5-2020

Exploring the Prevalence of Suboptimal Effort Among Children and Adolescents on Psychoeducational Evaluations

Kayla Jo Singleton

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

EXPLORING THE PREVALENCE OF SUBOPTIMAL EFFORT AMONG CHILDREN AND ADOLESCENTS ON PSYCHOEDUCATIONAL EVALUATIONS

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Kayla Jo Singleton

College of Educational and Behavioral Sciences
Department of School Psychology
School Psychology

May 2020

This Dissertation by: Kayla Jo Singleton

Entitled: *Exploring the Prevalence of Suboptimal Effort Among Children and Adolescents on Psychoeducational Evaluations*

has been approved as meeting the requirement for the Degree of Doctor of Philosophy in College of Education and Behavioral Sciences in Department of School Psychology, Program of School Psychology.

Accepted by the Doctoral Committee

_____
Robyn S. Hess, Ph.D., ABPP, Research Advisor

_____
Michelle Athanasiou, Ph.D., BCBA-D, Committee Member

_____
Thom Dunn, Ph.D., Committee Member

_____
Corey D. Pierce, Ph.D., Faculty Representative

Date of Dissertation Defense: March 31st, 2020

Accepted by the Graduate School

_____

Cindy Wesley
Interim Associate Provost and Dean
The Graduate School and International Admissions

ABSTRACT

Singleton, Kayla. *Exploring the Prevalence of Suboptimal Effort among Children and Adolescents on Psychoeducational Evaluations*. Published Doctor of Philosophy, dissertation, University of Northern Colorado, 2020.

This study represents one of the first known studies to explore suboptimal effort in children and adolescents as part of psychoeducational evaluations conducted within a school setting. Only recently has attention been given to pediatric performance validity testing. With the assistance of five credentialed school psychologists across two midwestern states, 52 students were administered the Test of Memory Malingering (TOMM) as part of their psychoeducational evaluation. The findings of the current study suggested that 19.2% of these students failed Trial 2 on the TOMM, a suggested indicator of suboptimal performance. Furthermore, school psychologists' ratings of observed effort did not correlate with failure on Trial 2 of the TOMM and there were no discernible patterns across disability area. Full scale ability scores provided a good predictor of performance on the TOMM. Overall, the findings from this study suggest the importance of including an objective performance validity measure for school psychologists in order to improve their ability to identify students who might be demonstrating suboptimal performance. Additional implications for practice and research are provided.

ACKNOWLEDGEMENTS

First and foremost, the greatest thank you goes to my husband, Jake Singleton, for his continuous support throughout my graduate training programs. The seemingly endless moves and sacrifices you have made to help to get to this point have not gone unnoticed. Through the highs and lows, you have pushed me to persevere and have ridden on this rollercoaster of life with me. I will forever be thankful for you. Thank you for believing in me, especially when I could not believe in myself.

Second, I would like to thank my parents for instilling the importance of education and providing me with an exceptional educational foundation and experience. My love for learning ultimately led me to seek and obtain a doctoral degree because of you. Thank you for your encouragement and love throughout this journey.

To my dissertation chair, Dr. Robyn Hess, and my dissertation committee, I am appreciative of your ongoing support and guidance. Thank you for running with my dissertation topic and seeing it through with me. I am beyond grateful for the knowledge, reassurance, and dedication you have shared with me. All of you have influenced me to become a stronger student, practitioner, and person. Thank you!

Finally, without Dr. Michael Kirkwood this dissertation would not exist. Dr. Kirkwood not only introduced me to pediatric validity testing, but also provided numerous recommendations and references to consider as part of this study. His expertise and commitment to student growth is extraordinary. Thank you, Dr. Kirkwood!

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

According to the U.S. Department of Education, during the 2018-2019 school year, approximately 7,130,238 students, ages 3 to 21, were served under the Individuals with Disabilities Education Act (IDEA; Part B) (U.S. Department of Education, 2019). This legislation provides multiple safeguards, policies, and procedures to protect the rights of students and their families. Further, each state department of education has developed its own policies and practices to guide practitioners in the identification of students with disabilities. These efforts are designed to facilitate the accurate identification of students who need additional academic support. Also, test developers undertake careful studies to ensure that their products yield valid and reliable scores. Despite the importance of these processes and the careful guidelines developed to assist with implementation, little has been done to ensure that the results of the psychoeducational evaluation reflect valid effort on the part of the test taker.

The first legislation that specifically outlined the responsibility of schools to identify, educate, and protect children with disabilities was originally known as the Education for All Handicapped Children Act of 1975 (Lechtenberger, 2010), and part of its content directed that federal funding would be provided to all states to educate individuals identified as having a disability (Weiss & Mattrick, 2010). Additionally, this Act outlined six major outcomes including (1) all students were entitled to a free and appropriate public education (FAPE), (2) students who were identified and evaluated for

special education services would not be discriminated against (e.g., assessed in their primary language), (3) students identified with disabilities would receive necessary supports through an Individualized Education Program (IEP), (4) students with disabilities would receive such services in the least restrictive environment (LRE), (5) due process was established to protect student and family rights throughout this process, and (6) family involvement was expected in the identification, evaluation, and IEP process (Lechtenberger, 2010). The Education for All Handicapped Children Act of 1975 has been reauthorized several times, including in 2004, when this law underwent several changes and was renamed as the Individuals with Disabilities Education Improvement Act (IDEA) (Weiss & Mattrick, 2010; Wright & Wright, 2018). Under IDEA, there are 13 categories that students with disabilities can be classified under in an educational setting.

There are some safeguards in place that address validity such as ensuring that a student has sufficient English language skills to participate in an assessment or that an appropriate alternative (e.g., use of a nonverbal test, use of a translated test with a bilingual examiner) is used. Additionally, the student should have had adequate educational opportunities. Finally, guidance is provided regarding the use of tests that have sound psychometric properties and are being used for their intended purpose. However, little has been written about how a student's effort might impact the validity of a psychoeducational assessment. Kirkwood, Yeates, Randolph, and Kirk (2012) noted that in order to be able to accurately and confidently interpret the results of any evaluation, the clinician must first be confident that the examinee has put forth adequate effort. Without optimal effort, the results potentially do not capture the true ability level

of the examinee (Kirkwood, Kirk, Blaha, & Wilson, 2010; Kirkwood et al., 2012). Therefore, more information is needed regarding children's effort during psychoeducational assessments in school settings to better understand the degree to which they are providing full effort.

Significance of the Problem

Each year, thousands of school children are assessed to determine whether an educational disability might be interfering with their ability to be successful in the classroom. Typically, school-based examiners tend to use their observations of apparent effort, attentiveness, and overall attitude to make a judgment about the validity of the assessment results. Without the inclusion of an objective measure of performance validity as part of a psychoeducational evaluation, there is no way to determine whether the student is putting forth optimal effort. Suboptimal effort could result in lower assessment scores and potentially lead a school-based multidisciplinary team (MDT) to incorrectly qualify a student for special education under one of the designated eligibility categories under IDEA. This false positive outcome is problematic for several reasons.

In order for students to receive special education services, they must first be found to meet eligibility criteria for a disability under IDEA. Without an objective measure of effort, students could be labelled with a disability that they do not actually have. Educational labels indicating disability can be stigmatizing (Ho, 2004) and misidentifying a disability may cause unnecessary harm. When determining a student's eligibility for special education services, the multidisciplinary team (MDT) must find that the benefits of special education outweigh the potential risks or harm of this placement.

Students could be unnecessarily labeled as a student with a disability without the incorporation of a performance validity measure.

There is a precarious balance between the burden and the value of identifying students with educational disabilities. Based on the number of students who are identified for special education services, schools receive an additional amount of federal funding to offset the costs of additional services. Unfortunately, the reimbursement rate from the federal government is not sufficient to cover the actual costs. By incorrectly identifying a student for special education, an unnecessary financial burden is accrued by the district.

Finally, students who are identified as eligible for special education could receive nonessential services or advantages. So much has been written about the disadvantages of special education (e.g., exclusion from one's peers, stigma), it might be hard to conceptualize the potential gains. On a day-to-day basis, students with disabilities are afforded additional time to complete their tests, allowed to complete their tests in a quiet environment, receive additional supports such as an extra study time, audio recordings of their texts, and instructor notes prior to class as well as other accommodations that are essential for helping them access needed course materials. Furthermore, when considering potential gains for suboptimal performance, secondary education students might be seeking additional time on ACT or SAT tests or seeking access to medications for their own use or to sell to others (Kirkwood, 2015a). These instances may provide motive for students to perform poorly for their own secondary gain.

Unfortunately, parents may also play a role in coaching their children towards suboptimal performance. If children are found to have an intellectual disability or some other serious disability, families may be found eligible for Medicaid funds. That is,

children who demonstrate suboptimal performance may not be acting on their own but complying with the direction received from their families. Lu and Boone (2002) published the first case study of malingering by proxy in a nine-year-old male who had sustained a traumatic brain injury as a result of a car accident. He was found to have failed all four of the performance validity tests administered, leading the examiners to question his effort. Therefore, developing a better understanding of how suboptimal performance or effort affects children and adolescent outcomes in psychoeducational evaluations within school settings is an important area for research.

Theoretical Basis

From a developmental standpoint, the idea of malingering and suboptimal performance has often been dismissed in child and adolescent populations because it was assumed that these individuals could not and would not exaggerate symptoms or "fake bad" during evaluations (Kirkwood, 2015b). Even children at a very young age are able to tell white lies (Talwar & Lee, 2002) and to make false statements (Ahern, Lyon, & Quas, 2011). However, the rationale as to why children might lie is not clearly understood.

From a moral development perspective, Kohlberg conceptualized moral development that spans from preschool-aged children to adulthood through three levels and six stages (Gibbs, 2014). Across the different levels and stages, the underlying reason for lying shifts to meet the developmental needs of the individual (Gibbs, 2014). For example, a school-aged child might engage in lying or false statements to fit in to a social group (conform) and/or maintain friendships with peers.

The truth of the matter is that the reasons behind children and adolescent suboptimal performance is largely unknown and likely varies on a case to case basis. Children and adolescents might not even recognize that they are not putting forth their best effort due to social emotional symptomology, physical symptoms, or the influence of their parents/caregivers (malingering by proxy) (Kirkwood, 2015b; Kirkwood et al., 2010; Lu & Boone, 2002). Although the reasoning behind suboptimal performance is unclear, ensuring accurate diagnosis, placement, and interventions is critical. Therefore, it seems that an important first step is to identify children and adolescents who are displaying suboptimal performance, as a safeguard to protect against possible interpretation of inaccurate information and the consequences of such.

Relevant Research

In recent years, it has been established that children and adolescents are capable of displaying suboptimal performance or effort during neuropsychological evaluations (Kirkwood & Kirk, 2010), with at least 17% (33 out of 193 participants) failing at least one measure of effort in an outpatient clinical setting and 7% of participants failing two performance validity measures in a psychiatric inpatient setting (Weber Ku et al., 2020). In a case series published by Kirkwood, Kirk, Blaha, and Wilson (2010) six cases of pediatric suboptimal performance were identified. Potential reasons for suboptimal performance that were considered among these patients included “attempts to get out of schoolwork” or an effort “to change a family or social situation” (Kirkwood et al., 2010, p. 607).

The scientific research base for suboptimal performance has established that children and adolescents are not only capable but do engage in meaningful rates of

suboptimal performance during neuropsychological evaluations (Kirkwood & Kirk, 2010). However, there is another important setting in which children and adolescents are routinely evaluated to determine whether additional resources are necessary to support students' inclusion within the general education setting. When students are struggling academically and do not respond to interventions designed to address their needs, they are often referred for a psychoeducational evaluation by a school psychologist to determine eligibility for special education. Similar to a neuropsychological evaluation, students are assessed in a number of areas including cognition, reasoning, memory, processing speed, and executive functioning. Therefore, given the similar types of measures and the likelihood of some gain (e.g., additional academic supports), it stands to reason that suboptimal performance may occur in school settings.

In most states, diagnoses according to the DSM-5 (APA, 2013) are not given in a school setting. Instead, students who are identified as qualifying for special education through the psychoeducational evaluation are classified with an educational disability in order to receive services. Unlike the motivations that adults might have to demonstrate poor effort and malingering, (e.g., monetary gain), children and adolescent may receive secondary gain such as additional academic help and support. Furthermore, with the implementation of many high stakes tests, students may experience additional pressure to gain any advantage possible. As noted, some children and adolescents might be influenced by their parent or caregiver to perform poorly on psychoeducational evaluations, which is referred to as "malingering by proxy" (Kirkwood, 2015b).

Despite a thorough search of studies published in English, the primary researcher was unable to locate any previous studies examining the number of students displaying

suboptimal performance on psychoeducational evaluations within a school setting. Given the findings within clinical settings (e.g., Kirkwood, 2015a) and the potential secondary gains that might motivate students and their parents, it stands to reason that there may be some students demonstrating suboptimal performance on their psychoeducational evaluations. The use of performance validity tests within a school setting with child and adolescent populations may yield important information regarding the prevalence of this phenomena within student populations.

