2016

Exploring The Construct Validity Of The Behavioral Rating Inventory Of Executive Function (Brief): Comparing Variance Associated With Common Method Versus Similar Constructs

Lauren K. Hoefling
Indiana State University

LAUREN K. HOEFLING

EDUCATION	PsyD, Clinical Psychology, 2016 Indiana State University, Terre Haute, IN
	MS, Clinical Psychology, 2013 Indiana State University, Terre Haute, IN
	BS, Clinical Psychology, 2011 University of Southern Indiana, Evansville, IN
CLINCIAL EXPERIENCE	Predoctoral Psychology Intern, 2015–Present Hennepin County Medical Center, Minneapolis, MN
	Graduate Clinician/Behavioral Health Provider, 2015 Valley Professionals Community Health Center, Clinton, IN
	Graduate Clinician, 2015 Union Hospital Family Medicine Center, Terre Haute, IN
	Graduate Clinician, 2013–2014 Murphy, Urban, and Associates, Terre Haute, IN
	Graduate Clinician, 2012–2014 Indiana State University Psychology Clinic, Terre Haute, IN
	ADHD Evaluator, 2012–2013 Indiana State University ADHD Clinic, Terre Haute, IN
	Social Skills Group Cofacilitator, Fall 2012 Ryves Hall Community Center, Terre Haute, IN
WORK EXPERIENCE	Peer Supervisor, 2014–2015 Indiana State University, Terre Haute, IN
	Course Instructor, PSY 384 Interviewing Skills, 2012–2013 Indiana State University, Terre Haute, IN
	Teaching Assistant: PSY 344 Cognitive Psychology, PSY 376 Advanced Research Methods, PSY 310 Learning, 2011–2012 Indiana State University, Terre Haute, IN

EXPLORING THE CONSTRUCT VALIDITY OF THE BEHAVIORAL RATING INVENTORY OF EXECUTIVE FUNCTION (BRIEF): COMPARING VARIANCE ASSOCIATED WITH COMMON METHOD VERSUS SIMILAR CONSTRUCTS

A Dissertation

Presented to

The College of Graduate and Professional Studies

Department of Psychology

Indiana State University

Terre Haute, Indiana

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Psychology

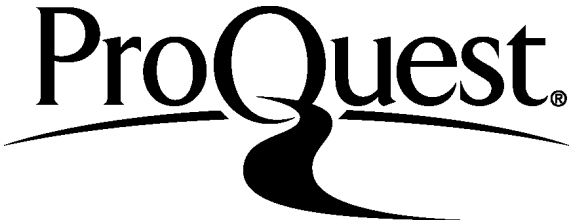
---

by

Lauren K. Hoefling

August 2016

Keywords: BRIEF, ADHD, executive function



COMMITTEE MEMBERS

Committee Chair: Liz O'Laughlin, Ph.D.

Professor of Clinical Psychology

Indiana State University

Committee Member: Jennifer Schriver, Ph.D.

Professor of Clinical Psychology

Indiana State University

Committee Member: Veanne Anderson, Ph.D.

Professor of Experimental Psychology

Indiana State University

ABSTRACT

The Behavioral Rating Inventory of Executive Function (BRIEF) is a teacher and parent questionnaire designed to measure executive functioning (EF). Due to the role of EF in ADHD, the BRIEF is often used in assessment of ADHD. The purpose of the current study was to explore the convergence and divergence of the BRIEF with a performance-based measure of EF, behavior rating scales, an observational measure of child behavior during testing, and longer-term academic outcomes. The current study used archival data representing 105 children referred to an ADHD evaluation clinic. It was hypothesized that the BRIEF would be significantly correlated with other rating scales assessing similar and dissimilar constructs (e.g., ADHD and depression) and longer-term goal pursuit, but would not be significantly correlated with a performance-based measure of EF. In addition, it was hypothesized that parent and teacher ratings on the BRIEF would cluster with other rating measures and measures of long-term goal pursuit, whereas the performance-based measure would group with other short-term performance measures (i.e., FSIQ, ratings from an observational measure completed during the assessment). As predicted, the BRIEF correlated with other rating scales (both those assessing ADHD and depression), as well measures of longer-term goal pursuit. Also as predicted, the BRIEF was minimally correlated with the performance-based measure. Factor analyses indicated that measures tended to cluster based on type of measure (i.e., self-report, performance measure). Results of this study provide support for the premise that rating scales and performance-based measures of EF are measuring distinct constructs. Results also suggest that ratings are influenced

by source (i.e., parent or teacher) and context. Future research should consider the role of performance-based measures in assessment ADHD in general, as well as the predictive validity of rating scales and performance-based measures in predicting real-world outcomes for children with ADHD.

ACKNOWLEDGMENTS

First, I must acknowledge my family and friends. To my parents, Andy and Patty Hoefling, I never would have finished this graduate school journey without your influence. Dad, thank you for teaching me how to be a hard worker through your daily example. Mom, thank you for your encouragement to reach my full potential and the endless hours of listening and support. To all of my family and friends, thank you for keeping me sane through this process and reminding me to have fun from time to time.

Second, thank you to the Indiana State University Psychology Department for your guidance and support both academically and clinically. My growth through this program is a direct reflection of the quality of ISU's training and the genuine kindness of the faculty. Thank you to my dissertation committee, Dr. O'Laughlin, Dr. Anderson, and Dr. Schriver, for the time and energy you devoted to this project. Dr. Anderson and Dr. Schriver, thank you for the classroom experiences that laid the foundation for this project and your feedback throughout the process. Finally, thank you to Dr. O'Laughlin. I appreciate the time and energy you dedicated to this project. You've greatly influenced my growth both as a clinician and researcher during my time at ISU and this influence is something I will take with me as I leave ISU.

TABLE OF CONTENTS

COMMITTEE ii

ABSTRACT iii

ACKNOWLEDGMENTS v

LIST OF TABLES viii

OVERVIEW 1

LITERATURE REVIEW 5

Overview of ADHD 5

Assessment of ADHD 8

Executive Function 10

Models of the Role of EF in ADHD 13

Associated Features of ADHD and EF 16

Academic Performance 16

Measures of EF 20

Present Study 34

Hypotheses 36

METHOD 38

Design 38

Participants 38

Measures 40

Behavioral Rating Inventory of Executive Function 40

Conners Continuous Performance Task....42
Attention Deficit Hyperactivity Disorder Rating Scale-IV....45
Teacher Ratings of Academic Performance....47
Test Observation Form....47
Behavioral Assessment System for Children-2....48
Wechsler Abbreviated Scale of Intelligence-II....50
Procedure....50
RESULTS....52
Overview....52
Preliminary Analyses....52
Correlational Analyses....53
Factor Analyses....55
DISCUSSION....58
Clinical Implications....64
Limitations....70
Future Research....73
REFERENCES....78
APPENDIX A: DIAGNOSTIC CRITERIA FOR ADHD (DSM-5)....89
APPENDIX B: BRIEF SAMPLE ITEMS....92
APPENDIX C: ADHD RATING SCALE-IV SAMPLE ITEMS....93
APPENDIX D: TEST OBSERVATION FORM SAMPLE ITEMS....94
APPENDIX E: TEACHER RATINGS OF ACADEMIC PERFORMANCE....95
APPENDIX F: TABLES....96

LIST OF TABLES

Table 1. Participant Demographic Information 96

Table 2. Mean, Standard Deviation, and Range for Parent/Teacher Questionnaires and Performance Measures 97

Table 3. Pearson Correlation Coefficients Between BRIEF Scales and Other Variables in Correlational Analyses (Parent) 98

Table 4. Pearson Correlation Coefficients Between BRIEF Scales and Other Variables in Correlational Analyses (Teacher) 99

Table 5. Pearson Correlation Coefficients Between Parent and Teacher Ratings 100

Table 6. Factor Analysis Including Teacher Ratings 101

Table 7. Factor Analysis Including Parent Ratings 102

CHAPTER 1

OVERVIEW

Attention-deficit hyperactivity disorder (ADHD) is a neurobiological disorder characterized by symptoms of inattention, hyperactivity, and impulsivity that occur in a minimum of two settings and cause impairment in academic, social, or occupational functioning (American Academy of Pediatrics, 2011; American Psychiatric Association, 2013). Despite substantial research regarding ADHD and how to assess it, there is currently no definitive test of the disorder (Jarratt, Ricio, & Siekierski, 2005). As such, diagnostic procedures typically include some combination of background information, diagnostic interview, direct observation, informant ratings (ideally from multiple sources), and performance-based measures of attention and impulse control (Sims & Lonigan, 2012). Diagnostic procedures also typically include measures that assess executive functioning.

Executive functioning (EF) has been defined as "the ability to maintain an appropriate problem-solving set for attainment of a future goal" (Welsh & Pennington, 1998, p. 201-202). Deficits in EF, predominantly those specific to working memory and response inhibition, have long been associated with ADHD (Barkley, 2011). However, numerous concerns have been brought forth regarding the role of EF in assessment of ADHD. One concern is whether the measures used to assess EF (i.e., rating measures and performance-based measures) actually assess the same underlying construct. Rating measures of executive function involve reports

from informants regarding difficulty with everyday tasks, whereas performance-based measures are standardized measures given by an examiner that typically assess response time and/or accuracy (Toplak, West, & Stanovich, 2013). A recent meta-analysis examined the relationship between rating scale measures of EF and performance-based measures of EF (Toplak et al., 2013). Overall, the researchers found that these measures were only minimally correlated and suggested that the two measures may be assessing distinct mental constructs. Specifically, it was suggested that rating measures might assess how well an individual can achieve goals in day-to-day life, whereas performance-based measures of EF assess cognitive efficiency under ideal conditions.

The Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000) is the most commonly used rating scale to assess EF (Toplak et al., 2013). The BRIEF asks parents and teachers to rate children's behavior in day-to-day situations over the previous six months. Thus, the BRIEF provides ecologically valid information regarding executive functioning in everyday, complex, problem-solving situations (Gioia et al., 2000).

McAuley, Chen, Goos, Schachar, and Crosbie (2010) explored the association between the BRIEF and performance-based measures of EF, teacher and parent rating scales assessing symptoms of ADHD, questionnaires assessing impairment in social-emotional and behavioral domains, and measures of academic proficiency. Interestingly, performance on the BRIEF was not related to results on the performance-based measures of EF but was strongly correlated with both teacher and parent ratings of impairment and behavioral disruption. McAuley et al. (2010) suggested that the relationship between the rating measure of EF and parent and teacher ratings might be due to a common core trait. However, they also note that the relationship could be due to the use of a common method of assessment (i.e., rating scales). In other words, the researchers

pose the question of whether the relationship is due to trait variance or to method variance. The researchers recommend that future research use a multitrait-multimethod design to further clarify whether the positive association between ratings of EF and parent and teacher ratings of impairment and behavioral problems is due to a common method, common trait, or both. Additionally, McAuley et al. recommend that future research confirm the validity of the BRIEF by exploring its relationship to naturalistic tasks that involve applying executive function skills in more realistic and complex settings.

Traditionally, discrepant information from raters and measures has been viewed as problematic and typically attributed to measurement error and/or rater bias (De Los Reyes & Kazdin, 2005). However, given that children may behave quite differently in different settings, it is possible that these measures are accurately reflecting true discrepancies in behavior across settings (Dirks, De Los Reyes, Briggs-Gowan, Cella, & Wakschlag, 2012). In addition, it is possible that rating measures of EF and performance-based measures of EF may be reflecting different aspects of the same construct (Toplak et al., 2013).

The goal of the current study was to explore the convergence and divergence of the BRIEF with performance-based measures of EF, as well as academic, behavioral, and cognitive measures. This was essentially a replication and extension of McAuley et al. (2010). In addition to exploring the relationship between the BRIEF and a performance-based measure of EF, the present study also considered measures of longer-term outcomes (e.g., teacher ratings of academic performance and work completion) as well as an observational measure of child behavior during testing. The purpose was to explore the convergence and divergence of measures considering common trait, common method (i.e., EF), and measures of EF across time in situations of varying structure versus measures of EF in a specific structured task. This research

also sought to increase understanding of the consistencies and inconsistencies that are frequently observed in ADHD assessment by considering the relations between measures that assess behavior across different types of settings.

CHAPTER 2

LITERATURE REVIEW

Overview of ADHD

Attention deficit hyperactivity disorder (ADHD) is a developmental disorder characterized by a pattern of symptoms related to inattention, hyperactivity, and impulsivity (American Psychiatric Association, 2013). The Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (DSM-5) reports prevalence rates of 5% for children and 2.5% for adults in most cultures. Similarly, a review article cited worldwide prevalence rates at 5.29% (Polanczyk, Silva de Lima, Horta, Biederman, & Rohde, 2007). A number of factors can impact diagnosis of ADHD. For example, Polanczyk et al. (2007) noted that source of information, diagnostic criteria, geographic origin, and whether impairment was required for diagnosis significantly influenced prevalence rates. In addition, prevalence rates typically vary substantially as a function of gender (Barkley, 2006). In general, boys are approximately three times as likely to be diagnosed with ADHD as compared to girls. In clinic-referred children, boys are approximately five to nine times more likely to have the disorder in comparison to girls. ADHD diagnosis decreases across development in both sexes (Barkley, 2006).

In 1798, Dr. Alexander Crichton published the first description of disorders related to attention (Barkley, 2009). Crichton viewed attention as central to awareness, as a deliberate activity, and as effortful, rather than automatic. However, George Still and Alfred Tredgold are

typically credited with highlighting the disorder we currently know as ADHD (Barkley, 2006). In his 1902 lecture, George Still described clinical cases of 43 children who had difficulty with attention and a "defect in moral control" that he considered relatively chronic (cited in Barkley, 2009, pg. 1). Still described moral control as behavior regulation for the moral good of everyone. In addition, he described these children as overactive, defiant, resistant, excessively emotional, and resistant to discipline (Barkley, 2009). In terms of etiology, Still believed that the disorder had a biological basis but also proposed that it was the result of complications at birth in some cases. He also hypothesized that the characteristics of the disorder (i.e., deficits in attention, moral control, and inhibitory volition) were interrelated, and he attributed them to a common underlying neurological deficiency (Barkley, 2006). According to Still, these cases could be treated through environmental modifications, special education environments, and medications. However, he noted that improvements from these treatments were often only temporary.

Since Still's time, the disorder we know as ADHD has been characterized in a variety of ways. During the 1930s, impulsivity, hyperkinesis, short attention span, and learning disabilities were described as "minimal brain damage" and then "minimal brain dysfunction" due to commonalities between these disorders and other central nervous system disorders (Spencer, Biederman, & Mick, 2007). In the 1950s, the label was amended to "hyperkinetic impulse disorder" and thought to be due to poor filtering of information that resulted in excess stimulation of the brain (Barkley, 2006).

In the early 1960s, many began to challenge the notion that the disorders described as "minimal brain dysfunction" were actually due to brain damage in children (Barkley, 2006). Slowly, the term "minimal brain damage" was exchanged for terms such as "dyslexia," "learning disabilities," "language disorders," and "hyperactivity." This change reflected the thought that

the disorders could be more accurately described by the deficits associated with them rather than the underlying etiological mechanism thought to be the cause. Simultaneously, researchers in the field of ADHD focused their attention on hyperactivity, which was thought to be the key feature of the disorder at the time. This is reflected in the DSM-II (American Psychiatric Association, 1968), which described ADHD as a disorder of restlessness, overactivity, short attention span, and distractibility that typically abated in adolescence. The DSM-III (American Psychiatric Association, 1980) represented a shift in the conceptualization of ADHD in that inattention and impulsivity were also recognized as important aspects of the disorder (Barkley, 2006). In addition, the DSM-III also noted the developmental variability of the disorder across the lifespan (Spencer et al., 2007).

Hyperactivity, impulsivity, and inattention have remained the defining features of ADHD in the current DSM-5 (American Psychiatric Association, 2013). The DSM-5 specifies three presentations of ADHD: combined presentation, predominantly inattentive presentation, and predominantly hyperactive/impulsive presentation. In all presentations, the symptoms must be present prior to age 12, be present for at least 6 months, be present in at least two settings (e.g., at school, work or home; with relatives or friends; or in other activities), cause impairment in academic, social, or occupational functioning, and not be the result of another mental disorder. For the combined presentation, the individual must have at least 6 symptoms of inattention and at least 6 symptoms of hyperactivity and impulsivity. For the predominately inattentive presentation, the individual must have at least 6 symptoms of inattention but less than 6 symptoms of hyperactivity/ impulsivity. For the predominately hyperactive/impulsive presentation, the individual must have at least 6 symptoms of hyperactivity/impulsivity but less than 6 symptoms of inattention. Five, rather than six, symptoms of inattention and/or

hyperactivity/impulsivity are required for diagnosis of individuals older than 17. The DSM-5 also includes specifiers regarding remission status and severity of the disorder (See Appendix A for a listing of DSM-5 symptoms of ADHD).

Assessment of ADHD

The primary goal of an ADHD evaluation in clinical settings is to determine the presence or absence of ADHD symptoms and gain enough information to make a differential diagnosis from other psychiatric disorders (Barkley, 2006). The underlying goal of assessment is generally to inform treatment and monitor progress or outcome of treatment (Pelham, Fabiano, & Massetti, 2005). Despite substantial research regarding evaluation of ADHD, no "gold standard" has been identified to diagnose the disorder (Jarratt et al., 2005). This is due, in part, to the lack of definitive information regarding the cognitive or biological aspects of ADHD (Pelham et al., 2005). For this reason, ADHD is currently defined by behaviors, and assessment procedures are centered on observable behaviors as noted by parent or teacher ratings of behavior in natural settings (home and school), ratings of behavior during assessment, or performance-based measures.

The American Academy of Child and Adolescent Psychiatry (American Academy of Child and Adolescent Psychiatry, 2007) makes three recommendations in their guidelines for evaluation of ADHD. First, it is recommended that clinicians complete a clinical interview with the parent to assess the duration, severity, and frequency of ADHD symptoms as listed in the DSM. Information should also be gathered regarding age of onset, settings of impairment, and academic progress. In addition, childhood psychiatric disorders, family history and family functioning, and the child's perinatal history, medical history, developmental milestones, and mental health history should be assessed. Secondly, the American Academy of Child and

Adolescent Psychiatry (AACAP) recommends the use of behavior rating scales to assess behaviors specific to ADHD as well as behaviors associated with other psychiatric disorders. Lastly, the AACAP recommends a child interview to confirm or disconfirm the diagnosis of ADHD.

In addition to areas of assessment suggested by AACAP, tools such as intellectual and achievement tests, neuropsychological tests, and observational measures are often included in evaluation of ADHD (Barkley, 2006). Previous research has indicated that behavioral observations during the evaluation provide unique variance in predicting hyperactivity/impulsivity beyond parent ratings (McConaughy et al., 2010). Pelham et al. (2005) indicated that evidence-based measures, (i.e., measures with research support) of assessment for ADHD include structured interviews, behavioral observations, global impairment measures, ADHD rating scales, and symptom rating scales consistent with the DSM. Although Pelham et al. report that structured interviews are useful for initial evaluation, they note that they might be less practical for assessing specific areas of impairment or when repeated measurements are needed. Additionally, Wolraich et al. (2003) suggest that structured interviews do not provide incremental validity beyond rating scales (parent and teacher). Pelham et al. note that the suggestion that structured interviews do not provide incremental validity is contrary to the general consensus regarding the importance of structured interviews and recommend that the utilization of rating scales rather than structured interviews could be equally valid when cost of services is a factor

Multi-rater, multi-method assessment of ADHD is recommended to provide a more complete representation of functioning (Pelham et al., 2005). This is especially important given the known influence of method of measurement and source of ratings in correlations between

criteria and measures (Campbell & Fiske, 1959). It is interesting to note that neuropsychological measures of attention and impulsivity are hardly mentioned in the review by Pelham et al. (2005), other than to note that they are not ecologically valid in diagnosis but may be useful in understanding the role of cognition in ADHD (Pelham et al., 2005). Pelham et al. appear to use the word cognition in reference to the role of cognitive efficiency or cognitive processing in ADHD.

