ADHD AND EXECUTIVE FUNCTIONING: A META-ANALYTIC STUDY ON SHIFTING AND COGNITIVE FLEXIBILITY

A Dissertation

by

RYAN HEATH HINOJOSA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Cynthia Riccio
Committee Members,	Steven Woltering
	Robert W. Heffer
	Lisa Bowman-Perrott
Head of Department,	Shanna Hagan-Burke

August 2019

Major Subject: School Psychology

ABSTRACT

Individuals with ADHD exhibit deficits in attention and hyperactivity/impulsivity. ADHD also is referred to as a disorder of executive function (EF). Theories of EF suggest behavioral inhibition and working memory are often areas of impairment for people with ADHD; however, less is known with regards to set shifting. The present study aimed to address this question. Using meta-analysis, this study examined studies across three standard assessments of EF and specifically, set shifting. Studies with child, adolescent, and adult participants were investigated, and factors such as age, intelligence (IQ), study quality, and test version (card sorting) were explored as potential moderators of the relationship between ADHD and impairments in set shifting.

Results indicated both performance measures and a rating scale of set shifting evidenced small to large effects that were statistically significant, suggesting these measures are sensitive to shifting impairments in people with ADHD. Meta-regression analysis indicated factors such as age and IQ, as well as test version in one instance, emerged as significant moderators of score variability, with additional variability in effects that remained unexplained. This suggests that, although these variables impact the degree of impairment as reflected by these measures, additional factors not accounted for may help to explain reasons for impairment. In summary, the relationship between poor shifting and ADHD is moderated by different factors. Within the current literature state, there is a paucity of studies targeting adolescents and adults with ADHD.

In order to develop measures of shifting for use with these age groups, future studies may want to continue investigating deficits related to shifting.

DEDICATION

To all the people in the world who struggle with ADHD. You are not alone and there is help.

ACKNOWLEDGEMENTS

I would like to thank my committee chair and mentor, Dr. Riccio, and my committee members, Dr. Bowman-Perrott, Dr. Heffer, and Dr. Woltering, for their guidance and support throughout the course of this research. Dr. Riccio, it has been a real pleasure working with you over the years and I will be forever grateful to you for helping me to realize my aspiration of becoming a neuropsychologist. I can only hope to approach the contributions you have made to our field. I would also like to thank Dr. Victor Willson for his professorship and guidance, as well as Margaret Foster at the Center for Health Systems and Design for aiding me in several key aspects of this study.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my family and friends for their encouragement. And to Chris, thank you for everything.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Cynthia Riccio [advisor], Professors Steven Woltering and Lisa Bowman-Perrott of the Department of Educational Psychology, and Professor Robert Heffer of the Department of Psychology and Brain Sciences.

Professor Margaret Foster of the Center for Health Systems and Design assisted in the design of the search methods. Professor Victor Willson provided consultation regarding statistical analyses for this study. Professor Thomas Meyer at the University of Texas Health Science Center Houston provided consultation regarding statistical analyses and presentation of the data. Jade Kestian, School Psychology doctoral student, participated in the data collection sequence.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

The 18th century physician, Sir Alexander Crichton, was the first to write about symptoms resembling the current disorder known as Attention-Deficit/Hyperactivity Disorder (ADHD; Lange, Reichl, Lange, Tucha, & Tucha, 2010). Crichton defined *inattention* as “the incapacity of attending with a necessary degree of constancy to any one object” (Crichton, 1798, p. 203). Crichton went on to indicate that inattention may be inherent or resulting from disease. The current diagnostic classification for ADHD (e.g., American Psychiatric Association [APA], 2013) reflects many of Crichton’s initial ideas, including the impact of attention problems as it relates to education.

The impact of ADHD can be problematic, often encompassing a host of negative and life-long outcomes for those afflicted (Biederman et al., 2004; Eme, 2016; Kawabata, Tseng, & Gau, 2012; Reid & Johnson, 2012). For many, a diagnosis of ADHD signifies impairments in executive function (EF; Biederman et al., 2007; Gau & Shang, 2010; Krieger & Amador-Campos, 2018; Lijffijt, Kenemans, Verbaten, & van Engeland, 2005; Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005). Miyake, Friedman, Emerson, Witzki, and Howerter (2000) refer to EF as higher-order logic that orients behavior toward goal setting and attainment. Their definition of EF comprises three distinct, yet connected abilities: “inhibition,” “updating,” and “shifting.” Within this model, updating refers to the monitoring, encoding, and active manipulation of information stored in working memory (Lehto, 1996, Miyake et al., 2000; Morris &

Jones, 1990). To be consistent with most of the research (e.g., Baddeley, 1986; Barkley, 1997, 2006), the term working memory (WM) will be used as synonymous with the construct of "updating."

Of these factors, shifting is the most salient to this study. Shifting between tasks or mental sets is termed in several ways including "attention switching," "task switching," and "cognitive flexibility" (Miyake et al., 2000, p. 55). Shifting ability is presumed to impact cognitive control (Monsell, 1996; Pennington & Ozonoff, 1996). Problems with inflexible thinking manifest through one's resistance to change, transitions, and fixation on a preferred activity or interest (Leung & Zakzanis, 2014). Research indicates that shifting deficits are presumed in ADHD and autism (Boonstra, Oosterlaan, Sergeant, & Buitelaar, 2005; Happé, Booth, Charlton, & Hughes, 2006; Hill, 2004; Lai et al., 2017; Leung & Zakzanis, 2014; Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005).

Current explorations of executive dysfunction in children with ADHD span across a myriad of topics related to individual EF constructs like inhibition and WM (e.g., Amorin & Marques, 2018; Demurie, Roeyers, Wiesema, & Sonuga-Barke, 2013; Fried, Hirshfield-Becker, Petty, Batchelder, & Biederman, 2015; Roberts, Martel, & Nigg, 2017); findings suggest these constructs are potential phenotypes in ADHD. A phenotype "is the observable expression of an individual's genotype" (Wojczynski & Tiwari, 2008, p. 75). Research on phenotypes of ADHD aims to identify common neurocognitive and behavioral profiles within subjects; nonetheless, it remains a heterogeneous disorder (Adams, Derefinko, Milich, & Fillmore, 2008; Castellanos, Sonuga-Barke, Milham, & Tannock, 2006; Kofler, Sarver, Spiegel, Day, Harmon, & Wells, 2017; Shang & Gau, 2011).

Efforts to adequately describe the level of heterogeneity in ADHD have resulted in three subtypes characterizing their overall symptom profile (e.g., predominantly inattentive, predominantly hyperactive, combined type; APA, 2013). With the transition from the *Diagnostic and Statistical Manual of Mental Disorders, Text Revision, Fourth Edition* (DSM-IV TR; APA, 2000) to the DSM-5 (APA, 2013) in the last decade, changes have occurred regarding some of the diagnostic criteria (e.g., age of onset) and increased consideration of ADHD in adults (Agosti, Chen, & Levin, 2011; Giacobini, Medin, Ahnemark, Russo, & Carlqvist, 2018; Klein et al., 2012; Turgay et al., 2012). Additionally, there is some evidence that ADHD symptoms manifest differently in males than they do in females (Gaub & Carlson, 1997; Hinshaw, Owens, Sami, & Fargeon, 2006). Together, these factors contribute to heterogeneity in ADHD.

Beyond the diagnostic criteria, there is considerable research to suggest associated, but nonspecific, EF deficits in individuals with ADHD that further adds to the heterogeneity (e.g., Egeland, 2010; Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005; Riccio, Homack, Jarratt, & Wolfe, 2006). The meta-analytic studies by Pennington and Ozonoff (1996) and Willcutt et al. (2005) examined the relationship between ADHD and EF. Both meta-analyses are important as they attempted to establish consistency across the literature. Specific to children and adolescents, these meta-analyses yielded generally congruent findings that ADHD predicts global EF deficits with evidence of more pronounced dysinhibition compared to typically developing children. The same holds true for the single meta-analysis focusing on adult studies (e.g., Boonstra et al., 2005). While these studies make substantial contributions to the knowledge base, the results have not been fractionated to more succinct theories of EF such as posited by Miyake et al. (2000).

Notably, at the individual study level, however, results are mixed, which may in part be related to issues in neuropsychological assessment like "task impurity" (Miyake & Friedman, 2012, p. 8) and generalized assumptions that disregard developmental changes in brain development (e.g., Pennington & Ozonoff, 1996; Willcutt et al., 2005). In addition, research suggests that performance-based measures and rating scales tap different aspects of EF (Barkley & Fischer, 2011; Ten Eycke & Dewey, 2015; Toplak, West, & Stanovich, 2013); however, studies have yet to specifically identify how patterns of impaired function differ based on the assessment method (i.e., performance-based measures versus rating scales).

Statement of the Problem

While existing research efforts have advanced the general understanding of EF in ADHD, little is known about performance differences on EF measures between individuals with ADHD and healthy controls (HC) for all three of the factors included in the model by Miyake et al. (2000). Additional meta-analyses suggest that individuals with ADHD yield significant deficits on measures of inhibition (Alderson, Rapport, & Kofler, 2007; Lijffijt et al., 2005) and WM (Kasper, Alderson, & Hudec, 2012; Martinussen et al., 2005) compared to HC; however, there is little understanding for deficits associated with performance measures of shifting. In tandem, research has been inundated with questions surrounding the lack of correlation between scores on performance-based measures and scores on rating scales of EF. Better understanding of the way in which deficits in shifting, consistent with Miyake et al.'s model as the theoretical framework, and how this factor is measured across the literature is important.

Significance of the Problem

As previously noted, understanding the development of EF in childhood and adolescence is an important consideration that has the capacity to contribute meaningfully to neuropsychological research and further benefit the fields of psychology, psychiatry, and education. Similar to child studies comparing the neuropsychological performance of ADHD groups against HC, prior research suggests adolescents with ADHD perform significantly lower on performance-based measures (PBM) than HCs (e.g., Martel, Nikolas, & Nigg, 2007; Skogli, Egeland, Andersen, Hovik, & Øie, 2014; Toplak, Bucciarelli, Jain, & Tannock, 2009). This may be due in part to the assumption that for many, adolescence is especially challenging due to the number of ongoing changes in development. For individuals with ADHD, these challenges may be more intensified given the struggles they face both academically and interpersonally (Poon & Ho, 2014). As these individuals transition from late adolescence into young adulthood, their symptomatology either remits or persists. In cases where symptoms are more chronic, problems begin to manifest in their occupational and relational endeavors (Reid & Johnson, 2012).

Because EF affects multiple aspects of daily life activity, its association with ADHD may help to explain these and other related challenges. EF can be predictive of educational and social/ emotional outcomes (e.g., Gioia, Isquith, Kenworthy, & Barton, 2002; Huang-Pollock, Mikami, Pfiffner, & McBurnett, 2009; Tseng & Gau, 2013). Specifically, these negative outcomes include school failure, substance abuse, and antisocial behavior (Moffitt et al., 2011; Molina & Pelham, 2014; Zulauf, Sprich, Safren, & Wilens, 2014). This is especially true for individuals with ADHD. Biederman et al. (2006) suggested that EF deficits in children with ADHD increase the likelihood that they will be retained, require tutoring services, and possibly be served by Special Education. The current study will review the literature describing these and other negative outcomes associated with ADHD, many of which have been linked to executive impairments with emphasis on shifting.

Purpose/Scope of the Current Study

The last meta-analysis investigating EF profiles in ADHD populations focused on adults (e.g., Boonstra et al., 2005). In the same year, Willcutt et al. (2005) replicated the findings of Pennington and Ozonoff (1996) in their synthesis of child and adolescent studies. Additional meta-analyses since have examined ADHD in relation to specific aspects of EF including inhibition (Alderson et al., 2007; Lijffijt et al., 2005) and WM (Kasper et al., 2012; Martinussen et al., 2005); however, there is no meta-analysis specific to shifting in relation to ADHD. A second issue regards the discrepancy between PBM and rating scales of EF. Toplak et al. (2013) indicated that PBM and rating scales assess different aspects of EF. Using meta-analytic techniques, the present study aims to examine and summarize findings specific to shifting deficits in ADHD since Pennington and Ozonoff's landmark study. Meta-analyses will compare the effects of group performances on PBM of shifting and examine the differences in effect sizes between PBM and rating scales of shifting.