Problem Statement

Practitioners often believe that they would be able to spot suboptimal performance. In fact, many graduate training programs instruct students to make a statement to the degree they believe the results are valid. However, previous research has suggested that subjective observations are inadequate at identifying children and adolescents displaying suboptimal performance (Faust, Hart, & Guilmette, 1988). Without incorporating a performance validity measure as part of a psychoeducational evaluation, there may be students who are being falsely identified as meeting criteria for special education or for Medicaid funding. It is impossible to predict the number of students who may be displaying suboptimal performance within educational settings without directly assessing for this possibility. With this information, school districts may be able to develop better criteria for identifying (or at least flagging) potential cases of suboptimal performance. The results of this study may serve to encourage school psychology practitioners to add another tool to their assessment battery that helps them, and their team members, to make the best decisions regarding the needs of students.

Research Questions

The purpose of this study was to identify rates of suboptimal performance among child and adolescent populations participating in psychoeducational evaluations in their school districts. Using quantitative methodology, the researcher asked school psychology practitioners to incorporate the Test of Memory Malingering (TOMM; Tombaugh, 1996) into their assessment batteries when conducting a special education evaluation. The practitioners then provided the age, sex, assessment results (including both Trials of the TOMM), and the practitioners' ratings of student effort. The following research questions were addressed:

Q1 What is the percentage of students who display suboptimal performance (45>) on Trial 2 of the TOMM during their psychoeducational evaluation conducted within a school setting?

Q2 What percentage of agreement exists between perceived student effort, as measured by practitioner observations and objective student effort, as measured by a stand-alone performance validity measure (TOMM Trial 2)?

Based on the number of students identified in this first question, additional follow up questions were posed.

Q3 Do differences exist among the different special education categories for which participants are identified (e.g., specific learning disability, traumatic brain injury, emotional disturbance) with regard to suboptimal performance?

Q4 Does the performance of students displaying suboptimal performance indicate more variable performance across all assessments (e.g., IQ and academic achievement) than students who are believed to be displaying optimal effort?

Definitions

Malingering: "the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military

duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" (DSM-5; APA, 2013 p. 726).

Performance Validity: "refers to the validity of actual ability task performance, assessed either by stand-alone tests such as Dot Counting or by atypical performance on neuropsychological tests such as Finger Tapping" (Larrabee, 2012, p. 626).

Sensitivity: "refers to the true positive (Hit) rate for a test" (Chafetz, Abrahams, & Kohlmaier, 2007, p. 9).

Specificity: "is the true negative rate" (Chafetz et al., 2007, p. 9).

Suboptimal Performance: "instance of an examinee not performing to the best of his or her ability as directed on tests" (Strauss, Sherman, & Spreen, 2006, p. 1145).

Symptom Validity: "refers to the accuracy of symptomatic complaint on self-report measures such as the MMPI-2" (Larrabee, 2012, p. 626).

Summary

Given that a number of psychoeducational evaluations are performed each year within school settings, and the identified occurrence of suboptimal performance by children in clinic settings, the purpose of this study was to identify rates of suboptimal performance on school-based evaluations. The potential for misidentification of students for special education services, based on suboptimal performance on school-based evaluations is likely and may carry negative consequences such as stigma, unnecessary accommodations, and added costs to district budgets. Historically, practitioners have displayed difficulties detecting suboptimal performance in youth adding to the potential for misidentification of students. Examining the rate at which suboptimal performance

occurs during psychoeducational evaluations has the potential to raise awareness of this possibility and allow practitioners to adopt strategies that reduce this potential.

CHAPTER II

REVIEW OF LITERATURE

The importance of detecting and identifying individuals who display malingering, exaggerated symptoms, and suboptimal effort has been a long-standing practice within the field of psychology. Historically, adults have been the primary population with whom these symptoms have been acknowledged and researched. The American Psychiatric Association (APA) first introduced malingering in the DSM-III and subsequently in the DSM-IV and DSM-5 under V65.2 Code and defined as "the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" (APA, 1980, 1994, 2013, p. 726). Malingering was not well-defined and exclusively directed towards an adult population when first released in the DSM-III.

Since its inclusion in the DSM-III, attempts have been made to better understand and differentiate malingering from other disorders that might have common elements. For example, two disorders that are commonly differentiated from malingering include Factitious Disorder (FD) and Conversion Disorder (CD). The distinguishing indicators are described in four ambiguous considerations when malingering is suspected (APA, 1980, 1994, 2013). While Factitious Disorder might also involve an exaggeration of symptoms, it differs from malingering, in that malingering is thought to be fueled by secondary gain. Secondary gain is not typically thought of as part of a FD diagnosis. Conversion Disorder might present as significant symptoms that cannot be medically

explained, such as seizures, numbness, or pain. These symptoms are thought to represent an involuntary expression, and thus differ from malingering, which is thought to be a conscious and voluntary exaggeration of symptoms. In fact, according to the DSM-5, "malingering should be strongly suspected if any combination of the following is noted: (1) medicolegal context of presentation (e.g., the person is referred by an attorney to the clinician for examination, or the individual self-refers while litigation or criminal charges are pending), (2) marked discrepancy between the individual's claimed stress or disability and the objective findings and observations, (3) lack of cooperation during the diagnostic evaluation and in complying with the prescribed treatment regimen, (4) the presence of antisocial personality disorder" (APA, 2013, p.727). Although these four suggested areas of concern help clinicians to be alert for malingering, there is still a gap in conceptualizing what malingering might look like as part of a psychological or neuropsychological evaluation.

In an attempt to provide guidance for clinicians, Slick, Sherman, and Iverson (1999) attempted to delineate diagnostic criteria for identifying malingering. They proposed a more detailed definition of Malingering of Neurocognitive Dysfunction (MND), as well as three independent categories including, definite, probable, and possible MND. Although their definition is comparable to that of the DSM, Slick et al. provided additional information within their definition. As such, Malingering of Neurocognitive Dysfunction was defined as,

> the volitional exaggeration or fabrication of cognitive dysfunction for the purpose of obtaining substantial material gain or avoiding or escaping formal duty or responsibility. Substantial material gain includes money, goods, or services of nontrivial value (e.g., financial compensation for personal injury). Formal duties are actions that people are legally obligated to perform (e.g., prison, military, or public service, or child support payments or other financial obligations). Formal

> responsibilities are those that involve accountability or liability in legal proceedings (e.g., competency to stand trial) (Slick et al., 1999, p. 552).

Furthermore, Slick et al. (1999) introduced the term, definite MND, which was defined as, "the presence of clear and compelling evidence of volitional exaggeration or fabrication of cognitive dysfunction and the absence of plausible alternative explanation" (p.552). Probable MND was defined as, "the presence of evidence strongly suggesting volitional exaggeration or fabrication of cognitive dysfunction and the absence of plausible alternative explanations" (Slick et al., 1999, p. 552), and possible MND was considered to be similar to both definite and probable MND, except that the individual may have other potential etiologies that could not be ruled out (Slick et al., 1999). Furthermore, Slick et al. (1999) created four overarching areas of qualifying criteria which included factors such as external incentive, performance across neuropsychological measures, self-reported symptoms, and the inability to better explain these symptoms within another psychological disorder.

In addition to the detailed explanations of the diagnostic criteria, Slick et al. (1999) proposed that clinicians incorporate supplemental considerations (e.g., informed consent, differential diagnosis, reliability, validity, standardized administration of diagnostic measures, individual differences, prior patient behavior, clinical judgement, and self-reported symptoms). If clinicians engaged in reflection in each of these areas, it was believed they could build a stronger information base for determining valid or invalid test performance throughout the evaluation. For example, Slick et al. explained that while providing consent for an evaluation, the examiner can specifically state that optimal effort throughout the evaluation is required to best capture the examinee's true

abilities. If suboptimal performance was detected, then an indication of such performance would be included within the results section of the report.

On the other hand, it is possible for an examinee to "pass" a performance validity measure and still demonstrate poor performance on other measures administered during the evaluation. Performance validity measures are designed to appear difficult, when in fact, they are very simple. Unfortunately, these measures only capture a moment in time and are not fluid throughout the evaluation. A chance that examinees can be misidentified still exists. Therefore, as with any other decision, no one instrument should be used in isolation. Clinical judgement should be incorporated alongside a performance validity measure, when deciding on suspected malingering (Slick et al., 1999). For instance, if different assessments are used to measure the exact same skill set and the examinee has vastly different performance across measures, the examiner would likely use observations, clinical judgment, consideration of different aspects of the test itself and when it was administered (e.g., at the end of a long assessment session) to determine whether or not the examinee was putting forth full effort on these measures. Likewise, an examiner might have previously evaluated the examinee and a comparison of the examinees' current and past behaviors might provide insight as to their effort level.

Malingering in Adults

As the definition of malingering evolved and additional criteria were proposed to aid in the detection of malingering, questions about the prevalence of this behavior emerged. In 2002, Mittenberg, Patton, Canyock, and Condit published a seminal article detailing probable malingering or symptom exaggeration among adults with differing diagnoses by setting, referral type, and litigating or compensation seeking cases. Data

were collected via survey, from 131 active clinicians, who were members of the American Board of Clinical Neuropsychology (ABCN) (Mittenberg et al., 2002). Participants were asked, “What percentage of your annual cases in each category involve probable symptom exaggeration or malingering (% of personal injury, % of disability or worker’s compensation cases, % of criminal cases, % of medical or psychiatric cases not involved in litigation or seeking compensation)” (Mittenberg et al., 2002; p. 1101). Based on annual cases, within the United States and Canada, civil cases had the highest mean at 29.85% and among different types of referrals, disability or worker’s compensation cases were estimated at 30.12%. Across litigating or compensation cases, malingering among those claiming mild head injury was estimated at 38.50%. The results of this study suggested that malingering was a factor in nearly a third of adults when there was something to be gained by having a disability (e.g., litigation, worker’s compensation).

Given these findings, heightened awareness in regard to suboptimal performance and potentially invalid assessment data sparked the release of a position statement from the National Association of Neuropsychology (NAN; Bush et al., 2005). Within the position paper, it was stated, “Clinical neuropsychologists are responsible for making determinations about the validity of the information and test data obtained during neuropsychological evaluations” (Bush et al., 2005, p. 419). In other words, clinicians were expected to make the determination as to whether or not the examinee displayed optimal effort, as well as whether adequate and accurate information about symptoms, social history, and other important variables were provided (Bush et al., 2005). In order to meet this responsibility, clinicians developed different methods to assess the validity of a client’s symptoms. As previously mentioned, Slick et al. (1999) proposed a number of

components (i.e., differential diagnosis and clinical judgment) to include as part of an evaluation to help clinicians detect possible malingering or symptom exaggeration.

One key component of these recommendations was a symptom validity assessment as a means to evaluate effort (Bush et al., 2005). Professional organizations such as NAN and the American Academy of Clinical Neuropsychology (AACN) have developed position statements to provide guidance to clinicians on this important topic. In fact, NAN clearly stated that clinicians who did not include a symptom validity assessment needed to be able to adequately justify why such an assessment was not included in their evaluation (Bush et al., 2005). With this publication, it was clear that the field was moving toward making this type of validity assessment the norm rather than the exception.

In 2009, the American Academy of Clinical Neuropsychology (AACN) published a consensus conference statement about the assessment of effort, response bias, and malingering (Heilbronner et al., 2009). Similar to the position paper released by NAN, the AACN provided recommendations surrounding definitions, information about ability issues, and types/methods of validity assessments to help guide the practice of neuropsychologists (Heilbronner et al., 2009). Overall, the AACN noted the importance of validity assessment, response bias, and effort in all evaluations and the ability of clinicians to be able to accurately interpret evaluation findings and make appropriate diagnosis and recommendations (Heilbronner et al., 2009).

More recently, the AACN sponsored another guidance document offering recommendations for change in the Social Security Administration (SSA) policy on validity testing (Chafetz et al., 2015). Surprisingly, despite the well-established and

growing research base in performance and symptoms validity testing in adults, the SSA had discouraged the use of such tests as part of consultative examinations (Chafetz et al., 2015). This position is also contrary to the disability determination procedures used by other agencies which routinely include performance validity tests (PVT) such as the Veteran's Administration, the Railroad Retirement Board, and other private disability insurers (Chafetz et al., 2015).

An independent evaluation was conducted by the Institute of Medicine (IOM) to determine the utility of performance validity tests (PVTs) and symptom validity tests (SVTs) given the scientific literature (Chafetz et al., 2015). Of note, PVTs "are used to determine the accuracy of measures of actual ability" (Chafetz et al., 2015, p. 729) and SVTs "help determine the accuracy of reporting of symptom experience" (Chafetz et al., 2015, p. 729). According to the report released by the Institute of Medicine (IOM), it was concluded, "that standardized psychological tests, including validity tests, are valuable and may increase the accuracy and consistency of SSA's disability determinations" (IOM, 2015, p. 4).

At this point, it is fairly evident that the scientific research base, along with two well-respected professional associations (NAN and AACN) have not only acknowledged the need for clinicians to assess whether examinees are putting forth their best effort on psychological and neuropsychological evaluations and determine the validity of the evaluation, but also highly encourage the use of at least one objective PVT during the evaluation. Perhaps the biggest reason to identify malingering within an adult population pertains to the secondary gain commonly associated with and included in the definition of malingering.

