (e.g., oppositional defiant disorder, schizophrenia) and a neurologic group (e.g., fetal alcohol syndrome, Asperger's). As part of the neuropsychological evaluation, each of these children received the WMT, California Verbal Learning Test, Children's Version (CVLT-C), and the Rey Complex Figure Test. Other aspects of the battery varied based on individual diagnostic considerations. Four adult groups were also utilized for comparison with the children's performance, these included a community sample of healthy adults, two head injury groups (moderate-severe and mild head injury), and a neurological group consisting of strokes, brain tumors, and multiple sclerosis.

Overall, Green and Flaro found that this measure has utility as a measure of effort with children. Interestingly, age, intelligence, and diagnostic category were not found to be related to overall performance on the effort portion of the WMT; however, reading level was related to lower effort scores, specifically when the child's reading level was below third-grade. Another interesting finding included the children performing better than the minimal brain injury group, which is consistent with past research that suggests that this group is particularly likely to engage in symptom exaggeration. In conclusion, this Green and Flaro's investigation of the WMT demonstrates promise for the use of this measure in determining the presence of response bias in children, although the authors' suggest that this measure needs further validation and highly recommended assessing reading level before making determinations of effort with this test.

*Malingered Impairment Literature*

The first attempt to address malingered neuropsychological performances within a child population was conducted by Faust, Hart, and Guilmette in 1988. In their study, three children between the ages of 9 and 12 were administered a battery of measures that included the Wechsler Intelligence Scale for Children-Revised (WISC-R) and the Halstead-Reitan Neuropsychological Test Battery for Older Children. Before completing these measures, the children were instructed to lower their cognitive performance but not so much as to be detected by the technician, who was blind to the study. Additionally, the participants were told that if they were successful in lowering their performance without being detected, they would receive extra compensation (an extra $5.00).

After the test batteries were conducted, the cognitive data were mailed to 240 judges for diagnostic interpretation. Also in the mailing was a brief vignette that described a child whose medical evaluation following an automobile accident was unremarkable but exhibited a decline in academic performance and reported experiencing memory difficulties. Judges were asked to complete a form that provided a number of multiple choice diagnostic options, including whether they felt the performance was normal and the type of abnormality that they felt best fit the data and history (cortical dysfunction, functional impairment, or malingering). Overall, only one-third of the mailings were returned with none of the respondents indicating a diagnosis of malingering. The authors argued that the incorrect diagnostic responses from the clinicians provided clear evidence that children are capable of feigning believable neuropsychological deficits.

Despite the authors' conclusions, there are a few notable concerns with this investigation, particularly with the study design that may raise concern regarding the generalizability of the findings. First, the children's intellectual functioning was not established prior to the feigned performance. Without first determining the actual intellectual capcity of the child, the authors cannot definitively establish that underperformance indeed occurred (that the child actually malingered). As a result, it is possible that the protocols did not actually represent poor effort in children.

Second, the assessors were not provided with behavioral observations regarding the children's presentations. As indicated in both the DSM-IV guidelines as well as Slick's criteria, behavioral observations are a critical element in establishing the consistency between cognitive functioning and observable symptomology. By only providing a summary of the children's cognitive scores and brief history, the assessors were not provided with enough information to gain a more complete conceptualization of the children, particularly with regard to the validity of their presentation.

Finally, there were also a few potential issues with the survey respondents. First, the low response rate from clinicians (23%) could possibly be evidence that some of the clinicians who did not respond had concerns regarding the study (e.g. methodology), although this is purely speculative. Second, the clinicians were broadly categorized as "neuropsychologists," although there was a wide range of reported specialization in this area, with some clinicians reporting minimal training in neuropsychological assessment. This raises the possibility that, for at least some of the respondents, the lack of malingering diagnoses may have been a function of inadequate experience assessing effort within the context of neuropsychological assessment.

Bigler (1988) offered a comment on Faust, Hart, and Guilmette's 1988 investigation and identified several potential weaknesses. First, Bigler argued that the history that was provided to the respondents did not offer the depth of information that is typically gained through standard clinical practice. For example, there was no opportunity to speak with the referring professional, conduct thorough diagnostic and neurologic interviews, or obtain medical reports/imaging results. In addition, Bigler noted that the nature of questionnaires may lead to higher rates of over-diagnosing. For instance, when provided with a high score on a measure of cognitive impairment, the judge may be likely to attribute the score to being related to cognitive impairment simply out of lack of alternative explanations. Bigler also commented on the qualifications of the respondents; only 17% of the clinicians had completed postdoctoral training in neuropsychology and little evidence was provided to verify the competence of these clinicians. Finally, Bigler contended that due to the relatively rare occurrence of malingering in children, the respondents should have been informed at the onset that this was a study of malingering and base rates of malingering should have been provided.

Faust, Hart, Guilmette, and Arkes (1988) conducted another investigation at roughly the same time as their child malingering study that replaced the children with adolescents. This study consisted of two independent evaluative conditions, the first with 3 adolescents who feigned cognitive impairment and 1 actual head injury case, while the second study only included 2 of the malingered protocols with an additional head injury protocol. Again, the clinicians were provided with a multiple choice form to determine the likely cause of the cognitive performance, with cortical dysfunction, functional impairment, or malingering as options. The clinicians were also instructed to rate their diagnostic confidence.

With the exception of the one legitimate protocol, the first study was essentially the same as their malingering study with children, with 3 adolescents being instructed to fake bad on a standard neuropsychological battery (e.g., Halstead Reitan, WISC-R). These results were summarized and mailed to neuropsychologists along with a brief vignette involving the adolescent being involved in an automobile accident. Roughly one- third of clinicians returned the forms. The malingered conditions combined resulted in “abnormal” ratings of 78%, while the injury case resulted in 100% abnormal ratings; however, out of the 3 diagnostic options given, none of the respondents indicated malingering in any of the 4 protocols, which is consistent with the original investigation.

In the second study, clinicians each received 2 protocols taken from the 2 malingered protocols and 2 head injury protocols (with 6 evenly distributed pair wise possibilities). In addition, they received a cover letter that provided details regarding the “50% base rate for malingering in the data sample.” The authors indicated that the reason for this letter was to see if forewarning clinicians regarding the possibility of malingering would increase the likelihood of their making this diagnosis. The respondents did indicate malingering in 10% of the 90% of the protocols that were determined to be abnormal. However, the authors noted that the forewarning regarding malingering resulted in less ratings of abnormality overall. They further argued that this warning likely increased the threshold for diagnosing abnormality, which ultimately could result in less accuracy in diagnosing more mildly impaired individuals (false negatives) while diagnosing the more exaggerated malingerers as impaired (false positives), which, did occur in their findings. Overall, this investigation suffered from some of the same methodological weaknesses as the original study (e.g., lack of behavioral observations, wide variance of respondent training), although the addition of the brain injury protocols did provide a comparison for the malingered performances.

In 2006, Nagle, Everhart, Durham, McCammon, and Walker conducted an investigation that attempted to compare “feigned” impairment with full effort children on forced choice recognition and on verbal learning. In that study, 35 children ages 6-12 were split into two groups with both of these groups completing a full-effort condition and a malingered condition (the only difference between the groups was that the order of the conditions were reversed, with one group receiving the full-effort condition prior to the malingered condition and the other group receiving the malingered condition prior to the full-effort condition). The TOMM was administered to capture performance on a forced-choice task, while the Hopkins Verbal Learning Test-Revised (HVLT-R) was administered to capture the children's verbal learning capacity. In addition, Reading from the Wide Range Achievement Test-3rd edition (WRAT-3) was also obtained prior to these conditions in order to ensure that the children possessed basic verbal academic ability. Finally, a checklist assessing knowledge of brain injury was devised to explore the possibility that children who are more aware of the potential symptoms of brain injuries may perform differently in the feigning condition than their less informed peers.