Research Questions/Hypotheses

With consideration of the three-factor structure of EF (Miyake et al., 2000), the first aim of this study is to add to our knowledge of the characteristics of EF, specifically shifting, in children and adults with ADHD. The second aim is to identify whether or not assumed effects, or phenotypes of ADHD specific to shifting significantly differ when EF is assessed with rating scales versus PBM. To answer these questions, the meta-analyses will investigate findings across child and adult studies specific to ADHD populations. The meta-analyses by Pennington and Ozonoff (1996) and Lijffijt et al. (2005) included studies with adult subjects. Given the dearth of research targeting shifting ability in ADHD, and evidence that adults continue to exhibit problems with inflexible thinking (Boonstra et al., 2005), meta-analysis across the lifespan is warranted. For this project, studies must contain a measure of shifting and a HC comparison group.

Research Question 1

Compared to healthy controls, do groups of participants identified as having ADHD yield lower outcomes on measures of shifting? It is hypothesized that shifting abilities are consistently deficient among individuals with ADHD (Pennington & Ozonoff 1996; Willcutt et al., 2005) regardless of measure.

Research Question 2

Does the pattern of impaired function, as evidenced by mean effect size for shifting, differ based on the method of EF assessment (i.e., PBMs versus rating scales)?

Because PBMs tap a narrower aspect of shifting (e.g., perseverative errors) while rating scales reflect a broader application of shifting such as problem solving and goal-directed behaviors. This hypothesis is formulated based on the assumption that rating scales and PBMs assess different constructs (Ten Eycke & Dewey, 2016; Toplak et al., 2013).

Implications for Practice

The main intent of this study is to update previous meta-analytic findings as they relate to ADHD and EF. The current state of the literature yields inconsistent results that name specific phenotypes associated with ADHD (e.g., Castellanos & Tannock, 2002; Pennington & Ozonoff, 1996; Thissen et al., 2013; Willcutt et al., 2005). An important extension of this study is to explain the association between ADHD and neuropsychological impairment within the context of an evidence-based factor structure of EF (Miyake et al., 2000). Given the noteworthy differences between neuropsychological performance and related EF ratings (e.g., Barkley & Murphy, 2010; Krieger & Amador-Campos, 2018; McAuley, Chen, Goos, Schachar, & Crosbie, 2010; Toplak et al., 2009; Ten Eycke & Dewey, 2015; Toplak, Bucciarelli, Jain, & Tannock, 2008; Toplak et al., 2013), the findings of this study will contribute to the emerging literature targeting the discriminant validity of scores on PBMs versus rating scales specific to the assessment of ADHD. Together, these considerations will afford researchers as well as practitioners, families, and educators a better and more concise understanding of executive impairment in ADHD. Expanding our knowledge on these topics also will inform and enhance efforts toward assessment and treatment planning (Toplak et al., 2013).

CHAPTER II

LITERATURE REVIEW

Attention-Deficit/Hyperactivity Disorder (ADHD)

ADHD is a complex and multifaceted disorder (Castellanos et al., 2006; Lange et al., 2010; Skogli et al., 2014; Weyandt, Swentosky, & Gudmundsdottir, 2013). In an epidemiological study, Visser et al. (2014) indicated that the prevalence rate of ADHD among school-aged children in the United States approximated 11% as of 2011 and that from 2003 to 2011, parent-reported history of ADHD increased by 42%. Clinically, Hibbs and Jensen (2004) identified ADHD as the most common type of child psychopathology, comprising 30-40% of referral problems in primary care settings. Research on children and adolescents suggests that ADHD predicts a host of negative outcomes including poor social functioning (Nijmeijer et al., 2008), conduct problems (Banaschewski et al., 2005; Hummer et al., 201; Poon & Ho, 2014), and substance use (Barkley, 2015; Derefinko & Pelham, 2014; Harstad & Levy, 2014).

Historical Perspective and Diagnostic Considerations

The idea of a “hyperkinetic impulse disorder” was conceptualized by Laufer, Denhoof, and Solomons (1957, p. 38), who used this term to explain a pattern of overly active behaviors, coupled with poor attentional control observable in childhood. Prior studies established the premise that these symptoms resulted from brain damage (e.g., Strauss & Lehtinen, 1947 as cited in Conners, 2000), a theory that led to the designation of minimal brain dysfunction (MBD) as an explanation for these symptoms. Clements and Peters (1962) forwarded this thinking in an influential paper emphasizing that neurological differences impact the ways in which children respond to their environment. Clinicians and researchers soon recognized, however, that instances of hyperactivity co-occur in multiple psychiatric conditions, and MBD did not always occur with accompanying hyperactive or impulsive behaviors (e.g., Herbert, 1964). In addition, MBD was an assumption that lacked empirical evidence (Rie, 1980) and perpetuated a debate that would continue through the next few decades. Because of this, a new disorder was created that would account for the level of heterogeneity in MBD and lack of specificity for hyperactivity, named "Hyperkinetic Reaction of Childhood" (DSM-II; APA, 1968). The second edition of the DSM described this diagnosis as being "characterized by overactivity, restlessness, distractibility, and short attention span" (p. 50 as cited by Lange et al., 2010); additionally, the prevalence was heightened in childhood with symptoms diminishing in adolescence (Lange et al., 2010).

The third edition of the Diagnostic and Statistical Manual (DSM-III; APA, 1980) renamed this cluster of symptoms Attention Deficit Disorder: With and Without Hyperactivity (ADD). The new label represented a shift in the 1970s from the focus on hyperactivity to a greater emphasis on the presence and impact of inattention and poor impulse control. One of the reasons for this shift was attributed to Douglas (1972) who indicated that those were the symptoms that responded best to stimulant medication. The DSM-III used three separate lists of symptoms to outline the tenets of ADD (i.e., inattention, impulsivity, and hyperactivity) and devised a set of guidelines and cutoff scores for counting symptoms. An additional diagnosis of ADD was referred to as the “residual type”, which was intended for those who previously met the diagnostic criteria for ADD with Hyperactivity, but experienced a remittance in symptoms (APA, 1980, p. 44).

The publication of the revised DSM-III (DSM-III-R; APA, 1987) incorporated changes that reflected some of the concerns surrounding the diagnosis for ADD. At the time ADD was conceptualized for the DSM-III (APA, 1980), there was a substantial dearth of literature in support of the classification of ADD into subtypes (Barkley, 2006). In the DSM-III, symptoms were divided into three groups (i.e., inattention, impulsivity, and hyperactivity); however, the DSM-III-R collapsed the groups into one symptom list. Another important addition to the DSM-III-R was the criteria for severity level (i.e., mild, moderate, or severe). These changes occurred mostly because researchers and clinicians alike had difficulty in differentiating between the subtypes of ADD with and without hyperactivity, resulting in the removal of the two subtypes and combining them under a new label Attention Deficit-Hyperactivity Disorder (ADHD; APA, 1987). A new category, “Undifferentiated ADD” was added to subsume disorders marked by persistent inattention, but for those who did not meet the full criteria for ADHD (p. 95). It also encompassed what was formerly known as ADD without hyperactivity (APA, 1987).

The end of the 1980s and early 1990s witnessed a surge in ADHD research, implicating a host of factors in the disorder (Barkley, 2006). Additional investigations lead to the identification of specific behaviors and deficits outlining subtype differences (e.g., Barkley, 2015; Lahey, Schaughency, Hynd, Carlson, & Nieves, 1987; Lahey et al.,

1994) as well as deficits related to motivation and reinforcement mechanisms. The study by Lahey and colleagues was conducted on a large scale that was the catalyst for the development of the classification system for ADHD subtypes subsequently in use (APA, 2013). Researchers also discovered the potential for genetics having some influence over the presence of ADHD (Biederman, Faraone, Keenan, Knee, & Tsuang, 1992). These findings yielded substantial changes in the fourth edition of the DSM (DSM-IV; APA, 1994), including an iterative classification system that differentiated predominantly inattentive subtypes of ADHD from predominantly hyperactive-impulsive types, as well as the combined type (inattentive plus hyperactive-impulsive symptoms).

The 1990s marks a time in ADHD research when a number of findings affected current practices for diagnosis and treatment (e.g., Achenbach, Howell, Quay, & Conners, 1991; Arcia & Conners, 1998; Wolraich, Hannah, Pinnock, Baumgaertel, & Brown, 1996). One of the important findings was that the symptoms of ADHD do not remit in adolescence but in many cases, persist into adulthood (e.g., Döpfner, Frölich, & Lhmkuhl, 2013; Klein & Mannuzza, 1991; Mannuzza, Klein, Bessler, Malloy, & Hynes, 1997).

The revised version of the DSM-IV (DSM-IV-TR; APA, 2000) was intended to bridge the gap between the DSM-IV (APA, 1994) and its follow-up edition while correcting for errors by making changes to the descriptive text but not diagnostic criteria (Lange et al., 2010). As a result, the definition for ADHD remained unchanged. Interestingly, related field trials prior to the release of the DSM-IV were restricted to children and adolescents under the age 17, a likely consideration for why the guidelines remained unspecified for adults.

Diagnostic Features and Exclusions

The shift from the DSM-IV-TR (APA, 2000) to DSM-5 (APA, 2013) reflects some important changes to the diagnostic criteria for ADHD. Under the DSM-IV diagnostic criteria, the prescribed age of onset for ADHD symptoms was seven years. Under the current DSM-5 requirements, it is now age 12. Second, the previous edition of the DSM did not specify the number of observed symptoms for an adult diagnosis of ADHD; however, the DSM-5 does specify a minimum of five symptoms per category instead of six. The symptoms within the inattention and hyperactivity/ impulsivity domains remained unchanged and the 18 symptoms have not changed. In children and adolescents under 17, the minimum number of symptoms required for a diagnosis remained the same (APA, 2013).

Additionally, symptoms must be present in two or more settings, and they need to have a substantial, negative impact on social, academic, or occupational functioning. To differentiate between ADHD and alternative disorders, the symptoms are not only present during episodes of psychosis or more indicative of another mental disorder (e.g., mood disorder, anxiety disorder, dissociative disorder, personality disorder, substance intoxication or withdrawal; APA, 2013).

The hallmark features associated with ADHD include inattention and hyperactivity/ impulsivity; they manifest differently for individuals but often persist throughout development (Adler, Barkley, & Newcorn, 2011; APA, 2013, Castellanos &

Tannock, 2002; Lundervold et al., 2010; Sims & Lonigan, 2012; Willcutt et al., 2012). Symptoms of inattention encompass difficulties paying attention to details; this means that individuals are prone to make careless errors in their work. Additionally, ADHD exacerbates difficulties with following through on instructions, schoolwork, and occupational duties. This is related to problems with sustained attention, which also suggests that individuals with ADHD struggle attending to prolonged tasks like reading, lectures, or conversations. Some of this may result from an inherent aversion to tasks requiring sustained mental effort; alternatively, a high sensitivity to extraneous distractions (unrelated thoughts in adolescents and adults). Additional problems with inattention include poor organizational skills, resulting in frequent loss of important items (i.e., wallet, keys, cell phone) and necessary supplies (i.e., school materials), and forgetfulness of daily activities like keeping appointments (APA, 2013).

The DSM-5 (APA, 2013) describes hyperactivity and impulsivity as being "driven by a motor" or feeling as if one is always "on the go" (APA, 2013, p. 60). These behaviors refer to any type of motor or vocal activity considered developmentally inappropriate or excessive. For school-age children, this includes fidgeting or tapping, running or climbing at inappropriate times, and getting out of their seats to engage in off-task, more preferred activities. In adolescents and adults, running or climbing may translate to feelings of extreme restlessness (i.e., fidgeting and tapping). Increases in activity levels can mean destructive and louder approaches to play, which exacerbates problems with social activities. Poor impulse control is evident when individuals have difficulty waiting their turn or problems blurting out answers; additionally, individuals with ADHD may talk excessively and interrupt others (APA, 2013).