Chafetz and Underhill (2013) conducted a study to examine the annual cost associated with individuals who displayed malingering in disability determination cases. They examined data released by the Social Security Administration (SSA) in 2011 involving mental disorder cases (Chafetz & Underhill, 2013). Malingering costs for adult cases in 2011 were estimated at $20.02 billion (Chafetz & Underhill, 2013). This astronomical number further demonstrates the need to include PVTs in psychological and neuropsychological evaluations, especially those where secondary gain is relevant. However, not all individuals who present with exaggerated symptoms or suboptimal effort are seeking secondary gain. Within a hospital setting, adult medical and psychiatric cases that did not involve any ligation or compensation were still found to have a mean base rate of 11.56 percent malingering (Mittenberg et al., 2002). This finding provides support for the idea that individuals do not necessarily have to be seeking monetary gain, in order to display malingering, exaggerated symptoms, or suboptimal performance. Secondary gain may look different to each individual based on what he or she hopes to attain.

The use of PVTs and SVTs with children and adolescents has not been formally mentioned in the position and conference statements released by the NAN and AACN. However, there is a growing awareness that suboptimal performance can apply to this younger population and the research in this area has been exponentially growing. Until recently, most malingering research primarily focused on adults. Kirkwood (2015a) proposed that child-based research was sparse due to the assumption that children and adolescents could and would not malinger, exaggerate symptoms, or display suboptimal

performance in evaluations. However, a review of developmental research strongly suggests otherwise as children learn to lie at a very early age (e.g., Talwar & Lee, 2002).

Development of Lying in Children

Talwar and Lee (2002) examined the development level in which children ages three to seven years were first able to tell white-lies without being detected by another individual. Specifically, in their study, the Reverse Rouge Procedure was employed in which the researcher purposefully placed on spot of rouge on her nose. The experimenter then asked the child "Before you take a picture of me, do I look okay for the picture?" (Talwar & Lee, 2002, p.165). The results of their study indicated that children as young as three years of age were able to tell white-lies with success in this politeness situation (Talwar & Lee, 2002). Consistent with these findings, Ahern et al. (2011) concluded from their study that children as young as two and one-half years of age are capable of making false statements, but it was not until age three and one-half years of age that children are able to consistently maintain the false statements.

In 2003, Wilson, Smith, and Ross published the first longitudinal study regarding the lying behavior of young children in a natural setting. Participants included 40 English-speaking Canadian families. Of all the children involved in the study, 96% were found to lie at least one time (Wilson et al., 2003). Older siblings displayed a higher rate of lying when compared to their younger siblings and both groups of siblings were found to lie more as they got older. Rates of lying varied across developmental age with a large increase in the number of children who lied between the ages of two and four years, but not as large of an increase between four and six years of age, possibly because children had reached or nearly reached a ceiling level. Once children reached the age of six years,

nearly all children lied, and it was proposed that the increase in lying behavior was associated with the growth in speech across the children's development (Wilson et al., 2003). Another fascinating finding of this study was that older males lied at a higher rate than older females; this difference was attributed to the idea that older males may engage in more transgressions resulting in lying behaviors (Wilson et al., 2003).

As might be expected, the types of lies told by children differed by age. Older children told more complex, detailed lies, when compared to younger children, yet the younger children were able to tell detailed lies, just not at the same level as the older children (Wilson et al., 2003). When examining the purpose or function of the lies told by the children in this study, three main functions were observed: "avoid responsibility, accuse their siblings, and gain control over another's behavior" (Wilson et al., 2003, p. 39). Older children were found to lie to their parents at twice the rate of their younger siblings, who lied at equal rates to their parents and other siblings. Overall, it was found that children lied more frequently to those who were in positions of power (Wilson et al., 2003). Surprisingly, parents paid little attention to the lies their children told suggesting that this behavior was often ignored or undetected by parents.

In fact, research indicates that adults are not very good at detecting when children are telling lies, with a rate of detection that is at or below the level of chance (Crossman & Lewis, 2006; Strömwall et al., 2007). There are some interesting consistencies between whether or not adults were able to accurately detect when children were telling the truth. For instance, both Strömwall et al. (2007) and Crossman and Lewis (2006) found that adults were able to accurately detect truths or reported more instances of detecting true statements than false statements. Another interesting component in the Strömwall et al.

(2007) study was that when children had time to plan the presentation of their lie, adults had a more difficult time identifying that the children were lying as compared to those who had no preparation time. Those children who did not prepare to present their lies were not that much easier to identify, as there was not a large effect size between the two categories.

Children as young as two and one-half years of age can produce false statements; their ability to lie increases and becomes more complex as they get older. Furthermore, adults demonstrate poor levels of detecting lying in children. Taken together, these findings lend additional support to the idea even children who are quite young are able to deceive which might be considered a form of suboptimal performance (i.e., saying you don't know an answer when you do or knowingly providing an incorrect answer). Furthermore, these findings suggest that subjective observations in psychological and neuropsychological evaluations are not sufficient to detect feigned symptoms in children and adolescents (Faust et al., 1988).

Pediatric Suboptimal Performance

Although pediatric malingering or suboptimal performance has not been studied as much as adult malingering, one of the first studies occurred more than 30 years ago. Faust et al. (1988) conducted a study in which they directed three, above-average functioning children to perform poorly on a neuropsychological evaluation. The participants were told that they could earn an additional $5 if they performed poorly, but at a level that was undetectable, in addition to the $15 for their participation in the evaluation. Along with a brief history, test protocols, answer sheets, and drawings were provided to a pool of judges to determine whether they could detect malingering from the

participants. Unfortunately, given this format, none of the judges was able to detect malingering providing support for the idea that clinical experience alone was not enough to accurately assess for malingering (Faust et al., 1988).

The ability to detect suboptimal performance in children and adolescent can be further complicated by the idea of malingering by proxy (Lu & Boone, 2002). According to Kirkwood (2015b), "malingering by proxy refers to when the incentive for the symptom production is clear external gain for the caregiver, rather than psychological benefit" (p.440). Caregivers might imply or suggest to a child to perform poorly and might even coach the child on how to give suboptimal performance (DeRight & Carone, 2015). In the Lu and Boone (2002) case study, a nine-year-old male, who had sustained a traumatic brain injury as a result of a car accident, was found to have failed all four of the performance validity tests he was administered and demonstrated an atypical pattern of performance across measures. This case study provided additional support for the need to use objective performance validity measures when testing children (Lu & Boone, 2002). Secondary gains, although they are likely to look different in children than in an adult population, can be more difficult to identify, especially if parents or caregivers are coaching or contributing to the suboptimal performance or symptom exaggeration displayed by children. As such, conceptualizing secondary gain in children and adolescents in terms of family benefit, rather than personal gain needs to be considered.

Similar to the adult literature on suboptimal performance, there are a number of case studies and reports providing support that suboptimal performance occurs among child and adolescent populations (Kirkwood & Kirk, 2010). In one of the first studies that examined the base rate of suboptimal performance in children and adolescents within a

specific population, those with mild traumatic brain injury (mTBI), 17% of the child and adolescent participants, ages 8 to 17 years, were found to display suboptimal performance, meaning that participants "failed at least one of the three primary effort indices of the MSVT" (Kirkwood & Kirk, 2010, p. 864). Even more surprising than the high base rate was the fact that at the time of the evaluations, none of the participants or their families were seeking any type of disability compensation (Kirkwood & Kirk, 2010). Similar to adults, children and adolescents completing psychological or neuropsychological evaluations with no potential monetary gains still displayed high rates of suboptimal performance (Kirkwood & Kirk, 2010; Mittenberg et al., 2002). As noted, secondary gain can look very different among individuals and it may be harder to detect the "gain" for younger populations.

Types of Performance Validity Measures for Children and Adolescents

Knowing that children and adolescents can and do display suboptimal performance, the next step for clinicians is to be able to utilize the appropriate means to assess for effort within this population. Since the focus of previous studies was directed towards assessing and detecting malingering among adults, there are few instruments that have been designed and studied with younger populations. Many child and adolescent focused studies have had to utilize adult normed PVTs.

In their systematic review of PVT measures, DeRight and Carone (2015) noted there were four instruments that are frequently used in a child and adolescent population including: Test of Memory Malingering (TOMM; Tombaugh, 1996), the Medical Symptom Validity Test (MSVT; Green, 2004), Word Memory Test (WMT; Green & Astner, 1995), the Reliable Digit Span (RDS; Greiffenstein, Baker, & Gola, 1994).

Children and adolescents, ages 5 to 19, were found to meet or exceed the adult established norms on the TOMM, MSVT, and WMT with 98% of children passing the TOMM, and 95% and 86% of children and adolescents passed the MSVT and WMT, respectively, using adult norms (DeRight & Carone, 2015). Of note, children and adolescents had a more difficult time passing the RDS using adult norms and even when norms were adjusted, sensitivity and specificity were not ideal, potentially leading to a large number of false positives on this measure (DeRight & Carone, 2015). Similarly, Kirk et al. (2011) reported corresponding results, revealing that 96% of children and adolescents, ages 5 to 16, with a number of presenting concerns including: Attention Deficit/Hyperactivity Disorder (ADHD), Learning Disabilities, Pervasive Developmental Disorders, Bipolar Disorder, and mild traumatic brain injury (mTBI) passed adult establishing cutoff norms. These studies provide support for the use of these adult based PVTs in pediatric populations.

The majority of studies of suboptimal performance in younger populations have used the TOMM (e.g., Blaskewitz, Merten, & Kathmann, 2008; Constantinou & McCaffrey, 2003; Kirk et al., 2011; Kirkwood & Kirk, 2010) likely because it has a longer history than many of the other measures and because the majority of children and adolescents show the ability to exceed the cut off score. The adult established cut off score of 45 or greater on Trial 2 of the TOMM is considered to be within normal limits or "passing" (Tombaugh, 1996).

Within a clinical, pediatric sample of children and adolescents ranging from 5 to 16 years of age, 97 out of 101 (96%) participants obtained passing scores on Trial 2 of the TOMM (Kirk et al., 2011). In another study examining the effort of 51 German

speaking students, ages 7 to 9, all 51 participants were found to pass Trial 2 on the TOMM utilizing the established adult norms (Blaskewitz et al., 2008). Another study examined the differences between Greek-Cypriot children (ages 5 to 12) and children from New York (ages 5 to 12) and found that all participants in both groups met or exceeded the designated adult cut off score (Constantinou & McCaffrey, 2003). In fact, Chafetz et al. (2007) found that 43 children, who had intelligence scores below 70, passed the second Trial of the TOMM with scores ranging from 48 to 50 (near perfect to perfect scores). Given the high rate of passing scores that children and adolescents obtain, regardless of nationality, language, and intelligence level, it seems the utility of the TOMM with a pediatric population is both credible and well-established.

More recently, the first and only PVT specifically normed with children and adolescents was developed. The Memory Validity Profile (MVP) has been normed for use with children and adolescents ages 5 to 21 years of age and was co-normed with the Child and Adolescent Memory Profile (ChAMP) (Sherman & Brooks, 2016). According to the test authors, 100% specificity and sensitivity were found during their initial study (Sherman & Brooks, 2016). Specificity refers to the “true negative rate”, or in other words, how certain can we be that an individual is not displaying suboptimal performance (Chafetz et al., 2007, p. 9). On the other hand, sensitivity refers to a “true positive rate”, meaning how certain we can be that an individual is truly displaying suboptimal performance (Chafetz et al., 2007, p. 9). Sensitivity and specificity are critical as clinicians want to be sure they are accurately detecting the targeted behavior or skill with the instrument they are using.

Only one recent study could be found that attempted to examine the sensitivity of the MVP (Dodd, Murphy, & Bosworth, 2020). The sample included only patients with mild traumatic brain injuries, ages 5 to 17. MVP cut off scores for this patient sample were reportedly not sensitive enough to detect true cases of suboptimal performance (Dodd et al., 2020). Therefore, future research needs to be conducted to confirm these results with other patient samples.

Psychoeducational Evaluations

Each year children and adolescents, who are struggling in their educational setting, are evaluated to determine whether they meet criteria for special educational services. Practitioners utilize laws and guidelines (e.g., IDEA) to help exclude children who have not had adequate educational exposure to learn, who speak another language, or who are economically disadvantaged, from being falsely identified for special education services. One way providers attempt to keep students from incorrectly being identified for services is through the use of norm-based assessments with high levels of reliability and validity. Two of the primary assessment measures administered to children and adolescents, as part of psychoeducational evaluations, cognitive and academic assessments, unfortunately, do not have "built in" validity measures to assess for student effort or inconsistencies.

Little is known about the inclusion of performance validity measures as part of psychoeducational evaluations. Given the lack of information of such measures, it is difficult to determine the validity of these evaluations. This is problematic as students could be misidentified as a student with a disability. Educational labels can be stigmatizing and create unnecessary harm to students (Ho, 2004). Also, schools receive

funding to aid in the education of students with disabilities. As such, if students are incorrectly identified, federal funding is unnecessarily dispersed. Furthermore, students could also be granted access to unnecessary supports if identified as a student with a disability. As previously noted, families may also play a part in coaching students to purposefully perform poorly during evaluations for family gain (Lu & Boone, 2002).