Measures of executive functioning (EF) are often included in ADHD assessment batteries due to the strong support for EF deficits among children with ADHD. It has been suggested that ratings of EF may provide more ecological validity as compared to performance-based measures of EF (Jarratt et al., 2005). More specifically, rating scales of EF provide information from parents and teachers about children in natural settings over extended periods of time. In contrast, neuropsychological measures tend to provide information on short-term performance in clinical settings characterized by minimal distraction and individual attention, which differs from many natural settings.

Executive Function

Executive function is a complex, multi-faceted construct. There is little consensus regarding the definition and primary components of EF (Barkley, 2011). This is especially true regarding the role of EF in ADHD. EF has been defined in a variety of ways. Welsh and Pennington (1998) define it as the ability to attain future goals through sustaining a particular problem-solving set. Barkley (2006) defines EF as "a specific class of self-directed actions by the individual that are being used for self-regulation toward the future" (Barkley, 2006, pg. 304-305). Gioia, Isquith, and Guy (2001) describe EF as a "conductor" that directs, controls, and organizes cognitive activity, behavior, and emotional responses. Thus, EF is comprised of

interrelated components that are associated with particular behaviors that move individuals toward future goals (Anderson, Jacobs, & Anderson, 2008).

While some have focused on defining EF, others have focused on describing the key elements associated with EF. Although EF is generally considered to be related to cognitive processes, it is important to note that it is also related to behavioral actions, emotional responses, and motivational aspects (Barkley, 2011). A review by Best, Miller, and Jones (2009) indicated that more than 15 components have been proposed as relating to EF in the literature. The researchers conclude that shifting attention, inhibition, working memory, and planning appear to be the fundamental components of EF. Barkley (2006) proposed that EF is comprised of nonverbal working memory, internalization of speech (verbal working memory), self-regulation of motivation/affect/arousal, and planning.

Impairments in EF, also known as executive dysfunction, present in a variety of ways. Symptoms of executive dysfunction often include impulsivity, disinhibition, inattention, reduced working memory, disorganization, difficulty planning actions in advance, poor reasoning ability, perseverative behavior, difficulties shifting from conflicting demands, failure to learn from mistakes, and a resistance to change activities (Anderson et al., 2008). In addition, executive dysfunction is associated with academic and workplace concerns (Barkley & Fischer, 2011). In terms of academic concerns, children with ADHD tend to perform worse in school than controls (Barkley, 2006). This underperformance goes beyond what would be expected based on academic achievement tests and intelligence. It has been suggested that EF deficits may contribute to the academic and intellectual underperformance frequently observed in those with ADHD (Langberg, Dvorsky, & Evans, 2013). In terms of occupational concerns, adults with ADHD are more likely to have worse job performance as rated by employers, more job

dismissals or experiences of being laid off, lower annual salaries and occupational status, and less adequacy in fulfilling job demands (Barkley, Murphy, & Fischer, 2008). Deficits in EF have been associated with mental health conditions other than ADHD including schizophrenia, traumatic brain injuries, substance use, psychopathy, obsessive-compulsive disorder, and pathological gambling (Toplak et al., 2013).

Regarding ADHD and EF, numerous studies have demonstrated increased executive dysfunction in those with ADHD as compared to controls (Biederman et al., 2006; Halperin, Trampush, Miller, Marks, & Newcorn, 2008). Frazier, Demaree, and Youngstrom (2004) reported average effect size differences of .59 standard deviations between groups of children with ADHD and control groups in terms of EF as measured by continuous performance tasks. Conversely, Jonsdottir, Bouma, Sergeant, and Scherder (2006) found no significant relationship between performance-based measures of EF and ADHD symptoms after controlling for intelligence. In addition, although EF deficits are more common in ADHD groups versus control groups, analyses at the individual level often suggest that EF deficits only occur in a minority (45%) of those with ADHD (Biederman et al., 2006). These findings have led some to suggest that ADHD is not associated with deficits in EF (Jonsdottir et al., 2006). However, others have suggested that limitations in measures of EF may be to blame (Barkley, 2011).

Barkley (2011) asserts that the heterogeneous results in regard to the relation between EF and ADHD found in the literature may be due to shortcomings of the methods used to assess EF rather than a lack of relation between ADHD and EF. Many studies assessing the role of ADHD in EF have used performance-based measures of EF (Biederman et al., 2006; Halperin et al., 2008; Jonsdottir et al., 2006). As mentioned, these tests have some serious limitations, especially it terms of ecological validity. More specifically, performance-based measures do not seem to

measure the long-term, future-oriented behaviors associated with EF. A detailed overview of the limitations of performance-based measures of EF will be addressed in a subsequent section (i.e., Performance-Based Measures).

As noted previously, further understanding of the construct of EF in ADHD can be gained through considering current definitions, elements associated with the construct, symptoms and mental health conditions related with executive dysfunction, and research in the field. However, it is also fundamental to explore theories underlying the construct, especially given the inconsistent results found in studies exploring the relationship between EF and ADHD. The subsequent section will explore models related to the role of EF in ADHD, with an emphasis on a model proposed by Barkley (2006).

Models of the Role of EF in ADHD

Barkley's model of ADHD posits that deficits in behavioral inhibition underlie the deficits in executive function and motor control typically observed in ADHD (Barkley, 2006). Broadly, behavioral inhibition is the first component of Barkley's model and thought to be the foundation for executive functions. Executive functions, in turn, influence motor control and behavioral responding. According to Barkley, behavioral inhibition refers to three primary processes: inhibiting the initial response, interrupting the ongoing response, and allowing for self-directed responses while simultaneously blocking out other responses and events (interference control). Barkley notes that the initial response is often the "prepotent response," which is defined as the response that results in instant reinforcement (Barkley, 2006, p. 320). According to this model, behavioral inhibition is necessary for the development, performance, and internalization of the executive functions.

According to Barkley's model, executive functions are the self-directed responses that occur during the period of delay in behavior inhibition (Barkley, 2006). The purpose of executive functions is to internalize behavior in order to anticipate change and inform behavior based on what is expected in the future. These processes may be publicly observable in young children but tend to become more covert and private over the course of development. It is hypothesized that the development and performance of executive functions is disrupted in ADHD due to deficits in behavioral inhibition. In turn, this impairs control of goal-directed motor responses. In other words, self-regulation in those with ADHD is impaired due to inhibitory deficits that disrupt the development and internalization of executive functions. As such, ADHD is primarily a disorder of behavioral inhibition but also a disorder of executive functioning and self-regulation due to these impairments in inhibition.

Barkley (2006) also posits that two dimensions of inattention should be recognized. The first is related to distractibility. This dimension is more closely related to how we typically conceptualize the construct and consistent with the symptoms listed for inattention in the DSM-5 (American Psychiatric Association, 2013). However, Barkley claims that there is a second dimension of inattention, which he refers to as sluggish cognitive tempo (SCT). This dimension includes children who are described as having a daydreamy quality that is lethargic and passive. Barkley argues that individuals with SCT may be experiencing a different cognitive deficit than those with ADHD. A recent study found that a three-factor model including hyperactivity/impulsivity, inattention, and SCT was the best fit for both parent and teacher ratings of symptoms of ADHD and SCT, thus suggesting that SCT characteristics differ from inattentive symptoms (Bauermeister, Barkley, Bauermeister, Martinez & McBurnett, 2012). SCT was not associated with deficits in EF but was associated with difficulties in math. Measures of

neuropsychological, psychosocial, and achievement difficulties were most strongly related to symptoms of inattention. Externalizing symptoms were most strongly related to hyperactive/impulsive symptoms, whereas internalizing symptoms correlated most strongly with teacher-rated inattention and parent-rated SCT.

Sergeant (2000) also proposed a model of ADHD. According to Sergeant's model, ideal states of arousal, effort, and activation are needed to complete goal-directed motor responses. In this model, arousal refers to attentiveness that impacts how a stimulus is encoded. Effort refers to the energy required to meet the demands of a particular task. Activation refers to variations in physiological activity that alter organization and/or speed of motor response. According to Sergeant's model, children with ADHD do not experience actual executive or attentional deficits but rather experience deficits in levels of arousal, motivation, or effort (i.e., response to contingencies) needed to meet task demands.

In order to assess Barkley and Sergeant's models, Fuggetta (2006) examined differences in efficiency of EF in children with ADHD as compared to matched controls. The goal of the study was to test Barkley's model of ADHD against Sergeant's model of ADHD. Participants included males between the ages of 8 and 11. Results of the study provide support for both Sergeant's and Barkley's models. Consistent with Barkley's (2006) model, participants with ADHD demonstrated greater deficits in EF as compared to matched controls without ADHD. More specifically, those with ADHD had difficulty in tasks requiring control and monitoring, which could reflect the deficits of behavioral inhibition proposed in Barkley's model. Consistent with Sergeant's (2000) model, participants with ADHD demonstrated deficits in processing speed as compared to children without ADHD. According to Fuggetta (2006), deficits in processing speed can be interpreted as deficits in the arousal portion of Sergeant's model.

Fuggetta (2006) noted that this interpretation has been made in other studies as well. Of note, there was a low correlation between various performance-based measures of EF in the study, which supports the view of EF as a multifaceted construct.

Associated Features of ADHD and EF

Children with ADHD demonstrate a higher likelihood of developing other developmental, academic, cognitive, and health related difficulties (Barkley, 2006). In addition, educational and occupational impairment are common in those with ADHD (Anderson et al., 2008). Lambek et al. (2010) found that children with ADHD and executive dysfunction performed significantly worse on intelligence tests as compared to those with ADHD without executive dysfunction. However, the groups were comparable in terms of symptoms of ADHD and functioning in school.

Academic Performance

One of the most common features associated with ADHD is poor academic performance (Frazier, Youngstrom, Glutting, & Watkins, 2007). Children with ADHD often perform more poorly in school than what would be expected based on academic achievement tests and intelligence tests (Barkley, 2006). Such underperformance is often attributed to their restless, impulsive, and inattentive behavior in the classroom.

Frazier et al. (2007) performed two related studies to explore the relationship between ADHD and academic achievement. First, a meta-analysis was conducted to explore the relationship between ADHD and achievement in the literature. Secondly, Frazier et al. (2007) explored the relationship between ADHD and achievement in college students and compared these results to those found in the first portion of the study. In terms of the meta-analytic portion of the study, Frazier et al. found a moderate to large discrepancy between those with ADHD and

controls on a measure of academic achievement. However, given that unpublished studies were not included, this may be an overestimate of the true effect size. Specifically, moderate to large effect sizes were also found for educational measures (e.g., ADHD symptoms associated with fewer years of education and lower GPA) and behavior rating scales (i.e., student and parent ratings of inattentiveness predicted student academic probation). Medium effect sizes were found for measures related to progress in school (e.g., repeating a grade, dropping out of school), such that individuals with ADHD were more likely to have repeated a grade or dropped out of school as compared to controls. Lastly, moderator analyses indicated a trend toward reduced academic impairment as age increased. Overall, the relationship between ADHD diagnosis and measures of academic impairment (e.g., years of education, GPA, repeating a grade, dropping out of school) suggests that the deficits observed in those with ADHD generalize beyond performance on standardized achievement tests into day-to-day difficulties that result in long-term performance deficits.

Although numerous studies have documented an inverse relationship between ADHD symptoms and academic performance (Mahone et al., 2002; Muir-Broaddus, Rosenstein, Medina, & Soderberg, 2002), the mechanism underlying this relationship is less clear. As mentioned, it has been hypothesized that deficits in EF may underlie the academic impairments commonly observed in those with ADHD. However, it has been noted that previous research in this area has some significant limitations (Langberg et al., 2013). One limitation in research regarding ADHD and academic performance is related to EF. First, most studies have measured EF using only performance-based measures. This is problematic, as performance-based measures may lack ecological validity because of the structured environment in which they are administered. Performance in a structured environment may differ substantially from

performance in a more typical, less structured setting (McAuley et al., 2010). In addition, a recent review found minimal correlations between rating measures of EF and performance-based measures of EF (Toplak et al., 2013). Lastly, EF tests do not seem to encompass the future-directed and long-term nature of EF in daily life as compared to rating measures (Barkley & Fischer, 2011). Taken together, these findings suggest that performance-based measures have some limitations in measuring EF, which has implications when discussing ADHD and academic impairment. More specifically, if EF deficits are responsible for the observed relationship between ADHD status and academic impairment, it is important to explore which aspects of EF correlate most closely with academic outcomes. However, it is unclear whether performance-based measures will be able to capture these aspects of EF due to their limitations.

A second set of limitations is related to how academic functioning has been operationalized in previous studies. More specifically, previous work has typically operationalized academic functioning primarily as performance on achievement tests (Langberg et al., 2013). This is problematic for a couple reasons. First, achievement tests and grades in school only correlate at the 0.5 to 0.6 level (Bowers, 2011). Thus, it appears that the two may be measuring different constructs. More specifically, grades not only incorporate academic knowledge but also encompass aspects such as classroom participation, motivation, and behavior (Bowers, 2011). Achievement tests may measure academic ability in ideal conditions, whereas grades may provide a more realistic picture of academic performance in long-term, ecologically valid situations.

A recent study explored the relation between academic functioning and rating measures of EF in middle school youth (Langberg et al., 2013). The study used rating scales rather than performance-based measures to operationalize EF and used parent ratings of homework

problems and school grades as reported by teachers to operationalize academic achievement. It should be noted that homework problems were assessed using a scale that considered homework completion as well as management of homework materials. The study included both parent and teacher ratings of EF using the Behavior Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000), and the various aspects of EF (e.g., inhibition versus working memory versus planning) were explored separately. Known covariates of academics were also included (e.g., intelligence, achievement scores, parent education level, and student gender) and statistically controlled.

Results revealed significant small to moderate correlations between school grades and standardized achievement tests (Langberg et al., 2013). This is consistent with the literature and highlights the importance of exploring school grades as well as achievement tests in order to provide a more complete profile of overall academic performance. Results also indicated that both parent and teacher ratings of the Planning and Organization subscale on the BRIEF predicted grades in school beyond ADHD symptoms and the covariates. In addition, parent ratings on the Shift subscale, which measures the ability to shift from one task to another, also predicted grades in school. In terms of parent ratings, ratings of hyperactivity/impulsivity, inattention, and abilities in planning and organizing significantly predicted homework problems. In terms of teacher ratings, symptoms of inattention and the Organization of Materials subscale were significant in predicting homework problems. Of note, Langberg et al. did not find a relationship between working memory and academic achievement. This finding is inconsistent with previous research indicating that working memory is predictive of academic functioning (Miller, Nevado-Montenegro, & Hinshaw, 2012). However, as mentioned, the majority of previous studies have used performance-based measures rather than rating measures to

operationalize EF. Thus, it seems probable that the relationship between working memory and academic functioning depends based on how working memory is measured. Additionally, achievement tests are typically used as measures of academic achievement rather than grades or other measures. It is possible that working memory is more closely related to performance in achievement tests versus measures that assess broader classroom performance (Langberg et al., 2013).

Measures of EF

In research and clinical settings, EF has been assessed in multiple ways (Toplak et al., 2013). In the past, performance-based measures were considered the gold standard in assessing EF (Barkley, 2011). However, rating scales of EF have become popular due to criticisms that performance-based measures are costly, time-consuming, and lack ecological validity.

Performance-based measures. Performance-based measures of EF typically measure response time and/or accuracy on a standardized task given by an examiner (Toplak et al., 2013). Performance-based measures of EF include the Wisconsin Card Sort Test, Stroop Word-Color Test, Hand Movements Test, Rey-Osterrieth Complex Figure Drawing, Trail Making Test (Parts A and B), and Continuous-Performance Tests, among others (Barkley, 2006). These tasks measure a variety of psychological functions including memory (visuospatial and verbal), attention, abstract reasoning, sequencing, language, problem-solving, inhibition, and planning. Barkley (2006) notes that many of these measures were not designed to measure EF and are not able to measure EF without considering the influence of the other psychological processes previously mentioned (i.e., memory, attention, abstract reasoning, sequencing, language, problem-solving, inhibition, and planning). Although performance-based measures of EF have

traditionally been considered the gold standard of EF assessment, current research is beginning to demonstrate that the measures have some limitations.

First, performance-based measures of EF seem to lack ecological validity (McAuley et al., 2010). These measures are administered in a manner that maximizes support, minimizes distractions, and provides a good deal of structure. For this reason, performance-based measures may be assessing a set of skills distinct from those required in naturalistic situations. Furthermore, given that EF is typically conceptualized as a set of behaviors that allow individuals to achieve future goals, it does not seem that EF can be captured through short-term performance-based measures (Barkley & Fischer, 2011). Additionally, many performance-based measures of EF are strongly influenced by cognitive processes, including cognitive ability, that do not necessarily reflect EF (Mahone et al., 2002).

Despite the limitations of performance-based measures in assessing EF, there are certain measures that have been widely used in assessing EF in ADHD. Willcutt, Doyle, Nigg, Faraone, and Pennington (2005) conducted a meta-analysis to explore differences between groups with and without ADHD in terms of EF. The researchers included the following measures: Stop-signal Reaction Time, Continuous Performance Test (commission and omission errors), Wisconsin Card Sorting Test perseveration errors, Trailmaking Test Part B, Tower of Hanoi/London, Porteus Mazes, Rey-Osterreith Complex Figure Test, Working Memory Sentence Span, Digits Backward, Self-ordered pointing, and CANTAB Spatial Working Memory. These measures were chosen based on frequency of administration in previous studies and/or relevance in theoretical models of ADHD. In order to decrease redundancy with other meta-analyses (e.g., Huang-Pollock and Nigg, 2003; Van Mourik Oosterlaan, & Sergeant, 2005), visuospatial orienting tasks and Stroop interference control were not included. Results indicated that children

with ADHD demonstrated more impairment than controls on all twelve tasks included in the study, with effect sizes falling into the medium range (.46-.69) for all measures. The most consistent and strongest effects were found for measures of vigilance, inhibition, planning, and spatial working memory. Similarly, Nigg (2005) notes that difficulties in visual working memory and response inhibition appear to be the most consistent EF deficits in children with ADHD found in the literature.