The authors hypothesized that when asked to do their best the children would perform on the TOMM at the same level as adults (which is consistent with past research examining this measure with children, see Constantinou & McCaffrey, 2003). They also hypothesized that when the children were in the feigned condition, the children would perform beneath the cut-off for this measure (<45). On verbal learning, it was hypothesized that full-effort condition overall would perform significantly better than the malingered condition.

Overall, it was found that regardless of condition, children performed above the cut-off on the TOMM. The authors suggest that this provides further evidence for the utility of this measure when attempting to address malingering in younger populations, although this raises concern regarding the sensitivity of this measure to malingering. On the HVLT, in the first administration there were few significant differences between the malingered and full-effort condition.  However, on the second administration, the children in malingered condition (who had already gone through the full effort condition) "remembered" significantly fewer words and evidenced less learning than the full effort condition.  The authors suggested that, at least for this measure, children may require prior exposure to the task before being able to successfully lower their memory performance.

There are some notable strengths and weaknesses to this study.  As with Constaniou and McCaffrey, this study provides further evidence supporting the use of the TOMM with children 6 and older).  One caveat, however, is that the children in the malingered condition did not score below the cut-off.  While the authors indicated that the failure to reach cut-off may suggest that children have difficulty "faking" this measure, it is also possible that the children were lowering their performance, just not enough to be identified as possibly malingering.  Assuming that the children were lowering their performance, this could suggest that the TOMM is actually a poor measure of validity in children.  In addition, it is also possible that the children did not reach cut-off because in this investigation they were only administered a few measures.  As a result, it is possible that they had not established a faking strategy which could have resulted in a higher score than may have been obtained in a full neuropsychological battery.  Their finding that children who first are administered the HVLT at full effort were able to lower their performance on the second administration offers some support for this hypothesis as this demonstrates that with some experience, the children were able to provide a more "impaired" performance, although without further exploration it is not possible to definitely establish why these performance patterns occurred.

In 2008, Blaskewitz, Merten, and Kathmann conducted an investigation in Germany with children who were assigned to either an experimental group (malingering) or a control group (full effort) and were administered a battery of malingering measures. This battery included the TOMM, the Medical Symptom Validity Test (MSVT), and the Rey Fifteen-Item Test. The children were also administered some standard neuropsychological tests. These included digit span from the WISC-III, form A of the Trail Making Test, and Raven's Coloured Progressive Matrices. In order to elicit sub-optimal effort in the experimental group, a script was read to the children involving a wizard coming to their school who was looking for an apprentice for his magic school. The script included the following:

"…the wizard definitely does not want children that are smarter than him, but the wizard definitely does not want children who are stupid either so he would be able to teach them appropriately."

After the script was read, a simple memory task was conducted to see if the children understood the concept of intentionally underperforming. If the child did not make a mistake following the first trial, then a second trial was conducted and if they did not make mistakes on the second trial they were excluded from the study. Only one child was excluded as a result of not understanding the task. Overall, Blaskewitz and colleagues found significant differences when comparing the full effort group from the malingering group on all measures except the Raven's and Trail's A. In addition, out of the 20 children in the malingered children, 19 failed at least one of the malingering measures and the majority failed three of these measures (n = 10), while only 2 children failed all effort measures. Interestingly, the authors also indicated that the majority of children did not appear to sustain their malingering effort throughout the entire battery. The authors noted that the children in the malingered condition tended to display an inconsistent pattern, underperforming on some measures while being apparently unaware of their role as a "sorcerer's apprentice" on other measures. They suggest that children may have difficulty maintaining their attentional awareness and/or motivation to successfully sustain a consistent level of underperformance over an extended period of time. In addition, the authors reported that the children had more success underperforming on some measures than others, as evidenced by more underperformances on the MSVT than on the TOMM (which is consistent with previous research which suggested that the TOMM may be problematic in detecting malingered effort in children). Finally, the authors offered the observation that the children had more difficulty lowering their performances in tasks where they were explicitly told to do something or when they were provided with immediate feedback, suggesting that children may struggle in maintaining their malingering when provided with feedback that is contrary to their underperforming.

As for the control group, all participants scored above the cut-off for the TOMM trial 2 and retention trial (which had not been previously established). In addition, almost all of the children were able to pass the MSVT. Also, all of the children passed the Rey 15 item test, which is contrary to the aforementioned study regarding the utility of this measure with younger children. Finally, the digit span task (RDS) was failed by the majority of the full effort group, suggesting that this measure is not appropriate for younger populations.

Overall, the Blaskewitz study provides a solid investigation on which to build future studies exploring response bias in children. As described by the authors, one area that may have contributed to the inconsistent underperformances in some of the children may involve the use of the sorcerer's apprentice scenario. Perhaps if a more plausible scenario had been provided or the children had been offered some external gain, then the children may have better maintained their motivation to underperform throughout the evaluation.

Finally, one other study that warrants consideration was conducted by McKinzey, Prieler, and Raven in 2003 that examined children's performance on the Raven's Standard Progressive Matrices. The purpose of this study was to cross-validate a formula that was developed by Gudjonsson and Shackleton (1986) for detecting malingering on the Raven's using a sample of children. In this study, 44 school aged children were administered the Raven's twice, first with the standard instructions to do their best and second with the instructions:

"We know that some people don't try their best on this test. We'd like to find a way to catch them. To help us, please do as badly on this test as you can, without getting caught."

Following these two administrations, the authors found that in the malingered condition the formula yielded a 64% false negative rate, suggesting that this formula may offer little when attempting to identify malingering in younger populations. Interestingly however, after closer examination of the children's response patterns in the malingered condition, the authors identified a subset of extremely easy items that increased the hit rate of malingering on this measure to 95%. This finding could suggest that children's patterns of malingering may be different from adults. Overall, this study provided the possibility of a new method for detecting sub-optimal effort in children. However, this investigation did suffer from a few weaknesses. First, providing the students a standard administration of the Raven's before administration in the malingering scenario potentially damages the external validity of this study as children who are being assessed in actual clinical settings are not going to have the opportunity to practice the measure before they underperform on instruments. In addition, the authors did not report if any effort was made to confirm that the children understood what was expected after they heard the malingering script. This could potentially confound the findings, particularly given that some of the children were very young and may not have understood their role in the study.

# CHAPTER III

## PRESENT STUDY

Malingering in adult populations has witnessed a rapid growth in the scientific literature over the past few decades. As a result of this increase in inquiry, certain facets of this behavior have emerged that allow for more precision when identifying malingered effort in older populations. These techniques include validity scales on personality measures, symptom validity tests that assess for simulated cognitive deficits, and patterns of performance on standard neuropsychological measures. One area that has been relatively neglected in the literature concerns malingered effort in younger populations. The extant literature that has attempted to provide clarity regarding children's performance on symptom validity and neuropsychological tests has documented that, to some extent, children are capable of reducing their effort when instructed to do so; however, these efforts to explore malingering in children represent only preliminary investigations.

The purpose of the present study was to expand the current literature in children's response bias in a few specific ways. First, children's malingered efforts on some common neuropsychological measures that have not been previously studied in this context were explored. Overall, it was hypothesized that the children in the control would perform better on all of these measures than children in malingered condition.

Second, the present investigation also attempted to expand the literature by including a measure of malingering, the Dot Counting Test, that has not previously been examined in the context of malingered effort in children (Rey, 1941). It was hypothesized that the children in the control condition would perform significantly better on this measure than the children in the malingered condition.