Consideration of suboptimal performance has been included in some types of assessment conducted in school setting. For example, VanDerHeyden and Witt (2008) developed an approach called "can't do, won't do" when screening for academic delays. A "can't do, won't do" assessment can be administered to students who are not performing at the expected level in academic and behavioral areas. The general premise is that if motivation (won't do) is thought to be driving the student's behavior, then that student is provided a motivating prize for increased performance. If the student is then able to carry out the task at the expected level, the examiner can rule out that the student is lacking skills in that area and in fact does possess the ability necessary to complete the task. However, if the student is still not able to complete the task at the appropriate level, despite a motivating prize, it is suggested that the student has a true skills deficit in this area (VanDerHeyden & Witt, 2008). This idea was initially implemented within academic assessments of student's skills but can also be implemented with behaviors as well.

Another example of suboptimal or in this case, exaggerated symptoms has been presented within the context of universal screening within a school setting. Furlong, Fullchange, and Dowdy (2017) conducted a study of student responses and found evidence of "mischievous responding" in adolescent populations. Their study was

designed to determine the rate at which adolescents endorsed multiple questions that have an extremely low base rate in isolation and are not likely to be answered with multiple affirmatives. As part of a universal mental health screening in a high school setting, the researchers included seven questions to assess if students would exaggerate their answers of these questions that would normally occur at a low frequency level. They found that about 2% of students endorsed explicit exaggerated answers to these seven questions. This study provided further support for the idea that students can and do, whether subconsciously or consciously, exaggerate symptoms.

Additionally, poor effort has been studied as part of post-concussion evaluations in high school athletes. Higgins, Denney, and Maerlender (2017) enlisted high school student athletes, who had previously completed baseline testing using the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT), to complete baseline testing under two different conditions. Student athletes were instructed to provide their best effort and also "sandbag" or provide suboptimal effort on baseline testing on the ImPACT. Students were randomly assigned to two groups to determine if they were instructed to provide best effort or suboptimal effort first. Baselines for both optimal and suboptimal performance were administered back to back. Results of the study indicated differences in baseline scores of the ImPACT when students were instructed to provide poor effort versus best effort. ImPACT composite and subtest scores differences were noted between the optimal and suboptimal performance baselines. The results of this study provide further evidence for the idea that students can be coached to provide optimal and suboptimal effort. However, administration of validity measures is not routinely incorporated as part of psychoeducational evaluations.

Therefore, given the potential for misidentification of students and subsequent negative impacts, it is important that school psychologists incorporate tools that will allow them and their team members to make the best decisions. Therefore, exploring the rate of suboptimal performance during psychoeducational evaluations is an important next step in determining whether and to what degree suboptimal performance occurs on psychoeducational assessments or whether it is a phenomenon that is unique to clinical settings. It is impossible to predict the number of students who may be displaying suboptimal performance within educational settings without directly assessing for this possibility. With this information, school districts may be able to develop better criteria for identifying (or flagging) potential cases of suboptimal performance.

CHAPTER III

METHODOLOGY

The literature on pediatric validity has gained increasing attention across inpatient and outpatient clinical settings (Kirkwood & Kirk, 2010; Weber Ku et al., 2020). However, there appears to be a gap in the research base between clinical and school settings, in that school-based practitioners rarely use PVTs. This study represents the first known attempt to examine the rate of suboptimal performance in children and adolescents who are participating in assessments as part of their initial and three-year psychoeducational evaluations. IRB approval from the University of Northern Colorado was received prior to data collection within school settings across Nebraska and Oklahoma. A discussion regarding participants, research design, measures, procedures, and data analysis can be found below.

Participants

The participants for this study included data from 54 students from various school districts in eastern and central Nebraska and central Oklahoma. Participants included students ranging from preschool to 11th grade, who were being evaluated for an initial or three-year psychoeducational evaluation. De-identified demographic information for each student participating in the study was collected and included: age, sex, grade, primary language, race, ethnicity, and medical and/or mental health diagnosis. The selected districts represent a convenience sample based on the researcher's preexisting relationships with school personnel. All data were collected by five school psychologists.

Cooperating school psychologists were appropriately credentialed for practice, as determined by their respective state requirements.

Because of the possibility that students who were very young or who had more serious disabilities might fail the TOMM, certain exclusionary criteria were implemented. These exclusion criteria included students under the age of 5 years 0 months, students with documented visual impairments, and students with moderate to severe intellectual disabilities (defined as an IQ score of 60 or lower). When student participants had a full-scale intelligence quotient (FSIQ) standard score of 60 or less, visual spatial index (VSI) or perceptual reasoning index (PRI) scores were also examined. If students also displayed visual spatial or perceptual reasoning weaknesses, identified as standard scores below 70, these students were excluded from this study. Using these criteria, only two participants were eliminated. One participant was under the age of five and the other participant had a very low IQ (FSIQ SS= 54; VSI SS=67) suggesting that failure on Trial 2 of the TOMM could be due to intellectual disabilities rather than suboptimal performance. Therefore, the data for 52 participants were included in the analysis.

Research Design

A quantitative research design was employed. A potential threat to external validity is generalizability. Since a convenience sample of participants was used in this study, generalizability to other students in different school settings might be more difficult. In order to manage this potential threat to external validity, attempts were made to gather information from a diverse group of students in multiple educational settings. There were no detectable threats to internal validity.

Measures

School psychologists within the cooperating school districts assisted with data collection. In addition to de-identified demographic information, school psychologists included overall composites and subtests scores on intelligence and academic achievement measures given as part of the psychoeducational evaluation, as well as, the eligibility category the student was classified under, if the student was found to meet state criteria for an educational verification. Since these other measures were not the focus of this study, only the intelligence and academic achievement composite scores and subtests scores were obtained when administered as part of the evaluation battery. Intelligence test standard scores were received from 37 of the 52 participants. Although academic scores were obtained for some participants, these scores varied greatly because many different measures were administered. Furthermore, some cooperating school psychologists provided subtest scale scores while others provided composite standard scores. School psychologists entered the deidentified demographic information and testing scores into an Excel spreadsheet (see Appendix C).

Test of Memory Malingering (TOMM)

Participants were administered the Test of Memory Malingering (TOMM) Trials 1 and 2 (Tombaugh, 1996). The TOMM is a visual recognition assessment that was originally designed for individuals ages 16 to 84 (Tombaugh, 1996). The TOMM is an effort test that is disguised as a picture memory test. It consists of 50 black and white pictures that are presented to the examinee in two Trials. In each Trial, the examinee is exposed to the 50 target pictures for 3 seconds each, and then asked to point to the target picture, when presented with a page containing the target and foil pictures (Tombaugh,

1996). An optional Retention Trial is also available for administration 15 minutes following the administration of Trial 2 (Tombaugh, 1996). The Retention Trial was not used in this study. The Retention Trial was not administered in the Chafetz et al. (2007) study and only administered to two participants, who had failed Trial 2 of the TOMM, in the study conducted by Constantinou and McCaffrey (2003).

Administration of Trials 1 and 2 of the TOMM takes approximately 15 to 20 minutes and an additional 5 to 10 minutes if the optional Retention Trial is administered (Tombaugh, 1996). This assessment was created to help examiners determine whether examinees had a true memory impairment as opposed to engaging in malingering or suboptimal performance (Tombaugh, 1996). The TOMM was normed within an adult population including individuals with dementia and traumatic brain injuries to ensure proper detection of malingering (Tombaugh, 1996). However, children and adolescents have demonstrated adequate ability to meet the established adult threshold necessary to "pass" or display optimal performance on the TOMM (Blaskewitz et al., 2008; Constantinou & McCaffrey, 2003; Kirk et al., 2011; Kirkwood & Kirk, 2010).

The adult established cut off score of 45 or greater on Trial 2 of the TOMM is considered to be within normal limits or "passing" (Tombaugh, 1996). Although these tasks appear difficult to the examinee, they are actually very simple. In fact, Chafetz et al. (2007) found that 43 children who had intelligence scores below 70, passed the second Trial of the TOMM with scores ranging from 48 to 50 (near perfect to perfect scores). Therefore, it is likely that even young children would be able to obtain passing scores on the TOMM. Given the number of published research studies (e.g., Blaskewitz et al., 2008; Constantinou, & McCaffrey, 2003) that have utilized with the TOMM as a tool to

assess for suboptimal performance within child and adolescent populations, this measure was selected and administered, as part of this study.

Student Effort Rating Scale

In addition to the TOMM, the primary researcher created a subjective rating scale for participating school psychologists to use to rate student effort based on practitioner observations. At the time this study was started, a subjective student effort rating scale was not available. School psychologists were asked to administer the first assessment in their psychoeducational battery and then rate student effort based on observed behaviors. Options included three potential ratings: 1 (Little Effort), 2 (Partial Effort), or 3 (Full Effort). In addition, spaces were provided on the student effort rating form for school psychologists to enter behaviors they observed that contributed to their rating of effort. Then, school psychologists were asked to administer Trials 1 and 2 of the TOMM to the student and proceed with any additional assessments, as they normally would in their evaluation battery.

Procedures

Prior to any data collection, permission to complete the study was obtained from the Institutional Review Board (IRB) from the University of Northern Colorado (UNC). The primary researcher contacted superintendents, special education directors, and school psychologists across school districts in Nebraska and Oklahoma to inquire about potential participation in the current study. All school-district policies and procedures were followed in gaining consent to collect data. Eighteen school districts between the 2018-2019 and 2019-2020 school years were contacted for participation in this study. The primary researcher emailed superintendents, directors of special education services, and

individual school psychologists, in an attempt to recruit their participation in the current study. As previously mentioned, the primary researcher primarily recruited individuals and schools for which previous relationships had been established. When previous relationships had not been established, the primary researcher collected administrative emails from school district websites to send recruitment emails to. Of the 18 school districts, 5 school districts agreed to participant. Data were collected over the 2018-2019 and 2019-2020 school years.

Once a district leader provided approval, the primary researcher contacted the individual school psychologists via emails provided by the district leader, or through arranged meetings with the district leader, who would be administering the assessment measures to the participating students. School psychologists were provided with information about the study, including its purpose, and provided an informed consent form. For all practitioners that chose to participate in this study, as determined by their signature on the informed consent form, they were asked to include the TOMM in their assessment battery to students, providing that they had also received parental consent for the student to participate in the study. Nine TOMM test kits were purchased by the researcher and provided to the five participating school districts and school psychologists. A total of ten school psychologists agreed to participate and signed informed consent forms. However, only five school psychologists provided data for this study.

Parental consent and student assent were obtained prior to any testing. Since school psychology practitioners were required to gain parental consent to administer a psychoeducational evaluation within a school setting, the participating practitioners were asked to provide all parents/guardians with the parental consent form for this study at the

same time that they collected parental consent for the evaluation. Student assent was either collected at the same time, or prior to administration of the psychoeducational evaluation. When necessary, the school psychology practitioner provided the student assent form to the student immediately before administration of the assessment battery.

To ensure that each participating school psychologist delivered the TOMM with fidelity, the primary researcher provided a live demonstration that described and demonstrated how to administer the TOMM in a standardized format. Part of the live training demonstration also described the purpose for administering the TOMM. Therefore, the participating school psychologists understood that this instrument was considered to be a measure of suboptimal performance or effort. It was not possible to have a "blinded" condition, instead participating school psychologists were instructed to administer the assessments in a specific order. School psychologists were provided with a fidelity checklist (see Appendix B), as a means to ensure all components of this study were given and administered in a standardized format. The fidelity checklist was reviewed at the live demonstration meeting with participating school psychologists.

School psychology practitioners were asked to administer the first assessment in their battery, then complete a short measure of student effort, based on their own observations during the administration of the first assessment. The student effort rating form asked that practitioners circle the level of effort they believe the student put forth on the first measure. Ratings are Full Effort (3), Partial Effort (2), and Little Effort (1). Additionally, practitioners were provided five blank spaces, in which they were encouraged to list behavior(s) contributing to their conclusion of effort displayed by the student.

Additionally, all cooperating school psychologists were asked to administer the TOMM as the second assessment in their psychoeducational battery. Both Trials 1 and 2 of the TOMM were administered to participants, taking approximately 15 to 20 minutes. After administration of the TOMM, all other aspects of the psychoeducational battery were completed as per the practitioner's customary process.

At the completion of their testing and after scoring the various instruments, the cooperating school psychologists entered participant data into an Excel spreadsheet provided by the primary researcher. The participating school psychologists then submitted the Excel spreadsheet to the primary researcher. Excel spreadsheets were saved, and password protected on the primary researcher's computer. Information collected includes de-identified demographic information for each student (age, sex, grade, primary language, race, ethnicity, and medical and/or mental health diagnosis), all intelligence and academic subtest and composite scores (when administered by practitioners as part of their typical battery), educational verification, initial or three year re-evaluation, TOMM Trial 1 and 2 scores, practitioner rating of overall student effort, and all self-reported BASC-3 validity ratings (when administered by practitioners as part of their typical battery).