Subtype and Gender Differences of ADHD

As noted, there are three subtypes of ADHD: ADHD predominantly inattentive type (ADHD-I), ADHD predominantly hyperactive-impulsive type (ADHD-H), and ADHD combined type (ADHD-C; APA, 2013). Given the level of integration between inattention and hyperactivity/ impulsivity, ADHD-I and ADHD-C seem to be the most common and stable subtype diagnosis (Ramtekkar, Reiersen, Todorov, & Todd, 2010; Reid & Johnson, 2012; Willcutt et al., 2012). Interestingly, Adams et al. (2010) challenged the notion that ADHD subtypes have distinct features, indicating that even purely inattentive groups can, and often do, show hyperactive/ impulsive symptoms. Relatedly, the specificity for the etiology of ADHD subtypes is unclear (Martel et al., 2007).

In Willcutt et al.'s (2012) meta-analysis, using the DSM-IV (APA, 1994) nosology, 59% of individuals diagnosed with ADHD continued to meet the diagnostic criteria for one of the three ADHD subtypes at follow-up assessment. Additionally, only 35% continued to meet the criteria for the same subtype as previously diagnosed. Specifically, 70% initially diagnosed with ADHD-C (n = 319) continued to meet diagnostic criteria for one of the three ADHD subtypes. At follow-up assessment, only 2.7% of those individuals met criteria for ADHD-H. Fifty percent of individuals initially diagnosed with ADHD-I (n = 218) continued to meet diagnostic criteria for one of the three ADHD subtypes. At follow-up assessment, 40% of this group continued to meet the criteria for ADHD-I, and only 1.6 % showed a diagnosis of ADHD-H. Of the 33% showing an initial diagnosis for ADHD-H (*n* = 64), only 14% continued to meet the same criteria. Furthermore, these results indicate that across development, the majority experience a shift from ADHD-C to ADHD-I (Willcutt et al., 2012), indicating some remittance of hyperactive tendencies like running and climbing (APA, 2013).

In the last decade, our knowledge on ADHD has been bolstered by studies investigating gender differences (Bruchmuller, Margraf, & Schneider, 2012; Lung-Cheng Huang, Weng, & Ho, 2016; Skogli et al., 2013). For example, there is an approximate two to one male-to-female ratio of children diagnosed with ADHD (APA, 2013). Generally, the topography of ADHD is more similar than different across males and females (Biederman et al., 2007; Gershon, 2002; Rucklidge, 2008; Skogli, Teicher, Andersen, Hovik, & Merete, 2013). Rucklidge (2010) reviewed some of the differences in psychosocial functioning, cognitive functioning, and response to treatment among the sexes. Across these aspects, males and females with ADHD are generally more similar than different because they encounter similar problems and exhibit no differences in response to treatment. To date, the few studies focusing on EF in adolescents have yielded mixed results (e.g., Krieger & Amador-Campos, 2018; Siedman, Biederman, Faraone, Weber, & Ouellette, 1997; Siedman et al., 2005). Uebel et al. (2010) showed marginal gender effects, namely that the response-style for females emphasized accuracy. Using nonparametric statistics, however, they showed that accuracy was characteristic of the response style for both genders.

Despite the general assumption that males and females are more similar than different in terms of ADHD presentation, Rucklidge and Tannock (2001) found that boys tend to be more aggressive than girls and have a higher self-esteem than females, strengthening their resilience in response to negative life experiences. Additionally, speech and language disorders or delays as well as deficient cognitive and intellectual abilities seem to be more prevalent among girls versus boys. In adolescents, males show slower processing speed whereas females display more vocabulary problems (Rucklidge & Tannock, 2001).

There are also some noteworthy gender differences related to ADHD subtypes (Graetz, Sawyer, & Baghurst, 2005; Papageorgiou et al., 2008; Ramtekkar et al., 2010). For example, Ramtekkar and colleagues found that ADHD-C or ADHD-I are the more prevalent diagnoses and that males tend to be diagnosed more than females; however, gender differences for ADHD-H diagnoses were statistically significant only in children. In children, the male-to-female ratio for ADHD-H was an estimated 2:1, with lower estimates across genders in adolescents and adults. Despite the disparity in diagnoses, the literature suggests that generally, males and females experience ADHD similarly. They face the same challenges and are equally deficient in tackling some of these issues (Rucklidge, 2010).

Long Term Outcome and ADHD

ADHD is linked to several related disorders, including Oppositional Defiant Disorder (ODD; Barkley, 1997; Golubchik & Weizman, 2019; Hummer et al., 2011), Conduct Disorder (CD; Banaschewski et al., 2005; Wilens et al., 2002), depression

(Angold, Costello, & Erkanli, 1999; Oddo, Knouse, Surman, & Safren, 2018), anxiety (Pliszka, 2019), Obsessive-Compulsive Disorder (OCD; Norman et al., 2018), and Bipolar Disorder (BD; Weintraub, Axelson, Kowatch, Schneck, & Miklowitz, 2019). In a United States sample ($N = 6{,}483$), Kessler et al. (2014) found that 79.2% of adolescents with ADHD presented comorbidity with an additional disorder(s). Approximately 47.5% showed a comorbid mood disorder, 35.1% presented with an anxiety disorder, 64.1% had a disruptive behavior disorder, and 22.7% reported a substance disorder. Excluding the domain of anxiety disorders, ADHD was the presumed primary diagnosis. Additionally, ADHD co-occurs with Autism Spectrum Disorders (ASD) at a rate of 30-50% (Davis & Kollins, 2012; Happé et al., 2006; Thomas, Sciberras, Lycett, Papadopoulos, & Rinehart; 2018), learning disabilities (LD) in 31-45% in children (DuPaul, Gormley, & Laracy, 2014), and language impairments (LI) at a rate of 3-5% (Mueller & Tomblin, 2012).

Children with ADHD also have been shown to experience a host of negative outcomes including peer rejection and negative teacher perceptions (DuPaul et al., 2014; Reid & Johnson, 2012). Academically, ADHD has been linked to lower grades, reading problems, and low test scores, leading to a heightened risk for dropping-out of high school and foregoing post-secondary education (DuPaul & Weyandt, 2009; Loe & Feldman, 2007). Adults with ADHD often exhibit difficulties with social functioning, psychological distress, and driving, as well as higher rates of drug dependence (Babinski et al., 2011; Lange et al., 2010; Murphy, Barkley, & Bush, 2002).

Many studies indicate that ADHD is a disorder overly represented among prison inmates (Edvinsson, Bingefors, Lindstrom, & Lewander, 2010; Maniadaki, Kakouros, & Karaba, 2010). In general, the prevalence rate of childhood ADHD among inmates ranges from 24% - 67%; adult ADHD among inmates ranges from 23% - 45% (Ghanizadeh, Mohammadi, Akhondzadeh, & Sanaei-Zadeh, 2011). As such, the converging potential for negative outcomes and the prevalence rate for ADHD substantiate a costly problem for society. For instance, the estimated annual cost for ADHD treatment in the United States is more than 50 billion dollars, signifying a major concern in public health (Pelham, Foster, & Robb, 2007). The negative outcomes give rise to questions about underlying cognitive deficits that may be associated with ADHD.

Executive Function

Across the myriad of theories attempting to explain the problems associated with ADHD, many point to differences in executive function (EF) compared to that of typically developing individuals (APA, 2013; Barkley, 1997; Biederman et al., 2007; Gau, Chiu, Shang, Cheng, & Soong, 2009; Gau & Shang, 2010). Executive function (EF) refers to one's capacity to independently execute behaviors that are purposeful, goal-directed, and self-serving (Lezak, Howieson, Bigler, & Tranel, 2012). It "modulates the operation of various cognitive sub processes and thereby regulates the dynamic of human cognition" (Miyake et al., 2000, p. 50). Problems with EF are typically linked to abnormalities in brain connectivity (Sharp, McQuillin, & Gurling, 2009; Weyandt, 2006) or damage in the frontal lobes, particularly the prefrontal areas (Bledsoe, Semrud-Clikeman, & Pliszka, 2013; Shaw et al., 2009; Tamnes et al., 2010).

Definitions and Theoretical Conceptualizations of EF

The theories described here and those listed in Table 1 are not an exhaustive account for the ways in which EF can be considered. Most theories, however, share the notion that EF refers to higher-order cognition that aids in the preparation and execution of complex goal-directed behaviors, disabling typical and automatic responses (Alvarez & Emory, 2006; Tseng & Gau, 2013; Zelazo & Cunningham, 2007). Both synonymous names for EF, executive control and cognitive control are hypothetical mental processes responsible for controlling thoughts and actions to achieve future goals (Jurado & Rosselli, 2007). Linked to the frontal – striatal – cerebellar region of the brain, EF is responsible for regulating thought processes, emotionality, instincts, and behavior (Posner & Rothbart, 2009). The regulation of these processes strengthens planning and organization efforts, as well as goal-setting capabilities and the persistence within oneself to attain those goals.

EF is a higher order set of abilities, different from global intelligence; it develops gradually, not linearly (Blakemore & Choudhury, 2006). Evidenced across some of the EF theories, current views in neuropsychology link frontal lobes to our innate abilities, or fluid intelligence (Gf). This assumes that EF has less relevance to acquired knowledge, referred to as crystalized intelligence (Gc; Duncan, 1995). These ideas developed out of the works of Goltz (1888) and Munk (1890), who were the first to argue that frontal lobe functioning was distinct from general intelligence.

Table 1 Theories of EF

Model	Domains of EF	Populations considered	Conceptualization (unitary or multi-component)	Empirical support
Gray (1982)	Fight/ flight response; behavioral activation system; behavioral inhibition system	Children with ADHD, CD, anxiety disorders, HC	Multi-component	None
Stuss & Benson (1986)	Drive; sequencing; executive control	Adults with TBI	Multi-component	Busch et al. (2005)
Barkley (1997, 2006, 2015)	Behavioral inhibition: WM; self-regulation of affect/ arousal/ speech; internalization of speech; reconstitution	Children with ADHD	Multi-component	Barkley (2006, 2015); Barkley & Murphy (2011)
Anderson, Anderson, Northam, Jacobs, & Catroppa (2001)	Attentional control; set shifting; goal setting	Healthy children and adolescents	Multi-component	None
Miyake et al. (2000)	Inhibition; WM; set shifting	Healthy college students	Multi-component	Fisk & Sharp (2004); Friedman & Miyake (2004); Friedman, Miyake, Corley, Young, DeFries, & Hewitt (2006); Huizinga, Dolan, van der & Molen (2006); Lee, Bull, & Ho (2013); Lehto, Juujarvi, Kooistra, & Pulkkinen (2003); Monette, Bigras, & Lafrenière (2015)

Table 1 Continued

Model	Domains of EF	Populations considered	Conceptualization (unitary or multi-component)	Empirical support
Stuss & Alexander (2000)	Memory; attention; verbal fluency; self-awareness; humor	Children and adults with frontal lobe lesions	Multi-component	None
Gioia et al. (2002)	Inhibit; shift; emotional control; initiate; WM; plan/ organize; monitor	Children and adolescents with reading disorders, ADHD, TBI, and ASD; normal controls	Multi-component	Donders, DenBraber, & Vos (2010); Egeland & Fallmyr (2010); Gioia, Isquith, Guy, & Kenworthy (2000); Gioia, Isquith, Retzlaff, & Espy (2002); Gioia & Isquith (2002); Slick, Lautzenhiser, Sherman, & Eyrl (2006)

Note. EF = Executive Function; ADHD = Attention Deficit/ Hyperactivity Disorder; ASD = Autism Spectrum Disorder; CD = Conduct Disorder; HC = Healthy Controls; TBI = Traumatic Brain Injury; WM = Working Memory.