Finally, this study attempted to further investigate the use of the TOMM with younger populations. As described previously, the TOMM has shown promise as being a valid measure of effort; however, it has not been demonstrated that children who are underperforming actually score below the adult cut-off for this measure. It was hypothesized that the children in the malingered condition would perform lower than the children in the control condition, however, it was expected that the majority of children would perform above the cut-off of 45 on both the 2$^{nd}$ trial and retention trial, regardless of condition.

# METHOD

## *Participants*

Participants included 38 children in the Stillwater, Oklahoma and greater Dallas, Texas communities. The average age of participants was 9 ($M$ = 9.72, $SD$ = 2.01), with 65% of the sample being girls and 35% being boys, although the sample included children as young as 6 and as old as 12. Nearly the entire sample was Caucasian (97%). Roughly half (46%) of the participant's families fell within the $50,000 to $75,000 annual income range, 35% were greater than $75,000, 13.5% fell within the $25,000 to $50,000 range, and 5.4% fell within the less than $25,000 range.

One participant was excluded from the study due to a previous diagnosis of mental retardation. Randomization resulted in 18 children in the control condition and 19 children in the treatment condition. All 18 children in the control condition reported that they understood that they were to “try their best” for the remainder of the experiment. In addition, all of the 19 participants in the poor effort condition were able to make at least one mistake on the second trial of the sample memory task, resulting in no children being excluded from the study due to not having a basic understanding of the underperforming role condition.

Prior to analyses, three additional participants were excluded due to potential inconsistencies in their presentation. In the treatment condition, one subject was excluded because they reported to the first examiner that they tried their best on all of the tests with the second examiner. This was consistent with this individual’s cognitive performance, as their KBIT-2 and WASI full scale intelligent quotients were highly consistent and their TOMM performances were not indicative of poor effort. The second examiner also reported that this individual appeared to be putting forth their best effort on all tests.

A second subject in the treatment condition was excluded due to their having a slightly higher WASI intelligence quotient than KBIT-2 quotient and because they received perfect scores on the TOMM, thus raising serious concern that this individual did not comply with the directions. The second examiner also noted that this individual appeared highly motivated to do well on all the tests.

One of the participants in the control condition was excluded due to possible contamination, as this participant reportedly knew children who had previously gone through the experiment in the treatment condition and appeared to be putting forth less than reasonable effort on a few of the subtests. For example, on block design, this participant received a raw score of zero, as they demonstrated significant difficulty in solving even the most basic two block configurations. However, following the experiment, the first examiner had the subject attempt the blocks again with encouragement to be sure to try their best, at which point this individual evidenced no difficulty in quickly solving some of the more complex four block designs. Due to this inconsistency, their data was excluded from analyses.

The removal of these participants resulted in both conditions having 17 children. The average age of the control condition was 10 (M = 10.06, SD = 1.85), while the average age in the treatment condition was 9 (M = 9.71, SD = 2.20). The groups did not differ significantly for age, grade, or family income. See table 1 for further detail regarding group characteristics.

*Measures*

The following measures were administered to all participants. The participants were first administered the KBIT-2 at full effort.

*Kaufman Brief Intelligence Test-2$^{nd}$ Edition (KBIT-2).* The KBIT-2 is a measure of intellectual functioning that yields an overall intelligence composite as well as a verbal and a nonverbal intelligence composite (Kaufman & Kaufman, 2004). Verbal intelligence is measured through two subtests, a measure of verbal knowledge and a measure that involves solving verbal riddles, while nonverbal intelligence is measured through a subtest of matrices. Administration time for the entire test is roughly 20 minutes.

*Wechsler Abbreviated Scale of Intelligence (WASI).* The WASI (The Psychological Corporation, 1999) is an abbreviated measure of intelligence that provides reliable estimates of verbal, performance, and full scale intelligence quotients (IQ). Performance ability is measured by two tasks; the first involves manipulating three-dimensional block designs while the second involves nonverbal abstract reasoning of matrices. Verbal ability is also measure by two tasks; the first requires defining words while the second involves verbal abstraction where the participant is required to find similarities between two words. Internal reliability for the WASI is reported to be .92 in adults according to the test publisher. Also, the full scale IQ's obtained from this measure correlate .92 with IQ's obtained from the *Wechsler Adult Intelligence Scale (WAIS-III)* (Psychological Corporation, 1999).

*Test of Memory Malingering (TOMM).* The TOMM is a forced choice recognition task that is one of the most frequently administered measures of malingering (Tombaugh, 1996). The TOMM has demonstrated some validity as an effort measure with children with even young children being able to perform at or above the cut-off; however, it remains unclear if this measure is sensitive to malingering in children. In adults, malingering is suspected on this measure if the individual scores below 45 correct on the second trial.

*Dot Counting Test.* The Dot Counting Test (Rey, 1941) is a brief measure of malingered effort that involves counting sets of dots that are presented on a computer. More specifically, the task requires the participant to count a set of dots as quickly as possible. The first six sets of dots are ungrouped, whereas the last six sets of dots are grouped in closer proximity to each other. Normal effort is evidenced by the participant counting the latter six sets more quickly than the original dot sets.

*Fingertapping.* Fingertapping is a simple test of motor speed and motor control that was originally part of the Halstead-Reitan Battery (Halstead, 1947). This procedure involves having the participant tap a mechanism that is fastened to a wooden board as quickly as they can. Previous research has indicated that underperforming in adults may be present when the total score for both hands (left performance + right performance) is less than 63 (Heaton, Smith, Lehman, & Vogt, 1978; Mittenberg, Rotholic, Russell, & Heilbronner, 1996).

*The Trail Making Test (Trails A & B).* The Trail-Making Test is a measure of mental flexibility, concentration, and visual-motor coordination. This measure is divided into two parts, part A and part B (Army Test Battery, 1944). Part A is simply a sheet of paper scattered with circles containing numbers. The participant is required to draw a line connecting the different circles in numeric order (1-25). Because in part A the subject is simply going in numeric order, this task does not have an executive component and rather is a measure of speed of information processing, specifically visual scanning, and visual-motor abilities. Part B has numbers and letters in circles scattered randomly across the instrument. As in part A, the participant draws a line connecting the circles in order; however, in part B the subject is required to alternate from number to letter, starting with the circle marked with the “1” next to the circle marked with “A” then to “2” then to “B”, etc. Part B is an executive functioning measure and it correlates positively with number of perseverative errors on the Wisconsin Card Sorting Task (Lezak, Howieson, & Loring, 2004), a classic executive functioning test. Many studies have examined the reliability for the Trail-Making Test, with most reporting reliability between .60 and .90 (Spreen & Strauss, 1998).

*Procedure*

IRB approval was obtained from Oklahoma State University (Appendix A, page 68) and the University of North Texas (Appendix B, page 69). In addition, all participants and their parents were treated in accordance with the American Psychological Association Ethical Code (APA, 2002). Prior to beginning the experiment, informed consent was obtained from the parent. This involved the overall purpose of the study, hypotheses of the study, information regarding the confidentiality of the data, and the possible benefits and risks of this study. To maintain the integrity of the children’s performances in both conditions, children were not present while informed consent was obtained from parents. Following study completion, parents were provided with an estimate of their child’s intellectual capacity via an administration of the Kaufman Brief Intelligence Test, 2$^{nd}$ edition (Kaufman & Kaufman, 2004). Parents were also provided with $20.00 for their time and effort in allowing their child to take part in the study. Children were also provided with a small prize at the end of the study for their participation.

Regarding the potential risks, parents were informed of the slight possibility, although highly unlikely, that children in the experimental condition would possibly exhibit sub-optimal effort in future tasks as a result of taking part in this study. However, in order to minimize this risk, the children were not provided with their prize until they completed a brief task with the examiner at full effort. There were no known risks for the control condition. Finally, the parents were reminded that their participation was voluntary and that they could withdraw their child from the study at any time.