Data Analysis

The purpose of this study was to examine potential suboptimal performance of children and adolescents among students completing a psychoeducational evaluation within an educational setting. Because no other studies were found specific to educational settings, it was difficult to predict what percentage of students would display suboptimal performance and subsequently what further analysis could be conducted. A large enough

percentage of suboptimal performance was discovered among students, leading to further analyses regarding age, sex, grade, primary language, race, ethnicity, medical and/or mental health diagnosis, and eligibility category.

Furthermore, school psychology practitioners' ratings of student effort were compared to TOMM Trial 2 cut off scores to determine the relationship between observed indicators of effort and potential suboptimal performance as indicated by Trial 2 of the TOMM. Full scale IQ standard scores were also examined as predictors of pass/fail rate on Trial 2 of the TOMM.

CHAPTER IV

RESULTS

An examination of effort put forth by children and adolescents being tested for special education services through a psychoeducational evaluation was studied. Suboptimal performance has been found among children and adolescents completing evaluations in clinical settings (Kirkwood & Kirk, 2010; Mittenberg et al., 2002), as well as psychiatric inpatient settings (Weber Ku et al., 2020), but there have been no known efforts to examine suboptimal performance in an educational setting. For this study, suboptimal performance was measured by a standalone validity measure, the TOMM, administered by school psychologists who were conducting the psychoeducational evaluation. Additionally, school psychologists were asked to report observable behaviors that each participant displayed providing support for their effort ratings for each participant.

The results are presented below following a presentation of demographic data on cooperating school psychologists and students. This information is followed by a review of the research questions and the analyses used for each of these questions.

Participant Demographics

There were two participant groups for this study. Although not the targeted participant group, volunteer school psychologists assisted in gathering these data. The second group consisted of the child population that was assessed by the participating school psychologists.

School psychologists

Data collection was conducted by five licensed school psychologists in central and eastern Nebraska, as well as, central Oklahoma. All participating school psychologists were female. Four of the five school psychologists reported having an education specialist (Ed.S.) degree in school psychology. One school psychologist had a master's degree (M.S.) in school psychology. They reported their years of practice as a school psychologist ranging from 7 to 18 years. Of the five participating clinicians, four maintained a designation as a Nationally Certified School Psychologist (NCSP) through the National Association of School Psychologists (NASP). Signed consent forms for each participating school psychologist were collected and copies were provided to each participant.

Students

Permission from each parent or caregiver to conduct a psychoeducational assessment was obtained by the participating school psychologists. Additionally, parents were notified of this study and permission to have their child's data included was gathered after providing informed consent (by the participating school psychologists). Parental consent for this study was obtained when parents gave permission for the psychoeducational evaluations, per customary school practice. Assent forms from each student participant were also collected. For each student who was consented into the study, de-identified demographic information was collected and included: age, sex, grade, primary language, race, ethnicity, and medical and/or mental health diagnosis (see Table 1).

Data were collected from a total of 54 participants; however, data provided by two students were excluded. Exclusionary criteria included students under the age of five years old, those with documented visual impairments, and students with significant intellectual disabilities. For student participants, who had a full-scale intelligence quotient (FSIQ) standard score of 60 or less, visual spatial index (VSI) or perceptual reasoning index (PRI) scores were also examined. Visual spatial or perceptual reasoning weaknesses were identified as standard scores below 70. When all relevant scores met these criteria, these students were excluded from this study. These exclusions were made to reduce the possibility of false positives on Trial 2 of the TOMM. Of the two participants excluded from this study, one participant was under the age of five and the other student had scores below the cutoffs described above. Therefore, the data from 52 students were analyzed in this study.

Student participants ranged in age from 5 to 17, with the majority of students falling between 7 and 8 years of age (20 students). Student participant sex included 40.4% females and 59.6% males. Of the data collected, information was obtained from students ranging from preschool to 11th grade. Students in 2nd grade made up 25% of the data collection sample. The primary spoken language reported by student participants was English, except for four participants who were reported to use Spanish as their primary language. Approximately 90% of the students identified as white, with reported ethnicity as non-Hispanic. Although the majority of students did not have a prior medical or mental health diagnosis, nearly 20% reported having a singular diagnosis of Attention Deficit Hyperactivity Disorder, while almost 4% were reported to have a diagnosis of Attention Deficit Hyperactivity Disorder comorbid with anxiety and depressive disorders.

Table 1
Demographic information for student participants in percentage and frequency.

	Percentage	Frequency
	%	*n*
Age		
5	3.8	2
6	7.7	4
7	19.2	10
8	19.2	10
9	1.9	1
10	11.5	6
11	15.4	8
12	7.7	4
13	5.8	3
14	0	0
15	5.8	3
16	0	0
17	1.9	1
Grade		
PK-Kindergarten	7.7	4
Elementary (1-5)	67.3	35
Middle School (6-8)	17.3	9
High School (9-11)	7.6	4
Primary Language		
English	92.3	48
Spanish	7.7	4
Race		
American Indian/Alaska Native	5.8	3
Asian or Pacific Islander	1.9	1
African American or Black	1.9	1
White	90.4	47
Ethnicity		
Hispanic/Latino	17.3	9
Non-Hispanic	82.7	43
Previous Medical/Mental Health Diagnosis		
Attention Deficit/Hyperactivity Disorder	19.2	10
Autism Spectrum Disorder	3.8	2
Epilepsy	1.9	1
Depressive Disorder/ Attention Deficit/Hyperactivity Disorder	1.9	1
Anxiety Disorder/Post Traumatic Stress Disorder/ Attention Deficit/Hyperactivity Disorder	1.9	1
Unknown, but medicated	1.9	1

Information regarding whether the student was participating in an initial psychoeducational evaluation or a three-year re-evaluation was gathered. Of the 52 participants, 23 of the psychoeducational evaluations conducted were initial evaluations and 29 were reevaluations, meaning that many of these students were already receiving special education services in some capacity. Table 2 shows the comprehensive break down of all the participants, regarding the primary educational disability, if present, or if the student did not meet criteria for special education services. In Nebraska and Oklahoma, there are 13 categories that students can potentially meet criteria for to receive special education services. These categories include: Autism (AU), Emotional Disturbance (BD/ED), Deaf Blindness (DB), Hearing Impairment (HI), Intellectual Disability (ID), Multiple Impairments/Disabilities (MULTI/MD), Orthopedic Impairment (OI), Other Health Impairment (OHI), Specific Learning Disability (SLD), Speech Language Impairment (SLI), Traumatic Brain Injury (TBI), Visual Impairment (VI), and Developmental Delay (DD). Twenty five percent (25%) of the students identified as meeting criteria for special education services were under the Specific Learning Disability category. It was surprising to note that 18 of the 52 participants, or 34.6%, did not meet verification criteria for special education services.

Table 2
Primary educational disability status for participating students in percentage and frequency.

	Percentage	Frequency
	%	*n*
Autism (AU)	5.8	3
Developmental Delay (DD)	3.8	2
Intellectual Disability (ID)	7.7	4
Other Health Impairment OHI)	19.2	10
Specific Learning Disability (SLD)	25.0	13
Speech/Language Impairment (SLI)	3.8	2
Did not qualify for special education services (DNQ)	34.6	18

Furthermore, participating school psychologists were asked to provide validity scale information for students who completed a self-report Behavior Assessment System for Children, Third Edition (BASC-3), as part of the psychoeducational evaluation. Unfortunately, only seven participants completed the BASC-3, as part of their evaluations, so no meaningful statistical analysis could be conducted.

However, from a qualitative perspective, of the seven completed self-report BASCs, all but one had acceptable validity. The participant who did not have acceptable validity across all indices, had an "extreme caution" warning for the *F*-Index and a "caution" warning on the consistency index. The other validity indices were acceptable for this participant. This participant passed both Trials 1 and 2 on the TOMM suggesting he/she put forth reasonable effort on his/her evaluation. Additionally, the school psychologist rated this individual's effort as a 3 (full effort).

On the other hand, one participant received acceptable ratings on BASC-3 self-reported validity indices yet failed both Trials 1 and 2 on the TOMM. The other five participants, who completed the BASC-3 self-report rating form received acceptable ratings in all validity indices and passed both Trials 1 and 2 on the TOMM.

Research Question Analysis

Q1 What is the percentage of students who display suboptimal performance (45>) on Trial 2 of the TOMM during their psychoeducational evaluation conducted within a school setting?

To answer research question 1, descriptive statistical analyses were conducted to determine the percentage of students who received a score of 44 or less on Trial 2 of the TOMM, a level that is defined as below expected performance and thus, considered to be potential evidence of suboptimal performance. Results of this analysis indicated that 19.2% of student participants failed Trial 2 of the TOMM. Below, Table 3 shows the minimum and maximum scores achieved on Trials 1 and 2 on the TOMM, as well as, means and standard deviations for each trial.

Table 3
TOMM Minimum and Maximum Scores, Means, and Standard Deviations.

	MIN	MAX	*M*	*SD*
TOMM Trial 1	21	50	42.69	6.332
TOMM Trial 2	19	50	46.81	6.701

Upon further investigation, a breakdown of age and pass/failure rate on Trial 2 on the TOMM was explored. The oldest participants to fail Trial 2 of the TOMM were 15 years old and the youngest were 5 years old, with 5, 6, 7, and 8-year-olds making up 7 of the 10 total students to fail Trial 2 on the TOMM (see Table 4 below).

Table 4
Student Age and TOMM Trial 2 pass and failure frequency.

Student Age	Fail	Pass (45>)
	n	*n*
5	2	0
6	1	3
7	3	7
8	1	9
9	0	1
10	0	6
11	1	7
12	0	4
13	0	3
15	2	1
17	0	1

Q2 What percentage of agreement exists between perceived student effort, as measured by practitioner observations and objective student effort, as measured by a stand-alone performance validity measure (TOMM Trial 2)?

For research question 2, school psychologists were asked to provide their rating of student effort on a scale with three anchor points. On the student effort rating scale, 1 indicated “Little Effort”, a rating of 2 indicated “Partial Effort” and finally, a rating of 3 denoted “Full Effort”. Again, the rating of student effort was a subjective rating provided by the participating school psychologists prior to TOMM administration. All five school psychologists indicated that they believed all students had given partial or full effort. Further, 76.9% of the school psychologists reported a rating of 3, full effort, for their student participants while the other 23.1% of ratings were a 2, indicating partial effort.

As noted, 10 students failed Trial 2 on the TOMM. Trial 2 is considered to be the identified method for determining possible suboptimal performance. However, it was interesting to note that on Trial 1 of the TOMM, 26 students (50%) failed meaning they received a score of 44 or less. Once more, none of the students were given an effort rating

of “Little Effort”. In fact, of the 26 students who failed Trial 1 on the TOMM, eight students received an effort rating of 2 and 18 students received an effort rating of 3.

Of the 10 students who failed Trial 2 of the TOMM, four students were given ratings of 2, and six students received ratings of 3 for perceived student effort. An examination of how school psychologists rated the effort of the 10 students who failed Trial 2 on the TOMM was broken down by student age (Table 5) and educational verification category (Table 6).

Given that the youngest students (age 5) had a 100% fail rate and were considered to put forth full effort, caution may be warranted in using the TOMM with this population despite previous research. Conversely, at least some suspicion (Effort rating of 2) was noted with four of the older student participants (ages 7 to 15). Participating school psychologists might have been reluctant to give effort ratings of 1, as this would have indicated suboptimal performance and potentially rendered the psychoeducational evaluation as invalid. However, something about the effort of 40% of students, who failed on Trial 2 of the TOMM, was noted as suboptimal. Unfortunately, participating school psychologists provided few written observations supporting their effort ratings.

Table 5
Student Age and Practitioner Rating of Effort for TOMM Trial 2 failure frequency.

	Effort Rating of 2	Effort Rating of 3
	n	*n*
5	0	2
6	0	1
7	2	1
8	1	0
11	0	1
15	1	1

Table 6
Educational verification category and Practitioner Rating of Effort for TOMM Trial 2 failure frequency.

	Effort Rating of 2	Effort Rating of 3
	n	*n*
Autism (AU)	0	1
Development Delays (DD)	0	1
Other Health Impairment (OHI)	0	1
Specific Learning Disability (SLD)	2	0
Speech Language Impairment (SLI)	1	0
Did Not Qualify for special education services (DNQ)	1	3

Q3 Do differences exist among the different special education categories for which participants are identified (e.g., specific learning disability, traumatic brain injury, emotional disturbance) with regard to suboptimal performance?

Because of the small number of students who were below the cutoff level of Trial 2 of the TOMM, this question was answered with descriptive analysis. In fact, very few conclusions can be derived from these data as the analysis of TOMM Trial 2 failure rates indicated that individuals were represented in 5 of the 13 categories of disability. As can be seen below in Table 7, 4 of the 10 students who were suspected of suboptimal effort (i.e., failing Trial 2 of the TOMM), were not ultimately verified for special education.

Table 7
Educational verification category and TOMM Trial 2 pass and failure frequency.

	Fail	Pass (45>)
	n	*n*
Autism (AU)	1	2
Developmental Delay (DD)	1	1
Intellectual Disability (ID)	0	4
Other Health Impairment OHI)	1	9
Specific Learning Disability (SLD)	2	11
Speech/Language Impairment (SLI)	1	1
Did not qualify for special education services (DNQ)	4	14

Q4 Does the performance of students displaying suboptimal performance indicate more variable performance across all assessments (e.g., IQ and academic achievement) than students who are believed to be displaying optimal effort?