Basing their claims on lesion experiments in animals, they influenced Hebb (1939, 1945) who presented several studies bolstering the evidence behind this notion of frontal lobe function. These studies involved patients who maintained high IQs post-surgery following the removal of their pre-frontal cortices (e.g., average to superior IQ ranges). While Hebb furthered the work of both Goltz and Munk, he found that frontal lobe functioning was not separate from intelligence or moral function. Rather, he highlighted the possibility that lesions result in problems with long term planning, initiative, creativity, and flexible thinking (Hebb, 1945).

The associations proposed by Hebb (1945) led to increasing research on the effects of frontal lesions. Studies began focusing on the evaluation of specific constructs like initiative and planning, incorporating tests that elicited some form of goal-directed behavior. These measures revealed that frontal lesions were linked to problems with perseveration, persistence, interfering stimuli, and initiative, distinct from issues related to perception, memory, or language comprehension (e.g., Fuster, 1989; Kolb & Wishaw, 1990; Shallice, 1988; Stuss & Benson, 1986). Furthermore, studies began to describe frontal lobe functioning as the "executive" or "supervisory" aspect of task performance (Pennington & Ozonoff, 1996, p. 52).

Specific to ADHD, Durston et al. (2003) examined differences in levels of hypoactivation/ hyperactivation in various brain regions between individuals with and without ADHD. Hypoactivation describes reductions in blood flow, whereas hyperactivation is the increase in blood flow (Weyandt et al., 2013). Their findings showed an association between ADHD and decreased activation in the frontal regions and fronto-striatal networks, as well as significantly reduced hypoactivation in the basal ganglia, ventral prefrontal cortex, and the anterior cingulated gyrus. Conversely, they found hyperactivation within the ADHD group specific to posterior regions of the posterior parietal lobe, posterior cingulate, and some regions of the dorsolateral prefrontal cortex.

Of interest are findings related to the basal ganglia and prefrontal cortex. The basal ganglia are implicated in frontal lobe functioning in that they receive input from most of the cortical regions and primarily, the prefrontal cortex. Once received, these circuits are segregated and then projected back to their region (Wasserman & Wasserman, 2012). Next, the basal ganglia select the cortical regions warranting activation based on feedback from the frontal cortices, a process involving several EFs (e.g., initiating, sustaining, switching, and inhibition; Koziol & Budding, 2009). These results suggest that individuals with ADHD may experience some form of disruption within these circuits, namely the basal ganglia. In other words, communication between the basal ganglia and cortical regions may be diffused in ADHD, further explaining the association between poor motor and inhibitory control (van Rooij et al., 2015; Zandbelt et al., 2013).

Many theories describe the interaction between frontal lobe functioning and cognitive control, while specifying EF constructs as primary and underlying mechanisms. For example, Barkley (1997, 2006) conceptualized ADHD as a deficit in self-regulation. Barkley hypothesized that EF is comprised of four distinct processes: WM, self-regulation of affect/ motivation/ arousal, internalization of speech, and reconstitution, all of which are influenced by behavioral inhibition. Similarly, Quay (1997) theorized ADHD as a primary deficit in behavioral inhibition. Pulling from Gray (1982) and his work on the behavioral inhibition system (BIS), Quay extended the theory to describe common deficits in children with ADHD. The BIS hypothesis of ADHD refers to three distinct, but interrelated systems: fight/ flight response, behavioral activation system (reward), and the BIS.

The extant research on the role of behavioral inhibition in ADHD reveals concerns regarding the assumptions of Barkley (1997) and Quay (1997). For example, how stable are inhibition deficits for individuals with ADHD? Studies show that adults with ADHD do not differ significantly from HC in their reaction times (stop tasks; Boonstra et al., 2005; Lijffijt et al., 2005). In addition, there is competing evidence that proposes WM as a possible phenotype for ADHD (e.g., Bolfer, Pacheco, Tsunemi, Carreira, Casella, & Caasella, 2016; Diamond, 2005; Rapport, Chung, Shore, & Isaacs, 2001; Skogli et al., 2014), supporting the notion that ADHD may be characterized by multiple executive impairments. Furthermore, the influence of WM extends to several EF theories (Gioia et al., 2002; Stuss & Alexander, 2000; Stuss & Benson, 1986).

Across the myriad of theories attempting to explain the structure of EF (e.g., Stuss & Alexander, 2000; Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001; Gioia et al., 2002), only a few have employed factor analysis to define their constructs. For example, Busch, McBride, Curtiss, and Vanderploeg (2005) used principal component factor analysis to test a three-factor model like the one proposed by Stuss and Benson (1986). These factors include drive, sequencing, and executive control, all of which associate with regions of the frontal lobes (Kane & Engle, 2002; Stuss, Bisschop, Alexander, Levine, & Izukawa, 2001). Specific to patients with traumatic brain injury (TBI), their findings yielded three factors that accounted for 52.7% of the variance.

Miyake's Model of EF

Within research pertinent to the structure of EFs, Miyake et al. (2000) is the most cited theoretical model (Monette et al., 2015), replicating findings in samples containing adults (Friedman et al., 2006), children (Lehto et al., 2003), and young adolescents (Lee et al., 2013). The development of the three-factor model by Miyake and colleagues was intended to provide a consensus specifically explaining the role of frontal lobe functioning. Their model focused on individual differences, providing additional evidence for a nonunitary basis of EF (e.g., Schachar, Tannock, & Logan, 1993; Welsh, Pennington, & Groisser, 1991). Prior to Miyake et al. (2000), studies comparing the neuropsychological performance of dichotomous groups (e.g., clinical versus healthy groups) oftentimes yielded within group correlations that were not statistically significant and relatively low (usually $r = .40$ or less; Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Lehto, 1996). This was (and still is) typically accomplished using EFA wherein the results reflect some indication that EF is not a unitary construct (Gioia & Isquith, 2002; Gorsuch, 1997; Skogan et al., 2016).

By examining the neuropsychological performance of healthy college students, Miyake et al. (2000) explored the extent to which EF components are interrelated (unitary) or act as separate entities, while distinguishing between contribution levels of those skills to the performance on tests used to assess EF. Miyake and colleagues designed their model to comprise functions heavily cited across the literature (e.g., Baddeley, 1996; Smith & Jonides, 1999). Their study focused on three factors to define EF: inhibition, updating (WM), and shifting. Relational interests in the functions as well as their uniqueness influenced their decision to select factors that would afford them distinct operational definitions. Moreover, these considerations increased the likelihood that the target functions would be tapped by widely used and readily available neuropsychological assessments.

Miyake et al. (2000) employed CFA to evaluate pre-existing EF models and the above-mentioned tests used to measure the three hypothesized factors (shifting, WM, and inhibition). Using CFA and fit statistics (χ^2), they determined the difference between the model's predictions and actual data pattern was not statistically significant ($\chi^2(24, N = 137) = 20.29, p > .65$). Further, when compared against one- and two-factor models (e.g., inhibition and shifting/ WM), the χ^2 difference tests showed that the three-factor was superior, yielding good fit indices. In terms of association and independence, the correlations among the three variables were moderate ($.42 \leq r \leq .65$). While they did not appear to equal one another, the model indicating complete independence from one another (i.e., separability) yielded a statistically significant $\chi^2(27, N = 137) = 47.03, p < .05$, and poor overall fit. Together, these findings suggest that the latent constructs are in fact separable, yet moderately correlated, and that there is a degree of unity and distinction throughout the model (Miyake et al., 2000, Monette et al., 2015). For instance, they used the Stroop task (Stroop, 1935), the antisaccade task (adapted from

Roberts, Hager, & Heron, 1994), and the stop-signal paradigm (Logan, 1994) to tap inhibition. Their WM tasks included the keep track task (Yntema, 1963), the letter memory task (Morris & Jones, 1990), and the tone-monitoring task (modified Mental Counters task; Larson, Merritt, & Williams, 1988). For shifting, they incorporated use of the plus-minus task (Jersild, 1927), the number-letter task (Rogers & Monsell, 1995), and the local-global task (Navon, 1977).

Review of Previous Meta-Analyses

Studies have investigated the effects of inhibition, WM, and shifting in the context of different models across both clinical and typical populations. Specific to ADHD, the three meta-analyses by Pennington and Ozonoff (1996), Willcutt et al. (2005), and Boonstra et al. (2005) are very important to this study as are the ones targeting a specific construct of EF (see Table 2). Pennington and Ozonoff systematically reviewed studies focusing on EF in the context of three different disorders (ADHD, CD, and ASD). Nine years later, Willcutt and colleagues replicated those efforts focusing on ADHD and typically developing groups; finally, Boonstra et al. targeted adult populations. These studies differ from the other meta-analyses described in this paper (e.g., Alderson et al., 2007; Lijffijt et al., 2005; Kasper et al., 2012; Martinussen et al., 2005) because they examined EF and ADHD from a broader perspective. In other words, their studies explored findings related to multiple measures and multiple constructs of EF across a wide age range.

Table 2 Prior Meta-analyses of EF in ADHD

Study	Years covered/ number of studies	Age range	Mediators and moderators (non-EF constructs)	Statistically significant mean ES for inhibition measures	Statistically significant mean ES for WM measures (auditory/spatial)	Statistically significant mean ES for shift measures
Pennington & Ozonnoff (1996)	1972-1996; 18 studies	6-20	Field dependence-independence; recall/ sequencing; verbal IQ; reading comprehension; short-term memory; attention; vigilance; processing speed; academic achievement conditional discrimination; visual-motor integration; language	0.69* (Stroop time) 0.85* (motor inhibition)	0.27* (letter fluency) 0.44* (spatial – time) 0.87* (spatial – errors)	0.45* (WCST – perseverations) 0.75* (TMT-B - time)

Table 2 Continued

Study	Years covered/ number of studies	Age range	Mediators and moderators (non-EF constructs)	Statistically significant mean ES for inhibition measures	Statistically significant mean ES for WM measures (auditory/spatial)	Statistically significant mean ES for shift measures
Boonstra et al. (2005)	1970-2003; 13 studies	Adults	Diagnostic procedures; comorbidity; gender; cognitive ability; medication status; lack of well-defined control comparisons (vague descriptions in some studies)	0.55** (CPT attentiveness) 0.64** (CPT commissions) 0.89* (Stroop color-word ratio)	0.62* (verbal fluency)	0.65** (TMT-B – time)
Martinussen et al. (2005)	1997-02/ 2003; 26 studies	4-18	Reading difficulties/ language impairments	NA	0.47** (verbal storage) 0.56** (verbal central executive) 0.85** (spatial storage) 1.06** (spatial CE)	NA

Table 2 Continued

Study	Years covered/ number of studies	Age range	Mediators and moderators (non-EF constructs)	Statistically significant mean ES for inhibition measures	Statistically significant mean ES for WM measures (auditory/spatial)	Statistically significant mean ES for shift measures
Willcutt et al., (2005)	1986-03/ 2004; 83 studies	Children and adolescents	Diagnostic considerations; ADHD subtypes (ADHD-I and ADHD-C)	0.61* (SSRT) 0.51* (CPT omission errors)	0.63* (spatial WM) 0.55* (verbal WM)	0.46* (WCST perseverative errors)
Lijffijt et al. (2005)	01/ 1998-02/ 2004; 29 studies	6.4-65	Age	0.52** (MRT – child) 0.72** (SDRT – child) 0.4** (SDRT – adult) 0.58** (SSRT – child) 0.79** (SSRT – adult) -0.56* (MRT vs. SSRT – adult)	NA	NA

Table 2 Continued

Study	Years covered/ number of studies	Age range	Mediators and moderators (non-EF constructs)	Statistically significant mean ES for inhibition measures	Statistically significant mean ES for WM measures (auditory/spatial)	Statistically significant mean ES for shift measures
Alderson et al. (2007)	1990-2004; 25 studies	7-12	Age; diagnostic evaluation; total experimental trials; go-stimulus modality	0.45** (MRT) 0.73* (SDRT)	NA	NA
Kasper et al. (2012)	1989-01/2012; 45 studies	8-16	Gender; response modality; trials per set size; CE demand	NA	0.74** (visuospatial WM) 0.69** (phonological WM)	NA

Note. EF = Executive Function; ADHD = Attention Deficit/ Hyperactivity Disorder; ADHD-I = Predominantly Inattentive Type; ADHD-C = Combined Type; CE = Central Executive; CPT = Continuous Performance Test; MRT = Mean Reaction Time; NA = Not Applicable; SSRT = Stop Signal Reaction Time; SDRT = Within-subject standard deviation of reaction time; TMT = Trail Making Test; WCST = Wisconsin Card Sorting Test; WM = Working Memory.
* $p < .05$. ** $p < .01$.