The first examiner next obtained assent from the child, which involved a brief summary of the experiment (e.g., you are going to be taking some tests, these are sort of like games). The child was first administered the Kaufman Brief Intelligence Test-2$^{nd}$ Edition (KBIT-2), which is a brief measure of overall intellectual functioning. All children were briefly encouraged before the administration of the KBIT-2 to put forth their best effort. Following this administration, the children were randomly assigned into either the experimental or the control condition. This was accomplished by having the child pull a slip of paper out of an envelope, which indicated whether they would participate in the full effort "control" condition or the poor effort "treatment" condition. In addition, two separate envelopes were used, one for children within the ages of 6-9 and one for children within the ages of 10-12. These separate envelopes were utilized in order to ensure roughly equal distribution of ages across the two conditions. If the child pulled a piece a paper indicating the control condition, they were read the following:

> "You did really well on this test and I can tell you tried your best. Now you are going to take some more tests with someone else and I want you to continue trying your best and you will get a present at the end of the session. Do you understand what you are supposed to do?"

They then proceeded to another room where they were tested by a second examiner who was blind to their condition. If the child's paper indicated the treatment condition, they were read the following:

> "You did really well on this test and I can tell you tried your best. Now we are going to play a little game. What I would like you to do in this game is pretend like you are ***not as smart*** as you really are. ***Do not try your best anymore.*** I want you to get some things wrong even though you know the answers. But, don't get so many wrong that the person giving you the tests will know that you are not trying your best. If you can get wrong answers but make the tester think you are trying your best, you will get a present at the end of the session. Do you understand what you are supposed to do?"

Following these instructions, to see if the child understood what was expected they were given the following simple memory task:

> "So, for example, I want you to remember these words, but don't try your best: apple, table, penny."

If the child did not make any mistakes, they were read the following:

> "Remember, you are not supposed to try your best, let's try again: apple, table, penny."

As with the control condition, the child next went to a separate room where they were given the remaining measures from a second examiner.

After the protocol was completed, the children returned to the room with the first administrator for debriefing. Children in the control condition were allowed to choose a toy and thanked for their hard work. The children in the experimental condition were asked if it was difficult for them to lower their effort and asked if they did anything specifically to lower their performance. They were next reminded that it was very important to always try their best and that they were only asked to not try their best this one time so that we could be able to know if children are not trying their best, so we can better help them. They then were asked to do a simple task at full effort, such as naming their three favorite foods and were allowed to choose a toy and thanked for participating in the study. While the first administrator was finishing with the children and parents, the second administrator completed a form in which they indicated which condition they thought the child was in, why specifically they thought this condition, and any other potentially relevant observations.

.

# CHAPTER IV

# RESULTS

## *Analyses*

A multivariate analysis of variance (MANOVA) was conducted in order to examine potential differences between the control group and the simulated poor effort group (treatment group) on the cognitive and symptom validity measures. A separate MANOVA was conducted to explore possible intellectual differences between the children in the control condition and children in the treatment condition on the KBIT-2. Two T-tests were also conducted in order to test for significance between the IQ's generated for the control group (both at full effort) and significance between the IQ's generated for the treatment group (one at full effort, one poor effort). It was anticipated that the children in the treatment condition would perform more poorly overall than the children in the full effort condition, particularly on symptom validity measures (e.g., TOMM). In addition, it was anticipated that IQ's generated in the control condition would not differ significantly, while the IQ's generated for the treatment group would differ significantly with the treatment IQ being significantly lower than the full effort IQ.

## *Results*

.

One-way multivariate analysis of variance did not yield a main effect for condition, $F(1, 32) = 1.026$, $p = .499$. However, several significant univariate effects did emerge, including significant effects for WASI full scale IQ ($p = .022$), WASI nonverbal IQ ($p = .033$), WASI Matrix Reasoning ($p = .008$), TOMM trial 1 ($p = .001$), TOMM trial 2 ($p = .001$), and TOMM Retention ($p = .001$). In addition, finger tapping total and finger tapping non-dominant were both approaching significance (*p*'s = .074 & .067, respectively) (Table 2). It is also worth noting that even though the other measures were not found to significantly differentiate the conditions; average performances on every measure were grossly lower in the treatment condition than the control condition (Table 3).

Due to finger tapping being entered as raw scores, an additional MANCOVA was conducted to rule out the possibility that this trend may have been confounded by age. This did reveal a significant main effect for age, $F(1, 32) = 6.231$, $p = .005$. However, univariate effects for finger tapping total and finger tapping non-dominant were still approaching significance with this control in place (*p*'s = .080 & .067, respectively).

A separate analysis of variance (ANOVA) that examined possible differences between the conditions on the KBIT-2 (which was full effort for all children) did not yield a main effect for condition, $F(1, 32) = 1.792$, $p = .170$. This includes no significant differences between full scale IQ's for the control children ($M = 109.94$, $SD = 10.48$) and treatment children ($M = 107.76$, $SD = 12.97$) ($p = .594$), verbal quotients (control $M = 102.65$, $SD = 10.67$; treatment $M = 106.12$, $SD = 11.38$, $p = .366$) and nonverbal intelligence (control $M = 114.05$, $SD = 10.78$; treatment $M = 106.94$, $SD = 14.78$, $p = .119$).

A dependent t-test revealed no significant differences between the control children's full scale IQ's on the KBIT-2 (*M* = 109.94, *SD* = 10.48) and WASI (*M* = 108.18, *SD* = 11.29), $t(16) = 1.072$, $p = .300$. Treatment children did demonstrate a significantly poorer IQ performance on the WASI (*M* = 97.94, *SD* = 13.45) relative to their full effort KBIT-2 IQ (*M* = 107.76, *SD* = 12.97), $t(16) = 2.598$, $p = .019$.

*Intellectual Measures*

Consistent with hypotheses, the treatment condition's full scale intelligence was significantly lower on the WASI than the control condition, although there were no significant differences between the conditions on the KBIT-2, which was administered at full effort for both groups. In addition, the children in the treatment condition also performed significantly lower on the WASI IQ than their own full effort KBIT-2 IQ. These findings clearly indicate that the children in the treatment condition were fully capable of lowering their intellectual performance with the minimal coaching that was provided. Interestingly, nonverbal IQ but not verbal IQ emerged as significantly different between groups and the only individual intellectual performance that was significantly different was Matrix Reasoning. This includes Block Design, Vocabulary, and Similarities not reaching significance. Overall, this pattern of the comprehensive full scale intelligence quotient and nonverbal quotient differentiating the conditions, but not 3 of the 4 individual tests reaching significance, suggests that perhaps the children were subtly lowering their effort across measures, with the exception of more pronounced lowered effort on Matrix Reasoning.

*Neuropsychological Measures*

Although not quite reaching statistical significance, a possible trend emerged for fingertapping, with the treatment condition nearly reaching significance for nondominant and combined performances when compared to the control condition.  This finding is consistent with the adult malingering literature, although the children in the present study performed on average marginally higher than what has been found in adults (combined average of 68, as opposed to the recommended combined cut-off of 63 for adults). Interestingly, dominant fingertapping performances were  not as near to significant as nondominant and combined performances, although the treatment condition did fewer taps on average than the treatment condition (control dominant $M = 39$; treatment dominant $M = 35$).

Neither trial of the Trial Making Test was significantly different between conditions, although average performances were lower for the treatment condition (Trails A control $X = 53$, Trails A treatment $X = 63$; Trails B control $X = 133$, Trails B treatment $X = 162$).  Errors on these measures also did not differ significantly between conditions.