Unfortunately, cognitive and academic achievement scores were not obtained from each student participant; therefore, it was not possible to examine differences between IQ and academic achievement scores. In the context of initial and three-year reevaluations, cognitive and academic scores are not always obtained for each student being evaluated for special education services. School multi-disciplinary teams (MDT) determine the necessity of these evaluations and measures to be administered on a case by case basis. However, IQ scores for 37 of the 52 participants were obtained.

Although research question 4 ultimately could not be answered, the cognitive scores that were collected were analyzed in relationship to the TOMM scores. In an attempt to determine if IQ scores impact a student's ability to pass or fail Trial 2 on the TOMM, a binary logistic regression was conducted. Below, the classification table displays predicted and observed outcomes of pass/failure rates on Trial 2 of the TOMM for the 37 student participants with IQ scores. As seen in Table 8, it can be determined that in six cases students, based on full scale cognitive scores, were predicted to pass but actually failed Trial 2 on the TOMM, indicating false positives. On the other hand, one student was predicted to fail Trial 2 but passed Trial 2 on the TOMM, indicating a false negative. Overall, the percentages of correct identification in the pass/fail categories on the TOMM Trial 2 suggest overall IQ scores could potentially impact a participant's ability to pass/fail Trial 2 on the TOMM. However, given that this model misidentified 7 out of 52 participants, caution is warranted if only using IQ scores as a predictor of overall pass/failure rates on the TOMM.

Table 8
Observed and Predicted frequency and percentage of TOMM Trial 2 pass and failure rates.

	Predicted TOMM Trial 2 P/F Frequency		Percentage Correct
Observed TOMM Trial 2 P/F Frequency	Fail	Pass (45>)	%
Fail	3	6	33.3
Pass (45>)	1	27	96.4
Overall Percentage			81.1

Additionally, a chi-squared test was conducted to determine if IQ is a predictor of pass/failure rates on Trial 2 on the TOMM. A full-scale IQ standard score was found to be a statistically significant predictor (p<0.45) of a student's ability to pass/fail Trial 2 on the TOMM. Therefore, it appears that caution is warranted when using the TOMM with students who have lower IQs as they may be falsely identified as putting forth suboptimal effort.

Post Hoc Analysis

Following the primary analyses used to address the research questions, an additional analysis was conducted to explore the pass/failure rate between Trial 1 and Trial 2 on the TOMM. A score of 45 or higher on Trial 2 of the TOMM is considered a passing score, according to the TOMM manual. A recent article published by Brooks, Sherman, and Krol (2012) suggested that a passing score on Trial 1 would also be predictive of a passing score of Trial 2 of the TOMM. Although adding to this sparse literature base was not the original intent of this study, data analysis of Trial 1 and Trial 2 pass/failure rates on the TOMM were also explored, as the data were readily available.

As previously mentioned, the overall failure rate for Trial 2 on the TOMM in this sample was found to be 19.2% or 10 participating students. When assessing the failure

rate for Trial 1 on the TOMM, this percentage increases to 50% or 26 participating students. Therefore, the failure rate for Trial 1 on the TOMM was nearly two and a half times greater than Trial 2 on the TOMM. Given this finding, it appears that administering Trials 1 and 2 of the TOMM is necessary for children and adolescents in school settings in order to minimize the number of false positives. If practitioners only utilize Trial 1 scores on the TOMM, they would likely be identifying too many students as giving suboptimal effort due to "failing" TOMM scores. It appears that inclusion of Trial 2 scores on the TOMM helps to give students an opportunity to learn the task.

Furthermore, Brooks et al. (2012) reported a 100% pass rate on Trial 2 on the TOMM for participants in their study who received a score of 36 or higher on Trial 1 on the TOMM. In the current study, 26 students failed Trial 1 on the TOMM. Of those 26 students, four did not achieve a Trial 1 score of 36 or higher. When looking at the remaining 22 students, five achieved scores of 36 or higher on Trial 1 on the TOMM, yet still failed Trial 2 on the TOMM. Therefore, 17 of the 22 participants, who failed Trial 1 on the TOMM and received a Trial 1 score of 36 or higher passed Trial 2 of the TOMM. Unlike the 100% pass rate on Trial 2 reported by Brooks et al. (2012), the current study found the pass rate to be 77.3% for students who obtained a Trial 1 score of 36 or higher and ultimately passed Trial 2 on the TOMM. Again, caution should be utilized when considering only administering Trial 1 of the TOMM, as misidentification appears to be far more likely than other studies had previously found.

Summary

To conclude, a sample of 54 student participants was gathered from five credentialed school psychologists in central and eastern Nebraska and central Oklahoma.

Data from 52 students were utilized to answer the proposed research questions, while data from two students were not used due to exclusionary criteria being met.

Overall, the current study found that 50% or 26 students were found to have failed Trial 1 on the TOMM, while 19.2% or 10 students failed Trial 2. This finding suggests the potential for suboptimal effort being given by students, as part of their psychoeducational evaluations. When comparing the scores students obtained on the TOMM (Trials 1 and 2) with practitioner ratings of observed, subjective observations, very little agreement was found to exist. Unfortunately, concludes surrounding educational verification categories that students, who failed Trial 2 on the TOMM were verified under, were unable to be made. Likewise, comparisons among academic, cognitive, and TOMM scores were not possible, due to the variation in scores provided in these areas. However, some support for full scale IQ scores being a predictor for pass/failure rates on Trial 2 on the TOMM was found. Furthermore, caution is recommended for individuals who are considering only administering Trial 1 of the TOMM, as a predictor of Trial 2 pass/failure rates.

CHAPTER V

DISCUSSION

Pediatric performance validity testing has recently become a focus of interest, largely among practitioners in clinical settings, especially those where neuropsychological evaluations are provided. There is growing awareness that children and adolescents are not only capable of putting forth suboptimal performance during these evaluations (Kirkwood & Kirk, 2010), but also that practitioners' subjective observations are not sufficient to spot suboptimal performance (Faust et al., 1988). This knowledge was the impetus for considering whether students might demonstrate suboptimal effort in other kinds of settings.

Each year, thousands of psychoeducational evaluations are conducted across the United States to determine students' eligibility for special education services. Although objective performance validity measures are commonly used in clinic settings, their use is limited in educational settings. The potential harm for misidentifying students for special education or Medicaid services includes provision of unneeded and costly services, greater cost burden on schools, and potential stigma to the student. Additionally, students may receive unfair advantages such as additional time on high stakes tests. The purpose of the current study was to better understand the rates of suboptimal performance in school settings and the ability of practitioners to accurately identify suboptimal efforts among their test taking students.

Findings

In answering the first research question, the results of this study supported previous work indicating that child and adolescent populations demonstrate suspected suboptimal effort on comprehensive assessments. In the current study, 19.2% of students completing psychoeducational evaluations failed Trial 2 of the TOMM (10 out of 52 students). This finding was consistent with those of Kirkwood and Kirk (2010) who reported a rate of 17% suspected suboptimal performance in children and adolescents with mild traumatic brain injuries, and the 18.5% of pediatric patients who failed at least one validity measure (Kirkwood et al., 2012). Rates of suboptimal performance may vary by setting as more recently, only 7% of children failed two PVTs, one being Trial 2 on the TOMM, in a psychiatric inpatient setting (Weber Ku et al., 2020). Therefore, it appears that suboptimal performance occurs among child and adolescent populations in both school and clinic or hospital contexts. Because research related to suboptimal effort in educational settings is so limited, further evaluation of this phenomena appears to be a relevant and timely topic.

To be clear, the aim of the current study was not to uncover or explain reasons that children and adolescents may provide suboptimal effort on psychoeducational evaluations. It is likely that there are any number of reasons for this apparent suboptimal effort. In fact, from a malingering perspective (attempting to access secondary gains), the greatest number of students (40%) who demonstrated potential suboptimal efforts, did not qualify for special education services. In fact, practitioners could be misinformed in believing that indications of suboptimal effort automatically mean children and adolescents were malingering. Kirkwood et al. (2010) addressed, through a case series, a

variety of suspected reasons participants had failed validity measures during evaluations. These reasons included social factors, school avoidance, sport-related factors (i.e., looking for a reason to stop playing a sport), family factors (i.e., attempts to keep a family together), psychogenic amnesia, and often times unknown etiology (Kirkwood et al., 2010). Furthermore, changes in attention, arousal, and overall cognitive ability levels may contribute to suboptimal effort in children and adolescents (Kirkwood et al., 2010). Best practice would be to consider all of these different factors as potential variables that could influence effort for students who fail an administered PVT.

Of note, although the results of the current study showed that two 5-year-old students and one 6-year-old student failed Trial 2 on the TOMM, previous research has reported that children as young as age 5 are able to pass Trial 2 on the TOMM (Constantinou & McCaffrey, 2003; Dodd et al., 2020; Kirk et al., 2011). However, it is unclear how many participants within this younger age group (i.e., ages 5 and 6) had participated in these studies. Therefore, additional research in this area, especially within school settings, would be beneficial in determining whether a broader sample of younger children are capable of meeting the adult norms of the TOMM or other PVTs.

Training programs often encourage the use of statements regarding student effort on psychoeducational evaluations based on subjective observations. However, research suggests that practitioner observations are not sufficiently sensitive to detect suboptimal performance (Faust et al., 1988). In the current study, none of the school psychologists rated students as putting forth little effort. Instead, they endorsed ratings of “partial” and “full effort” suggesting that their observations of effort did not align with the objective measure of suboptimal effort. Although a more scientific measurement of student effort

was not used, the findings still lend support for the idea that subjective observations are not sufficient to detect underlying suboptimal performance. It is also possible that school psychologist participants were reluctant to provide a student rating effort below a level of 2 (Partial Effort), as it would mean that their assessment was not valid and could not be used to determine eligibility for that particular student.

Qualitatively speaking, practitioners who rated student effort at 2 (Partial Effort) noted behaviors such as, eagerness to participate in memory game, good eye contact, impulsive answers, and restless behavior as administration continued, as contributing factors to their effort rating of 2. It was interesting to note that both positive (e.g., good eye contact, eager) and more negative (e.g., impulsive, restless) behaviors were used to define Partial Effort. Unfortunately, no specific behaviors were described for the participants who had received ratings of 3 (Full Effort).

The findings of research question 3 did not indicate any discernible pattern of performance among students who were verified under the different categories, suggesting that many students have the potential to provide suboptimal effort on Trial 2 of the TOMM. Because there were only 10 student participants who failed Trial 2 of the TOMM, meaningful interpretation of these data were not possible. However, no patterns across educational disability categories were discovered. It is unlikely that failures were due to specific disabilities (e.g., Autism Spectrum Disorder) as previous research has reported that students with many different types of presenting concerns were able to pass Trial 2 of the TOMM at high rates (Kirk et al., 2011). Some of these presenting concerns included mild traumatic brain injury, Pervasive Developmental Disorders, Specific Learning Disorders, and other mental health disorders (e.g., Bipolar Disorder and

Attention Deficit Hyperactivity Disorder). The findings from this study also suggest that students with a wide variety of educational disabilities are able to demonstrate a passing score on a performance validity measure such as the TOMM.

Given the limited number of individuals who indicated suspected suboptimal effort and inconsistent use of a full battery (both ability and achievement scores), it was not possible to answer the original research question related to variability in scores. Instead, research question four was adapted to examine the relationship between ability and pass rates on the TOMM and to examine estimates of false positive and negative rates. Ability scores showed a high rate of correctly identifying those students who should pass the TOMM at a rate of 96.4%. However, overall ability did not make a good predictor of those who failed, with a correct rate of only 33.3%. In some ways, this supports earlier findings that students with low cognitive functioning can pass the TOMM (e.g., Chafetz et al., 2007). However, it does suggest that for some students who have low cognitive ability, the TOMM may not be an appropriate instrument to use.

Overall, full scale IQ standard scores were found to be a significant predictor of student ability to pass/fail Trial 2 on the TOMM. This finding is surprising given that two students had a FSIQ score of 60. These results suggest the importance of considering all information (e.g., social history, performance on other tests) when making a determination whether a student with low cognitive ability performs poorly on a PVT. For example, Kirkwood et al. (2012) noted performance on neuropsychological tasks within normal limits but poor performance on ability-based tests were twice as likely for participants who had failed a validity measure.

Implications

The findings from this study support several different avenues for enhancing practice and training in the field of school psychology. As anticipated, school psychology practitioners who participated in this study endorsed little to no training or education in the area of performance validity measures. It is possible that a lack of awareness of these measures and of the broader topic of suboptimal performance may contribute to the continued use of examiner observation to assess effort. Although the concepts of malingering and suboptimal performance are more frequently explored within neuropsychological evaluations, the results of this study suggest that it is important to consider the occurrence of suboptimal performance through the use of PVTs in educational settings as well.