Pennington and Ozonoff (1996) analyzed 18 studies. Of the 18 studies reviewed, 15 reported a significant difference between clinical (ADHD and comorbid reading and language disorders) and HC groups on multiple measures of EF. Across the 18 studies, scores from 60 EF measures and 54 non-EF measures (e.g., receptive language) were analyzed in which they found the ADHD group performed significantly lower than HCs on 40 of the EF measures. Moreover, the ADHD groups did not yield a significantly better performance on any of the EF measures.

In terms of specificity, group differences emerged on 19 non-EF measures (35%). Overall, the ADHD group performed worse on measures of vigilance and perceptual speed (e.g., coding and digit span), which aligns with related research (e.g., Corkum & Siegel, 1993). This was not the case for their performance on tasks measuring verbal IQ, naming, phoneme awareness, receptive language, or storytelling. In sum, EF deficits appeared to be more prevalent in ADHD groups over non-EF deficits with the potential for motor inhibition being a core feature; in children, deficits may not be specific to EF but generally reflect instances of cognitive inefficiency.

The results of Willcutt et al. (2005) generally favor those of Pennington and Ozonoff (1996). Unique to Willcutt and colleagues' study was the addition of weighted mean effect sizes as a way of demonstrating the differences between the ADHD and HC groups. Across 83 studies and 13 EF tasks (*N* = 3734 with ADHD and 2969 without ADHD), weighted, the mean effect sizes reflected a medium effect (*d* = 0.43-0.69; Cohen, 1988).

External validity for the measures in Willcutt et al. (2005) was evidenced by studies (e.g., Matáro, Garcia-Sánchez, Junqué, Estévez-González, & Pujol, 1997; Owen, Morris, Sahakian, Polkey, & Robbins, 1996; Rowe, Owen, Johnsrude, & Passingham, 2001) containing neuroimaging data for event-related potentials (ERP), functional and structural magnetic resonance imaging (fMRI and sMRI), and positron emission tomography (PET). Additional studies showed that a number of tasks loaded on multiple factors (e.g., CPT, WCST, digits backward; Willcutt et al., 2005; Willcutt et al., 2001), and that lesions caused impaired performance on several of the measures (e.g., Carlin et al., 2000; Demakis, 2003; Levin, Song Ewing-Cobbs, & Roberson, 2001). Using these empirically supported measures, Willcutt and colleagues indicated that the relationship between EF and ADHD spanned across multiple groups (i.e., clinic and community-based samples).

In short, the efforts of Willcutt et al. (2005) and Pennington and Ozonoff (1996) evidence a myriad of executive impairments linked to ADHD, and they also imply a causal relationship between these constructs. While the study by Willcutt and colleagues yielded medium effect sizes for all EF measures, the effects were most consistent on measures of response inhibition, vigilance, spatial WM, and some measures of planning. Rather, the biological basis and overall etiology for ADHD encompasses a host of weaknesses wherein EF deficits represent a subset of characteristics found in most individuals with ADHD.

The meta-analysis by Boonstra et al. (2005) was limited to 13 studies ($N = 662$ with ADHD and 417 without ADHD). Across six tasks, they reported scores for both

EF and non-EF dependent variables to describe the level of specificity for EF deficits in ADHD. Overall, the results were congruent with that of Pennington and Ozonoff (1996) and Willcutt et al. (2005), specifically highlighting differences in verbal fluency and inhibition. Within the non-EF domain, Boonstra et al. (2005) showed variability in reaction times, further supporting the idea that performance variability, regardless of the task, is a phenotype in ADHD (Castellanos & Tannock, 2002). Since this is a consistent theme across child and adult studies, it is possible that ADHD infers a general cognitive slowing, and not motor slowing, as the CPT hit reaction time for both groups was comparable. In adult ADHD, the strong similarity between ESs for both EF and non-EF domains indicates that deficits are diffused across multiple aspects of cognitive functioning. Furthermore, the specificity between EF deficits and ADHD cannot always be assumed.

EF in ADHD

Results from the meta-analyses indicate that the association between EF and ADHD is generally significant, with deficits persisting into adulthood. Although ADHD has been investigated in the context of various EF models, Miyake et al. (2000) is empirically supported and has been replicated across the lifespan. As such, it may provide the best framework for conceptualizing EF in children and adults with ADHD. The previous literature specific to inhibition, WM, and shift in ADHD will be summarized. Mixed findings as well as possible within-group differences, including specific EF deficits postulated by some to associate with certain ADHD subtypes (e.g., Becker & Langberg, 2014; Diamond, 2005) will be highlighted. There are many studies that have looked at inhibition, working memory, and shift, as well as other EF. There also have been previous meta-analyses on EF in ADHD (see Table 2). What is known about EF in ADHD for these three domains is summarized here.

Inhibition

Inhibition is the ability to control one's impulses and stop behavior; moreover, the act of controlling one's behavior is a deliberate effort to combat dominant, automatic, or prepotent responses (Barkley, 2006; Miyake et al., 2000; Monette et al., 2015). It should be noted that inhibition describes different functions at varying levels of complexity (Kok, 1999). For instance, Miyake and colleagues identified inhibition as a lower level function; in other cases, it is considered higher-order logic (e.g., Grafman & Litvan, 1999). With that in mind, it is important to discern the definition proposed by Miyake et al. from others involving activation systems and connectionist networks. For example, Gray's (1982) theory of brain-behavior processes describes uninhibited responses resulting from the mechanism's failure to elicit anxiety and fear, which then manifests through the initiation or continuation of unwanted behaviors (Quay, 1997). Throughout these processes, the BIS can override the activation system (BAS), but it is not an indication that the behavioral response is deliberate or controlled. Instead, it may reflect a continuation or increase in activation levels due to possible misfiring or poor connections within the septo-hippocampal system and frontal cortex (Gray, 1991; Quay, 1997). Both concepts share the notion that inhibitory control is associated with functioning of the prefrontal cortex (Aron et al., 2011; Kiefer, Marzinzik, Weisbrod, Scherg, & Spitzer, 1998; van Rooij et al., 2015).

Dysinhibition and delayed behavioral responses are cardinal deficits linked to ADHD (e.g., Barkley, 1997; Frazier-Wood et al., 2012; McAuley, Crosbie, Charach, & Schachar, 2014). Like other EF variables, inhibition is assessed in terms of performance and ecology. PBM like the Stroop test (Stroop, 1935) and Stop Task (Schacar, Mota, Logan, & Tannock, 2000) are frequently referenced within the literature. Using these measures, Rucklidge and Tannock (2002) pioneered the exploration of inhibition specifically in adolescents with and without ADHD. Compared to healthy controls, the ADHD group was notably slower in naming colors and incongruent color/ words on the Stroop test. For the Stop Task, they observed an association between ADHD and greater problems inhibiting responses and slower reaction times (MRT). As a group, the ADHD sample also demonstrated greater variability in their responses.

Conversely, there is evidence that these facets of inhibition improve throughout development. For instance, Lijffijt et al. (2005) found that adult comparison groups (i.e., ADHD versus HCs) did not yield statistically significant differences in MRT or stop signal reaction time (SSRT). The meta-analysis conducted by Lijffijt and colleagues indicated that the amount of variability was larger in the ADHD groups and the most discernable difference between the clinical and HC groups. In adults, however, the difference in MRT yielded an effect size (ES) of 0, meaning that there were no differences between the groups. These findings are consistent with those in Willcutt et al. (2005), who found that the most obvious differences were scores for the SSRT and omission errors on continuous performance tests (CPT).

Although specific to children and adolescents, Alderson et al. (2007) explored the range of subject and task variable moderator effects on behavioral inhibition processes germane to the stop-signal paradigm. Noted moderators for reaction times included age, diagnostic evaluation, delay schedule, total experimental trials, and go-stimulus modality. Interestingly, Alderson and colleagues did not indicate a statistically significant between-group stop signal delay (SSD) metric.

In adults, the analyses by Boonstra et al. (2005) revealed poorer performance for the ADHD group on all three types of Stroop cards (i.e., read color names, naming colors, color word), yielding a significant ES for interference control. Controlling for performance on the Color card (C), however, resulted in a non-significant ES for the Color Word card (CW). Further, this changed the *f*-value to a medium effect (Cohen, 1988). In adult ADHD, this finding suggests that response inhibition may not signify problems with selective visual attention; furthermore, a better interpretation of interference scores warrants correcting for performance on card C.

Studies attempting to bolster these findings through use of ecologically sound measures yield similar results. Skogli et al. (2014) indicated that the combined type group (ADHD-C) demonstrated lower scores than the predominantly inattentive (ADHD-I) group on the Inhibition scale of the BRIEF. These findings replicated those in previous studies (e.g., Semrud-Clikeman et al., 2010).

The findings of Rucklidge and Tannock (2002), Lijffijt et al. (2005), and Alderson et al. (2007), with support from studies using ecological data (e.g., Skogli et al., 2014), suggest that differences in the SSRT are better explained by a more general deficit in attention and cognitive processing. Furthermore, the basic clinical features of ADHD (e.g., inattention) may be transient, and that changes in variability reflect the inconsistent nature of ADHD (i.e., performance can change in a given moment).

Despite slower MRT remitting in adulthood, some individuals with ADHD may continue to exhibit slower reaction times throughout their development (Alderson et al., 2007; Cubillo et al., 2010). Relatively few studies speculate explanations for a persistent, slower SSRT (e.g., Lijffijt et al., 2005); however, age seems to be a plausible moderator for ADHD and slow MRT (Alderson et al., 2007; Shen, Lee, & Chen, 2014). For example, it is possible that attention problems and poor inhibitory motor control and attention are related to WM capacity (e.g., Mecklinger, Weber, Gunter, & Engle, 2003). In other words, the stop-task paradigm may present more challenges to children in that their performance requires a greater use of WM, which may be underdeveloped in children and adolescents compared to adults (e.g., Tamm, Menon, & Reiss, 2002).

Working Memory (WM)

WM is the capacity to which one can hold information in their mind and manipulate that knowledge in some fashion (Baddeley, 1966). It involves updating, or the act of monitoring and coding knowledge based on what the task requires. Changes to items held in WM then replace irrelevant information with new knowledge (Baddeley, 1986, 2012; Monette et al., 2015). The encoding of new information occurs visually and aurally; hence, there are two main substrates of WM – spatial and verbal (Awh, Anllo-Vento, & Hillyard, 2000; Morris & Jones, 1990). The storage and maintenance of information is associated with premotor areas of the frontal cortex and parietal lobes, whereas updating is linked to the dorsolateral portion of the prefrontal cortex and caudate nucleus (Levy, Friedman, Davachi, & Goldman-Rakic, 1997; Roman-Urrestarazu et al., 2016; Smith & Jonides, 1999).

The meta-analysis by Pennington & Ozonoff (1996) showed that the ADHD groups performed significantly lower on the sequential memory and self-ordered pointing tasks. Conversely, effect sizes suggested that verbal tasks were not sensitive to ADHD symptomatology, a finding that was later confirmed by Martinussen et al. (2005). Similarly, Willcutt et al. (2005) found significant group differences on tasks measuring spatial WM as well as a significant level of heterogeneity for spatial and verbal WM. In adults, Boonstra et al. (2005) indicated that interpretation of scores on DS Forwards was more accurate following a correction for performance on DS Backwards.