*Symptom Validity Performances*

The participants' performances on the TOMM were promising with regard to specificity, as every child in the control condition performed above the cut-off of 45 on both the second administration and the retention (delayed) trials, with an average second administration performance of 49.76 and average retention also of 49.71.  In addition, all but 4 control children performed above 44 on trial 1, with an average performance of 45.71.

Performances in the treatment effort condition on the TOMM raise sensitivity concerns regarding the use of this measure in younger populations. This is due to only slightly over half (10 of 17) of the treatment children performing below the cut-off of 45 on trial 2 and the same number performing below the cut-off for the retention trial (these were not all the same children), with average performances of 38.12 and 35.35 for trial 2 and the retention trial respectively. It should be noted that 12 of 17 (71%) of the children in the treatment condition performed below 45 on the 1$^{st}$ trial of the TOMM, with an average performance of 37.29, although this trial is typically considered a measure of recognition memory rather than symptom validity.

Performances on the Dot Counting Test were not significant between groups. This includes no significant effects for response time to grouped dots, ungrouped dots, average response time across groups, total correct dots, and difference between response time of grouped and ungrouped dots. There does, however, appear to be a slight trend with the latter as the average difference in response time between the grouped and ungrouped dots was slightly shorter for the treatment children ($p$ = .11) when one control outlier was removed from the analysis.

*Examiner Observations*

As part of collecting data in this study, examiners were instructed to be mindful of any behavioral indicators that could assist in discerning if the child was not putting forth a reasonable effort on the measures. Overall, the child's TOMM performances were the most consistently reported determinant in establishing whether the child was in the control or treatment condition. Block design was the second most identified task establishing effort, although this test did not significantly differentiate the groups. Matrix Reasoning and finger tapping were the two neuropsychological measures least reported by examiners, although they were found to be the most effective in identifying underperforming. However, the lack of attention these measures received could be related to the relatively passive role the examiner plays in their administration, as little feedback or instructions are provided once the task has commenced. Other reported behavioral indicators included mistakes on easy items while getting more difficult items correct, reports of confusion on relatively simple tasks, and the child giving up prematurely.

Examiners also noted that some children appeared to fluctuate with regard to their underperforming; appearing to put forth good effort on some tests while appearing to intentionally underperform on other tests. Younger children were observed to engage in more obvious behaviors suggestive of underperforming (e.g., laughing after a grossly incorrect response on Vocabulary, asking the second examiner if they were not supposed to try their best). In general, older children were observed to engage in more sophisticated methods of underperforming, such as intermittently getting items incorrect or providing answers that were nearly correct on tests with open ended questions. It should be noted, however, that some of the older children engaged in unusual behaviors as well, just not as frequently as the younger children. For example, one of the older children reportedly got lost while trying to find the restroom which was straight down the hall approximately 15 feet.

*Examiner Accuracy*

Overall, the examiners were highly accurate in identifying the condition of the children. As mentioned previously, the TOMM was the most typically cited instrument in assisting with this determination. In general, accuracy was slightly greater for the full effort children (15 of 17 correctly identified), although the majority of the children in the treatment condition were also correctly identified (13 of 17 correctly identified). Of the 6 cases that were misidentified, 4 were false negatives while 2 were false positives. Interestingly, when examining these by age, the false negatives were all above the age of 10, while the false positives were below the age of 10, which again suggests that the older children were more convincing with regard to their faking performance. This also suggests that younger children are possibly more prone to engage in erratic behaviors that may be perceived as underperforming, although these findings should both be considered preliminary due to the small sample size.

*Reported Strategies*

Consistent with prior research, older children tended to report more sophisticated strategies in underperforming than younger children, which was also consistent with examiners' observations. For example, older children reported alternating between periods of full effort and poor effort throughout the testing, such as giving an incorrect response following a certain number of correct responses. Younger children tended to report simplistic strategies, such as simply not trying their best. However, it should be noted that the majority of children had difficulty articulating a precise strategy for their underperforming.

.

# CHAPTER V

# CONCLUSION

The present investigation sought to further elucidate malingered presentations of children on neuropsychological and symptom validity measures. This was accomplished through the administration of the WASI, TOMM, Dot Counting Test, Trail Making Test, and fingertapping to children in either a full effort control condition or a simulated malingering treatment condition. This study also assessed the full effort intellectual capacity of all participants prior to randomization, thus allowing for further inquiry regarding possible intellectual differences between the groups and to provide further evidence that the children were truly underperforming.

The first hypothesis regarding the control children performing significantly better than the children in the treatment group on measures of neuropsychological and intellectual functioning was partially supported, as the groups differed significantly on Matrix Reasoning and nearly reached significance on finger tapping. In addition, the WASI full scale intelligence quotients as well as nonverbal quotients both significantly differentiated the conditions. The remaining individual WASI subtests (Block Design, Vocabulary, and Similarities) did not reach significance. This pattern of significance for the intellectual quotients but not for all of the individual subtests likely suggests that the -

intellectual quotients but not for all of the individual subtests likely suggests that the children were putting forth a consistent, yet subtle, pattern of underperforming across subtests. The Trail Making Test (Trails A & B) did not significantly differ between the groups.

The next hypothesis regarded the use of the Dot Counting Test for assessing effort in children; however, this measure was not found to differentiate the conditions. It was hypothesized that perhaps this was due to the younger children's performances erasing differences with the older children, as younger children are likely not as skilled at counting as older children and also may be more prone to count the grouped dots individually. However, controlling for age still did not yield significant differences between groups. As Blaskewitz et al (year) described, another possibility is that this task may be overly engaging to children, as it requires active participation which may interfere with the child's capacity to intentionally underperform. The computerized administration of this task also may have resulted in a tendency for children to break from their malingered presentations, as they were instructed to play a game with the examiner and perhaps would have approached this task differently had the examiner administered this test. Regardless, at this time it is not recommended that the Dot Counting Test be administered for the purpose of effort detection in children.

This study also sought to further explore the TOMM as a valid measure of effort with younger individuals, although the findings were inconclusive. Consistent with past research (Blaskewitz, et al, 2008; Constantinou, et al, 2003; Donders, 2005; Nagle et al, 2006), it is apparent that children, even as young as six, are capable of passing both the second trial and the retention trial of this test. However, problems with sensitivity clearly emerged with the treatment children, as slightly less than half of this treatment sample passed either the second trial or retention trial, and five of the children passed both trials. Due to this lack of consistency with the treatment condition, it is not recommended that the TOMM be utilized as a standalone assessment of effort in child populations. Rather, as is recommended with adults, the TOMM should be utilized as one of many potential sources that address effort, with attention on both multiple symptom validity measures as well as performances on standard cognitive tests (Larrabee, 2003). However, when considering the past findings of the TOMM (Blaskewitz, et al, 2008; Constantinou, et al, 2003; Donders, 2005; Nagle et al, 2006) with the present finding of all the control children passing both trial 2 and the retention trial, failure of this measure by children 6 and older should raise serious concerns regarding effort validity.

In general, the children appeared to have more difficulty lowering their effort on tasks when they were explicitly asked to do something and provided feedback regarding their performance (e.g., Trails being corrected, being queued regarding responses on Vocabulary). It is also possible that tasks that are more cognitively engaging may be more difficult for younger individuals to exhibit sub-optimal effort. As described by Blaskewitz et al., who also found this trend, children may have difficulty underperforming when the cognitive demands of the task exceeded their ability to effectively monitor their underperforming. For example, given the attention and visual scanning demands of the Trail Making Test, it is possible that the children may have reduced their underperforming in an effort to meet the high cognitive demands of this task. It is also worth noting that the two individual measures that significantly differentiated children, (TOMM and Matrix Reasoning) are both multiple choice format.