A meta-analysis by Pauli-Pott and Becker (2011) explored the relations between a range of performance-based measures and ADHD symptoms in a preschool population. Studies were included in the meta-analysis if they assessed at least one of the six domains of EF known to be related to ADHD in school age children and adolescents (i.e., response inhibition, interference control, working memory, delay aversion, vigilance/arousal, and flexibility; Willcutt et al., 2005). Studies including children between ages three to six were included. Overall, Pauli-Pott and Becker found that performance on these measures was predictive of ADHD symptoms. In addition, effect sizes varied based on the type of deficit assessed. Medium to large effect sizes were found for delay aversion and response inhibition as predictors of ADHD symptoms. Working memory was found to have a small mean effect sizes, and interference and vigilance/arousal were somewhere between small and medium effects. Of note, vigilance/arousal was assessed through use of computerized continuous-performance tasks, which will be included in the current study. The researchers note that weighted mean effect size was highly significant for vigilance/arousal. In addition, the mean age of the sample was related to effect size for vigilance/arousal, with smaller effects found for younger versus older children (i.e., more impairment in vigilance/arousal for four year old children than five year old children).

Given the wide range of performance-based measures and breadth of research in the area, the current section will focus on continuous performance tests (CPTs), which were of particular interest in the current study. CPTs are tasks of vigilance that require attention to different stimuli as they are presented. When a certain stimulus appears, the participant must respond by pressing a key. Although CPTs are similar to other performance-based measures in that they may have low ecological validity, Barkley (2006) states that they are superior to all other performance-based measures in assessing the core features of ADHD (i.e., inattention and hyperactivity). However, the utility of CPTs in ADHD evaluations remains somewhat unclear (Edwards et al., 2007).

A recent meta-analysis explored the relation between performance on CPTs and ADHD symptoms (Huang-Pollock, Karalunas, Tam, & Moore, 2012). After correcting for both measurement and sampling error, Huang-Pollock et al. (2012) found that children with ADHD had slower and/or more variable reaction times and made more errors than those without ADHD on CPTs. More specifically, large effect sizes were found when comparing children with ADHD to controls for reaction time, omissions, commissions, and standard deviation of reaction time. Huang-Pollock et al. note that the effect sizes in their study were much smaller and more similar to findings reported in previous studies when only sampling error (not measurement error) was accounted for.

Vaughn et al. (2011) explored relations between CPT performance and parent/teacher ratings on a rating scale in groups of children with ADHD and control group of children without ADHD. Participants included children between the ages of seven and nine. The Conners' Continuous Performance Task (CCPT) was used in the study, with percent errors of commission, percent errors of omission, reaction time, and reaction time standard deviation being the primary

variables of interest. Vaugh et al. (2011) found that children with ADHD demonstrated more variability in responding and more errors of commission and omission in comparison to the control group. Although both groups demonstrated improved performance on the CPT when tested after a year, the ADHD group continued to make more errors and demonstrate more response variability as compared to the control group. In addition, specific CPT variables were found to be related to ADHD symptoms. More specifically, reaction time variability on the CPT was correlated with all ADHD subtypes (Inattentive, Hyperactive, and Combined) at both 24-month and 36-month time points. Interestingly, parent and teacher ratings of inattention remained constant over time, whereas ratings of hyperactivity/impulsivity decreased over time.

Although some studies have supported the utility of CPTs in ADHD evaluation, findings from other research is less clear. For example, Edwards et al. (2007) did not find significant correlations between the overall score on a CPT and ratings of hyperactive/impulsive and inattentive symptoms. Collings (2003) found that children with ADHD-Combined Type performed significantly worse on a CPT task than children diagnosed with ADHD-Inattentive Type and children without ADHD. Additionally, the inattentive group did not significantly differ from the group without ADHD. Overall, Collings concludes that CPTs may under identify children with ADHD inattentive type and noted that this should be considered when CPTs are used for ADHD evaluations.

Egeland, Johansen, and Ueland (2009) explored differences in CPT performance in participants with ADHD-Inattentive type and ADHD-Combined type. Results indicated that participants with ADHD-Inattentive type performed below controls on a measure of sustained attention (i.e., Hit Reaction Time Block Change) and that participants with ADHD-Combined

type performed significantly worse than controls on a measure of vigilance (i.e., Hit Reaction Time Inter-Stimulus-Interval).

In the current study, the Conners' Continuous Performance Task-II (CCPT-2) was used as a measure of sustained attention and impulse control (Conners, 2000). Four of the 11 subscales of the CCPT were considered in the present study: Omissions, Reaction Time Variability, Commissions, and Hit Reaction Time. Based on factor loadings from prior factor analyses (e.g., Murphy, Barkley, & Bush, 2001), Omissions (targets that were not responded to) and Reaction Time Variability scores were used to reflect sustained attention and Commissions (response to non-targets) and Hit Reaction Time scores were used to reflect response inhibition. Similarly, Conners (2000) reported that Omissions and Reaction Time Variability are measures of inattention, whereas Commissions and Hit Reaction Time are measures of impulsivity.

Rating scale measures of EF. Rating scales of EF were developed to provide an indicator of ability to perform in everyday, complex, and problem-solving situations (Gioia et al., 2000). EF rating scales typically ask informants, such as parents and teachers, to rate difficulties related to carrying out daily tasks. Rating measures often assess behavior over an extended period of time and may better reflect the cross-temporal organization of EF as compared to performance-based measures (Barkley, 2011). In addition, rating scales provide information regarding behavior in more natural settings (e.g., home, school), which may provide substantially different information than a performance-based measure, which assesses behavior in an artificial clinical setting (Barkley, 2006).

The BRIEF is one of the most commonly used rating scales of EF (Toplak et al., 2013). The BRIEF is a teacher and parent questionnaire designed to measure executive functioning in individuals 5 to 18 years of age (Gioia et al., 2000). The BRIEF assesses eight subdomains of EF

that are grouped into two composite scales, the Behavioral Regulation Index (BRI) and the Metacognitive Index (MI). The Behavioral Regulation Index (BRI) is comprised of Shift, Emotional Control, and Inhibit subdomains. The Metacognitive Index (MI) is comprised of the subdomains Working Memory, Initiate, Organization of Materials, Monitor, and Plan/Organize. Together, the BRI and MI form the Global Executive Composite (GEC). The BRI scales typically correlate with symptoms of hyperactivity/impulsivity, whereas the MI scales correlate with symptoms of inattention (Gioia et al., 2000).

Numerous studies have explored the BRIEF's discriminant validity for children with ADHD. Reddy, Hale, and Brodzinsky (2011) explored the validity of the BRIEF in distinguishing children with ADHD from matched controls. Overall, the groups differed significantly on all BRIEF scales. Discriminant function analyses demonstrated that the BRI and MI were able to correctly classify 82% of the sample. The BRI was most useful in distinguishing the ADHD group from the non-ADHD group. A related study by Mares, McLuckie, Schwartz, and Saini (2007) found that EF deficits on the BRIEF were associated with ADHD. In addition, teachers rated higher EF impairment in comparison to parents. Teacher ratings on the BRIEF reflected more severity and variety of EF impairment and tended to identify more children as having EF deficits in the clinically significant range. Mares et al. (2007) conclude that teachers may be better able to recognize deficits in EF in comparison to parents due to their familiarity with age-appropriate behavior.

In terms of predictive validity on the BRIEF, Gioia et al. (2000) reported that parent and teacher ratings on the Working Memory scale were significantly higher for children diagnosed with ADHD-Inattentive type and ADHD-Combined type versus control children. Parent and teacher ratings on the Inhibit scale were found to be useful in differentiating children with

ADHD-Combined type from those with ADHD-Inattentive type and controls in 65-68% of cases. Gioia et al. (2000) reported that the MI on the BRIEF is more closely related to inattention, whereas the BRI is more strongly related to hyperactivity and impulsivity.

Jarratt et al. (2005) explored the relation between the BRIEF and a broad range rating scale of behavioral and emotional functioning. Scores on the parent and teacher BRIEF indicated that the ADHD group demonstrated more deficits in EF as measured by the BRIEF in comparison to the control group. Jarratt et al. (2005) conclude that the BRIEF and the broad range measure seem to measure similar constructs but note that the BRIEF is useful in providing additional information regarding working memory and metacognition, whereas the broad range rating scale provided more information about internalizing disorders.

Mahone et al. (2002) examined the concurrent validity of the BRIEF by comparing it to performance-based measures of EF as well as to interviews and behavior rating scales. The BRIEF's relation to scores on intelligence and achievement was also explored. Mahone et al. compared four groups: children with ADHD, children with Tourette syndrome (TS), children with ADHD and Tourette syndrome, and a control group. Items of interest in the BRIEF included the three index scores (Global Executive Composite, Behavioral Regulation Index, and Metacognitive Index). The Working Memory and Inhibit scales were also included due to their overlap with the hyperactive/impulsive and inattentive symptoms of ADHD. Performance-based measures included measures of planning (i.e., Tower of London) and inhibition, readiness to respond, consistency of response time, and persistence (Test of Variables of Attention; TOVA). Of note, the TOVA is a continuous performance task and thus most similar to what will be used in the current study. TOVA variables that were considered include: commission, omission, variability in reaction time, and response time for correct responses. A task assessing fluency,

initiation, and rapid lexical retrieval (i.e., Controlled Oral Word Association Test) was also included. Results indicated that both the ADHD and combined ADHD/TS group had higher ratings of impairment on all five BRIEF variables (i.e., Working Memory scale, Inhibit scale, Metacognitive Index, Behavioral Regulation Index, Global Executive Composite). The BRIEF was not significantly related to any of the performance-based measures, scores on intelligence tests, or scores on reading and spelling portions of an achievement test. The BRIEF was correlated with math achievement. However, the BRIEF did correlate with parent interview information and rating measures of hyperactive/impulsive and inattentive behaviors. Overall, previous research supports the use of the BRIEF in ADHD assessment. In addition, results of the Mahone et al. study suggest that the BRIEF may correlate with same source/method measures (e.g. parent interviews and rating measures of symptoms of ADHD) rather than performance-based measures of EF (Mahone et al., 2002).

Rating scales and performance-based measures. As mentioned, EF is typically operationalized in clinical settings through rating scales or performance-based measures. As noted, the extent to which rating scales of EF and performance-based measures of EF assess the same underlying construct is unclear. In order to explore this issue, a recent meta-analysis examined 20 empirical studies that assessed the relationship between rating measures and performance-based measure of EF (Toplak et al., 2013). Of the 20 studies included in the meta-analysis, thirteen included children and seven included adults. Seven studies used clinical samples, two used nonclinical samples, and 11 used combined nonclinical and clinical samples. Of the 20 studies, 13 of the studies used the BRIEF as the rating measure of EF. This is noteworthy given that the BRIEF was the focus of the current study. For those studies that reported correlations, 19% of the correlations between the BRIEF questionnaire and

performance-based measures of EF were statistically significant. The mean and median correlation coefficients between the BRIEF and performance-based measures were .15 and .18 respectively. It should be noted that these values are most likely overestimates due to most studies not reporting values for non-significant correlations. Overall, these results suggest that the BRIEF and performance-based measures are only minimally correlated. Similar results were found for the Behavioral Assessment of the Dysexecutive Syndrome-Dysexecutive Questionnaire (BADS-DEX; Wilson, Alderman, Burgess, Emslie, & Evans, 1996), the other rating scale of executive function assessed in the study. Toplak et al. 2013 also included three studies that explored the relation between ratings of impulsivity and performance-based measures of inhibition. Results indicated modest correlations between ratings of impulsivity on multiple measures and performance-based measures of inhibition.

The results of the Toplak et al. (2013) meta-analysis are consistent with the general trend found when examining individual studies. For example, McAuley et al. (2010) found that the subscales on the BRIEF were not significantly associated with performance-based measures of executive function (i.e., working memory, inhibition, and performance monitoring). However, there was a strong relationship between the BRIEF indices and teacher and parent ratings of behavioral impairment and disruption. The researchers concluded that the BRIEF is a useful tool to assess broad concerns but suggested that it may not relate to performance-based measures of executive function.

Biederman et al. (2008) found a modest but non-significant overlap between rating measures and performance-based measures of EF. Performance-based measures of EF included variables assessing response inhibition, sustained attention/vigilance, set shifting/categorization, planning/organization, memory, and visual and verbal learning. Additionally, only 14% of

individuals who reported executive function deficits as assessed by rating measures also exhibited impairment on performance-based measures. Performance-based measures of EF primarily identified participants with lower achievement test scores and lower IQ, whereas the rating measure primarily identified participants with increased levels of psychiatric comorbidity, ADHD symptoms, and interpersonal deficits. Bodnar, Prahme, Cutting, Denckla and Mahone (2007) found that ratings of inhibitory control on the BRIEF by parents were minimally correlated with variables of inhibition on two performance-based measures, the Conners' Continuous Performance Test-II (CCPT-II; Conners, 2000) and the Test of Variables of Attention-Visual (TOVA; Greenberg, 1996). In addition, Bodnar et al. found that factor analyses revealed that the BRIEF subscales load on a distinct, single factor that was separate from any performance based variables (i.e., CCPT-II and TOVA). These results suggest that method of assessment (i.e., rating scale versus performance-based measure) may explain more variance in latent variables in measures of EF than the construct (i.e., inhibition).

Although the previously mentioned studies did not find a relationship between rating scales and performance-based measures of EF, other studies have found significant correlations between ratings on the BRIEF and performance based measures of EF (McCandless & O'Laughlin, 2007; Toplak, Bucciarelli, Jain, & Tannock, 2009). For example, Toplak et al. (2009) explored the relation between parent and teacher ratings of EF (i.e., BRIEF) and performance-based measures of EF in a sample of adolescents. On the BRIEF, the variables of interest were the Working Memory, Inhibit, Plan/Organize, and Shift subscales. Performance-based measures were selected to mirror the subscales on the BRIEF and included measures of working memory, inhibition, planning, and set shifting. In comparison to controls, those with ADHD demonstrated deficits on all performance-based measures of EF and significantly higher

scores on both the parent and teacher BRIEF subscales. In addition, virtually all ratings on the BRIEF (parent and teacher) were significantly correlated with the performance-based measures of EF, with one exception. More specifically, results indicated significant but modest correlations between all subscales on the BRIEF and performance-based measures of EF, with the exception of the performance-based measure assessing inhibition (i.e., the stop task) and parent ratings of Working Memory, Plan/Organize, and Shift subscales. The researchers also assessed whether each BRIEF subscale was correlated with its respective performance-based measure. Results suggested that there are not unique relations between subscales on the BRIEF and specific performance-based measures. Lastly, logistic regression analyses revealed that ratings measures (both parent and teacher) were superior to performance-based measures in terms of predicting ADHD status.

McCandless and O'Laughlin (2007) found significant correlations between the parent and teacher ratings on the BRIEF and the Integrated Visual and Auditory Continuous Performance Test (IVA-CPT; Sanford & Turner, 1995), a performance-based measure of EF. On the IVA-CPT, variables of interest were the Full Scale Response Control Quotient and the Full Scale Attention Quotient (FSAQ). The Full Scale Response Control Quotient is a measure of response inhibition, consistency in responding, and sustained effort. The FSAQ is a measure of attention difficulties, slow processing speed, and loss of focus. In terms of parent ratings on the BRIEF, the Plan/Organize scale was significantly correlated with measures of response control on the IVA-CPT. In terms of teacher ratings, the Initiate subscale on the BRIEF was significantly correlated with a measure of attention of the IVA-CPT. In addition, teacher ratings on the Plan/Organize subscale on the BRIEF were correlated with performance on a measure of response control on the IVA-CPT. Interestingly, significant correlations between teacher and

parent ratings were only found on three of the eight subscales on the BRIEF, suggesting the possibility that parent and teacher reports may vary due to different child behavior across settings (McCandless & O'Laughlin, 2007).

Overall, both Toplak et al. (2009) and McCandless and O'Laughlin (2007) found support for the relationship between rating scales and performance-based measures of EF. Toplak et al. (2009) found support for the relationship between performance-based measures of EF assessing working memory, inhibition, planning, and set shifting, and the Working Memory, Inhibit, Plan/Organize, and Shift subscales on the BRIEF. However, unique correlations were not found between performance-based measures and subscales on the BRIEF thought to reflect similar constructs. McCandless and O'Laughlin (2007) found that response control as measured by a CPT was correlated with parent ratings on the Plan/Organize subscale on the BRIEF. Teacher ratings on the Plan/Organize subscale were also associated with deficits in response control and teacher ratings on the Initiate subscale were associated with sustained attention on the CPT. These results suggest that there is a complex relationship between rating scales and performance-based measures of EF and that further research is needed to better clarify the relationship between rating scales and performance based measures of EF, as well as between rating scale measures of EF and other assessment measures commonly used in evaluation for ADHD.

Recent research seems to suggest that rating measures of EF and performance-based measures may be capturing unique aspects of EF (Toplak, et al., 2013). Given that children behave differently in different settings, it may be that these measures are accurately reflecting discrepancies in behavior in various contexts (Dirks et al., 2012). It seems that both measures may provide useful but distinct types of information in assessment of ADHD. More specifically, it may be possible that ratings of EF capture the future-directed and cross-temporal aspects of EF

in a way that performance-based measures of EF do not (Barkley & Fischer, 2011). This makes sense given that ratings of EF measure performance over time in daily life, whereas performance-based measures measure functioning at one point in time on a specific, structured task. Furthermore, Barkley and Fischer (2011) assert that EF should be viewed as a meta-construct that is hierarchically organized with various levels, with each level functioning in an increasingly complex manner. The lower levels serve as a foundation for the higher levels, from which more complex and longer-term goals arise. Following from this, it seems likely that ratings measures of EF and performance-based measures of EF may assess different levels of the meta-construct. Thus, the two should not be pitted against one another but rather viewed as providing different types of information regarding EF. More specifically, it seems as if performance-based measures of EF may measure lower levels of EF, whereas ratings may measure higher levels that assess longer-term performance. For example, lower levels of EF might include processes such as inhibitory control, whereas higher levels of EF might include academic performance. One way to test this claim is to measure the relationship between the two measures of EF and impairment in different types of settings (i.e., measures reflecting performance in short-term, structured domains as well as measures reflecting performance in longer-term, day-to-day domains).

Overall, previous research comparing rating measures to performance-based measures is inconsistent. While some studies have found support for a relationship between the two measures (i.e., McCandless & O'Laughlin, 2007; Toplak et al., 2009), other studies have not (Biederman et al., 2008; Bodnar et al., 2007; McAuley et al., 2010). In addition, multiple studies have found that the BRIEF tends to correlate with other rating scales rather than performance-based measures assessing the same construct (Mahone et al., 2002; McAuley et al., 2010). As

mentioned, it has been suggested that the validity of the BRIEF may be further explored by examining its relationship to naturalistic tasks that require the application of executive function skills in more complex and realistic settings (McAuley et al., 2010).

Present Study

Previous research has suggested that the BRIEF tends to correlate with other rating scales, even those assessing related but distinct constructs, more closely than it correlates with performance-based measures of executive functioning (Mahone et al., 2002; McAuley et al., 2010). It has been suggested that rating measures of EF may assess achievement of long-term goals in day-to-day life, whereas performance-based measures of EF may assess cognitive efficiency under ideal situations (Toplak et al., 2013). Furthermore, Barkley and Fisher (2011) stated that EF could be conceptualized as a meta-construct with various levels. The levels are organized by complexity, with the lower levels of EF serving as a foundation for higher levels. Higher levels of EF serve as the basis for more complex and longer-term goals. Based on this model, it seems that rating measures of EF and performance-based measures of EF may be measuring different levels of the same meta-construct.