The findings in Pennington and Ozonoff (1996) and Willcutt et al. (2005) support a number of related studies that specifically propose visual memory as a potential phenotype for ADHD (Rhodes, Coghill, & Matthews, 2004; Roman-Urrestarazu et al., 2016; Shang & Gau, 2011). The meta-analysis by Martinussen et al. (2005) revealed that children with ADHD performed lower on tasks measuring spatial storage and the spatial central executive components of WM relative to healthy controls. Although the ADHD group was deficient across the verbal domains of WM, performance differences from the healthy controls were modest and not statistically significant.

Martinussen et al. (2005) identified possible moderators to explain the discrepancy in performance between verbal and spatial tasks including localization of brain functioning, task difficulty levels, and comorbidity. For instance, there is some evidence suggesting that ADHD is associated with predominantly right hemispheric responses (Giedd, Blumenthal, Molloy, & Castellanos, 2001), which is where spatial processing is localized (Fletcher & Henson, 2001; Kwon, Reiss, & Menon, 2002). Second, they posited that spatial tasks are more challenging than verbal tasks, often requiring less automated processes. Finally, given high comorbidity rates for individuals with ADHD, it is only natural to assume that instances of comorbidity impact WM. Further, they identified reading and language problems as significant moderators for spatial WM deficits, which was not the case for verbal tasks (Martinussen et al., 2005).

In a later meta-analysis, Kasper et al. (2012) updated the findings of Martinussen et al. (2005), noting statistically significant and large effects for phonological (verbal; $g = 0.69$) and visuospatial WM ($g = 0.74$) tasks. Specific to verbal WM, Kasper and colleagues reported larger ES estimates relative to Martinussen et al., likely due to nuances in the organization of tasks. For instance, Martinussen and colleagues grouped tasks to reflect either storage or central executive (CE); conversely, Kasper et al. grouped tasks by modality (phonological or visuospatial). They included the CE demand variable as a moderator of between-study ES heterogeneity. Scores on both types of tasks were moderated by the underrepresentation of females in studies, greater trial numbers and recall tasks, and greater CE demands.

Despite congruent findings between Martinussen et al. (2005) and Kasper et al. (2012), evidence suggesting that spatial WM is a potential phenotype for ADHD is inconclusive and yields a large amount of variability across studies (Claesdotter, Cervin, Åkerlund, Råstam, & Lindvall, 2018; Kebir, Tabbane, Sengupta, & Joober, 2009; Kofler et al., 2017). Additionally, there are mixed results regarding the impact of WM across child and adolescent development. Tseng and Gau (2013) indicated that poor WM, independent of age, gender, and IQ, mediated the effects of ADHD and social problems. Miller and Hinshaw (2010) specifically observed a relationship between WM and peer acceptance; similarly, Kofler et al. (2011) noted an association between WM and social competence in ADHD. Chiang and Gau (2014) substantiated these findings, indicating that spatial WM deficits combined with an ADHD diagnosis negatively impacted peer functions.

These findings make sense, given the involvement of WM as a mechanism needed to encode social language and then evaluate the appropriateness of one's response. Relatedly, Barkley's (1997, 2006) concept of reconstitution interacts heavily with WM as it involves the analysis of verbal and/ or behavioral responses in the moment. Nonetheless, there are conflicting studies disassociating EF deficits with social functioning. Across clinic samples, Biederman et al. (2004) cited age as a possible explanation for the lack of association between ADHD and poor social functioning, in which they described as a delayed manifestation occurring in adulthood. For individuals with ADHD, adaptive concerns exclusive of academic functioning are less overt in younger samples and may heighten over time.

Shifting

Synonymous names exist for the shifting of mental sets, including "attention switching", "task switching," "set shifting," and "cognitive flexibility," all of which refer to the voluntary change of thoughts or actions, either self-initiated or in response to changes or cues in a given situation (Blijd-Hoogewys, Bezemer, & van Geert, 2014; Miyake et al., 2000; Monsell, 1996). The concept of shifting is not restricted to merely changing tasks, but it also involves the ability to phase out distractors (i.e., interference) and focus on transitive sets of directions, while adjusting visual attention. Posner and Raichle (1994) described how different neural circuits mediate visual attention shifting and deliberate responses to new sets of instructions, suggesting that visual attention is the product of the parietal lobes and mid-brain functioning (i.e., posterior attention network), while conscious acts of adherence are regulated by the frontal lobes, including the anterior cingulate (i.e., anterior attention network).

Research on shifting deficits in childhood ADHD yields mixed results (Amorim & Marques, 2018; Gau & Shang, 2010; Toplak, Bucciarelli, Jain, & Tannock, 2008; Tseng & Gau, 2013; Willcutt et al., 2005); however, some evidence shows that it may be more problematic in adult ADHD (e.g., Boonstra et al., 2005). Issues of inflexibility often are perceived as a core deficit of ASD, (Blijd-Hoogewys et al., 2014; Gioia et al., 2002; Hill, 2004; Lai et al., 2017). For example, measures of shifting differentiated high functioning children with Autism from the ADHD and control groups, a less obvious comparison than between the ADHD and control groups (Sergeant, Geurts, & Oosterlaan, 2002).

Meta-analytic findings by Pennington and Ozonoff (1996) and Willcutt et al. (2005) showed statistically significant effects across a wide range of shifting measures ($d = 0.45$), particularly perseverative errors on the WCST. In adults, Boonstra et al. (2005) indicated that scores on the TMT-B yielded a larger ES than the TMT-A, suggesting that deficits in adult ADHD may not be limited to just novelty, but also set shifting.

Skogli et al. (2014) reported that the ADHD groups demonstrated an overall lower performance on measures of controlled attention and shifting compared to the HC group. Similar to findings in Sergeant et al. (2002), their study did not yield differences between the subtypes. This is supported by another study, in which correlations between perseverative errors and levels of inattention and hyperactivity were statistically significant (Martel et al., 2007). Compared to research on inhibition and WM in ADHD, less is known about shifting deficits in ADHD. To date, there is no known meta-analysis to explore the effects of ADHD across measures of shifting; moreover, no study has attempted to conceptualize ADHD using the three-factor model of EF (Miyake et al., 2000).

Within-Group Differences

In terms of symptom presentation, severity level, and prognosis, ADHD is a highly variable disorder (Banaschewski, Neale, Rothenberger, & Roessner, 2007; McAuley, Crosbie, Charach, & Schachar, 2014; Rajendran, O'Neill, Marks, & Halperin, 2015; Thapar, Harrington, Ross, & McGuffin, 2000). Such heterogeneity in EF has been linked to differences in ADHD subtypes and gender; however, the literature yields mixed results (Rucklidge & Tannock, 2002; Skogli et al., 2013). The following sections highlight related literature.

ADHD Subtypes

In general, child and adolescent studies do not support the specificity of EF deficits for subtype distinction (e.g., Riccio, Homack, Jarratt, & Wolfe, 2006; Skogli et al., 2014; Willcutt et al., 2005). For example, Diamond (2005) posited that WM impairments are a hallmark feature of ADHD-I. This finding conflicts with that of Willcutt et al. (2005), who did not report noteworthy differences between ADHD-C and ADHD-I. Rucklidge and Tannock (2002) suggested that poor inhibitory control is exclusively linked to hyperactive/ impulsive symptoms because neither reaction time nor variability in reaction times specifically associated with any one subtype. Becker and Langberg (2014), however, suggested that hyperactive symptoms are specific to poor behavioral regulation (i.e., behavioral inhibition, shifting, emotional control). Hence, research on EF and the nosology of ADHD is inundated with conflicting findings.

Methodological Issues

Limitations

Pennington and Ozonoff (1996) noted several methodological issues linked to their study, which are discussed here. First, while the potential for confounding variables was not a concerning factor, including the influence of IQ, socioeconomic status, ethnicity, and/ or comorbid conditions, they referenced instances of possible ascertainment bias. Ascertainment bias describes data that has been distorted because of measuring the true frequency of a given phenomenon (Pennington & Ozonoff, 1996). In this case, only one study failed to detect EF deficits associated with ADHD, suggesting that the relationship between these variables may be an artifact of selection.

Secondly, the results from this study challenged the assumed level of accuracy that EF measures discriminate ADHD from typical development. In other words, to what extent do individuals with ADHD exhibit executive impairment more so than healthy controls? In general, measures of EF seemed to be better at classifying subtypes of ADHD and less sensitive to clinical versus healthy status. Using discriminant function analyses, however, only six of the studies reviewed found significant results linking EF deficits specifically to subjects with ADHD. Moreover, these analyses revealed that the Stroop, errors on the MFFT, and the GDS vigilance scores best discriminated ADHD from healthy controls. Since Pennington and Ozonoff (1996), studies have addressed issues related to discriminant validity (e.g., Alderson et al., 2007; Homack & Riccio, 2004; van Mourik, Oosterlaan, & Sergeant, 2005); however, additional research is needed.

Relatedly, the third issue regards evidence suggesting that ADHD is better explained by cognitive deficits outside the scope of EF. In other words, varying differences in difficulty levels ensues a differential deficit deemed a resource artifact. To dispel or confirm this possibility, Pennington and Ozonoff (1996) called for a within-task manipulation wherein two versions of the same task differed in their executive requirements and everything else remained intact. This was executed in only three studies, in which they yielded evidence to support a specific association between ADHD and EF deficits. Nonetheless, this was only done in three studies, and more research is needed to concretize the evidence for a differential EF deficit in ADHD.

The implications and future directions cited by Willcutt et al. (2005) include the need to test multiple models of EF, especially those related to ADHD (e.g., Barkley, 1997, 2006, 2015). Similarly, they noted a dearth of studies that tested specific theories of EF and furthermore, compared theories. Like Pennington and Ozonoff (1996), they emphasized the importance of examining the reliability of EF tasks specifically for use among children. At best, reliability estimates across studies were moderate but, in many cases, unknown. Once this data becomes available, studies can employ more stringent criteria for the inclusion of EF tasks, which may minimize problems related to task impurity.

In Boonstra et al. (2005), noted limitations include moderator effects for certain variables (e.g., diagnostic procedures, comorbidity, gender, and IQ) and selection bias. For example, their study selection was limited to those published; hence, all studies showed statistically significant effects. As a result, the "file drawer problem" may have inflated some of the findings (Boonstra et al., 2005, p. 1106). This is similar to the problem with ascertainment bias in Pennington and Ozonoff (1996). Relatedly, the inclusion criteria may have been too stringent, given their exclusion of studies wherein the total *N* was less than 50.

Measurement Issues

As previously noted, the three-factor structure of EF by Miyake et al. (2000) is the most cited model within neuropsychological research (Huizinga et al., 2006; Lee et al., 2013; Lehto et al., 2003; Monette et al., 2015). Nonetheless, there are alternative EF models that yield noteworthy differences related to organization and structure of factors (see Table 1). Moreover, there is some evidence these discrepancies are largely due to age differences (Monette et al., 2015). Alternative views, however, suggest that EF is relatively stable throughout development (e.g., Mischel et al., 2011; Miyake & Friedman, 2012; Moffitt et al., 2011). At any rate, these issues impact the assessment of EF. Additional investigation into these conflicting views can inform research towards developing more precise measures that account for developmental changes in EF. Further, issues related to EF assessment also involve discrepancies in the level of separability in EF constructs. The following sections discuss these issues as well as differences between PBM and informant ratings of EF.

Moderators

General Cognitive Ability (IQ)

The importance of controlling for IQ is implied often, given the tests' intent to explicitly tap neuropsychological functions (Lahey et al., 1988). This is problematic, however, because intelligence and EF in many cases interface with one another. Barkley (1997, 2006) indicated that both EF deficits, as well as other neurocognitive weaknesses, impact performance on standardized measures of cognitive ability (McAuley et al., 2010; Pennington & Ozonoff, 1996; Toplak et al., 2008; Toplak et al., 2013).