It is therefore possible that the reduction in concentrative demand that is inherent in multiple choice tests may lead to the greater sensitivity found with these measures, at least with children.

Block Design provided an interesting case, as this study was the first to examine this measure within the context of malingering in children and was the second most reported instrument by the second examiner in attempting to determine children's effort, and yet did not reach statistical significance. As with the Trail Making Test, it is certainly possible that the high cognitive demands of Block Design may have reduced children's capacity to perform poorly, resulting in the null hypothesis. However, what remains unclear is why this measure was so frequently identified as abnormal by examiners during the battery administration, yet failed to emerge as significant.

It is also worth noting that finger tapping nearly differentiated the two conditions. To the author's knowledge, this is the first investigation that has explored this test within the context of a simulation study of younger children. Given that finger tapping has received considerable support within the adult literature as an identifier of sub-reasonable effort (Heaton et al., 1978; Mittenberg, et al, 1996), the present finding warrants further inquiry in establishing more definitive guidelines regarding children's performances. One question that is immediately apparent based on the present findings is whether the expected average of underperforming in children is higher than adults or if the slightly higher average found in child simulators is simply an artifact of this small sample size.

In addition to these hypotheses, also explored was examiners accuracy in identifying children who are underperforming. Although not surprising given that the expected malingering base rate of children in the study was 50%, the examiners demonstrated a high degree of accuracy in identifying the children's condition, as they identified 88% of the children in the control condition and 76% of the treatment children. Interestingly, false positives were only found in children under the age of 10, while false negatives were only found in children 10 and above. This suggests that older children's malingered presentations were more convincing than their younger counterparts, while younger children may exhibit behaviors that may raise validity concerns even though they theoretically are putting forth reasonable effort. As with Blaskewitz et al., examiners also briefly interviewed the children regarding malingering strategies and found that older children were more likely to engage in more sophisticated strategy than younger children (e.g., providing incorrect responses intermittently). However, the majority of children had significant difficulty clearly articulating a faking strategy.

As with all simulation studies, this study lacks external validity due to there not being a true malingering sample (Bernard, 1990; Rogers et al, 1993). Simulation designs also have multiple potential confounds with regard to how individuals with a more substantial secondary gain may perform, as well as potential personality confounds that could lead one to be more likely to engage in deceptive behaviors (e.g., anti-social personality disorder, which has been linked to specific cognitive profiles; Dinn & Harris, 2000). Given that children's base rates for malingering on neuropsychological measures are not known, but are presumed to be substantially lower than adults, it would be exceedingly difficult to identify and effectively measure this behavior as it exists within the normal populous. Although not completely out of the realm of possibilities for a known-group design to be conducted, it is likely that these examinations will continue within the more controllable simulated design.

Future investigations may benefit from exploring the impact of more directive coaching on children's ability to realistically feign cognitive impairment. For example, it is certainly possible that within the context of pending litigation a child may receive more substantial explanation regarding strategies that could lead to a more believable cognitive performance, possibly even including information regarding specific instruments. This is particularly relevant given that recent studies have highlighted the increasing accessibility of the public to information pertaining to symptom validity measures (Ruiz, Drake, Glass, Marcotte, & van Gorp, 2002) and have found that adults evidence more sophisticated malingering when provided with information regarding symptom validity methods and techniques (Youngjohn, Lees-Haley, & Binder, 1999). As a result, it may be helpful for future investigations to provide children with a more thorough understanding of the test process, specifically with respect to symptom validity techniques. This could be accomplished by conducting a similar investigation that simply raises the children's awareness of effort detection methods. In addition, given that in this investigation children were encouraged by strangers to underperform, it may be worthwhile to explore how children would perform when instructed by someone who would be more likely to make this suggestion in the "real world" (e.g., parent), as this may increase the child's commitment to the fake bad role.

# CHAPTER VI

## TABLES

**Table 1 Demographics**

| Condition: | Control | | Treatment | |
|---|---|---|---|---|
| **Age** | ***X*** | **SD** | ***X*** | **SD** |
| | 10.06 | 1.83 | 9.71 | 2.20 |
| **Gender** | ***n*** | | ***n*** | |
| Boys | 5 | | 6 | |
| Girls | 12 | | 11 | |
| **Grade** | | | | |
| Kindergarten | 2 | | 5 | |
| First | 2 | | 1 | |
| Second | 2 | | 1 | |
| Third | 3 | | 1 | |
| Fourth | 3 | | 5 | |
| Fifth | 4 | | 3 | |
| Sixth | 1 | | 1 | |

| **Family Income** | | |
|---|---|---|
| 10,000 -25,000 | 1 | 1 |
| 25,000-50,000 | 3 | 2 |
| 50,000-75,000 | 7 | 8 |
| Over 75,000 | 6 | 6 |

**Table 2 Univariate Effects**

| | Measure | Significance |
|---|---|---|
| All Full Effort | | |
| | KBIT-2 Verbal | .366 |
| | KBIT-2 Nonverbal | .119 |
| | KBIT-2 Full Scale | .594 |
| Experiment | | |
| | WASI Verbal | .161 |
| | WASI Performance | .033 |
| | WASI Full Scale IQ | .022 |
| | Vocabulary | .210 |
| | Similarities | .130 |
| | Block Design | .284 |
| | Matrix Reasoning | .008 |
| | Finger Tapping Dominant | .110 |
| | Finger Tapping Non-dominant | .067 |
| | Finger Tapping Total | .074 |
| | Trails A | .427 |
| | Trails A errors | .227 |
| | Trails B | .239 |
| | Trails B errors | .520 |
| | TOMM Trial 1 | .001 |
| | TOMM Trial 2 | .001 |
| | TOMM Retention | .001 |
| | Dot Counting Overall | .768 |
| | Dot Counting Grouped | .727 |
| | Dot Counting Ungrouped | .978 |
| | Dot Counting Difference | .576 |
| | Dot Counting Counted Dots | .506 |

.

**Table 3 Test Performances**

| Condition: | Control | | Treatment | |
|---|---|---|---|---|
| **KBIT*** | **Standard Score** | **SD** | **Standard Score** | **SD** |
| Verbal IQ | 102.65 | 10.67 | 106.12 | 11.38 |
| Nonverbal IQ | 114.06 | 10.78 | 106.95 | 14.78 |
| Full Scale IQ | 109.94 | 10.48 | 107.76 | 12.98 |
| **WASI** | | | | |
| Verbal IQ | 105.72 | 13.92 | 97.35 | 17.22 |
| Performance IQ | 109.35 | 11.14 | 100.06 | 13.14 |
| Full Scale IQ | 108.17 | 11.29 | 97.94 | 13.44 |
| **WASI Subtests** | T | SD | T | SD |
| Vocabulary | 50.53 | 9.44 | 45.29 | 13.97 |
| Similarities | 55.88 | 9.19 | 50.35 | 11.43 |
| Block Design | 53.76 | 8.80 | 49.94 | 11.47 |
| Matrix Reasoning | 57.64 | 7.56 | 49.52 | 8.97 |
| **Symptom Validity Tests** | *X* | SD | *X* | SD |
| TOMM Trial 1 | 45.71 | 4.43 | 37.29 | 8.90 |
| TOMM Trial 2 | 49.76 | .75 | 38.12 | 12.96 |
| TOMM Retention | 49.71 | .47 | 35.35 | 15.64 |
| Dot Counting Total Time | .64 | .21 | .67 | .25 |
| Dot Counting Grouped | .49 | .19 | .52 | .26 |
| Dot Counting Ungrouped | .81 | .25 | .80 | .25 |
| Dot Counting Difference | .32 | .17 | .29 | .15 |

| | | | | |
|---|---|---|---|---|
| Dot Counting Total Dots | 431.64** | 118.34** | 411.06 | 43.76 |
| **Neuropsychological Measures** | *X* | SD | *X* | SD |
| Trails A Time | 53.82 | 20.26 | 63.00 | 42.44 |
| Trails A Errors | .24 | .44 | 1.00 | 2.52 |
| Trails B Time | 133.24 | 45.37 | 161.94 | 87.63 |
| Trails B Errors | .65 | 1.06 | .88 | 1.05 |
| Finger Tapping D | 39.73 | 5.65 | 35.67 | 8.45 |
| Finger Tapping ND | 37.16 | 6.19 | 32.48 | 8.08 |
| Finger Tapping Total | 76.89 | 11.19 | 68.15 | 16.01 |