Given the inconsistencies in prior research regarding the relationship between rating scales and performance-based measures of EF, the present study examined the convergent and divergent validity of the BRIEF by exploring its relationship with a performance-based measure of EF, as well as other measures related to EF in an ADHD battery. More specifically, associations between the BRIEF and a performance-based measure of EF, diagnostic rating scales of ADHD, and an observational measure of inattentive and hyperactive behavior during testing were examined. In addition, longer-term academic outcomes thought to be associated with EF abilities were considered.

It was hypothesized that the BRIEF would correlate with parent and teacher ratings on an ADHD rating scale due to a common method and related construct being assessed. Conversely, it was hypothesized that ratings on the BRIEF would not significantly correlate with performance on a continuous performance-task despite the fact that both measures assess aspects of executive functioning. Consistent with Toplak et al. (2013), it was hypothesized that ratings on the BRIEF scales would be significantly associated with measures of longer-term goal pursuit, reflecting the idea that rating scale measures may assess long-term functioning in daily life. In addition, it was hypothesized that an observational measure of inattentive and hyperactive symptoms during testing would correlate more strongly with performance on a continuous performance task as compared to ratings on the BRIEF. There is little previous research that examines construct validity of structured observations in general, and only one that examines the construct validity of the Test Observation Form (i.e., McConaughy et al., 2010). Thus, this hypothesis is based on the notion that both the continuous performance task and the observational measure of behavior during testing should reflect behavior in a short-term, structured setting, whereas ratings on the BRIEF would be based on more long-term performance. In addition, given that the TOF is a rating scale based on performance during testing, it was hypothesized to be more strongly correlated with performance on the Conners' Continuous Performance Task-II versus other behavior rating scales due to similar setting (i.e., structured/ ideal setting).

Regarding the influence of common method versus common construct, it was hypothesized that ratings on the BRIEF would be more highly correlated to a dissimilar construct (i.e., parent and teacher ratings of child behavior associated with depression) assessed using a common method (i.e., rating scale) than to a measure of the same construct assessed using a different method (i.e., the continuous performance task). Lastly, it was hypothesized that results

of a factor analysis would indicate that parent and teacher ratings on the BRIEF cluster with other rating measures, whereas the continuous performance task would group with other short-term performance measures.

Hypotheses

1. The BRIEF Metacognitive Index (parent and teacher) will be significantly positively correlated with ratings of Attention problems (parent and teacher) on the ADHD-IV Rating scale.
2. The BRIEF Metacognitive Index and Behavioral Regulation Index will not be significantly associated with performance on the Conners' Continuous Performance Task-II (as measured by Omissions, Reaction Time Variability, Commissions, and Hit Reaction Time measures).
3. The BRIEF Metacognitive and Behavioral Regulation Indices (parent and teacher) will be negatively correlated with measures of longer-term goal pursuit (as measured by teacher questionnaires).
4. Clinician ratings of inattention on the TOF (Inattention and Hyperactive/Impulsive subscale) will be more strongly related with performance on the Conners' Continuous Performance Task-II (Omission and Variability scales) as compared to the BRIEF Metacognitive and Behavioral Regulation Indices (parent and teacher).
5. The BRIEF Metacognitive and Behavioral Regulation Indices will be more positively correlated with parent and teacher ratings of depression (dissimilar construct) than with the Conners' Continuous Performance Task-II Omission scale (similar construct), supporting the influence of common method as opposed to common construct.

6. Based on results of a factor analysis, parent and teacher ratings on the BRIEF will group with parent and teacher ratings of inattention and hyperactivity, long-term academic outcomes, and ratings of depression whereas the Conners' Continuous Performance Task, which measures a similar construct, will group with other short-term performance measures (i.e., a measure of cognitive ability and an observational measure).

CHAPTER 3

METHOD

Design

A correlational design was used to examine the relationships among the BRIEF, a performance-based measure of EF, diagnostic rating scales of ADHD, an observational measure of inattentive and hyperactive behavior during testing, and longer-term variables related to academic outcomes. In addition, factor analysis was used to further examine the relationship between measures assessing similar constructs and measures using a similar method (i.e., behavior rating scale, performance-based measure).

Participants

The present study used archival data representing 105 children (74 male and 31 female) seen for evaluation at a Midwestern University ADHD clinic between 2012 and 2015. Children ranged from five to eleven years old (M = 95 months, SD = 18 months). More boys are evaluated in the clinic compared to girls with approximately 78% of boys receiving an ADHD diagnosis compared with 81% of girls. Regarding ethnic background, the majority (89.5%) of participants were identified as Caucasian, with the remaining children identified as African American (4.8%), Hispanic (1.9%), or Biracial (1.0%). Of the 98 participants whose parent/guardian reported information related to household income, almost half (48.5%) reported an annual income of

$30,000 or less, while 31.4% reported a yearly income between $30,000 and $60,000, and 13.4% reported yearly income above $60,000.

Approximately 76% of children in the current study met the DSM-5 criteria for some type of ADHD diagnosis (i.e., ADHD-Combined, ADHD-Inattentive type, ADHD- Not Otherwise specified). Additionally 72.2% met criteria for a second diagnosis including: Oppositional Defiant Disorder (37.3%), an anxiety or mood disorder (9.6%), a learning disorder (12.1%), an Autism Spectrum Disorder (1.2%), or an "Other" DSM-5 disorder (12%). Table 1 contains frequencies and percentages for participant demographic information.

Children were referred to the ADHD Evaluation clinic by parents, teachers, or primary care physicians due to observed difficulties with inattention and/or hyperactivity/impulsivity. Parent questionnaires were typically completed by the child's mother. However, forms were completed by the child's father or legal guardian (e.g., grandparent) in a small number of cases. Teacher questionnaires were typically completed by the child's regular education teacher. If the child had both special education and regular education teachers, a letter sent with the questionnaires asked that the teacher who spends the greatest amount of time with the child in the classroom complete the questionnaires. When children were taking medication for ADHD, it was requested that parents and teachers rate the child's behavior when they were not taking medication, if possible.

Consistent with previous research in this area, children with a full-scale intelligence quotient of less than 70 were excluded from the current study as cognitive ability has been a powerful moderator in ADHD assessment in general and performance-based measures in particular (Mahone et al., 2002). One child was removed from the study due to IQ below 70. Children taking medication for ADHD were excluded from the study if they were not on

medication on the day of the assessment but taking medication on a daily basis. This was done in order to reduce inconsistencies between various types of measures (e.g., rating scales and performance based measures). Eight children were removed from the study due to inconsistency in medication (e.g., on medication, but not on meds the day of the evaluation). There were ten children taking ADHD medication on a daily basis as well as on the day of the evaluation that were not excluded from the study, given that ratings scales and performance based measures were both influenced by medication.

The presence of comorbid disorders did not warrant exclusion from the study, given that a high prevalence rate of comorbid disorders is often part of the diagnostic picture when evaluating ADHD in clinical settings (i.e., Barkley, 2006). Comorbid diagnoses included Oppositional Defiant Disorder (37.3%), Anxiety/Depression (9.6%), Learning Disorder (12.1%), Autism Spectrum Disorder (1.2%), or Other (12%). Finally, given that all children had been referred to be assessed for possible ADHD, the study included both children diagnosed with ADHD and those who did not receive a diagnosis of ADHD. The scores reflected both higher and lower levels of functioning on performance-based measures and ratings of symptom endorsement.

Measures

Behavioral Rating Inventory of Executive Function

The Behavioral Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000) is a teacher and parent questionnaire designed to measure executive functioning in individuals 5 to 18 years of age. The parent and teacher questionnaires each have 86 items that are rated using a three-point scale (never, sometimes, often) and require 10-15 minutes to complete (See Appendix B for sample items from the BRIEF). The BRIEF is made up of two indices: the

Behavioral Regulation Index (BRI) and the Metacognitive Index (MI). The BRI is comprised of Shift, Emotional Control, and Inhibit subdomains. The MI is comprised of the subdomains Working Memory, Initiate, Organization of Materials, Monitor, and Plan/Organize. Together, the BRI and MI form the Global Executive Composite (GEC). The BRIEF has a hidden carbon copy inside the questionnaire that aids in adding item scores to compute T-scores for each scale. Results are reported using t-scores relative to the normative sample ($M = 50$; $SD = 10$). T-scores greater than 65 are considered clinically significant. The normative sample was based on ratings of children from 1,419 parent forms and 720 teacher forms from urban, suburban, and rural areas (Gioia et al., 2000). The normative sample reflected a full socioeconomic status distribution, and gender and ethnicity distribution was weighted based on the census. The BRIEF also includes two validity scales, an inconsistency scale (acceptable, questionable, inconsistent) and a negativity scale (acceptable, elevated, highly elevated).

Gioia et al. (2000) reported satisfactory reliability with the BRIEF. Test-retest reliability statistics over a 2-week period ranged from .79 to .88. Test-retest reliability demonstrated t-score stability over 2-3 week intervals and supported the use of the measure for repeated administrations. In terms of internal consistency, the Cronbach α coefficient ranged from .80 to .98 for the teacher and parent form and normative and clinical samples. In terms of inter-rater agreement, moderate inter-rater correlations were found for parent and teachers ($r = .32$) but this seems to reflect the variation of children's behavior in different settings.

Ratings on the BRIEF and comparable behavioral rating scales have been demonstrated to have good concurrent validity (Gioia et al., 2000). Gioia et al. (2000) reported that the BRIEF demonstrates convergent validity with the ADHD Rating Scale-IV, the Child Behavior Checklist, the Behavior Assessment for Children, the Conners' Rating Scale, and Teacher's

Report Form. Mahone et al. (2002) found that the BRIEF demonstrated good concurrent validity with other rating measures of ADHD behaviors (i.e., the ADHD Rating Scale IV-Home Version, Child Behavior Checklist-Parent Report Form) and with an interview (i.e., the Diagnostic Interview for Children and Adolescents, Fourth Edition). Mahone et al. (2002) found that the BRIEF did not significantly correlate with performance-based measures of EF or with intellectual or achievement tests (i.e., Wechsler Intelligence Scale for Children, Third Edition and Wechsler Individual Achievement Test). However, the BRIEF did correlate with the math composite of the achievement test.

Gioia et al. (2000) reported that the MI on the BRIEF is more closely related to inattention, whereas the BRI is more strongly related to hyperactivity and impulsivity. In the current study, the BRI and MI were the variables of interest, with total scores on these indices used in analyses.

Conners Continuous Performance Task

The Conners Continuous Performance Task (Conners, 2014a, 2014b; Conners, 2000) is a performance-based measure designed to assess attention and concentration. In the current study, approximately two-thirds of the participants (n = 64) were given the Conners Continuous Performance Task II (CCPT-II) and the remaining were given the Conners Continuous Performance Test III (Conners CPT-3; n = 17) or the Conners Kiddie Continuous Performance Task II (K-CPT 2; n = 10). Administration of the CCPT-II takes approximately 15 minutes. Consistent with Barkley's aforementioned theory of ADHD, the CCPT-II is thought to measure behavioral inhibition (Barkley, 2006; Conners, 2000). The CCPT-II yields eleven measures including: Omissions, Commissions, Hit Reaction Time, Hit Reaction Time Standard Error, Variability of Standard Error, Attentiveness, Perseverations, Hit Reaction Time Block Change,

Hit Standard Error Block Change, Hit Reaction Time ISI Change, and Hit Standard Error ISI change. Omissions, Variability of Standard Error, Commissions, and Hit Reaction Time were the variables used in the current study. An omission error is scored when the respondent fails to respond to target stimuli (non-Xs). Variability of Standard Error is a measure of the amount of variability shown in the separate sections of the test as compared to the individual's overall standard error. The commission variable is calculated based on the number of time the child responds to non-targets (i.e., Xs) stimuli. Hit Reaction Time is a measure of the average speed of accurate responses for the full test.

The administration time for the Conners CPT-3 also takes 15 minutes (14 minutes for the test with a 1-minute practice test). Whereas the CCPT-II is recommended for children 6 years of age and older, the Conners CPT-3 is recommended for children 8 years of age and older and the Conners Kiddie Continuous Performance Test, 2nd edition (Conners K-CPT 2) is recommended for use with 6 and 7 year olds. The Conners K-CPT 2 is based on the Conners CPT but has a shorter administration time (i.e., 7 minutes) and uses pictures rather than letters for stimuli. Analyses revealed that the K-CPT 2 and Conners CPT-3 produce comparable results (Conners, 2014a, 2014b). Other changes in the Conners CPT-3, as compared to the CCPT-II, include the color of the letters and background (i.e., changed to black letters with a white background to improve visibility) and the number of trials presented (i.e., 361 trials in the Conners CPT-3 as compared to 360 trials in the CCPT-2). On the CCPT-II, targets are displayed in 90% of the trials and non-targets appear in 10% of trials. On the Conners CPT-3, the proportion has been shifted to 80% targets to 20% non-targets. This change was made in order to improve the Commissions variable's psychometric properties.

The Conners CPT-3 yields fourteen scales as compared to the eleven yielded from the CCPT-II. The new or modified variables include: Response style, Perseverations, Commissions, d-prime, and Response speed consistency. Of note, the variables of interest in the current study are all included in both the CCPT-II and Conners CPT-3. However, as noted, the Commissions variable was changed to be more able to detect clinical differences. For the current study, means were compared across the three CPT's for the Omission, Commission, Variability of Standard Error, and Hit Reaction Time variables. Results revealed a significant difference in Commission errors based on type of CCPT. No significant differences were found for the other three variables based on type of CCPT. Post hoc tests indicated significant differences in commission errors between the CCPT-2 and the K-CPT 2 and the CCPT-2 and CPT-3. Thus, type of CPT (i.e., CCPT-II, CPT-3, K-CPT 2) was included as a covariate for analyses that considered the commission variable.

The Conners CPT-3 included both normative and clinical populations in the standardization process (Conners, 2014a, 2014b). The normative sample included 1,400 individuals from the general population. Median split-half reliability in the normative sample was .93 for male youths and .92 for other groups. The median split-half reliability estimate for the clinical sample ranged from .92 to .95 for all groups. The median corrected test-retest coefficient (based on administrations 1 to 5 weeks apart) was $r = .67$.

The Conners K-CPT 2 standardization process included both a normative and clinical sample (Conners, 2015). The normative sample included 320 children between ages 4 to 7. The median split-half reliability estimates for the normative sample was .88 for the female sample, .89 for the male sample, and .87 for the combined sample. The median split-half reliability estimate for the clinical sample was .89 for the female sample, .86 for the male sample, and .86

for the combined sample. Overall, the K-CPT 2 exhibited strong internal consistency for the normative and clinical samples. Test-retest reliability for the K-CPT 2 was assessed using 1 to 5 week intervals. The median test-retest reliability coefficient was $r = .57$.

The CCPT-II normative sample included both clinical and nonclinical populations. Individuals age 6 and older were included in the normative sample. The clinical sample included 378 cases with ADHD diagnosis and 223 adults with neurological impairment. The non-clinical sample was based on 1,920 individuals from the general population. In terms of reliability, split-half reliability correlations ranged from .66 to .95 for the measures. Test-retest correlation coefficients ranged from $r = .43$ to.84, with 3 months being the average period of time between administrations.

Edwards et al. (2007) explored the classification utility and validity of the CCPT. The CCPT was not significantly correlated with ratings of hyperactive/impulsive and inattentive behaviors. In terms of classification utility, both the CCPT Overall Index and Omission Errors out performed what could be predicted based only on chance (.50) in identifying children with ADHD. The CCPT Overall Index ranged from poor to fair in identifying ADHD subtypes (i.e., Inattentive, Hyperactive, or Combined). The classification utility of the CCPT Omission score (84^{th} percentile cutoff) ranged from slight to fair in identifying subtypes of ADHD and showed improvement over the Overall Index.

Attention Deficit Hyperactivity Disorder Rating Scale-IV

The Attention Deficit Hyperactivity Disorder Rating Scale-IV (ADHD-RS-IV; DuPaul et. al., 1997) is a brief questionnaire created to gather information regarding symptoms of ADHD in the previous 6 months. It is composed of two subscales (Inattention and Hyperactivity/ Impulsivity) with 9 items each. Each item is rated on a 4-point Likert scale ranging from 0 to 3

points (0 = never or rarely, 1 = sometimes, 2 = often, and 3 = very often). It has two versions (e.g., Home version and School version), which are completed by parents or teachers and are identical (See Appendix C for sample items from the ADHD-RS-IV). The standardization sample was nationally representative and included 2,000 children and adolescents between ages 4 and 20. The scale is designed for use with individuals between the ages of 5 and 18. In both the Home and School versions, the support for validity, stability, and internal consistency of the ADHD-RS-IV are generally strong (DuPaul et al., 1997). In the current study, both the Inattentive and Hyperactive/Impulsive subscales were of interest.

The measure yields a score in the form of a percentile. For a diagnosis of ADHD, Predominately Inattentive Presentation, DuPaul et al. recommend a cut-off in the 93rd percentile for parent reports of inattention and a cut-off in the 90th percentile for teacher reports of inattention. In order to receive a diagnosis of ADHD, Combined Presentation, DuPaul et al. recommend that the aforementioned criteria for inattention be met as well as teacher reports of hyperactivity/impulsivity in the 98th percentile or greater. No cut-off score is recommended for parent reports of hyperactivity/impulsivity.

Zhang, Faries, Vowles, and Michelson (2005) examined the psychometric properties of the ADHD-RS-IV with a sample of over 600 children and adolescents diagnosed with ADHD in 14 countries. Results indicated that the scale has acceptable test-retest reliability across a one week interval for all scales (r = .78 to .89). The scale also had acceptable internal consistency (Cronbach's alpha ranging from .80 to .84) and factor structure, with a two-factor solution with items reflecting inattention (odd-numbered items) loading on Factor 2 and items reflecting hyperactivity (even-numbered items) loading on factor 1. In addition, the scale demonstrated adequate inter-rater reliability (Kappa statistics ranging from .58 to .63) and responsiveness.

Responsiveness is the scale's capacity to detect changes in symptom severity that are small but clinically significant. Zhang et al. (2005) also found moderate to high correlations between the ADHD-RS-IV and other parent and clinician-rated measures of ADHD symptoms and low to moderate correlations with teacher-rated measures (e.g., the Conners Parent Rating Scale-Revised: Short Form and the Conners Teacher Rating Scale-Revised: Short Form; Conners, 1997). The scale also has acceptable discriminant validity. The ADHD-RS-IV had low and non-significant correlations with measures assessing symptoms of other disorders (i.e., child measures of depression and anxiety). There was also a low correlation between the ADHD-RS-IV Home Version and other teacher ratings. However, this is consistent with other research comparing parent and teacher ratings. In addition, the scale was able to discriminate between children whose diagnosis suggested clinically significant inattentive symptoms or hyperactive/impulsive symptoms and children whose diagnosis did not.

Teacher Ratings of Academic Performance

Teachers of children referred for an ADHD evaluation are asked to rate the work completion and work accuracy for math and language arts on a scale from 1 to 100%. For the purpose of this study, the ratings for work completion and work accuracy for math were averaged. Similarly, ratings for work completion and work accuracy for language arts were averaged.