ADHD Subtypes

Willcutt et al. (2005) found significant differences between subtypes for SSRT scores (ADHD-IA $d = 0.68 +/- 0.26$; ADHD-C $d = 0.86 +/- 0.25$). Although the majority of the other EF measures yielded medium effect sizes for both subtypes (d across measures = 0.58 +/- 0.14 for ADHD-IA and 0.57 +/- 0.12 for ADHD-C), differences between ADHD-C and ADHD-IA across all EF measures were not consistent (mean *d* across measures = 0.09 +/- 0.10). Interestingly, predominantly hyperactive types were not associated with EF deficits (mean *d* = 0.14), but small sample sizes in many of the studies warrant cautionary interpretation. Nonetheless, this finding suggests that the etiology for ADHD-HI differs from that of other ADHD subtypes. Examination of the effects of additional covariates (e.g., intelligence, academic achievement, comorbidity), group differences in scores for SSRT, CPT, omission errors, planning tasks, and spatial and verbal WM tasks remained statistically significant.

Age of Subjects

The number of factor analytic studies describing Miyake et al. (2000) across the lifespan suggest that EF continues to develop well into adulthood (Blakemore & Choudhury, 2006; Lee et al., 2013; Tamnes et al., 2010); however, findings related to the organization of EF are inconclusive (e.g., Lee et al., 2013, Wiebe, Espy, & Charak, 2008; Willoughby, Wirth, & Blair, 2012). At age three, children seem to function within a unidimensional structure characterized by an aggregate of inhibitory control, WM, and attention-shifting abilities (Wiebe et al., 2011; Willoughby, Blair, Wirth, & Greenberg, 2010; Willoughby et al., 2012). In four- and five-year-olds, studies dually support a unitary (Fuhs & Day, 2011; Shing, Lindenberger, Diamond, Li, & Davidson, 2010) and two-factor model (Lee et al., 2013; Miller, Giesbrech, Müller, McInerney, & Kerns, 2012; Usai, Viterbori, Traverso, & De Franchis, 2014). The cross-sectional study by Miller and colleagues identified inhibition and WM as the two latent variables wherein shifting ability and WM loaded onto the same factor.

Similarly, Monette et al. (2015) tested one-, two-, and three-factor models of EF on a large sample of kindergarteners. Their efforts indicated that the variance across EF tests is best explained by a two-factor structure comprised of inhibition and WM-flexibility (shifting). In school-age children, two- or three- latent variables constitute the overall EF structure (i.e., inhibition/ shift and WM for two latent variables; inhibition, WM, and shift for three latent variables Lee et al., 2013), and a three-factor model for adults (i.e., inhibition, WM, and shift; Friedman et al., 2006). In adolescents and adults, however, there lacks a consensus regarding the age at which the organization of EF factors shifts from two to three latent variables. For example, Lehto et al. (2003) found that the tripartite structure by Miyake et al. (2000) best described data for children aged 8-13 years, and Lee and colleagues observed the shift to a three-factor model occurred around age 15. In some cases, the shift from two- to three-factor functioning can begin as early as age 11 and is protracted over several years. It is not a linear process and varies based on the individual. In addition, there is some evidence that the efficiency of EF is affected by processing speed (Rose, Feldman, & Jankowski, 2011).

Performance-Based Measures (PBM) v Self/ Other Ratings

Most studies use PBMs to test models of EF (e.g., Busch et al., 2005; Lee et al., 2013; Miyake et al., 2000). A PBM is a context-specific task that generally requires the examinee to act or select "in the face of strongly competing, but context-inappropriate responses" (Pennington & Ozonoff, 1996, p. 55). The use of PBM came about when research indicated that intelligence tests were not sensitive to planning and problem-solving abilities in patients with prefrontal lesions (Kolb & Wishaw, 1990).

Tests like the Wisconsin Card Sorting Test (WCST; Berg, 1948) and Trail Making Test (TMT; Partington, 1949) are prototypical in neuropsychological assessment (Pennington & Ozonoff, 1996). Interestingly, the TMT originated as a test of intelligence in 1938 and was later added to the Army Individual Test Battery (1944); however, later studies established its validity and sensitivity to brain damage in children and adults (Reitan, 1955, 1966, 1967).

Findings related to EF assessment in ADHD have been mixed (e.g., Piek, Dyck, Francis, & Conwell, 2007), which may result from discrepant operational definitions of EF models as well as their individual constructs (e.g., inhibition – activation systems versus connectionist networks; McAuley et al., 2010; Miyake & Friedman, 2012; Ten Eycke & Dewey, 2015; Toplak et al., 2013). Given the number of ways EF has been conceptualized, studies may incorporate specific theories that exclude factors salient to other EF models. Unlike Welsh and Pennington (1988), who include planning in their three-factor model, Miyake et al. (2000) does not; rather, planning permeates throughout the constructs. In other words, inhibition, shifting, and WM influence one's planning ability, but planning is not the underlying mechanism.

Furthermore, the results of Miyake et al. (2000) support the idea that within EF are individual constructs that should be evaluated as separate entities, but are interrelated. In other words, despite their interrelationship, each construct is distinctive from one another. This expectation, however, is met with challenges related to task impurity. "Task impurity" refers to a derived score for a given construct that also reflects "systematic variance" explained by secondary EF and non-EF processes tapped by a given task (Miyake & Friedman, 2012, p. 8). A prime example includes Stroop-like tasks because they simultaneously tap inhibition and shifting (Homack & Riccio, 2004, Miyake & Friedman, 2012).

Relatedly, score variability on inhibition tasks may be attributed a host of internal factors not accounted for by PBM of EF. For instance, Alderson et al. (2007) and Castellanos et al. (2006) suggested that high score variability MRT on the stop-signal paradigm was indicative of both deficient attentional processes and slower motor speed, in addition to highly variable profiles of behavioral inhibition (Lijffijt et al., 2005). Similarly, Martinussen et al. (2005) attributed inconsistent WM performance in children and adolescents to differences in linguistic and number knowledge. In other words, neither comorbid status nor IQ explained the variance in neither the verbal storage nor verbal CE domains of WM.

The results of Martinussen et al. (2005) highlight two important, interpretative considerations linked to discriminant validity of scores on EF measures. Although comorbid status (i.e., reading disability/ language impairment [RD/ LI]) did not mediate the variance for verbal aspects of WM, this was not the case for either of the spatial domains of WM (i.e., storage and CE). Further, their results yielded larger ES pertinent to studies that controlled for comorbid RD/ LI, adding to the inconclusive nature of this issue. Where some authors link these problems to spatial WM deficits (e.g., McInnes, Humphries, Hogg-Johnson, & Tannock, 2003), others do not (e.g., Williams, Stott, Goodyer, & Sahakian, 2000). These issues are germane to reasons why the literature presents conflicting views on the use of PBM in EF assessment (McAuley et al., 2010;

Pennington & Ozonoff, 1996; Toplak et al., 2008; Toplak et al., 2013). Their utility is met with several criticisms, which are discussed here.

The consensus is that performance on a standardized, laboratory test is too narrow, and does not capture the broad sense of EF (Barkley & Murphy, 2011; Salthouse, Atkinson, & Berish, 2003; Toplak et al., 2013). For example, perseverative errors on the WCST do not sufficiently describe one's ability to problem solve and shift behaviors toward goal setting and attainment. In other cases, PBM have shown to assess abilities that are parallel to EF, but may be more indicative of general cognitive ability. For example, digit span (DS) backward tasks are typically used to measure WM; however, Rosenthal, Riccio, Gsanger, and Jarratt (2006) described these measures as also reflecting one's sequencing ability.

Standardized, neuropsychological tests also have been described as overly sensitive to score variability (Manchester, Priestley, & Jackson, 2004) and having high within-person variability (Salthouse, 2007). The dynamic nature of EF tasks (Barkley, 1997), and the examinee's vulnerability to practice effects (Salthouse & Tucker-Drob, 2008) both have been cited as considerable problems. These considerations may contribute to reasons why PBM yield minimal correlations with EF rating scales (Barkley, 2011; McAuley et al., 2010; Ten Eycke & Dewey, 2015; Toplak et al., 2013). Alternatively, rating scales aim to describe EF in an ecologically valid way (Gioia et al., 2000; Toplak et al., 2013).

Ecological validity is "the extent to which research findings would generalize to settings typical of everyday life" (Lewis-Beck, Bryman, & Liao, 2003, p. 2).

Traditional neuropsychological tests have been perceived as insufficient measures of EF and contrived, thus providing a mere snapshot of performance in one setting (Burgess, 1997; Eslinger & Damasio, 1985; Levine et al., 1998; Mesulam, 1986). As a result, rating scales were designed to inform individuals about the real-world application of EF, and ways in which deficits may interfere with their daily functioning (Isquith, Roth, & Gioia, 2013; McAuley et al., 2010; Silver, 2000; Toplak, West, & Stanovich, 2017). As such, rating scales have become highly accepted methods for evaluating EF across multiple clinical populations (e.g., Azouvi et al., 2016; Blijd-Hoogewys, Bezemer, & van Geert, 2014; Sadeh, Burns, & Sullivan, 2012).

Rating scales differ from PBM in the sense that the former represents the "algorithmic mind," while the latter is the "reflective mind" (Toplak et al., 2013, p. 137). Together, they afford practitioners and researchers the opportunity to examine different aspects of cognitive functioning. The algorithm analogy refers to input coding mechanisms, perceptual reasoning, and memory; conversely, the reflective aspect describes interactions between goal-oriented responses, beliefs, and rational control (Stanovich, 2011).

Specific to the BRIEF, studies generally show that the Behavioral Regulation and Metacognition Indices significantly associate with parent and teacher ratings of childhood inattention and behavioral/ social-emotional problems (McAuley et al., 2010). Low ratings also correlate significantly with ADHD symptoms (Mahone et al., 2002; Sullivan & Riccio, 2006; Toplak et al., 2009). In children with ADHD and comorbid Tourette's Syndrome (TS), Mahone et al. (2002) indicated that neither the Behavioral

Regulation nor Metacognition Indices of the BRIEF correlated with PBM of EF, intellectual ability, or reading abilities; however, they did predict math achievement while correlating with scores on the ADHD Rating Scales IV – Home Version (DuPaul, Power, Anastopoulos, & Reid, 1998). McAuley et al. (2010) partially corroborated these findings, specifically that the Metacognition Index associated with both reading and math performance. Although these findings are mixed, it may be due to differences in the measurement of academic performance. Interestingly, Toplak et al. (2009) yielded significant, yet modest correlations between the performance measures and parent and teacher ratings on the BRIEF, suggesting a convergence of the neuropsychological and behavioral ratings of executive functions in ADHD.

While studies use rating scales to elucidate developmental aspects of factor structures, the findings are mixed. For instances, some studies show that for clinical and healthy groups using the BRIEF, a two-factor model best fits child and adolescent data: the Behavioral Regulation Index (BRI) and Metacognition Index (MI). The BRI encompasses shifting abilities, emotion/ behavior regulation, and inhibitory control; the MI refers to WM, plan initiation, problem-solving/ organizational skills, and self-monitoring capabilities. In a few instances, the BRI has been fractionated into two separate entities of regulation: emotional and behavioral, hence results indicated a three-factor solution (Egeland & Fallmyr, 2010; Gioia & Isquith, 2002). Additionally, the BRIEF rating form has been adapted for preschool-aged children (BRIEF-P; Gioia, Espy, & Isquith, 2003) and adults (BRIEF-A; Roth, Isquith, Gioia, 2005). In healthy preschool children, results indicated that a three-factor solution best fit the data;

however, the difference in fit compared to the second-order one-factor model was marginal (Skogan et al., 2016). In mixed sample of adults (ADHD and healthy controls), CFA suggested a three-factor solution (Roth, Lance, Isquith, Fischer, & Giancola, 2013).