**all at full effort, **including outlier*

REFERENCES

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

Allen, L. M., Conder R. L., Green, P., & Cox, D. R. (1997). *CARB 97: Manual for the computerized assessment of response bias.* Durham, NC: CogniSyst.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition, Text Revision (DSM-IV-TR).* Washington D.C.: American Psychiatric Association.

American Psychological Association (2002). Ethical principles of psychologists and codes of conduct. *American Psychologist, 47,* 1060-1073. Bernard, L. C. (1990). Prospectives of faking believable memory deficits on neuropsychological tests and the use of incentives in simulation research. *Journal of Clinical and Experimental Neuropsychology, 12,* 715-728.

Bernard, L. C., McGrath, M. J., & Houston, W. (1996). The differential effects of simulating malingering, closed head injury, and other CNS pathology on the

Wisconsin Card Sorting Test: Support for the "Pattern of Performance" hypothesis. *Archives of Clinical Neuropsychology, 11*, 231-245.

Bigler, E. D. (1990). Neuropsychology and malingering: Comment on Faust, Hart, and Guilmette (1988). *Journal of Consulting and Clinical Psychology, 58,* 244-247.

Blaskewitz, N., Merten, T., & Kathmann, N. (2008). Performance of children on symptom validity tests: TOMM, MSVT, and FIT. *Archives of Clinical Neuropsychology, 23,* 379-391.

Carroll, S., Abrahamse, A., & Vaiana, M. (1995). *The costs for excess medical claims for automobile personal injuries.* Santa Monica, CA: Random House.

Chafetz, M., & Abrahams, J. (2005). *Green's MACT helps identify internal predictors of effort in the Social Security disability exam*. Paper presented at the annual meeting of the Annual Committee of Neuropsychology, Tampa, FL.

Chestermann, L. P., Terbeck, S., & Vaughan, F. (2008). Malingered Psychosis. *The Journal of Forensic Psychiatry and Psychology, 19,* 275-300.

Constantinou, M., & McCaffrey, R. J. (2003). Using the TOMM for evaluating children's effort to perform optimally on neuropsychological measures. *Child Neuropsychology, 9,* 81-90.

Courtney, J. C., Dinkins, J. P., Allen, L. M., & Kuroski, K. (2003). Age related effects in children taking the Computerized Assessment of Response Bias and Word Memory Test. *Child Neuropsychology, 9*, 109-116.

Dinn, W. & Harris, C. (2000). Neurocognitive functioning in antisocial personality disorder. *Psychiatry Research*, *97*, 173-190.

Donders, J. (2005). Performance on the Test of Memory Malingering in a mixed pediatric sample. *Child Neuropsychology, 11,* 221-227.

Faust, D., Hart, K, & Guilmette, T. J. (1998). Pediatric malingering: The capacity of children to fake believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 56,* 578-582.

Faust, D., Hart, K., Guilmette, T. J., & Arkes, H. R. (1988). Neuropsychologists' capacity to detect adolescent malingerers. *Professional Psychology: Research and Practice, 19,* 508-515.

Flaro, L., Green, P., Allen. L. M. (2000). Symptom validity test results with children: CARB and WMT. *Archives of Clinical Neuropsychology, 15,* 840.

Geller J. L., Erlen J., Kaye N. S., Fisher W. H.(1990). Feigned insanity in nineteenth century America: Tactics, trials and truth. *Behavioral Science and the Law, 8,* 2-26.

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

Green, P. & Flaro, L. (2003). Word Memory Test performance in children. *Child Neuropsychology, 9,* 189-207.

Green, P., & Iverson, G. (2001). Effects of injury severity and cognitive exaggeration on olfactory deficits in head injury compensation claims, *NeuroRehabilitation, 16,* 237–243.

Green, P., Rohling, M. L. Lees-Haley, P. R., & Allen, L. M. (2001). Effort has greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury, 15,* 1045-1060.

Griffenstein, M. F., Baker, W. J., & Gola, T. (1996). Motor dysfunction profiles in traumatic brain injury and post-concussion syndrome. *Journal of the International Neuropsychological Society*, 2, 477-485.

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

Gudjonsson, G., & Shackleton, H. (1986). The pattern of scores on Raven's Matrices during "faking bad" and "non-faking" performance. *British Journal of Clinical Psychology, 25,* 35-41.

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

Halstead, W.C. (1947). *Brain and Intelligence.* Chicago: University of Chicago Press.

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

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

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

Iverson, G. L. (2003). Detecting malingering in civil forensic evaluations. In A. M. Horton & L. C. Harlage (Eds.), *Handbook of forensic neuropsychology* (pp.137-177). New York: Springer.

Kaufman, A. S. & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test, 2nd edition (KBIT-2)*. Circle Pines, MN: American Guidance Service.

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

Larrabee, G. J. (2007). *Assessment of Malingered Neuropsychological Deficits*. USA: Oxford University Press.

Lezak, M.D., Howieson, D.B., & Loring, D. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.

McCann, J. T. (1998). *Malingering and deception in adolescents: Assessing credibility in clinical and forensic settings.* Washington DC: American Psychological Association.

McKinzey, R. K., Prieler, J., & Raven, J. (2003). Detection of children's malingering on Raven's Standard Progressive Matrices. *British Journal of Clinical Psychology, 42,* 95-99.

*Merriam-Webster collegiate dictionary* (10th ed.). (2002). Springfield, MA: Merriam-Webster.

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

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

Mittenberg, W., Rotholic, A., Russell, E. & Heilbronner, R. (1996). Identification of malingered head injury on the Halstead-Reitan Battery. *Archives of Clinical Neuropsychology, 11,* 271-281.

Mittenberg, W., Arin, R., Millsaps, C. & Heilbronner, R. (1993). Identification of malingered head injury on the Wechsler Memory Scale – Revised. *Psychological Assessment, 5,* 34-40.

Nagle, A. M., Everhart, D. E., Durham, T. W., McCammon, S. L., & Walker, M. (2006). Deception Strategies in Children: Examination of forced choice recognition and verbal learning and memory techniques. *Archives of Clinical Neuropsychology, 21,* 777-785.

Newton, P., Reddy, V., Bull, R. (2000). Children's everyday deception and performance on false-beliefs tasks. *British Journal of Developmental Psychology, 18,* 297-317.

Nichols, E., Harrison, A. G., McCloskey, L., Weintraub, L. (2002). *Learning opportunities task force 1997–2002: Final report*. Richmond Hill: Learning Opportunities Task Force.

Polak, A., & Harris, P. L. (1999). Deception in young children following noncompliance. *Developmental Psychology, 35*, 561-568.

Psychological Corporation (1999). *Wechsler Abbreviated Scale of Intelligence (WASI).* San Antonio, TX: The Psychological Corporation.