Test Observation Form

The Test Observation Form (TOF; McConaughy & Achenbach, 2004) is a questionnaire designed to assess child behavior, test-taking style, and affect during assessment or over an observational period. The TOF is designed for children between the ages of 2 and 18. The TOF has 125 items that are rated by the examiner on a 4-point scale (i.e., no occurrence; very

slight/ambiguous occurrence; definite occurrence with mild to moderate intensity/frequency and less than 3 minutes total duration; definite occurrence with severe intensity, high frequency, or 3 or more minutes total duration).

The TOF is comprised of five scales: Anxious, Language/Thought Problems, Oppositional, Withdrawn/Depression, and Attention Problems (which consists of items associated with both inattentive and hyperactive/impulsive behaviors). Additionally, the TOF includes an ADHD Problems scale, which is composed of Inattention and Hyperactivity/Impulsivity subscales that mirror the DSM-IV-TR criteria for ADHD. Results are reported using t-scores relative to age and gender. T-scores greater than 70 are considered clinically significant. T-scores falling between 64-69 are considered to be in the borderline range. In the current study, t-scores for the Inattention and Hyperactive/Impulsive subscales on the ADHD Problems scale were the variables of interest.

In terms of reliability, McConaughy and Achenbach (2004) reported test-retest reliability ranging from .53 to .87 and internal consistency ranging from .74 to .94 for all TOF scales. With the exception of the anxious scale, the TOF scales all have moderate to strong inter-rater reliability. For the ADHD Problems Scale, test-retest reliability = .87, internal consistency = .88, and inter-rater reliability = .73.

Behavioral Assessment System for Children-2

The Behavioral Assessment System for Children-2 (BASC-2; Reynolds & Kamphaus, 2004) are rating scales completed by parents and teachers to assess child behaviors in behavioral, emotional, and executive functioning domains (Jarratt et al., 2005). The parent version has 160 items and the teacher version has 139 items (Reynolds & Kamphaus, 2004). The BASC-2 is comprised of the following scales: Attention Problems, Conduct Problems, Atypicality,

Depression, Hyperactivity, Learning Problems, Leadership, Social Skills, Study Skills, Withdrawal, and Somatization. The parent version does not include the Learning Problems items or subscale and includes several adaptive functioning subscales that are not included in the teacher version. The scales in the measure make up five composite scores: Internalizing Problems, Externalizing Problems, Behavioral Symptoms Index, Adaptive Skills (parent version only), and School Problems (teacher version only). The BASC-2 has four normative samples: clinical norms, learning disability norms, ADHD norms, and general norms. Results are interpreted using T-scores with a T-score greater than 70 as the suggested cutoff for clinical significance.

The BASC-2 parent scale has been found to have strong concurrent validity with other measures of child behavior (Reynolds & Kamphaus, 2004). The BASC-2 parent scale also appears to have strong test-retest reliability over a one-week period (ranging from .77 to .91) and strong internal consistency for the general norms (ranging from .89 to .94). The BASC-2 teacher scale has also been found to have strong concurrent validity with other rating scales assessing child behavior. Test-retest reliability over a one-week period ranged from .84 to .90 and internal consistency coefficients ranged from .87 to .97 for the general norms. In terms of inter-rater reliability, both the parent and teacher version of the scale have adequate inter-rater reliability, although it is lower for the parent version (Reynolds & Kamphaus, 2004).

In the current study, only the Depression subscale from the BASC-2 was used. Correlations between the BASC-2 Depression subscale (parent and teacher) and Metacognitive and Behavior Regulation composite scales on the BRIEF will be calculated in considering divergent validity between two ratings scale variables based on dissimilar constructs.

Wechsler Abbreviated Scale of Intelligence-II

The Wechsler Abbreviated Scale of Intelligence-II (WASI-II; Wechsler, 2011) is an individually administered assessment of intellectual functioning for individuals between the ages of 6 through 90. It includes 4 subtests: Block Design, Vocabulary, Matrix Reasoning, and Similarities. The composite scores obtained include: FSIQ-4, FSIQ-2, verbal comprehension index (VCI) score, and perceptual reasoning index (PRI). The VCI score is derived from vocabulary and similarities, and the PRI score is derived from block design and matrix reasoning. The Full Scale IQ (FSIQ) was the variable of interest in the current study.

In terms of reliability, test-retest reliability ranged from .79 to .90 for the subtests and .87 to .95 for the composite scores for the child sample. For the adult sample, average test-retest reliability coefficients ranged from .83 to .94 for the subtests and .90 to .96 for the composite scores. In terms of inter-rater agreement, inter-rater reliabilities ranged from .94 to .99. In terms of validity, the FSIQ on the WASI-II is highly consistent with the FSIQ on full Wechsler batteries. However, the FSIQ-2 on the WASI-II accounts for less variance and has less differential ability on the WISC-IV and WAIS-IV FSIQ scores. This suggests that the use of four subtests may be a better screening tool than the use of two subtests.

Procedure

The present study made use of archival data collected over a period of approximately three years (2012-2015). Parents or legal guardians provided consent allowing assessment data to be used for research purposes. Data was only included in the current study from parents who completed the consent for research participation form. When the appointment was scheduled for the ADHD evaluation, parents were mailed a packet of questionnaires for the child's teacher. This included a consent form that allowed teachers to send the completed questionnaires to the

ADHD Evaluation clinic. On the day of the assessment, the child and parent/guardian came to the ADHD Evaluation clinic for a two to three hour assessment. Parents provided information during a semi-structured diagnostic interview and were also asked to complete several questionnaires. Children were administered the WASI-II, WIAT-III Abbreviated, and CCPT-II (CPT-3 or K-CPT-2 for children seen after October 2014). Evaluators completed the structured observation measure (TOF) based on the child's behavior during testing.

Following the evaluation, each participant was assigned a unique subject number. Data was entered based on subject number rather than client name or identifying information in order to assure confidentiality. Data was entered into a database by graduate and undergraduate research assistants who had completed training through the IRB as well as additional training and supervision from the researcher. All information obtained from the evaluation was kept in a locked filing cabinet in a locked office in the ADHD Evaluation Clinic.

CHAPTER 4

RESULTS

Overview

Results are presented in the following manner. First, preliminary analyses and descriptive statistics are presented. Following this, correlational analyses addressing the first five hypotheses related to individual relationships between measures in the ADHD battery are presented. Lastly, a factor analysis exploring the relationship between the BRIEF and other measures in an ADHD battery will be presented.

Preliminary Analyses

Table 2 presents means, standard deviations, and ranges for the primary research variables. Parent and teacher mean ratings on the BRIEF were in the above average range for the Metacognitive Index (MI) but fell within the average range for the mean ratings on the Behavior Regulation Index (BRI). Mean parent ratings for ADHD-RS-IV were just below for the 90% cut-off for clinical significance for both the teacher mean (M = 89th percentile) and parent mean (M = 89th percentile) on the Inattention subscale. The teacher mean on the hyperactivity scale fell within the high average range (M = 87th percentile) and parent rating was also in the average range (M = 74th percentile).

Correlational Analyses

Bivariate correlations were used to examine the relationship between the BRIEF Metacognitive Index and ratings of attention problems on the ADHD-RS-IV. The first hypothesis predicted that the BRIEF MI would be correlated with ratings of Attention problems on the ADHD-RS-IV. As predicted, both parent ratings and teacher ratings on the BRIEF MI were correlated with the corresponding ratings on the ADHD-RS-IV (See Table 3 and 4). There was a smaller effect across raters, with teacher ratings on the MI significantly correlating with parent ratings of inattention but a nonsignificant relationship between parent ratings on the BRIEF MI and teacher ratings of inattention on the ADHD-RS-IV (See Table 5).

There was also support for the second hypothesis, which predicted that the BRIEF indices (MI and BRI) would not be significantly correlated with performance on the CCPT (Omissions, Variability, Commissions, and Hit Reaction Time). Neither the BRIEF MI nor the BRI scales were significantly correlated with performance on the CCPT (all *p*-values > .23; See Tables 3 and 4). Given the variation in type of CPT task (i.e., CCPT-II, CPT-3, KCPT) across participants, a partial correlation was conducted controlling for type of CPT administered. The results were very similar, with the BRIEF indices not significantly correlating with CCPT performance (all *p*-values > .26).

The third hypothesis predicted that the BRIEF indices would be negatively correlated with teacher questionnaires assessing longer-term goal pursuit. Teacher ratings on the BRIEF MI were significantly negatively correlated with work completion/accuracy in both language arts and mathematics, providing support for this hypothesis. Teacher ratings on the BRI and parent and teacher ratings on the MI and BRI indices were not significantly correlated with work completion/accuracy in math or language arts (all *p* values > .15) (See Tables 3 and 4).

Additional correlational analyses were completed to assess the relationship between the Test Observation Form (TOF) and the BRIEF and the TOF and the CCPT. It was hypothesized that the TOF inattention subscale would be more strongly correlated with the CCPT (Omissions and Variability) than with the BRIEF indices. Results indicated that clinician ratings on the TOF inattention subscale were significantly correlated with the CCPT Omission scale and CCPT Variability scale. Clinician ratings on the TOF hyperactive/impulsive subscale were not significantly correlated with any CCPT variables (all *p*-values > .16). Clinician ratings on the TOF were not significantly correlated with parent or teacher ratings on the BRIEF MI or BRI indices (all *p*- values > .08). Tables 3 and 4 present correlation coefficients for BRIEF, TOF, and CCPT scales. A partial correlation was conducted in order to control for type of CCPT administered. The partial correlation yielded similar results as the initial analyses, with clinician ratings on the TOF inattention subscale significantly correlated with CCPT Omission and Variability scales. In addition, when controlling for type of CPT, the TOF Inattention subscale was significantly correlated with CCPT Hit RT. Results controlling for type of CPT revealed a trend towards significance between the TOF Hyperactive/Impulsive subscale and the CCPT Variability scale. Consistent with the initial analyses (i.e., not controlling for type of CPT), parent and teacher ratings on the BRIEF indices were not significantly correlated with any CCPT variables (all *p*-values > .12).

The fifth hypothesis predicted that the BRIEF indices would be more positively correlated with parent and teacher ratings of Depression (on BASC-2) than with the CCPT (as measured by the Omissions variable). Parent BRI and MI ratings were significantly correlated with parent ratings of Depression and; See Table 3). Teacher BRI and MI ratings were also significantly correlated with teacher ratings of depression (See Table 4). Parent ratings on the MI

were significantly correlated with teacher ratings of depression. Parent ratings on the BRI were not significantly correlated with teacher ratings of depression (p = .07), and teacher ratings on the BRI and MI were not significantly correlated with parent ratings of depression. See Table 5 for correlation coefficients between parent and teacher ratings. As predicted, the BRIEF indices were not significantly correlated with the CCPT variable (all p- values > .23; See Tables 3 and 4). A partial correlation controlling for type of CCPT was also conducted for this hypothesis. Similar to the initial analysis, parent BRI and MI ratings were significantly correlated with parent ratings of depression. Teacher BRI and MI ratings were also significantly correlated with teacher ratings of depression. Parent MI was significantly correlated with teacher ratings of depression. No significant correlations were found between parent BRI and teacher reported depression or teacher BRI/MI and parent reported depression.

Factor Analyses

A principle components factor analysis was conducted to explore the relationship between the BRIEF MI and BRI scales and other performance-based and rating scale measures used in the ADHD assessment battery. A total of twelve variables were entered in the factor analysis. It was predicted that ratings on the BRIEF indices would cluster with ratings of ADHD symptoms, ratings of depression, and teacher ratings of academic outcomes, whereas the CCPT would cluster with the observational measure and full-scale intelligence quotient. Variables included scales from the following measures: the BRIEF (MI and BRI), ADHD-RS-IV (Inattentive subscale and Hyperactive/Impulsive subscale), CCPT (Omissions and Commissions), TOF (Inattention and Hyperactive/Impulsive subscales), Teacher Ratings of Academic Performance (Work completion/accuracy for math, work completion/accuracy for language arts), Wechsler Abbreviated Scale of Intelligence-II (Full Scale Intelligence Quotient)

and BASC-2 (Depression subscale). Two factor analyses were conducted in order to consider parent and teacher ratings separately.

For teacher ratings, four factors with roots greater than one were extracted using varimax rotation, accounting for 67% of the variance. Factor loading coefficients across the four factors ranged from a low of .43 to a high of .93. Variables loading on the first factor, 'Rating Scales,' which accounted for 28% of variance and included teacher ratings on the BRIEF, ADHD-RS-IV Hyperactivity scale, and BASC-2 Depression Scale. The average factor loading for the first factor was .74 (range = .61 to .89). In addition, although teacher ratings of inattention on the ADHD-RS-IV loaded most strongly on the second factor (-.60), this variable also loaded on the first factor with a coefficient of .56. The second factor, 'Long-Term Academic Outcomes,' accounted for 15% of the variance and included work completion/accuracy ratings for language arts and math as well as the ADHD-RS-IV inattentive scale in a negative direction. The average factor loading for the second factor was .81 (range = .60 to .93). Although the BRIEF MI loaded most strongly on the first factor (.68), this variable also loaded on the second factor with a coefficient of -.46. Variables loading on the third factor, 'TOF/CCPT Omissions', included the two TOF scales and the CCPT Omissions scale, accounting for 13% of the variance. The average factor loading for the third factor was .74 (range = .47 to .89). The fourth factor, 'FSIQ/CCPT Commissions,' included the WASI-II FSIQ and the CCPT Commissions scale and accounted for 11% of the variance. The average factor loading for the fourth factor was .62 (range = .43 to .80). Although teacher ratings of depression loaded most strongly on the first scale (.61), this variable loaded to a lesser degree on the fourth factor also (-.49). In addition, the CCPT Omissions scale also loaded on the fourth factor (-.40), in addition to its primary loading on the third factor (.47; See Table 6).

For parent ratings, four factors with roots greater than one were extracted using varimax rotation, explaining 65% of the variance (See Table 7). Similar to the teacher factor analysis, the first factor, 'Rating Scales,' included the BRIEF subscales, ratings on the ADHD-RS-IV subscales, and parent ratings of depression on the BASC-2, and accounted for 25% of the variance. The average factor loading for the first factor was .75 (range .62 to .87). Teacher ratings of math and language arts work completion/work accuracy loaded on the second factor, 'Long-Term Academic Outcomes,' explaining 16% of the variance. The average factor loading for the second factor was .94 (range .94 to .95). The third factor, 'TOF/CCPT Omissions,' explained 14% of the variance and included the TOF subscales and the CCPT Omissions scale. The average factor loading for the third factor was .75 (range .53 to .89). The fourth factor, FSIQ/CCPT Commissions,' included the FSIQ from the WASI-II and the CCPT Commissions scale and explained 10% of the variance. The average factor loading for the fourth factor was .71 (range .65 to .77).

CHAPTER 5

DISCUSSION

Despite ample research related to assessment of ADHD, there is no gold standard for diagnosing the disorder (Jarratt et al., 2005). Given the requirement of impairment in more than one setting, multi-rater multi-method assessment that includes background information, a diagnostic interview, and use of various methods to assess symptoms and behavior (e.g., informant ratings, direct observation, and performance-based measures of impulse control and attention) is typically recommended (Sims & Lonigan, 2012). Measures that assess executive functioning (EF) are often included in ADHD assessment batteries as well. Previous research has indicated that rating scales assessing EF do not tend to correlate with performance-based measures of EF, suggesting that rating scales and performance-based measures may be assessing different components of the same construct.

Given these findings, the current study sought to replicate and extend previous research by examining the relationship of the BRIEF with a variety of measures typically used in an ADHD battery, including an ADHD rating scale, a performance-based measure of EF, measures of academic performance, an observational measure completed during testing, and another rating scale assessing a dissimilar construct (i.e., depression).

As predicted in the first hypothesis, the BRIEF Metacognitive Index was correlated with ratings of inattention (on the ADHD-RS). This is consistent with previous research (e.g., Mahone

et al., 2002; McAuley et al., 2010) that has found the BRIEF MCI to correlate with ratings of inattention. Correlations were much stronger when considering the same rater for both scales (i.e., parent rating on Metacognitive Index with parent rating of inattention) as opposed to different raters (i.e., teacher ratings on the Metacognitive Index with parent ratings of inattention), reflecting common construct, same method (i.e., rating scale), as well as influence of the same rater completing both measures.

There was also support for the second hypothesis that the BRIEF (Metacognitive and Behavioral Regulation Indices) would not significantly correlate with a performance-based measure (as measured by Omissions, Variability, Commissions, and Hit Reaction Time on the CCPT) due to differences in method (i.e., rating of EF versus performance-based measure). This was true in the initial analyses, as well as additional analyses controlling for type of CCPT administered. This finding is consistent with previous research, which has found minimal relationships between rating scales and performance-based measures of EF (Biederman et al., 2008; Bodnar et al., 2007; McAuley et al., 2010; Topak et al., 2013).

The third hypothesis posited that the BRIEF (Metacognitive and Behavioral Regulation Indices) would be negatively correlated with teacher questionnaires assessing longer-term goal pursuit. This hypothesis was designed to test the hypothesis that performance-based measures of EF assess cognitive efficiency or lower levels of EF, whereas rating scales assess longer-term goal pursuit, or higher levels of EF (Barkley & Fischer, 2011; Toplak et al., 2013). As predicted, teacher ratings on the Metacognitive Index were significantly correlated with ratings of work completion/work accuracy both in language arts and mathematics. Teacher ratings on the Behavioral Regulation Index were not correlated with academic ratings. In addition, parent ratings on the BRIEF were not correlated with academic ratings. McAuley et al. (2010) found a

strong correlation between parent ratings on the Metacognitive Index and child performance in reading and mathematics on a measure of academic achievement. Mahone et al. (2002) found that parent ratings on the Behavioral Regulation Index and Metacognitive Index were correlated with math scores on an achievement test but not with reading scores. Teacher ratings of academic performance were used in the current study, rather than performance on an academic achievement measure, in order to assess longer-term goal pursuit rather than short-term performance on a structured task. It may be that teacher ratings on the BRIEF correlated with academic performance, whereas parent ratings did not, due to the influence of same source. It is also possible that teachers, rather than parents, are better able to identify EF deficits due to greater familiarity with age appropriate behavior (Mares et al., 2007). However, it should also be noted that teacher ratings on the BRIEF MI, but not BRI index, were significantly correlated with academic ratings. Previous research has found the BRIEF MI Index to be more strongly correlated with academic achievement than the BRI Index (McAuley et al., 2010), thus providing support for similar construct.

The fourth hypothesis predicted that an observational measure (TOF Inattention subscale) would be more strongly correlated with a performance-based measure (CCPT Omissions and Variability) than with the BRIEF (Metacognitive and Behavioral Regulation indices). Both the performance-based measure and observational data (i.e., the TOF) reflect performance in short-term, structured settings. Conversely, the BRIEF reflects long-term performance. Results supported this hypothesis; with the observational measure correlating with the performance-based measure but not correlating with the BRIEF indices (parent or teacher ratings). A similar pattern of correlations, supporting the hypothesis, was also found when controlling for type of CCPT administered. This supports an influence for setting (e.g., short-term performance during

the assessment versus long-term day-to-day performance) and suggests that children might perform differently in various settings.