Specifically, school-based practitioners are encouraged to incorporate the use of performance validity measures and for training programs to teach and support the use of these instruments. However, a recent study examining the frequency of PVT use in documentation of pediatric neuropsychological evaluations was conducted (MacAllister, Vasserman, & Armstrong, 2019). When reviewing reports submitted to their practice from neuropsychologists in the surrounding region, MacAllister et al. (2019) reported that only six reports they reviewed from six different clinicians documented use of PVTs (4.58% of reports). Although Brooks, Ploetz, and Kirkwood (2016) found that 92% of practitioners report using at least one PVT, these data seem to suggest that reported practice may not match actual practice. Perhaps this finding suggests there is more need for training across all providers administering psychoeducational and neuropsychological evaluations on the importance of including effort assessments.

In the field of school psychology, motivation is commonly explored in the context of “can’t do” versus “won’t do” behaviors (VanDerHeyden & Witt, 2008). Practitioners have learned a variety of techniques to encourage the best performance from students and to differentiate skills deficits from motivational deficits. This type of work suggests an awareness that children and adolescents are capable of modifying their effort based on how motivated they are to achieve a particular goal or task. Therefore, incorporating awareness of suboptimal performance on psychoeducational evaluations is a topic that needs to be considered and further explored.

Incorporating training specific to performance validity measures as part of training programs would not only help raise awareness to the issue of suboptimal effort, but also increase practitioner’s comfort level with administering and interpreting results of objective performance validity measures. Teaching practitioners about performance validity measures alongside cognitive measures would be ideal considering they should both be administered collectively. Since existing cognitive measures do not inherently include a validity component and often carry heavy implications, ensuring optimal effort is crucial.

It is also important to note how terminology differences may play a role in limiting school psychologists’ knowledge of suboptimal performance. In clinical settings, the term malingering is used to identify suboptimal performance. VanDerHeyden and Witt (2008) identified suboptimal performance as “won’t do” and more recently, Furlong et al. (2017) described “mischievous responding” as occurring when students exaggerated their answers to questions that would typically have a low base rate, as part of a universal health screening in a high school. That same year, Higgins et al. (2017) used the term

"sandbagging" to refer to suboptimal effort in their study of high school student athlete performance on the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT). The use of these different terms may make it difficult to develop a cohesive body of research specific to suboptimal performance in child and adolescent populations across different contexts. If consensus could be reached in reference to an agreed upon term to utilize when studying these similar and related concepts of suboptimal effort, then there might be greater awareness and knowledge of this important concept among practitioners.

Limitations

As the first known study examining suboptimal performance using the TOMM as part of a psychoeducational evaluation in a school setting, there are several limitations to consider. First, the sample size was smaller than expected. Given the number of students who participate in psychoeducational evaluations each year, it was believed that it would be easier to obtain a larger sample. The number of child and adolescent participants in previous research studies varied from 61 participants (Constantinou & McCaffrey, 2003) to as many as 193 in clinical settings (Kirkwood & Kirk, 2010). The current study had 52 student participants who were assessed by five different school psychologists. Expanding this study to a larger participant pool with greater diversity that is consistent with the demographics of the general population would allow for greater generalizability. Due to the smaller sample size, follow up analyses to detect any patterns within educational verification or ability level were not possible. With a larger sample size, the potential to further explore these areas would be more feasible.

Data were collected within a limited number of school districts across two states. Gaining approval from school district leaders represented one hurdle, subsequent to recruiting school psychologists who would agree to participate and follow through with data collection. Finally, there was the difficulty of obtaining parental/caregiver consent. Although it is not possible to estimate the total number of parents who refused to consent to this study, school psychology participants shared this refusal as a common reason for having few student participants. This same issue is not likely to be present in clinic settings as research is frequently conducted and parents sign a general consent to have their child tested. It is possible that a separate form to request permission from parents, limited the potential sample. If administration of PVTs became common practice within school settings, general data could be maintained and accessed for research purposes under the category of existing data.

The student effort rating provided by school psychologists was limited in that only three options for student effort were given. Expanding the student effort rating scale to include additional anchor points would help get a more accurate perception of student effort as rated by practitioners. Also, providing school psychologists with operational definitions of what effort would look like (i.e., behavioral descriptions) for each anchor points might allow practitioners to more accurately rate student effort.

Unfortunately, psychoeducational evaluations are often completed across different sessions due to several time restrictions. Therefore, a potential limitation of this study is that testing was not always completed in one setting on the same day. Oftentimes, testing in an educational setting occurs across different days for a multitude of reasons such as testing fatigue for the student, reduction of time out of core classes, and the schedule of

the school psychologist. In clinical settings, all evaluation measures are typically administered on a single day. It is possible that breaking up the testing time frames could have potentially affected both student scores on the TOMM and practitioner ratings of observed effort.

Finally, the Retention Trial on the TOMM was not utilized as part of this study. Asking school psychologists and parent/caregivers to consent to additional time out of the classroom for psychoeducational evaluation completion was difficult and by eliminating the Retention Trial, the total assessment time was reduced. Although consistent with previous research studies indicating administration of the Retention Trial was not necessary to determine a passing score on the TOMM and detect overall suboptimal performance, it might have provided additional information.

Future Research

Future research is recommended to address some of the noted limitations and further, to evaluate the clinical utility of established child and adolescent based performance validity measures. For example, school-based practitioners might be more confident to utilize an assessment normed for children and adolescents, such as the Memory Validity Profile (MVP) (Sherman & Brooks, 2016). Administering both the TOMM and MVP to a sample population (ages 5-17) would help to ensure appropriate sensitivity and specificity of the MVP. A recent study attempted to establish these components of the MVP and reported concerns regarding cut scores as being “insensitive to non-credible performance” (Dodd et al., 2020, p. 141). These findings suggest the importance of additional study is warranted on the MVP before it can be offered as an alternative to the TOMM (Dodd et al., 2020).

Additionally, the idea of "can't do" versus "won't do" behaviors were discussed as a procedure for eliciting maximum effort from students in testing situations. When considering future efforts, the idea of applying this concept to a psychoeducational evaluation is intriguing. Although it is not possible to "re do" the assessment with an incentive, there are options of testing the limits in evaluations. That is, practitioners can attempt to determine the ability level of the examinee by extending beyond a ceiling once the standardized assessment is complete. Further investigation as to how to apply this concept and whether it reveals potential patterns of suboptimal performance is needed. For example, practitioners might administer an objective performance validity measure and test the limits of specific assessments with individuals to determine whether there is an inverse relationship between motivated performance (e.g., with encouragement) and a PVT. This process might provide important information as to whether a relationship exists between PVTs and how individuals respond when limits are tested.

In light of the small sample of participants who failed Trial 2 on the TOMM, additional exploration into whether or not primary language and racial/ethnic differences exist among participants who fail Trial 2 would be helpful. In general, no information was found that disaggregated samples by these demographic variables and such research would add to the literature base. Again, meaningful findings in these areas were not able to be made, due to sample size, but would certainly be meaningful in using and interpreting PVTs across different groups.

Conclusions

This study represented one of the first known study to explore suboptimal effort in children and adolescents as part of psychoeducational evaluations within a school

setting. These preliminary findings suggested that students may demonstrate suboptimal effort at a rate that is consistent with those identified in clinic settings. The apparent alignment of percentages of suboptimal performance in clinical and school settings among children and adolescents is both affirming and alarming. The idea that the rates of suboptimal performance does not vary greatly across settings suggests a potential baseline rate for suboptimal performance among child and adolescent populations. This information is also alarming given the high rate and potentially negative consequences of misidentifying students for special education services. Overall, this study provides further support for the inclusion of an objective performance validity measure for school psychologists in order to improve their ability to identify when students might be demonstrating suboptimal performance.

REFERENCES

Ahern, E. C., Lyon, T. D., & Quas, J. A. (2011). Young children's emerging ability to make false statements. *Developmental Psychology*, 47, 61-66.

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, DC: American Psychiatric Association.

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

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: American Psychiatric Association.

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.

Brooks, B. L., Ploetz, D. M., & Kirkwood, M. W. (2016) A survey of neuropsychologists' use of validity tests with children and adolescents, *Child Neuropsychology, 22*, 1001-1020. DOI: 10.1080/09297049.2015.1075491

Brooks, B. L., Sherman, E. M. S., & Krol, A. L. (2012). Utility of TOMM Trial 1 as an indicator of effort in children and adolescents. *Archives of Clinical Neuropsychology, 27*, 23-29.

Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., Reynolds, C. R., & Silver, C. H. (2005). Symptom validity assessment: Practice

issues and medical necessity NAN policy & planning committee. *Archives of Clinical Neuropsychology, 20*, 419-426.

Chafetz, M., & Underhill, J. (2013). Estimated costs of malingered disability. *Archives of Clinical Neuropsychology, 28*, 633-639.

Chafetz, M. D., Abrahams, J. P., & Kohlmaier, J. (2007). Malingering on the Social Security Disability Consultative Exam: A new rating scale. *Archives of Clinical Neuropsychology, 22*, 1-14. doi:10.1016/j.acn.2006.10.003

Chafetz, M. D., Williams, M. A., Ben-Porath, Y. S., Bianchini, K. J., Boone, K. B., Kirkwood, M. W., Larrabee, G. J., & Ord, J. S. (2015). Official position of the American Academy of Clinical Neuropsychology Social Security Administration policy on validity testing: Guidance and recommendations for change. *The Clinical Neuropsychologist*, 29, 723-740.

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.

Crossman, A. M., & Lewis, M. (2006). Adults' ability to detect children's lying. *Behavioral Sciences and the Law, 24*, 703-715.

DeRight, J., & Carone, D. A. (2015). Assessment of effort in children: a systematic review. *Child Neuropsychology, 21*, 1-12.

Dodd, J. N., Murphy, S., & Bosworth, C. (2020). Sensitivity of the Memory Validity Profile (MVP): Raising the bar. *Child Neuropsychology, 26*, 137-144.

Faust, D., Hart, K., & Guilmette, T. J. (1988). Pediatric malingering: the capacity of children to fake believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 56*, 578-582.

Furlong, M. J., Fullchange, A., & Dowdy, E. (2017). Effects of mischievous responding on universal mental health screening: I love rum raisin ice cream, really I do! *School Psychology Quarterly, 32*, 320-335.

Gibbs, J. C. (2014) *Moral Development and Reality: Beyond the Theories of Kohlberg, Hoffman, and Haidt, Third Edition* (pp. 3-21). New York, New York: Oxford University Press.

Green, P. (2004). *Manual for Green's Medical Symptom Validity Test (MSVT).* Edmonton: Green's Publishing Inc.

Green, P., & Astner, K. (1995). *The Word Memory Test*. Edmonton: Neurobehavioural Associates.

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

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference participants. (2009). American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, 23, 1093-1129.

Higgins, K. L., Denney, R. L., & Maerlender, A. (2017). Sandbagging on the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) in a high school athlete population. *Archives of Clinical Neuropsychology*, *32*, 259-266.

Ho, A. (2004). To be labelled, or not to be labelled: that is the question. *British Journal of Learning Disabilities, 32,* 86-92.

Institute of Medicine. (2015). Psychological testing in the service of disability determination. Washington, DC: Institute of Medicine. Retrieved from http://www.nationalacademies.org/hmd/~/media/Files/Report%20Files/2015/SSA_RB.pdf

Kirk, J. W., Harris, B., Hutaff-Lee, C. F., Koelemay, S. W., Dinkins, J. P., & Kirkwood, M. W. (2011). Performance on the Test of Memory Malingering (TOMM) among a large clinic-referred pediatric sample. *Child Neuropsychology, 17,* 242-254.

Kirkwood, M. W. (2015a). A rationale for performance validity testing in child and adolescent assessment. In M. W. Kirkwood (Ed.), *Validity testing in child and adolescent assessment: Evaluating exaggeration, feigning, and noncredible effort* (pp. 3-21). New York: Guilford.

Kirkwood, M. W. (2015b). Pediatric validity assessment. *NeuroRehabilitation, 36,* 439-450.

Kirkwood, M. W., & Kirk, J. W. (2010). The base rate of suboptimal effort in a pediatric mild TBI sample: Performance on the medical symptom validity test. *The Clinical Neuropsychologist, 24,* 860-872.

Kirkwood, M. W., Kirk, J. W., Blaha, R. Z., & Wilson, P. (2010). Noncredible effort during pediatric neuropsychological exam: A case series and literature review. Child Neuropsychology, 16, 604-618. doi: 10.1080/09297049.2010.495059

Kirkwood, M. W., Yeates, K. O., Randolph, C., & Kirk, J. W. (2012). The implications of symptom validity test failure for ability-based test performance in a pediatric sample. *Psychological Assessment, 24*, 36-45.

Larrabee, G. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society, 18*, 625-631.

Lechtenberger, D. (2010). Education for All Handicapped Children Act of 1975. In C. S. Clauss-Ehlers (Ed.), *Encyclopedia of Cross-Cultural School Psychology* (pp. 412-413). New York: Springer Science+Business Media, LLC.

Lu, P., & Boone, K. B. (2002). Suspect cognitive symptoms in a 9-year-old child: Malingering by proxy? *The Clinical Neuropsychologist, 16*, 90-96.

MacAllister, W. S., Vasserman, M., & Armstrong, K. (2019). Are we documenting performance validity testing in pediatric neuropsychological assessments? A brief report. *Child Neuropsychology, 25*, 1035-1042.