Psychological Corporation (1999). *Manual for the Wechsler Abbreviated Scale of Intelligence.* San Antonio, TX: The Psychological Corporation.

Rey, A. (1941). *L'examen clinique en psychologie.* Paris: Presses Universitaires de France.

Rogers, R., & Cavanaugh, J. L. (1983). "Nothing but the truth"…A re-examination of malingering. *Journal of Psychiatry and Law, 11,* 443-460.

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

Ruiz, M., Drake, E., Glass, A., Marcotte, D., & van Gorp, W. (2002). Trying to beat the system: Misuse of the internet to assist in avoiding the detection of psychological symptom dissimulation. *Professional Psychology: Research and Practice, 33,* 294-299.

Ruffman, T., Olsen, D. R., Ash, T., & Keenan, T. (1993). The ABC's of deception: Do young children understand deception in the same way as adults? *Developmental Psychology, 29,* 74-87.

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

Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Psychology, 19,* 468-483.

Sodian, D., Taylor, C. Harris, P. L., & Perner, J. (1991). Early deception and the children theory of mind: False traits and genuine markers. *Child Development, 62,* 468-483.

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

Suhr, J. A., & Boyer, D. (1999). Use of the Wisconsin Card Sorting Test in the detection of malingering in student simulator and patient samples. *Journal of Clinical and Experimental Neuropsychology, 21*, 701-708.

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM).* North Tonawanda, NY: Multi-Health Systems.

Whitney, K., Hook, J., Steiner, A., Shepard, P, & Callaway, S. (2008). Is the Rey 15-Item Memory Test II (Rey II) a Valid Symptom Validity Test? Comparison with the TOMM. *Applied Neuropsychology, 15,* 287-292.

Youngjohn, J. R. (1995). Confirmed attorney coaching prior to neuropsychological evaluation. A*ssessment, 2*, 279-283.

Youngjohn, J. R., Lees-Haley, P. R., & Binder, L. M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology, 14,* 511-515.

# APPPENDICES

## Appendix A

**Oklahoma State University Institutional Review Board**

| | | | |
|---|---|---|---|
| Date | Tuesday, April 26, 2011 | **Protocol Expires:** | **5/11/2012** |
| IRB Application No: | AS1042 | | |
| Proposal Title: | Effort Testing in Children on Neuropsychological and Symptom Validity Measures | | |
| Reviewed and Processed as: | Full Board<br>**Continuation** | | |

Status Recommended by Reviewer(s) **Approved**

Principal Investigator(s)

| | | |
|---|---|---|
| Phillip L. Rambo<br>391 Mountain Rd<br>Cheshire, CT 06410 | Jennifer L. Callahan<br>Univ. of North Texas<br>Denton, TX 76203 | Larry L. Mullins<br>116 North Murray<br>Stillwater, OK 74078 |

---

Approvals are valid for one calendar year, after which time a request for continuation must be submitted. Any modifications to the research project approved by the IRB must be submitted for approval with the advisor's signature. The IRB office MUST be notified in writing when a project is complete. Approved projects are subject to monitoring by the IRB. Expedited and exempt projects may be reviewed by the full Institutional Review Board.

X The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

The reviewer(s) had these comments:

Should additional data collected at University of North Texas be needed after August 11, 2011, a modification request including a continuation approval from UNT IRB must be submitted for OSU IRB review and approval before such data can be included.

Signature :

Shelia Kennison, Chair, Institutional Review Board

Tuesday, April 26, 2011
Date

Appendix B

UNT
UNIVERSITY OF
NORTH TEXAS
Discover the power of ideas.

OFFICE OF THE VICE PRESIDENT FOR RESEARCH AND ECONOMIC DEVELOPMENT
Research Services

August 5, 2010

Jennifer Callahan
Department of Psychology
University of North Texas

Re: Human Subjects Application No. 10219

Dear Dr. Callahan:

As permitted by federal law and regulations governing the use of human subjects in research projects (45 CFR 46), the UNT Institutional Review Board has reviewed your proposed project titled "Effort Testing in Children on Malingering and Neuropsychological Measures." The risks inherent in this research are minimal, and the potential benefits to the subject outweigh those risks. The submitted protocol is hereby approved for the use of human subjects in this study. **Federal Policy 45 CFR 46.109(e) stipulates that IRB approval is for one year only, August 5, 2010 to August 4, 2011.**

Enclosed is the consent document with stamped IRB approval. Please copy and use this form only for your study subjects.

It is your responsibility according to U.S. Department of Health and Human Services regulations to submit annual and terminal progress reports to the IRB for this project. The IRB must also review this project prior to any modifications.

Please contact Shelia Bourns, Research Compliance Administrator, or Boyd Herndon, Director of Research Compliance, at extension 3940, if you wish to make changes or need additional information.

Sincerely,

P.L. Kaminski, PhD

Patricia L. Kaminski, Ph.D.
Associate Professor
Chair, Institutional Review Board

PK:sb

1155 Union Circle #305250 | Denton, Texas 76203-5017 | TEL 940.565.3940 | FAX 940.565.4277
TTY 940.369.8652 | research.unt.edu

# VITA

Philip Louis Rambo

Candidate for the Degree of

Doctor of Philosophy

Thesis: EFFORT TESTING IN CHILDREN ON NEUROPSYCHOLOGICAL AND SYMPTOM VALIDITY MEASURES

Major Field: Clinical Psychology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in clinical psychology at Oklahoma State University, Stillwater, Oklahoma in December, 2011.

Completed the requirements for the Master of Science in clinical psychology at Oklahoma State University, Stillwater, Oklahoma in 2008.

Completed the requirements for the Bachelor of Science in psychology at Texas A&M University, College Station, Texas in 2002.

Experience:

Philip has research experience in a wide range of settings, including, but not limited to, neuropsychological assessment of children, psychophysiology, and pediatric psychology.

Professional Memberships:

Philip is a member of the American Psychological Association. He is also a member of the Golden Key Honor Society, as well as Phi Kappa Phi, and has served in multiple department and graduate student organizations.

Name: Philip Louis Rambo Date of Degree: December, 2011

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: Effort Testing in Children on Neuropsychological and Symptom Validity Measures

Pages in Study: 69 Candidate for the Degree of Doctor of Philosophy

Major Field: Clinical Psychology

Scope and Method of Study: Recent efforts have contributed to significant advances in the detection of malingered performances in adults during cognitive testing. However, children's ability to purposefully underperform has received relatively little attention. The purpose of the present investigation was to examine children's performances on some of the more common intellectual and neuropsychological measures, as well as two symptom validity measures: the Test of Memory Malingering (TOMM) and the Dot Counting Test. This was accomplished through the administration of measures to children ages 6-12 in a full effort control condition and a simulated poor effort treatment condition. In addition, prior to randomization, children's general intellectual functioning was estimated via administration of the Kaufman Brief Intellectual Battery, 2nd edition (KBIT-2).

Findings and Conclusions: Univariate analyses revealed that the control group and treatment group differed significantly on some, but not all measures. Also, children's estimated IQ's in the treatment condition significantly differed from the IQ's obtained in the full effort condition and from the full effort IQ's obtained from the treatment children on the KBIT-2, when they were putting forth full effort. These findings suggest that children are fully capable of willfully underperforming during cognitive testing; however, consistent with prior investigations, some measures evidence greater sensitivity in identifying poor effort in younger populations. For instance, WASI Matrix Reasoning and the TOMM were significantly poorer in the treatment condition. In addition, fingertapping also evidenced promise as a test of effort, as both non-dominant tapping and combined tapping performances approached significance in differentiating the two conditions ($p < .10$).

ADVISER'S APPROVAL: Larry L. Mullins