In order to assess the influence of common method versus common construct, the fifth hypothesis proposed that the BRIEF (Metacognitive and Behavioral Regulation Indices) would be more strongly correlated with ratings of depression than with a performance-based measure of EF (i.e., the CCPT). The prediction was that if the BRIEF correlated more strongly with a rating scale assessing a dissimilar construct (i.e., depression) than with a performance-based measure designed to assess the same construct (i.e., executive functioning), there would be greater support for method of measurement as opposed to construct being measured. Support for this hypothesis was found in initial analyses as well in additional analyses controlling for type of CCPT. Consistent with previous research (e.g., De Los Reyes, 2011), ratings across informants were more discrepant than ratings from the same source. More specifically, parent ratings on the Metacognitive and Behavioral Regulation Indices were correlated with parent ratings of depression and teacher ratings on the Metacognitive and Behavioral Regulation Indices were correlated with teacher ratings of depression. Although parent ratings on the Metacognitive Index were correlated with teacher ratings of depression, there were no other significant correlations between ratings from different sources (i.e., teacher ratings on Metacognitive or Behavioral Regulation Indices not significantly correlated with parent ratings of depression). As predicted, neither parent nor teacher ratings on the BRIEF indices were significantly correlated with the performance-based measure (CCPT Omissions).

Overall, significant correlations were found between variables assessing dissimilar constructs (i.e., depression and executive functioning) using the same method (rating scale), whereas no significant correlations were found between variables assessing the same construct

(executive functioning) assessed using a different method (rating scale versus performance based measure). Although this finding seems to support the influence of method over construct, it is also possible that this finding is due commonalities between symptoms of ADHD, or executive functioning in particular, and depression, as previous research has found that 14 to 18% of children with ADHD experience comorbid symptoms of anxiety and/or depression, based on parent report (Larson, Russ, Kahn, & Halfon, 2011). In addition, the pattern of strong correlations between ratings from the same source despite dissimilar constructs (i.e., parent rating of depression and parent ratings on BRIEF) as compared to minimal relationship across sources (parent rating of depression and teacher ratings on BRIEF) suggests support for possible informant discrepancies. More specifically, parents as well as teachers, might have rated children as higher or lower on both measures based on subjective impressions overall rather than child's actual behavior.

The sixth hypothesis predicted parent and teacher ratings on the BRIEF would group with other rating scales and longer-term outcomes (i.e., parent and teacher ratings on the ADHD-RS-IV, ratings of academic performance, and ratings of depression on the BASC-2), whereas the CCPT would group with other short-term performance measures (i.e., FSIQ on the WASI-II and the TOF). Two factor analyses were conducted, one with parent ratings and another with teacher ratings to evaluate this hypothesis.

For teacher ratings, four factors with roots greater than one were extracted using varimax rotation. Although four factors emerged, rather than two as predicted, the four factors tended to parallel the type of assessment method used. For example, the first factor labeled 'Rating Scales' was composed of teacher ratings on the BRIEF, ratings of hyperactivity, and ratings of depression. The second factor, labeled 'Long-Term Academic Outcomes' consisted of longer-

term academic outcomes in language arts and math and also included teacher ratings of inattention. The third factor was labeled 'TOF/CCPT Omissions' and consisted of both variables from the observational measure (i.e., TOF Inattention and TOF Hyperactivity/Impulsivity) and one variable from the performance-based measure (i.e., CCPT Omissions). The fourth factor, 'FSIQ/CCPT Commissions' consisted of Full Scale IQ as well as one variable from the performance-based measure (i.e., CCPT Commissions). It is not entirely clear why the third and fourth factors clustered in the manner they did, given that all four variables reflect performance in short-term, structured settings. However, it is possible that children displaying greater CCPT Commission errors (a measure of impulsivity) experienced more difficulty on the IQ test whereas Omission errors (a measure of inattention) may be more prone to be noticed by examiners and reflected in observer's ratings of Inattention and Hyperactivity/Impulsivity on the observational measure. Thus, although the TOF was considered a shorter-term "performance" measure for this study, it is technically a rating scale and thus influence of possible rater bias (i.e., the bias of the clinician completing the evaluation) must be considered. More specifically, examiners may be more apt to consider the child's behavior during the CCPT rather than during intelligence or achievement testing when completing the TOF. One potential reason for this is the order of measures in the battery (i.e., the CCPT is first, followed by IQ and achievement testing, and then a child interview), suggesting that a primacy effect might be influencing clinicians. Secondly, clinicians have much more time to observe the child and take notes during the CCPT when little is required of them during the task, versus IQ and achievement tests in which the examiner is active in administering the assessments.

The factor analysis that considered parent ratings clustered in a similar manner. One primary difference is that teacher ratings of inattention loaded with 'Long-Term Academic

Outcomes' whereas parent ratings of inattention loaded with the other rating scales. Thus, for parents, there was clear support for the influence of method, as all of the parent ratings clustered on the same factor. Given that teachers completed both the ADHD-RS-IV and the questionnaires assessing academic outcomes, it is possible that these results could also be explained by a rater effect. For example, teachers may have been influenced by their ratings in regard to the child's average level of work completion and work accuracy when later responding to items on the ADHD-RS-IV rating scale, or vice versa.

Clinical Implications

This study considered the relationships between parent and teacher ratings on the BRIEF and other measures included in an ADHD battery. The relationship between the BRIEF and some of these measures has been well researched (i.e., performance-based measures, rating scales of ADHD symptoms). However, additional measures were included in the study (e.g., an observational measure, longer-term academic outcomes, a subscale assessing a dissimilar construct) in order to gather more information regarding the influence of common method and common construct in an ADHD battery. Results of the current study are consistent with prior research indicating that rating scales and performance-based measures of EF are minimally correlated and potentially assessing different constructs but likely both contributing useful information. As mentioned, Toplak et al. (2013) proposed that performance-based measures assess cognitive efficiency whereas rating scales assess longer-term goal pursuit. Results of this study run parallel to this idea, given that ratings on the BRIEF were related to academic outcomes, which reflect long-term goal pursuit. Similarly, Barkley and Fisher (2011) proposed that EF be viewed as a meta-construct, with lower levels (those targeted by performance-based measures) serving as a foundation for higher levels (those targeted by rating scales). Stated

differently, the results of this study support the premise that performance-based measures may measure underlying skills, whereas rating scales measure the ability to apply those skills at school and at home (McAuley et al., 2010). McAuley et al. (2010) note that previous research has found associations between performance-based measures and a variety of school-related outcomes (e.g., teacher perception of student functioning, risk of grade retention, and performance on state-wide exams). Rating scales may predict academic functioning to an even greater degree than performance-based measures due to their ability to take into account fluctuations in child behavior from day to day and across different settings. Thus, future studies should compare the relative merits of the BRIEF versus performance-based measures in predicting school-related and other real-world outcomes (e.g., interpersonal relationships, impairment at home).

The minimal relationship between performance-based measures and rating scales of EF found in the current study has numerous clinical implications in terms of assessment, diagnosis, recommendations, and treatment. Given that rating scales and performance-based measures provide unique information, evaluations should ideally include both rating scales and performance-based measures. Furthermore, these measures should not be used interchangeably in ADHD evaluations. For example, it may problematic to diagnose a child with ADHD based only on a performance-based measure, given that these measures do not consider contextual factors and are influenced by other factors (e.g., medication and IQ). Conversely, a diagnosis based primarily on rating scales, without consideration of performance-based measures, may miss information related to how well the child can perform in highly structured settings.

Secondly, clinicians should consider the relationship between performance-based measures and rating scales when making recommendations and considering treatment options.

For example, if a child performs well on the performance-based measure but the rating scales suggest EF deficits, the clinician might hypothesize that the child could benefit from increased structure in the home and school environments, given their relative strength on the performance-based measure (which is highly structured) as compared to their day to day functioning. Furthermore, it is possible that children who are able to perform well on performance-based measures may experience greater benefit from increased support and structure in their environment, as compared to children who experience EF deficits as evidenced by both performance-based measures and rating scales. Conversely, if a child performs poorly on a performance-based measure but is rated in the average range on rating scales, the clinician might consider other possibilities (e.g., something about the testing situation was less than ideal for the child) or diagnoses, such as anxiety or a learning disorder.

Results of the present study also suggest that clinicians carefully consider patterns of assessment findings within and across sources (i.e., parent and teacher). In the present study, ratings of inattention clustered more closely with the 'Long-Term Academic Outcomes' in the teacher factor analysis and more closely with the 'Rating Scale' factor in the parent factor analysis. The BRIEF Metacognitve Index also loaded more highly on the 'Long-Term Academic Outcomes' variable in the teacher factor analysis compared to parent factor analysis; although its primary loading was with the 'Rating Scales' factor in both analyses. It is possible that teacher ratings clustered more closely with 'Long-Term Academic Outcomes' due to teacher's ability to better predict how symptoms influence schoolwork. However, it seems likely that this relationship was due, at least in part, to same source. More specifically, teachers completed questionnaires assessing academic outcomes as well as the BRIEF and ratings of inattention.

Post-hoc correlations between parent and teacher ratings were conducted for key measures in the study (i.e., BRIEF, ADHD-RS-IV, ratings of depression on BASC-2) to further examine influence of source. Correlation coefficient for parent and teacher ratings ranged from r = -.07 (for ratings of depression) to r = .42 (for ratings of hyperactivity on the ADHD-RS) with a mean of r = .27 for the five variables. The modest correlations between parent and teacher ratings are consistent with previous research suggesting that informant discrepancies seem to be the norm when considering ratings of child behavior, with correlation coefficients typically in the .20s (Achenbach, McConaughy & Howell, 1987). As mentioned, it has been proposed that these discrepancies arise due to true differences in child behavior in various settings rather than lack of accuracy from one rater (De Los Reyes, 2011). Recent work by Burns, Servera, Bernad, Carrillo & Geiser (2014) further explored this issue by examining ratings of ADHD symptoms and academic factors from two sources at school (i.e., a teacher and an aide) and two sources at home (i.e., mother and father) in order to examine whether informant discrepancies could be attributed to source effects (e.g., parent versus teacher) or context effects (home vs. school). Results indicated that teacher/aide ratings and mother/father ratings were interchangeable but correlations of ADHD symptoms between home and school were small, providing support for the premise that ADHD symptoms vary based on context more than source. In regards to the current study, findings by Burns et al. may suggest that teacher ratings clustered more closely with long-term academic outcomes due to true differences in child behavior in the context of school rather a bias in teacher ratings.

In addition to an influence for source, the present study also found strong support for the influence of method of measurement. More specifically, the BRIEF was found to correlate with rating scales assessing ADHD symptoms (ADHD-RS) as well as rating scales assessing

symptoms of depression (BASC-2). Given the similarity in method of measurement and similarity of construct in the BRIEF and ADHD-RS, the correlation between these measures is not surprising. However, the correlation between the BRIEF and ratings of depression was somewhat more unexpected. The Behavioral Regulation Index in particular was strongly correlated with ratings of depression (r = .61 for teacher and r = .63 for parent). This could be attributed to two factors. First, it is possible that the strong correlation between the BRIEF and ratings of depression was due to common method (i.e., both rating scales). However, this does not explain why the Metacognitive Index was only moderately correlated with ratings of depression whereas the Behavioral Regulation Index was strongly correlated. A second, and potentially more plausible explanation is that the BRIEF variables were correlated with ratings of depression due to similarities in the constructs assessed. Barkley (2014) proposed that children with ADHD often experience deficient emotional self-regulation (DESR), which is defined as the inability to inhibit the primary emotional reaction in order to self-soothe and refocus attention (typically toward an external goal). Given that the Behavioral Regulation Index is comprised of subdomains such as inhibition, emotional control, and shifting, it makes sense that these two measures would correlate in the current study. Clinically, it is useful for examiners to be aware of the relationship between some aspects of EF (particularly inhibition, emotional control, and shifting) and how these constructs might manifest in child behavior (e.g., low frustration tolerance or emotional excitability in general). This information could also be useful in diagnosis, as well as in recommendations and treatment. In addition, clinicians should carefully consider whether elevations on scales assessing emotional difficulties (i.e., anxiety or depression scales) are truly suggestive of comorbid depressive disorders or rather associated with DESR associated with ADHD.

An additional implication is based on findings that ratings on the TOF were largely independent of other measures in an ADHD battery and raises questions regarding the clinical utility of structured observational measures in ADHD evaluations. The finding that the TOF is independent of other measures may reflect variability in child behavior across settings, which suggests the TOF could provide valuable incremental validity for better understanding patterns of child behavior across various types of settings, particularly for cases in which parent and teacher ratings are strongly discrepant. There is little research exploring the construct validity of structured observation in ADHD evaluations and only one study that examines the construct validity of the TOF in particular (McConaughy et al., 2010). Thus, additional research is needed to better understand how information obtained from the TOF may provide clinical utility.

From a theoretical standpoint, the results of this study are not necessarily consistent with Barkley's model of ADHD. As mentioned, Barkley's model of ADHD proposed that behavioral regulation is the foundation for executive functions (Barkley, 2006). He also proposes that behavioral inhibition is necessary for the development and implementation of executive functions. In the current study, short-term, performance-based measures may place more emphasis on assessment of behavioral regulation, whereas longer-term measures may assess the executive functions that are built on behavioral inhibition. From this perspective, we would expect a moderate correlation between the two types of measures, as individuals struggling with behavioral inhibition would also struggle with the wider range of executive functioning skills. A moderate rather than strong relationship is expected as some individuals may display good ability to inhibit behavior in the short term, but experience greater difficulty on longer term tasks due to the cumulative influence of various executive functioning deficits. In the present study, there was no significant correlation between child performance on short term performance-based

measures and longer term academic outcomes. However, this lack of relationship may be due to differences in constructs and measurement as well as possible rater bias.

An additional model of ADHD, proposed by Sergeant (2000) stated that children with ADHD struggle with deficit levels of motivation, arousal, or effort rather than executive or attentional deficits. As mentioned, Fuggetta (2006) used processing speed as a measure of arousal when examining Sergeant's model and found that children with ADHD demonstrated deficits as compared to children without ADHD. Given that a neither processing speed, nor motivation or effort were examined in the current study, it is unclear whether results support or do not support Sergeant's model.

Limitations

One limitation of this study is related to the representativeness of the sample. The majority of children included in the current study were male (70.5%), Caucasian (89.5%), and from families who reported an annual income of less than $60,000 (79.9%). Although the gender differences are representative of true gender differences in the prevalence of ADHD, generalizability is limited in terms of ethnicity and socioeconomic status. A second limitation of the study involves the possible influence of child medication status on measures in the battery. The original sample contained 113 participants. However, eight participants were excluded from the study due to inconsistency in medication influence on assessment results (i.e., on medication on a daily basis but not taking medications at the time of the assessment). This was done in order to minimize differences in ratings across measures due effects of medication. An additional group of children (n = 10) taking medication both during the assessment and prior to assessment were included in the study. Analyses were conducted to compare assessment results for those ten participants to other participants in the study and no significant differences between the groups

were found. Although analyses found no significant differences between groups, it is possible that medication could influence parent and teacher ratings. For example, if a child had recently begun a trial of medication prior to the evaluation, teachers and/or parents may have endorsed fewer symptoms or less impairment due to a placebo effect. Conversely, parents and/or teachers may have been influenced by knowledge of the child being on medication to rate symptoms more strongly to send a message that the current medication was not effective in managing perceived symptoms. Previous research indicates that CPTs are especially sensitive to the effects of stimulant medication (Epstein et al., 2006), but research is less clear regarding the effects of medication on rating scales. Thus, including children on medication in the current study may have resulted in greater inconsistency between different types of measures (e.g., rating scales and CCPT), as well as perhaps greater rater bias in select cases (e.g., placebo effect).

Since this study made use of archival data, there were limitations related to measures used/available. For example, only one performance-based measure of executive functioning (i.e., CCPT) was available for inclusion. Given that the CCPT is just one type of performance-based measure targeting a few facets of EF (i.e., inattention and inhibition), it would have been preferable to have various performance-based measures of EF to compare to the BRIEF. Incorporating measures such as the Wisconsin Card Sort Test, Rey-Osterreith Complex Figure Test, Stroop Word-Color Test, Tower of London, or Trail Making Test, might have provided more information about other facets of EF (e.g., initiation, organization, planning, inhibition, shifting of thought or attention). Ideally, measures would be included to assess each of the EF deficits most consistently associated with ADHD (i.e., inhibition, vigilance, planning, and working memory). In addition, it might have been helpful to compare specific subscales of the BRIEF (e.g., Working Memory, Inhibit, Shift, Plan/Organize) with performance-based measures

thought to mirror these subscales, as done by Toplak et al. (2009). For example, subtests such as Digit Span and Spatial Span from the Wechsler Intelligence Scale for Children (WISC) could have been used as measures of working memory and compared directly to the Working Memory subscale on the BRIEF.

An additional limitation related to the performance-based measure is that the version of the CCPT used in the test battery was changed midway through the data collection period. Although statistical analyses were included to control for type of CCPT administered, it would have been preferable if the same version of CCPT was used throughout the study. In comparing means for the variables of interest (Omissions, Commissions, Hit RT, and Variability) a significant difference was found for Commission errors between the two newer CCPTs (i.e., Conners CPT-3 and KCPT-2) and the previous CCPT (i.e., CCPT-2). This difference was likely due to improvements made to the Commissions variable in the two newer versions to address ceiling issues. It may be that the two newer CCPTs provided a more valid measure of Commission. It is also possible that measures of Omissions and Commissions failed to group together on the factor analysis due to changes made in the Commissions variable on the newer versions of the CCPT. Consideration should also be given to that fact that the K-CPT 2 has a shorter administration time (i.e., 7 minutes as compared to 15 minutes) and uses pictures rather than letters for stimuli. Thus, 6 and 7 year olds taking the K-CPT 2 were exposed to a shorter and likely more developmentally appropriate measure than 6 and 7 year olds taking the CCPT-2.

There were also limitations related to the measure used to operationalize longer-term goal pursuit. As mentioned, Toplak et al. (2013) proposed that rating scales of EF assess longer-term goal pursuit whereas performance-based measures of EF assess EF in ideal, short-term, structured settings. In the current study, teacher ratings of work completion and work accuracy

were used to operationalize longer-term goal pursuit. However, this is somewhat flawed in that this measure is essentially a type of informant rating. Given the known influence of source, it is possible that correlations were found between teacher rating scales and measures of longer-term goal pursuit due to the influence of same source rather than the relationship between the two measures. Ideally the present study would have included a more objective measure of longer-term goal pursuit, such children's grades based on a copy of the child's report card or transcript.

Lastly, the manner in which the rating scales were filled out should be considered. More specifically, parents were asked to fill out all of the rating scales in the waiting room of the clinic prior to the appointment. However, teachers were given the rating scales and then asked to fill them out and mail them back to the clinic. Thus, teachers had more freedom to take breaks and consider questions for a longer period of time. This discrepancy may have influenced the way in which parents and teachers filled out the forms. In addition, given that parents had already made the decision to have their child evaluated when they completed the ratings scales, they may have been more likely to endorse ADHD symptoms, as compared to teachers, due to a confirmatory bias. In other words, parents may have been more likely than teachers to endorse ADHD symptoms on the rating scales in order to "confirm" their previous bias that the child has ADHD. Future research might explore this issue by conducting a factor analysis with a larger sample size that includes both parent and teacher ratings in the same analysis in order to explore the influence of source in addition to the influence of common method.