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

Sherman, E., & Brooks, B. (2016). *Memory validity profile (MVP)*. Lutz, FL: PAR.

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

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). A compendium of neuropsychological tests: administration, norms, and commentary, Third Edition. New York, New York: Oxford University Press.

Strömwall, L. A., Anders Granhag, P., & Landström, S. (2007). Children's prepared and unprepared lies: Can adults see through their strategies? *Applied Cognitive Psychology, 21*, 457-471.

Talwar, V., & Lee, K. (2002). Emergence of white-lie telling in children between 3 and 7 years of age. *Merrill-Palmer Quarterly, 48*, 160-181.

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

U.S. Department of Education. (2019, July 10). *ED Facts Data Warehouse (EDW): IDEA Part B Child Count and Educational Environments Collection, 2018-19.* https://www2.ed.gov/programs/osepidea/618-data/static-tables/index.html

VanDerHeyden, A. M., & Witt, J. C. (2008). Best practices in can't do/won't do assessment. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology, fifth edition* (pp. 131-140). Bethesda, MD: National Association of School Psychologists.

Weber Ku, E., Oliveira, J. S., Cook, N. E., McCurdy, K., Kavanaugh, B., Cancilliere, M. K., & Holler, K. A. (2020). Assessing performance validity with the TOMM and automatized sequences task in a pediatric psychiatric inpatient setting. *Child Neuropsychology*, DOI: 10.1080/09297049.2020.1712345.

Weiss, C. L. A., & Mattrick, J. E. (2010). Individuals with Disabilities Education Act (IDEA). In C. S. Clauss-Ehlers (Ed.), *Encyclopedia of Cross-Cultural School Psychology* (pp. 542-545). New York: Springer Science+Business Media, LLC.

Wilson, A. E., Smith, M. D., & Ross, H. S. (2003). The nature and effects of young children's lies. *Social Development*, 12, 21-45.

Wright, P. W. D., & Wright, P. D. (2018). *The history of special education law*. Retrieved from http://www.wrightslaw.com/law/art/history.spec.ed.law.htm

APPENDIX A

Student Effort Rating Form

Please circle the level of effort you believe this student exerted as part of this psychoeducational battery.

Full Effort	Partial Effort	Little Effort
3	2	1

Please list behavior(s) contributing to your conclusion of effort:

1.

2.

3.

4.

5.

Student Effort Rating Form

Please circle the level of effort you believe this student exerted as part of this psychoeducational battery.

Full Effort	Partial Effort	Little Effort
3	2	1

Please list behavior(s) contributing to your conclusion of effort:

1.

2.

3.

4.

5.

APPENDIX B

Fidelity Checklist

— Obtain parental/caregiver consent *(Parents/caregiver keeps one copy, you keep the other)*

— Obtain student assent *(Verbal/Written, depending on age. See Student Assent Forms.)*

— Administer first assessment of your psychoeducational evaluation

— Complete rating of student effort

— Administer Trial 1 of TOMM

— Administer Trial 2 of TOMM

— Complete additional testing, as part of psychoeducational evaluation

— Enter deidentified student demographic information in provided Excel document

— Enter all scores (subtest and composite) for all intelligence and academic achievement tests in Excel

— Enter BASC-3 validity scores (if given as part of the psychoeducational evaluation)

— Retain possession of parental/caregiver consent and student assent forms in provided file

— Email Kayla Singleton (sing5329@bears.unco.edu) Excel file

Please contact me with any questions or concerns at sing5329@bears.unco.edu.

APPENDIX C

Participant Number
Age (year, month)
Sex
Grade
Race
Ethnicity
Primary Language
Prior Medical or Mental Health Diagnosis
Initial or Re-Evaluation
Educational Verification
Practitioner Rating of Effort
TOMM: Trial 1
TOMM: Trial 2
WISC V (VCI)
WISC V (VSI)
WISC V (FRI)
WISC V (WMI)
WISC V (PSI)
WISC V (FSIQ)
WISC V BD
WISC V VP
WISC V SI
WISC V MR
WISC V FW
WISC V DS
WISC V CD
WISC V VC
WISC V PS
WISC V SS
BASC Validity Scales for Self-Report
WIAT III Scores or WJ-IV Scores

APPENDIX D

Institutional Review Board

CONSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH
UNIVERSITY OF NORTHERN COLORADO

Project Title: Exploring the Prevalence of Suboptimal Effort Among Children and Adolescents on Psychoeducational Evaluations

Researcher: Kayla J. Singleton, Ed.S., NCSP
E-mail: sing5329@bears.unco.edu
Research Advisor: Robyn S. Hess, Ph.D., ABPP (School)
Email: robyn.hess@unco.edu

Hello,

I am researching the usefulness of measures of student effort, as part of a psychoeducational evaluation. This test was developed for use with adults and more recently has been utilized by providers working with children and adolescents. The support for the use of this assessment with children and adolescents has been well established within a clinical setting (i.e., hospital). However, at this time there are no studies that have attempted to evaluate its usefulness as part of a psychoeducational evaluation in a school setting. This test helps evaluators to ensure that students are putting forth their best effort.

Your student has been referred for an initial or three-year reevaluation through their school team. If you grant permission and if your student indicates a willingness to participate, your student will be administered one additional assessment as part of the psychoeducational evaluation. This assessment is very similar in its presentation to a memory test and will be described to your student as an activity similar to a memory game. It will require an additional 15 to 20 minutes to administer. The evaluator within your student's school will administer this assessment as part of the selected battery. Since we do not know what the results of this assessment could mean at this time, we would not want to misinterpret the data. Therefore, the results of this assessment will not be provided with the rest of the assessment results.

The school evaluator will not provide identifying information about your child. Rather, deidentified demographic and assessment information will be shared. Your student's name will be withheld, and a numerical value will be used in its place. All information exchanged between myself and the school evaluator will be shared in a password protected document, as an added layer of security.

_____ initials, pg. 1 of 2

I foresee very little risk to your student; no more risk than typically encountered on a school day that includes testing. The only discomfort could potentially be boredom or mild fatigue from the extended amount of time (15 to 20 minutes) to administer the test.

Participants will indirectly benefit from this study, as the information will add to the research base for student effort and potentially spark a new line of research in student effort research within a psychoeducational evaluation in a school setting.

Please feel free to email me if you have any questions or concerns about this research and please retain one copy of this letter for your records.

Thank you for assisting me with my research.

Sincerely,

Kayla J. Singleton

Participation is voluntary. You may decide not to allow your student to participate in this study and if (s)he begins participation you may still decide to stop and withdraw at any time. Your decision will be respected and will not result in loss of benefits to which you are otherwise entitled. Having read the above and having had an opportunity to ask any questions, please sign below if you would like to participate in this research. A copy of this form will be given to you to retain for future reference. If you have any concerns about your selection or treatment as a research participant, please contact Nicole Morse, IRB Administrator, Office of Sponsored Programs, Kepner Hall, University of Northern Colorado Greeley, CO 80639; 970-351-1910.

_____

Child's Full Name (please print)

_____

Parent/Guardian's Signature Date

_____

Researcher's Signature Date

APPENDIX E

Institutional Review Board

ASSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH
UNIVERSITY OF NORTHERN COLORADO

Hello,

My name is Kayla Singleton and I'm a student at the University of Northern Colorado. I am going to college to learn more about how psychologists can help students learn. In order to help me learn more about how to help students, like you, I am asking for your help to do an activity. This activity is like a memory game. You will be shown pictures, one at a time. Then, you will be asked to point to the picture that you saw before.

If you want to complete this activity, it will take about 15 to 20 minutes. Your answers will not be shared with anyone at your school, other than me and the person who will do the activity with you. I won't even have them write your name down. We will ask your teacher for the best time to talk with you so that you don't miss anything too important.

Your parent(s) have said it's okay for you to complete the activity, but you do not have to. It's up to you. Also, if you say "yes" but then change your mind, you can stop any time you want to.

If you want to be in my research, sign your name below and write today's date next to it.

Thank you!

_____

Student's Name Date

_____

Witness Signature Date

_____

Researcher's Signature Date

APPENDIX F

UNIVERSITY OF NORTHERN COLORADO

Institutional Review Board

ASSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH
UNIVERSITY OF NORTHERN COLORADO

Hello,

My name is Kayla Singleton and I'm a student at the University of Northern Colorado. I am going to college to learn more about how psychologists can help students learn. In order to help me learn more about how to help students, like you, I am asking for your help to do an activity. This activity is like a memory game. You will be shown pictures, one at a time. Then, you will be asked to point to the picture that you saw before.

If you want to complete this activity, it will take about 15 to 20 minutes. Your answers will not be shared with anyone at your school, other than me and the person who will do the activity with you. I won't even have them write your name down. We will ask your teacher for the best time to talk with you so that you don't miss anything too important.

Your parent(s) have said it's okay for you to complete the activity, but you do not have to. It's up to you. Also, if you say "yes" but then change your mind, you can stop any time you want to.

If you want to be in my research, sign your name below and write today's date next to it.

Thank you!

_____

Student's Name Date

_____

Researcher's Signature Date

APPENDIX G

Institutional Review Board

CONSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH
UNIVERSITY OF NORTHERN COLORADO

Project Title: Exploring the Prevalence of Suboptimal Effort Among Children and Adolescents on Psychoeducational Evaluations

Researcher: Kayla J. Singleton, Ed.S., NCSP
E-mail: sing5329@bears.unco.edu
Research Advisor: Robyn S. Hess, Ph.D., ABPP (School)
Email: robyn.hess@unco.edu

Purpose and Description: The primary purpose of this study is to determine the rate at which children and adolescents completing a psychoeducational evaluation (initial or three-year re-evaluation) display suboptimal performance utilizing the Test of Memory Malingering (TOMM). The TOMM is a widely known and utilized assessment that was designed to detect suboptimal performance within an adult population. However, recent research in clinical pediatric populations has shown that children and adolescents, as young as 5 years of age, are also able to pass the TOMM using the adult normed criteria. This is the first known study to examine suboptimal performance, as part of a psychoeducational evaluation within a school setting. As such, the clinical utility of the TOMM is unknown within a school setting, as part of a psychoeducational evaluation. Therefore, the interpretation of the findings of this assessment will be unclear until all data is collected and analyzed. This data should not be interpreted or included in the findings of the psychoeducational evaluations, due to the unknown utility at this time.

As licensed school psychologists, interns, and practicum students, you will be asked to include the TOMM in your routine psychoeducational evaluations. Either a video or in person training with the researcher will take place to ensure you have been trained to administer the TOMM in a standardized manner. You will be asked to administer the TOMM second in your battery. Following the administration of the first test in your battery, you will be asked to complete a rating form. The rating form is based on your observations of the student you just completed a test with, and to what degree of effort you believe they put forth on that test. There are three potential options 1) little effort, 2)

partial effort, or 3) full effort. In addition to your rating, please list any behavior(s) that led you to conclude the level of effort displayed by the student. Then, you will administer the TOMM, which consists of two trials. Both trials will be administered to all students (to whom consent, and assent has been provided). The administration of both trials 1 and 2 of the TOMM should not exceed 20 minutes. Then, I ask that you complete your routine battery as you normally would.

Additionally, you will be asked to enter the results of the psychoeducational evaluation into an Excel spreadsheet provided by the researcher. Deidentified information including age, gender, grade, primary language, race, ethnicity, and medical and/or mental health diagnosis, all intelligence and academic subtest and composite scores, TOMM trial 1 and 2 scores, practitioner rating of overall student effort, and all self-reported BASC validity ratings (when administered as part of your typical battery), will be typed into the password protected Excel document. On the first of each month, you will email the Excel spreadsheet to the researcher. The researcher will retain possession of the password protected Excel spreadsheets. A detailed protocol will be provided as a checklist to ensure all steps have been completed for each participating student.

The researcher will provide you with the stimulus materials needed to administer the TOMM, as well as the rating form for student effort. At the conclusion of this study, you may retain ownership of the TOMM, as a gift for your participation in the study. Additionally, you will receive a $25 gift card for your help and participation in this study.

The potential risks of this study are minimal. The risks inherent in this study are no greater than the potential benefits of the data that will be collected. There is potential for some discomfort in adding an additional measure to a psychoeducational evaluation, as the additional measure will increase the amount of time needed to complete the battery. However, the need for the information that this additional measure will provide outweighs any potential discomforts, as the additional time should only be increased by a maximum of twenty minutes. This information could be used to better serve the needs of practitioners in a school setting. Information collected in this study might help school teams to develop guidelines or strategies to detect suboptimal performance of students. As a participant in this study you may benefit from learning a new assessment (e.g., TOMM).

Participation is voluntary. You may decide not to participate in this study and if you begin participation you may still decide to stop and withdraw at any time. Your decision will be respected and will not result in loss of benefits to which you are otherwise entitled. Having read the above and having had an opportunity to ask any questions, please sign below if you would like to participate in this research. A copy of this form will be given to you to retain for future reference. If you have any concerns about your selection or treatment as a research participant, please contact Sherry May, IRB

Administrator, Office of Sponsored Programs, Kepner Hall, University of Northern Colorado Greeley, CO 80639; 970-351-1910.

_____

Subject's Signature Date

_____

Researcher's Signature Date