Future Research

Future research should continue to explore the relationship between performance-based measures and rating scales of EF in order to determine how we can best use information from both types of measures in evaluations. More specifically, future research should compare the

relative merits of rating scales and performance-based measures in predicting real-world outcomes. For example, Barkley and Fischer (2011) examined the predictive validity of the two approaches to measuring EF among adults with ADHD, and found that EF ratings predicted real-world outcomes (e.g., occupational outcomes) whereas performance-based measures did not. Currently, there is minimal research examining the predictive validity of ADHD assessment outcomes for children. Future research might explore the relative ability of the BRIEF and performance-based measures to predict real-world outcomes relative to child functioning, such as grade retention, performance on statewide exams, interpersonal relationships, and completion of tasks at home.

In addition, future research should consider the impact of medication on ADHD evaluations, given that it is not uncommon for children to be referred for an evaluation after starting a trial of medication. The effects of medication on performance-based measures and rating scales should be considered. In future studies, it will be important to consider the effects of medication on informant ratings on rating scales. Based on the current study, three broad areas might be useful for future research to address. First, future research should examine the influence of children beginning a trial of medication prior to the evaluation and whether parents/teachers rate the child differently due the expectation that the medication will have an effect (i.e., the placebo effect), rather than rating the child's actual behavior. In addition, for children who have not yet been prescribed ADHD medication, consideration should be given to the possibility that parents/teachers might inflate ratings in order to press for medication intervention. Lastly, it would be useful to explore the effects of medication side effects on rating scales. More specifically, children taking medication prior to the evaluation might experience side effects (e.g., irritability, nervousness, insomnia, stomach aches) that increase ratings of emotionality or

somatic complaints in the evaluation. In sum, additional studies are needed investigate the differential effects of medication on performance-based measures and rating scales in order to provide more concrete guidelines for clinicians completing evaluations.

Additional research exploring the role of the TOF and other observational measures in ADHD evaluations would also be beneficial. In the current study, the TOF was found to be largely independent of other measures in the ADHD battery. Consistent with De Los Reyes' (2011) proposal regarding parent/teacher discrepancies, it is possible that the TOF is independent of other measures because it represents child behavior in another setting (e.g., parent ratings representing behavior at home, teacher ratings representing behavior at school, TOF representing child behavior in the clinic) and provides valuable information. However, it is unclear how this information should be integrated in the evaluation. One possibility is that the TOF may provide information regarding how well children can perform in short-term, structured settings, which makes sense given that the TOF correlated more highly with the CCPT (which is also administered in short-term, structured setting) than with the BRIEF (which is thought to reflect long-term outcomes). If this is the case, the TOF might provide information regarding how much children could benefit from increased support and structure in the classroom or at home. Future studies should continue to examine the convergent, divergent, and predictive validity of the TOF in order to determine how to best use the measure in ADHD evaluations.

Future research should also continue to explore the role of informant discrepancies in ADHD evaluations in general, as well as in measures associated with EF. Much of the previous research in this area has focused on parent or teacher ratings. Thus, a strength of the current study was that both parent and teacher ratings were included. However, these ratings were considered separately in factor analyses in order to focus on the influence of method of

measurement and construct. As mentioned, future research may wish to could involve a factor analysis with a larger sample including both parent and teacher ratings in order to explore the influence of method of measurement and source simultaneously.

Lastly, future research should continue to address the variability in assessment batteries used to diagnose ADHD. The present study contributes to previous research suggesting that multi-method, multi-rater assessment of ADHD is essential. The American Academy of Child and Adolescent Psychiatry (AACAP, 2007) guidelines recommend that ADHD evaluations include clinical interviews, behavior rating scales, and a child interview. The American Academy of Pediatrics (AAP) published updated guidelines in 2011 on evaluation and treatment of ADHD in children and also recommends obtaining information from both parents and teachers but makes no mention of performance-based measures (AAP, 2011). In a summary of evidence-based assessment of ADHD for children and adolescents, Pelham et al. (2005) recommend structured interviews, behavioral observations, DSM-based rating scales, and global impairment measures be included in assessment of ADHD. It is noteworthy that none of the aforementioned guidelines include performance-based measures in their recommendations. Thus, future research should continue to explore the role of performance-based measures in ADHD evaluations.

Assessment procedures also vary based on the setting of the evaluation and type of psychologist (e.g., clinical, counseling, school) conducting the evaluation (Handler & DuPaul, 2002). For example, school psychologists completing assessments in academic settings might rely more on teacher reports and naturalistic observation, whereas clinical or counseling psychologists might rely more on parent reports. Overall, there seems to be a large discrepancy between practice guidelines and actual diagnostic practices. Thus, future research should

continue to explore the ideal test battery, balancing breadth of coverage with the limitations of time and resources often encountered in practice.

REFERENCES

Achenbach, T., McConaughy, S., & Howell, C. (1987). Child/Adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213–232. http://dx.doi.org/10.1037/0033-2909.101.2.213

American Academy of Child and Adolescent Psychiatry. (2007). Practice parameters for assessment and treatment of children and adolescents with attention deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(7), 894–921.

American Academy of Pediatrics, Subcommittee on Attention-Deficit/Hyperactivity Disorder, Steering Committee on Quality Improvement and Management. (2011). ADHD: Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics*, 128(5), 1–16. http://dx.doi.org/10.1542/peds.2011-2654

American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric.

Anderson, V., Jacobs, R., & Anderson, P. J. (2008). *Executive functions and the frontal lobes: A lifespan perspective*. New York, NY: Taylor & Francis.

Barkley, R. A. (2006). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (3rd ed.). New York, NY: Guliford Press.

Barkley, R. A. (2009). History of adult ADHD pushed back to 1798! *The ADHD Report*, *17*(1), 1–5.

Barkley, R. A. (2011). Is executive functioning deficient in ADHD? It depends on your definitions and your measures. *The ADHD Report*, *19*(4), 1–10.

Barkley, R. A. (2014). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (4th ed.). New York, NY: Guilford Press.

Barkley, R. A., & Fischer, M. (2011). Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported executive function (EF) deficits versus EF tests. *Developmental Neuropsychology*, *36*(2), 137–161.

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

Bauermeister, J. J., Barkley, R. A., Bauermeister, J. A., Martinez, J. V., & McBurnett, K. (2012). Validity of the sluggish cognitive tempo, inattention, and hyperactivity symptom dimensions: Neuropsychological and psychosocial correlates. *Journal of Abnormal Child Psychology*, *40*, 683–697.

Best, J. R., Miller, P. H., & Jones, L. J. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*, 180–200.

Biederman, J., Petty, C., Fried, R., Black, S., Faneuil, A., Doyle, A. E., . . . Faraone, S. V. (2008). Discordance between psychometric testing and questionnaire-based definitions of executive function deficits in individuals with ADHD. *Journal of Attention Disorders*, *12*, 92–102.

Biederman, J., Petty, C., Fried, R., Fontanella, J., Doyle, A. E., Seidman, L. J., & Faraone, S. V. (2006). Impact of psychometrically defined deficits of executive functioning in adults with attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *163*, 1730–1738.

Bodnar, L. E., Prahme, M. C., Cutting, L. E., Denckla, M. B., & Mahone, E. M. (2007). Construct validity of parent ratings of inhibitory control. *Child Neuropsychology*, *13*(4), 345–362. http://dx.doi.org/10.1080/09297040600899867

Bowers, A. J. (2011). What's in a grade? The multidimensional nature of what teacher-assigned grades assess in high school. *Educational Research and Evaluation*, *17*(3), 141–159.

Burns, G. L., Servera, M., Bernad, M. d. M., Carrillo, J. M., & Geiser, C. (2014). Ratings of ADHD symptoms and academic impairment by mothers, fathers, teachers, and aides: Construct validity within and across settings as well as occasions. *Psychological Assessment.* Advance online publication. http://dx.doi.org/10.1037/pas0000008

Campbell, D. T., & Fiske. D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Collings, R. D. (2003). Differences between ADHD inattentive and combined types on the CPT. *Journal of Psychopathology and Behavioral Assessment*, *25*, 177–189.

Conners, C. K. (1997). *Conners' rating scales: Revised technical manual.* North Towanda, NY: Multi-Health Systems.

Conners, C. K. (2000). *Conners' Continuous Performance Test II*. Toronto, Canada: Multi-Health Systems.

Conners, C. K. (2014a). *Conners' Continuous Auditory Test of Attention (Conners CATA)*. North York, Canada: Multi-Health Systems.

Conners, C. K. (2014b). *Conners' Continuous Performance Test (Conners CPT 3)* (3rd ed.). North York, Canada: Multi-Health Systems.

Conners, C. K. (2015). *Conners Kiddie Continuous Performance Test (Conners K-CPT 2)* (2nd ed., North York, Canada: Multi-Health Systems.

De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *40*(1), 1–9. http://dx.doi.org/10.1080/15374416.2011.533405

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychology Bulletin*, *131*, 483–509.

Dirks, M. A., De Los Reyes, A., Briggs-Gowan, M., Cella, D., & Wakschlag, L. S. (2012). Annual research review: Embracing not erasing contextual variability in children's behavior: Theory and utility in the selection and use of methods and informants in developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *53*(5), 558–574.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., Reid, R., McGoey, K. E., & Ikeda, M. J. (1997). Teacher ratings of attention deficit hyperactivity disorder symptoms: Factor structure and normative data. *Psychological Assessment*, *9*(4), 436–444. http://dx.doi.org/10.1037/1040-3590.9.4.436

Edwards, M. C., Gardner, E. S., Chelonis, J. J., Schulz, E. G., Flake, R. A., & Diaz, P. F. (2007). Estimates of the validity and utility of the Conner's continuous performance test in the assessment of inattentive and/or hyperactive-impulsive behaviors in children. *Journal of Abnormal Child Psychology*, *35*, 393–404.

Egeland, J., Johansen, S. N., & Ueland, T. (2009). Differentiating between ADHD sub-types on CCPT measures of sustained attention and vigilance. *Scandanavian Journal of Psychology*, *50*, 347–354.

Epstein, J. N., Conners, C. K., Hervey, A. S., Tonev, S. T., Arnold, L. E., Abikoff, H. B., Elliot, G., . . . Wigal, T. (2006). Assessing medication effects in the MTA study using neuropsychological outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(5), 446–456.

Frazier, T. W., Demaree, H. A., & Youngstrom, E. A. (2004). Meta-analysis of intellectual and neuropsychological test performance in attention-deficit/ hyperactivity disorder. *Neuropsychology*, *18*, 543–555.

Frazier, T. W., Youngstrom, E. A., Glutting, J. J., & Watkins, M. W. (2007). ADHD and achievement: Meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. *Journal of Learning Disabilities*, *40*(1), 49–65.

Fuggetta, G. P. (2006). Impairment of executive functions in boys with attention deficit/hyperactivity disorder. *Child Neuropsychology*, *12*(1), 1–21. http://dx.doi.org/10.1080/09297040500203418

Gioia, G. A., Isquith, P. K., & Guy, S. C. (2001). Assessment of executive functions in children with neurological impairment. In R. J. Simeonsson & L. Rosenthal (Eds.), *Psychological and developmental assessment: Children with disabilities and chronic conditions* (pp. 317–356). New York, NY: Guilford Press.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function, *Child Neuropsychology*, *6*(3), 235–238.

Greenberg, L. M. (1996). *Test of variables of attention.* Los Alamitos, CA: Universal Attention Disorders.

Halperin, J. M., Trampush, J. W., Miller, C. J., Marks, D. J., & Newcorn, J. H. (2008). Neuropsychological outcome in adolescents/young adults with childhood ADHD: Profiles of persisters, remitters and controls. *Journal of Child Psychology & Psychiatry*, *49*(9), 958–966.

Handler, M., & DuPaul, P. (2002). Diagnosis of Childhood ADHD: Differences across psychology speciality areas. The ADHD Report, 7–9.

Huang-Pollock, C. L., Karalunas, S. L., Tam, H., & Moore, A. N. (2012). Evaluating vigilance deficits in ADHD: A meta-analysis of CPT performance. *Journal of Abnormal Psychology*, *121*(2), 360–371.

Huang-Pollock, C. L., & Nigg, J. T. (2003). Searching for attention deficit in attention deficit hyperactivity disorder: The case of visuospatial orienting. *Clinical Psychology Review*, *23*, 801–830. http://dx.doi.org/10.1016/S0272-7358(03)00073-4

Jarratt, K. P., Riccio, C. A., & Siekierski, B. M. (2005). Assessment of attention deficit hyperactivity disorder (ADHD) using the BASC and BRIEF. *Applied Neuropsychology*, *12*(2), 83–93. http://dx.doi.org/10.1207/s15324826an1202_4

Jonsdottir, S., Bouma, A., Sergeant, J. A., & Scherder, E. J. (2006). Relationship between neuropsychological measures of executive function and behavioral measures of ADHD symptoms and comorbid behavior. *Archives of Clinical Neuropsychology*, *21*, 383–394.

Lambek, R., Tannock, R., Dalsgaard, S., Trillingsgaard, A., Damm, D., & Thomsen, P. H. (2010). Validating neuropsychological subtypes of ADHD: How do children "with" and "without" an executive function deficit differ? *Journal of Child Psychology and Psychiatry*, *51*(8), 895–904.

Langberg, J. M., Dvorsky, M. R., & Evans, S. W. (2013). What specific facets of executive function are associated with academic functioning in youth with attention-deficit/hyperactivity disorder? *Journal of Abnormal Child Psychology*, *41*, 1145–1159.

Larson, K., Russ, S. A., Kahn, R. S., & Halfon, N. (2011). Patterns of comorbidity, functioning, and service use for US children with ADHD, *Pediatrics*, *127*(3), 462–470.

Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., & Denckla, M. B. (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, *17*(7), 643.

Mares, D., McLuckie, A., Schwartz, M., & Saini, M. (2007). Executive function impairments in children with attention-deficit hyperactivity disorder: Do they differ between school and home environments? *Canadian Journal of Psychiatry*, *52*(8), 527–534.

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*, *16*, 495–505.

McCandless, S., & O'Laughlin, L. (2007). The clinical utility of the behavior rating inventory of executive function (BRIEF) in the diagnosis of ADHD. *Journal of Attention Disorders*, *10*(4), 381–389.

McConaughy, S. H., & Achenbach, T. M. (2004). *Manual for the Test Observation Form for Ages 2–18*. Burlington: University of Vermont, Research Center for Children, Youth, & Families.

McConaughy, S. H., Harder, V. S., Antshel, K. M., Gordon, M., Eiraldi, R., & Dumenci, L. (2010). Incremental validity of test session and classroom observations in a multimethod assessment of attention deficit/hyperactivity disorder. *Journal of Clinical Child and Adolescent Psychology*, *39*, 650–666.

Miller, M., Nevado-Montenegro, A. J., & Hinshaw, S. P. (2012). Childhood executive function continues to predict outcomes in young adult females with and without childhood-diagnosed ADHD. *Journal of Abnormal Child Psychology*, *40*, 657–668.

Muir-Broaddus, J. E., Rosenstein, L. D., Medina, D. E., & Soderberg, C. (2002). Neuropsychological test performance of children with ADHD relative to test norms and parent behavior ratings. *Archives of Clinical Neuropsychology*, *17*, 671–689.

Murphy, K. R., Barkley, R. A., & Bush, T. (2001). Executive functioning and olfactory identification in young adults with attention deficit-hyperactivity disorder, *Neuropsychology*, *15*(2), 211–220. http://dx.doi.org/10.1037//0894-4105.15.2.211

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

Pauli-Pott, U., & Becker, K. (2011). Neuropsychological basic deficits in preschoolers at risk for ADHD: A meta-analysis. *Clinical Psychology Review*, *31*, 626–637.

Pelham, W. R., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 449–476.

Polanczyk, G., Silva de Lima, M., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and meta regression analysis. *The American Journal of Psychiatry*, *164*(6), 942–948.

Reddy, L. A., Hale, J. B., & Brodzinsky, L. K. (2011). Discriminant validity of the Behavior Rating Inventory of Executive Function Parent Form for children with attention-deficit/hyperactivity disorder. *School Psychology Quarterly*, *26*(1), 45–55. http://dx.doi.org/10.1037/a0022585

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children–2 Manual*. Circle Pines, MN: AGS.

Sanford, J. A., & Turner, A. (1995). *Integrated Visual and Auditory Continuous Performance Test: Administration and interpretation manual* (Version 1.2). Richmond, VA: Braintrain.

Sergeant, J. (2000). The cognitive-energetic model: An empirical approach to attention-deficit hyperactivity disorder. *Neuroscience and Biobehavioral Reviews*, *24*, 7–12.

Sims, D. M., & Lonigan, C. J. (2012). Multi-method assessment of ADHD characteristics in preschool children: Relations between measures. *Early Childhood Research Quarterly*, *27*, 329–337.

Spencer, T. J., Biederman, J., & Mick, E. (2007). Attention-deficit/hyperactivity disorder: Diagnosis, lifespan, comorbidities, and neurobiology, *Journal of Pediatric Psychology*, *32*(6), 631–642.

Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2009). Executive functions: Performance-based measures and the Behavior Rating Inventory of Executive Function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *15*(1), 53–72. http://dx.doi.org/10.1080/09297040802070929

Toplak, M. E., West, R. F., & Stanovish, K. E. (2013). Practitioner review: Do performance-based measures and rating of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*(2), 131–143.

Van Mourik, R., Oosterlaan, J., & Sergeant, J. A. (2005). The Stroop revisited: A meta-analysis of interference control in AD/HD. *Journal of Child Psychology and Psychiatry*, *46*(2), 150–165. http://dx.doi.org/10.1111/j.1469-7610.2004.00345.x

Vaughn, A. J., Epstein, J. N., Rausch, J., Altaye, M., Langberg, J., Newcorn, J. H., Hinshaw, S., . . . Wigal, T. (2011). Relation of outcomes of continuous performance test and ADHD symptoms over time. *Journal of Abnormal Children Psychology*, *39*, 853–864.

Wechsler, D. (2011). *WASI-II manual.* Bloomington, MN: NCS Pearson.

Welsh, M. C., & Pennington, B. F. (1998). Assessing frontal lobe functioning in children: Views from developmental psychology. *Developmental Neuropsychology*, *4*, 199–230.

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

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *Behavioural Assessment of the Dysexecutive Syndrome (BADS).* UK: Thames Valley Test Company.

Wolraich, M. L., Lambert, W. Doffing, M. A., Bickman, L., Simmons. T., & Worley. K. (2003). Psychometric properties of the Vanderbilt ADHD diagnostic parent rating scale in a referred population. *Journal of Pediatric Psychology*, *28*, 559–568.

Zhang, S., Faries, D., Vowles, M., & Michelson, D. (2005). ADHD Rating Scale IV: Sychometric properties from a multinational study as a clinician-admnistered instrument. *International Journal of Methods in Psychiatric Research*, *14*(4), 186–201.

APPENDIX A: DIAGNOSTIC CRITERIA FOR ADHD (DSM-5)

A. A persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development, as characterized by (1) and/or (2):

(1) **Inattention**: Six (or more) of the following symptoms have persisted for at least 6 months to a degree that is maladaptive and inconsistent with developmental level and that negatively impacts directly on social and academic/occupational activities:

Note: The symptoms are not solely a manifestation of oppositional behavior, defiance, hostility, or failure to understand tasks or instructions. For older adolescents and adults (age 17 and older), at least five symptoms are required.

(a) Often fails to give close attention to details or makes careless mistakes in schoolwork, at work, or during other activities (e.g., overlooks or misses details, work is inaccurate)

(b) Often has difficulty sustaining attention in tasks or play activities (e.g., has difficulty remaining focused during lectures, conversations, or lengthy reading).

(c) Often does not seem to listen when spoken to directly (e.g., mind seems elsewhere, even in the absence of any obvious distraction)

(d) Often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (e.g., starts tasks but quickly loses focus and is easily sidetracked)

(e) Often has difficulty organizing tasks and activities (e.g., difficulty managing sequential tasks; difficulty keeping materials and belongings in order; messy, disorganized work; has poor time management; fails to meet deadlines)

(f) Often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (e.g., schoolwork or homework; for older adolescents and adults, preparing reports, completing forms, reviewing lengthy papers)

(g) Often loses things necessary for tasks or activities (e.g., school materials, pencils, books, tools, wallets, keys, paperwork, eyeglasses, mobile telephones)

(h) Is often easily distracted by extraneous stimuli (for older adolescents and adults, may include unrelated thoughts)

(i) Is often forgetful in daily activities (e.g., doing chores, running errands; for older adolescents and adults, returning calls, paying bills, keeping appointments)

(2) **Hyperactivity and Impulsivity**: Six (or more) of the following symptoms of have persisted for at least 6 months to a degree that is inconsistent with developmental level and that negatively impacts directly on social and academic/occupational activities:

Note: The symptoms are not solely a manifestation of oppositional behavior, defiance, hostility, or a failure to understand tasks or instructions. For older adolescents and adults (age 17 and older), at least five symptoms are required.

(a) Often fidgets with or taps hands or feet or squirms in seat
(b) Often leaves seat in situations when remaining seated is expected (e.g., leaves his or her place in the classroom, in the office or other workplace, or in other situations that require remaining in place)
(c) Often runs about or climbs excessively in situations in which it is inappropriate (**Note**: in adolescents or adults, may be limited to feeling of restlessness)
(d) Often unable to play or engage in leisure activities quietly
(e) Is often "on the go" acting as if "driven by a motor" (e.g., is unable to be or uncomfortable being still for extended time, as in restaurants, meetings; may be experienced by others as being restless or difficulty to keep up with)
(f) Often talks excessively
(g) Often blurts out answers before questions have been completed (e.g., completes people's sentences; cannot wait for turn in conversation)
(h) Often has difficulty waiting his or her turn (e.g., while waiting in line)
(i) Often interrupts or intrudes on others (e.g., butts into conversations, games, or activities; may start using other people's things without asking or receiving permission; for adolescents and adults, may intrude into or take over what others are doing)

B. Several inattentive or hyperactive-impulsive symptoms were present prior to the age of 12 years
C. Several inattentive or hyperactive-impulsive symptoms are present in two or more settings (e.g., at home, school, or work; with friends or relatives; in other activities)
D. There is clear evidence that the symptoms interfere with, or reduce the quality of, social, academic, or occupational functioning.
E. The symptoms do not occur exclusively during the course of schizophrenia or other psychotic disorder and are not better accounted for by another mental disorder (e.g., mood disorder, anxiety disorder, dissociative disorder, personality disorder, substance intoxication or withdrawal).

Specify whether:

314.01 (F90.2) Combined presentation: If both Criteria A1 (inattention) and A2 (hyperactivity-impulsivity) are met for the past 6 months

314.00 (F90.0) Predominantly inattentive presentation: If Criterion A1 (inattention) is met but Criterion A2 (hyperactivity-impulsivity) is not met for the past 6 months

314.01 (F90.1) Predominantly hyperactive-impulsive presentation: If Criterion A2 (hyperactivity-impulsivity) is met but Criterion A1 (inattention) is not met for the past 6 months

Specify if:

In partial remission: When full criteria were previously met, fewer than the full criteria have been met for the past 6 months, and the symptoms still result in impairment in social, academic, or occupational functioning.

Specify current severity:

Mild: Few, if any, symptoms in excess of those required to make the diagnosis are present, and symptoms result in no more than minor impairments in social or occupational functioning.

Moderate: Symptoms or functional impairment between "mild" and "severe" are present.

Severe: Many symptoms in excess of those required to make the diagnosis, or several symptoms that are particularly severe, are present, or the symptoms result in marked impairment in social or occupational functioning.

(American Psychiatric Association, 2013, pp. 59-61)

APPENDIX B: BRIEF SAMPLE ITEMS

Rating Scale: N = Never; S = Sometimes; O = Often

Behavioral Regulation Index

8. Tries the same approach to a problem over and over even when it does not work

25. Has outbursts for little reason

54. Acts too wild or "out of control"

62. Angry or tearful outbursts are intense but end suddenly

Metacognitive Index

2. When given three things to do, remembers only the first or last

14. Does not check work for mistakes

22. Forgets to hand in homework, even when completed

57. Has trouble remembering things even for a few minutes

APPENDIX C: ADHD RATING SCALE-IV SAMPLE ITEMS

Rating scale: 0-Never or Rarely, 1-Sometimes, 2-Often, 3-Very Often

Inattentive scale sample items

1. Fails to give close attention to details or makes careless mistakes in schoolwork

9. Has difficulty organizing tasks or activities

Hyperactive/Impulsive scale sample items

2. Fidgets with hands or feet or squirms in seat

10. Is “on the go” or acts as if “driven by a motor”

APPENDIX D: TEST OBSERVATION FORM SAMPLE ITEMS

Rating scale: 0 – no occurrence, 1 – very slight or ambiguous occurrence, 2 – definite occurrence with mild to moderate intensity/frequency and less than 3 minutes total duration, 3 – definite occurrence with severe intensity, high frequency, or 3 or more minutes total duration.

ADHD Problems

Inattention subscale sample items

16. Doesn't seem to listen to what is being said

31. Doesn't concentrate or pay attention for long on tasks, questions, topics

53. Lapses in attention

Hyperactivity-Impulsivity subscale sample items

32. Doesn't sit still, restless, or hyperactive

67. Out of seat

100. Talks too much

APPENDIX E: TEACHER RATINGS OF ACADEMIC PERFORMANCE

Sample Work Completion and Work Accuracy items

Rating Scale: 1 – 0-49%, 2 – 50-69%, 3 – 70-79%, 4 –80-89%, 5 – 90-100%

Estimate the percentage of written math work/language arts work completed (regardless or accuracy) relative to classmates.

Estimate the accuracy of completed written math work/language arts work (i.e.. percent correct work done).

Sample Ratings of Child Performance on Core Subjects (i.e., Mathematics and Reading)

Rating Scale: 1 – far below grade, 2 – somewhat below grade, 3 – at grade level, 4 – somewhat above grade, 5 – far above grade

Current School Performance (list academic subjects and indicate performance based on scale)

APPENDIX F: TABLES

Table 1

Participant Demographic Information

Characteristic	*N*	%	*M*	*SD*
Child gender				
Male	74	70.5		
Female	31	29.5		
Child age in months			94.6	17.9
Child ethnicity				
Caucasian	94	89.5		
African-American	5	4.8		
Hispanic	2	1.9		
Biracial	1	1.0		
Annual family income				
Less than $30,000	51	48.5		
$30,000 to $60,000	33	31.4		
More than $60,000	14	13.4		
Primary diagnosis				
ADHD – Inattentive Type	8	7.6		
ADHD – Hyperactive Type	0	0.0		
ADHD – Combined Type	53	50.5		
ADHD—Not Otherwise Specified	18	17.1		
ADHD Provisional/Rule Out	3	2.9		
Anxiety/Depression	3	2.9		
Learning Disability	5	4.8		
Autism Spectrum Disorder	3	2.9		
Adjustment Disorder	2	1.9		
Other	3	2.9		
No diagnosis given	7	6.7		
ADHD Comorbid Diagnoses				
ODD	31	37.3		
Anxiety/Depression	8	9.6		
Learning Disorder	10	12.1		
Autism Spectrum Disorder	1	1.2		
Other	10	12.0		

Note. ADHD = Attention Deficit/Hyperactivity Disorder; ODD = Oppositional Defiant Disorder; Other = DSM-5 diagnosis not previously listed.

Table 2

Mean, Standard Deviation, and Range for Parent/Teacher Questionnaires and Performance Measures

Questionnaire	*M*	*SD*	Range
Parent Questionnaires			
BRIEF			
Behavior Regulation	62.40	12.46	41–97
Metacognitve	66.16	12.35	34–94
ADHD-RS-IV			
Inattention	89.46	15.82	10–99
Hyperactive/Impulsive	87.47	17.97	10–99
BASC-2			
Depression	55.92	13.71	33–102
Teacher Questionnaires			
BRIEF			
Behavior Regulation	60.66	12.99	41–101
Metacognitve	69.84	12.74	42–104
ADHD-RS-IV			
Inattention	82.82	16.01	25–99
Hyperactive/Impulsive	74.34	24.47	10–99
BASC-2			
Depression	54.72	13.69	40–103
Performance Measures			
CCPT			
Omissions	58.63	16.51	39–113
Variability	56.36	9.71	35–82
Commissions	51.25	8.79	27–72
Hit RT	55.38	10.77	34–84
TOF			
Inattention	63.47	7.19	55–85
Hyperactive/impulsive	63.61	7.82	55–88
WASI-II FSIQ	91.90	10.61	72–119

Note. BRIEF = Behavior Rating Inventory of Executive Function. BRIEF scores are reported above as *T* Scores and scores greater than 65 are clinically significant; ADHD-RS-IV = Attention Deficit Hyperactive Disorder Rating Scale-IV. ADHD-RS-IV scores are reported as percentiles. BASC-2 = Behavior Assessment System for Children–Second Edition. BASC-2 scores are reported above as *T* Scores and scores greater than 70 are clinically significant

Table 3

Pearson Correlation Coefficients Between BRIEF Scales and Other Variables in Correlational Analyses (Parent)

Scales	BRI	MI	ADHD-IN	CCPT-O	CCPT-C	CCPT-V	CCPT-H	MATH	LANG	TOF-IN	TOF-H	BASC-DEP
BRI	—											
MI	.67**	—										
ADHD-IN	.43**	.67**	—									
CCPT-O	-.08	-.13	-.10	—								
CCPT-C	-.01	.10	-.03	-.10	—							
CCPT-V	.01	-.02	-.08	.57**	.15	—						
CCPT-H	.04	-.05	.02	.45**	-.35**	.33**	—					
MATH	-.12	-.15	-.19	-.03	.05	-.06	-.07	—				
LANG	-.03	-.12	-.09	-.08	.07	-.07	-.08	.82**	—			
TOF-IN	.06	-.01	.04	.31**	.01	.23*	.23*	-.08	-.02	—		
TOF-H	.17	.11	.23*	.16	.09	.20	.08	-.15	-.08	.68**	—	
BASC-DEP	.63**	.37**	.26**	.04	.04	.03	.05	.04	.08	.06	.17	—

Note. BRI = Behavior Rating Inventory of Executive Function (BRIEF) Behavior Regulation Index, MI = BRIEF Metacognitive Index, ADHD-IN = ADHD-IV Rating Scale Inattention subscale, CCPT-O = Conners Continuous Performance Task (CCPT) Omissions subscale, CCPT-C = CCPT Commissions subscale, CCPT-V = CCPT Variability subscale, CCPT-H = CCPT Hit Reaction Time subscale, Math = mean teacher rating of work completion/work accuracy in math, Lang = mean teacher rating of work completion/work accuracy in language arts, TOF-IN = Test Observation Form Inattentive Subscale, TOF-H = Test Observation Form Hyperactive/Impulsive Subscale, BASC-DEP = Behavior Assessment System for Children–Second Edition (BASC-2) Depression subscale.

Table 4

Pearson Correlation Coefficients Between BRIEF Scales and Other Variables in Correlational Analyses (Teacher)

Scales	BRI	MI	ADHD-IN	CCPT-O	CCPT-C	CCPT-V	CCPT-H	MATH	LANG	TOF-IN	TOF-H	BASC-DEP
BRI	—											
MI	.57**	—										
ADHD-IN	.41**	.64**	—									
CCPT-O	.12	.01	.11	—								
CCPT-C	-.13	-.20	-.07	-.01	—							
CCPT-V	.06	-.03	.14	.57**	.15	—						
CCPT-H	.15	.10	.16	.45**	-.35	.33**	—					
MATH	-.11	-.40**	-.47**	-.03	.05	-.06	-.07	—				
LANG	-.09	-.38**	-.52**	-.08	.07	-.07	-.08	.82**	—			
TOF-IN	.14	.13	.05	.31**	.01	.23*	.23*	-.08	-.02	—		
TOF-H	.10	.12	.16	.16	.09	.20	.08	-.15	-.08	.68**	—	
BASC-DEP	.61**	.36**	.34**	.24*	-.05	.10	.17	-.19	.20*	.14	.04	—

Note. BRI = Behavior Rating Inventory of Executive Function (BRIEF) Behavior Regulation Index, MI = BRIEF Metacognitive Index, ADHD-IN = ADHD-IV Rating Scale Inattention subscale, CCPT-O = Conners Continuous Performance Task (CCPT) Omissions subscale, CCPT-C = CCPT Commissions subscale, CCPT-V = CCPT Variability subscale, CCPT-H = CCPT Hit Reaction Time subscale, Math = mean teacher rating of work completion/work accuracy in math, Lang = mean teacher rating of work completion/work accuracy in language arts, TOF-IN = Test Observation Form Inattentive Subscale, TOF-H = Test Observation Form Hyperactive/Impulsive Subscale, BASC-DEP = Behavior Assessment System for Children–Second Edition (BASC-2) Depression subscale.

Table 5

Pearson Correlation Coefficients Between Parent and Teacher Ratings

Scales	BRI-P	MI-P	BRI-T	MI-T	ADHD-IN-P	ADHD-H-P	ADHD-IN-T	ADHD-H-T	DEP-P	DEP-T
BRI-P	—									
MI-P	.67**	—								
BRI-T	.36**	.26**	—							
MI-T	.21*	.26**	.57**	—						
ADHD-IN-P	.43**	.67**	.11	.22*	—					
ADHD-H-P	.47**	.41**	.28**	.22*	.39**	—				
ADHD-IN-T	.02	.15	.41**	.64**	.23*	.10	—			
ADHD-H-T	.19	.05	.58**	.38**	.09	.42	.35**	—		
DEP-P	.63**	.37**	.06	.06	.26**	.26	-.20*	-.04	—	
DEP-T	.18	.22*	.61**	.36**	.15	.12	.34**	.29**	-.07	—

Note. BRI = Behavior Rating Inventory of Executive Function (BRIEF) Behavior Regulation Index, MI = BRIEF Metacognitive Index, ADHD-IN = ADHD-IV Rating Scale Inattention subscale, ADHD-H = ADHD-IV-RS Hyperactive/Impulsive subscale DEP = Behavior Assessment System for Children–Second Edition (BASC-2) Depression subscale, P = Parent, T = Teacher.

Table 6

Factor Analysis Including Teacher Ratings

Assessment Measure	Type	Factor 1 Rating Scale	Factor 2 Long-Term	Factor 3 TOF/ CCPT-O	Factor 4 FSIQ/ CCPT-C
BRIEF BRI	Rating scale	**.89**	-.01	.05	-.17
ADHD-IV-RS Hyperactive	Rating scale	**.77**	.04	.15	.23
BRIEF MI	Rating scale	**.68**	-.46	.01	-.05
BASC-2 Depression	Rating scale	**.61**	-.11	.10	-.50
Language Arts	Long-term outcome	-.04	**.93**	-.01	.10
Math	Long-term outcome	-.04	**.92**	-.07	.02
ADHD-IV-RS Inattention	Rating scale	.56	**-.60**	.03	.03
TOF Inattention	Observation	.06	-.02	**.89**	-.04
TOF Hyperactivity	Observation	.11	-.11	**.86**	.22
CCPT Omissions	Performance based	.07	.02	**.47**	-.40
WASI-II FSIQ	Performance based	.18	.04	-.09	**.80**
CCPT Commissions	Performance based	-.14	.04	.15	**.43**

Note. Factor analysis was performed according to the principal components rotated varimax procedure. BRIEF BRI = Behavior Rating Inventory of Executive Function (BRIEF) Behavior Regulation Index, BRIEF MI = BRIEF Metacognitive Index, ADHD-IV-RS Inattention = ADHD-IV Rating Scale Inattention subscale, ADHD-IV-RS Hyperactive = ADHD-IV Rating Scale Hyperactive/Impulsive subscale, CCPT-Omissions = Conners Continuous Performance Task (CCPT) Omissions subscale, CCPT-C = CCPT Commissions subscale, Math = mean teacher rating of work completion/work accuracy in math, Lang = mean teacher rating of work completion/work accuracy in language arts, TOF-Inattention = Test Observation Form Inattentive Subscale, TOF-Hyperactivity = TOF Hyperactive/Impulsive Subscale, BASC-2 Depression = Behavior Assessment System for Children–Second Edition (BASC-2) Depression subscale; WASI-II FSIQ = Weschler Abbreviated Scale of Intelligence –Second Edition (WASI-II) Full Scale Intelligence Quotient.

Table 7

Factor Analysis Including Parent Ratings

Assessment Measure	Type	Factor 1 Rating Scale	Factor 2 Long-Term	Factor 3 TOF/ CCPT-O	Factor 4 FSIQ/ CCPT-C
BRIEF BRI	Rating scale	**.87**	.00	.01	-.06
BRIEF MI	Rating scale	**.84**	-.14	-.11	.09
ADHD-IV-RS Inattention	Rating scale	**.72**	-.17	.00	.03
BASC-2 Depression	Rating scale	**.67**	.13	.09	-.12
ADHD-IV-RS Hyperactive	Rating scale	**.62**	.03	.18	.30
Language Arts	Long-term outcome	.00	**.95**	-.02	.07
Math	Long-term outcome	-.10	**.94**	-.06	.02
TOF Inattention	Observation	.04	-.01	**.89**	.02
TOF Hyperactivity	Observation	.24	-.09	**.82**	.19
CCPT Omissions	Performance based	-.13	-.00	**.53**	-.25
WASI-II FSIQ	Performance based	.06	.04	-.10	**.77**
CCPT Commissions	Performance based	-.02	.03	.05	**.65**

Note. Factor analysis was performed according to the principal components rotated varimax procedure. BRIEF BRI = Behavior Rating Inventory of Executive Function (BRIEF) Behavior Regulation Index, BRIEF MI = BRIEF Metacognitive Index, ADHD-IV-RS Inattention = ADHD-IV Rating Scale Inattention subscale, ADHD-IV-RS Hyperactive = ADHD-IV Rating Scale Hyperactive/Impulsive subscale, CCPT-Omissions = Conners Continuous Performance Task (CCPT) Omissions subscale, CCPT-C = CCPT Commissions subscale, Math = mean teacher rating of work completion/work accuracy in math, Lang = mean teacher rating of work completion/work accuracy in language arts, TOF-Inattention = Test Observation Form Inattentive Subscale, TOF-Hyperactivity = TOF Hyperactive/Impulsive Subscale, BASC-2 Depression = Behavior Assessment System for Children–Second Edition (BASC-2) Depression subscale; WASI-II FSIQ = Weschler Abbreviated Scale of Intelligence –Second Edition (WASI-II) Full Scale Intelligence Quotient.