In adults, Barkley and Murphy (2011) found that self-ratings across ADHD and subclinical (i.e., individuals who did not meet DSM-IV criteria for ADHD but showed symptoms related to other types of pathology) groups exhibited impairments across five dimensions of EF, including inhibition, nonverbal WM, verbal WM, motivational self-regulation, and planning/ problem-solving. Like child and adolescent studies, they did not detect a significant relationship between self-reported EF ratings and PBM but that each measured different aspects of EF. In terms of adaptive impairments and real-world functioning, their findings showed that CPT commission score yielded a modest yet significant contribution to the crime diversity score, and poor performance on the WCST was linked to the frequency of being in jail. Specific to self-reported driving scores, both the Speeding and Driving Under the Influence (DUI) domains were associated with the Stroop Interference Score. EF ratings yielded better predictive validity over PBM. For instance, low scores on the EF scales associated with aspects of antisocial functioning (e.g., crime diversity, arrests, jailed) and self-reported driving scores. In summary, the literature is characterized by a myriad of issues related to EF assessment.

Both PBM and rating scales purport to measure EF; however, their association with one another is inconclusive. The study by Toplak et al. (2009) contributes to this discrepancy in a meaningful way. Their efforts accounted for potential issues related to sample size and statistical power. Compared to previous studies exploring different populations (e.g., brain disease and injury; Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002; Vriezen & Pigott, 2002), their study encompassed a larger sample size ($N = 90$). Additionally, they factored considerations such as the use of well-defined groups and measures sensitive to ADHD. Despite these differences, furthermore, their study bolsters evidence for the importance of using both PBM and rating scales in clinical practice (Barkley, 2011). The current study is interested in these differences, particularly as they relate to the assessment of shifting ability in ADHD. Additionally, there is a paucity of literature focusing on the longitudinal trajectory of shifting in individuals with ADHD. A closer look at the developmental aspect of shifting may help elucidate changes in the reflective and algorithmic mind specific to neurodevelopmental impairment (Ten Eycke & Dewey, 2016).

Statement of the Problem

ADHD is a commonly diagnosed disorder in children with a myriad of potentially long-term negative outcomes (APA, 2013; Kessler et al., 2014; Visser et al., 2014). As it stands, the literature presents several challenges related to the conceptualization and assessment of EF. Our knowledge on inhibition and WM in ADHD is comprehensive and evidenced by several meta-analyses (e.g., Alderson et al., 2007; Kasper et al., 2012; Lijffijt et al., 2005; Martinussen et al., 2005), yet less is known about shifting in ADHD. Adults with ADHD seem to continue showing shifting deficits (Boonstra et al., 2005); however, it is unknown as to whether those presumed deficits are pervasive or merely specific to performance on a single task. Finally, discrepancies related to the assessment of EF create challenges in understanding the results across studies; in tandem, they do not allow for a thorough investigation into ADHD within a specified, empirical model of EF (e.g., Miyake et al., 2000).

Because the tripartite model by Miyake et al. (2000) is empirically supported and has been replicated across the lifespan, it provides the best framework for conceptualizing EF in children and adults with ADHD. The current study aims to rely on previous literature describing each individual construct (i.e., inhibition, WM, shifting) to bolster evidence that links shifting deficits to ADHD. This will allow for future comparisons between the presumed phenotypes associated with ADHD (i.e., inhibition and WM) as well as other types of pathology specifically linked to cognitive inflexibility (e.g., autism).

A second issue is tied to methodology and regards the assessment of EF. Overall, research suggests that EF continues to develop well into adulthood (Lee et al., 2013; Miyake et al., 2000), and that multicomponent structures are specific to older children and adults (Wiebe et al., 2011; Willoughby et al., 2010; Willoughby, Wirth, & Blair, 2012). The latter assumption is somewhat inconclusive. Lee et al. (2013) found that data for younger children (six- and seven-year-olds) best fits a two-factor model encompassing the WM and combined inhibition-shift factors. Naturally, discrepancies in the ways the factor structure of EF is conceptualized impacts assessment methods and drives decisions regarding the specific measures administered in studies and the targets of intervention. Furthermore, research is inundated with several conflicting views regarding the nature of assessment methods for EF (PBMs v. rating scales). The current study explored these issues as they relate to the shift factor in Miyake et al. (2000).

Third, the model by Miyake et al. (2000) has not been systematically tested in individuals with ADHD. Given the many discrepancies in the literature describing certain EF traits as possible phenotypes, an updated meta-analysis will reflect findings specific to shifting deficits and bolster the current state of the literature as it relates to EF and ADHD. Assuming there are noteworthy differences in shifting between ADHD and HC groups, and within group differences moderated by different variables (e.g., IQ, age, etc.), it will be important to determine whether those differences are affected by the assessment method (PBMs versus rating scales). The current study is seeking to do just that. A systematic review of studies focusing on individuals with ADHD will afford us that better and more concise understanding of shifting as a possible phenotype as well as differences resulting from changes in assessment methods.

CHAPTER III

METHODS

The current study used meta-analysis to aggregate research within the scope of ADHD and the EF component of "shift." Meta-analysis refers to "the statistical synthesis of results from a series of studies" (Borenstein, Hedges, Higgins, & Rothstein, 2009, p. 2). Lipsey and Wilson (2001) described meta-analysis as a survey of studies instead of people. Over time, it has evolved into the gold standard for summating and synthesizing research into one study. It is methodology that originated out of psychotherapeutic literature and subsequently has been adopted across several disciplines, including educational and medical sciences (Cheung & Vijayakumar, 2016; Cooper, Hedges, & Valentine, 2009; Pierce, 2007). The intent behind the present analyses was to update the findings of Pennington and Ozonoff (1996) and Willcutt et al. (2005) by exploring group mean differences on measures of shifting between ADHD and HC groups. A systematic review was performed, which can aid in answering many types of questions (McKenzie, Beller, & Forbes, 2016; Steenbergen-Hu & Olszewski-Kubilius, 2016).

Eligibility Criteria

Studies eligible for this review included at least one ADHD and one healthy control (HC) group, along with results pertinent to set shifting ability as defined by Miyake et al. (2000) and indicated by frequently used measures of EF. That is, study results need to have evidenced each group's overall performance on task-based measures of set shifting and/or ratings for a shift index on a behavioral checklist. The selection criteria expanded beyond literature solely focused on PBMs of EF because it is well documented that PBMs and rating scales assess different aspects of EF (Davidson, Cherry, & Corkum, 2016; McAuley et al., 2010; Toplak et al., 2009; Toplak et al., 2013), which is one reason why it is important for clinicians to use PBMs and rating scales in conjunction with, rather than exclusive of, one another (Barkley, 2011).

To use meta-analytic procedures, it was essential for results to be expressed as means and standard deviations so that group mean differences and resulting effects could be computed for each study. Because EF continues to develop well into adulthood (Lee et al., 2013), this study aimed to capture that lineage by including studies with adult participants as well as children. As a way of corroborating past findings, these considerations remained consistent with criteria outlined in other relevant studies (e.g., Barkley & Murphy, 2011; Lijffijt et al., 2005, Pennington & Ozonoff, 1996; Willcutt et al., 2005).

To capture the breadth of available research, published studies and unpublished dissertations were included, which also helped in minimizing the potential for publication bias (Banks et al., 2012). A host of different comparisons also were considered, including those involving pharmacological effects on EF (e.g., Yilmaz et al., 2013) and other factors like age and gender (e.g., Seidman et al., 2005). Given the comorbidity rates of ADHD and other disorders (e.g., Kessler et al., 2014), a consideration that was difficult to avoid, studies using clinical groups with some degree of comorbidity also were included. Both the nature and amount of comorbidity were accounted for in the coding form (see Appendix A). This study did not impose any restrictions regarding sample size, setting (i.e., hospital, school, community, etc.), inclusion or exclusion of measures of cognitive functioning, or geographical region.

Study Selection

Design and execution of the search strategy occurred over three days. The author consulted with a librarian specialist to develop and organize search terms. Consultation took place over two meetings. The final search incorporated text words related to ADHD, EF, set shifting, and measures of set shifting. To maintain a comprehensive search strategy, the proposed terms outlined multiple considerations, including the investigated neurodevelopmental pathology (e.g., "ADHD"), primary construct (e.g., "set shifting"), and outcome measure(s) (e.g., "card sort"). Imposing the above-mentioned limits, only quantitative studies were sought.

The literature search also was limited to studies reported in English, involving human subjects, and conducted between 1994 and 2018. The rationale for using studies published after Pennington and Ozonoff (1996) and before Willcutt et al. (2005) was to create a comprehensive search, which is important given that no other study has broadly focused on shifting deficits in ADHD. To further promote this, reference lists of included studies, along with relevant reviews such as previous meta-analyses (e.g., Frazier, Demaree, & Youngstrom, 2004; Romine et al., 2004, Toplak et al., 2013) were scanned. These steps are consistent with those outlined in Tricco et al. (2012).

The literature search was conducted on July 21, 2018, using two online databases accessed through Texas A&M University's university library: PsycINFO and MEDLINE

(PubMed). The use of two databases is consistent with practices outlined by Willcutt et al. (2005). The initial search yielded 808 records.

Study Records

Records were organized using Rayyan (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016), an Internet based software program and mobile application for systematic reviews. Rayyan allows users to screen and designate records of abstracts to a filing system customized to fit the study design. One of the key features in Rayyan is that users can filter duplicate records that are retrieved from multiple databases. Screening questions were developed and based on the above-mentioned inclusion and exclusion criteria:

1) Is the record in English?

2) Does the record mention the use of human subjects?

3) Does the record mention and describe the sample as having at least one ADHD and one HC group?

4) Does the record include the age range of participants?

5) Is the age range within five and 50 years?

6) Does the record mention some variant of set shifting as a target variable?

Using these questions, study records were then assessed in two phases. In phase one, records were screened and assessed based on the information provided in the abstract and title. Records that either met the inclusion criteria or required further review were assigned to an "Included" folder. Those that did not meet the criteria were designated to an "Excluded" folder, meaning records required no further review and were excluded from the study. In phase two, records in the "Included" folder underwent a full-text assessment wherein those meeting inclusion criteria were retained for the study and those that did not were excluded. If the eligibility of a given record was unclear, advice from the committee chair was sought or the author of the record in question was contacted through e-mail. Excluded records, either based on a single screening or full-text assessment, were labeled with the reason(s) for not meeting the inclusion criteria. Exclusionary factors were labeled as "wrong language," "wrong population," "wrong study design," and/ or "wrong outcome." All records in the "Included" folder underwent full-text assessment. Throughout this process, titles and authors were reviewed to identify and remove all duplicated records. This protocol is consistent with that outlined in Macdonald and McCartan (2014). Based on full text assessment, 66 studies remained in the "Included" folder.

Data Collection Process

Using a standardized form in Microsoft Excel (2016) and detailed coding form (see Appendix A), a number of study characteristics were extracted, including publication type, representative age range and mean age of participants, race/ethnicity, gender ratio, prevalence of comorbid disorders, medication status, recruitment strategies (i.e., type of setting), study location, ADHD subtype specification (if indicated), diagnostic criteria utilized, mean IQ scores and indication of its utility as a covariate, name and version of the PBM (or rating scale), and outcome variables reported in the study. For instances in which effect size calculation required additional information outside of the study, authors were contacted.

To ensure consistency and reliability across the data extraction process, a second reviewer recoded 20 articles from the total pool of studies. Recoding a minimum of 20 articles from a "moderate" size sample yields "a relatively stable reliability estimate" (Lipsey & Wilson, 2001, p. 86). Recoding procedures took place following a 30-day time lapse from completion of the initial coding phase. The second coder was trained in one session. At this session, the author reviewed the coding form with the second coder and discussed definitions for each code. Discussions were had regarding some of the challenges experienced in the initial coding process. One calibration exercise was completed between coders. Disagreements pertaining to objective variables (e.g., sample size) were resolved through discussion between the two coders.

An aim of this study was to examine the impact of study quality on the relationship between symptoms of ADHD and shifting impairments. Decisions for developing a rubric for study quality were guided by themes outlined in GRADE (Grading of Recommendations Assessment, Development and Evaluation; Dijkers, 2013). GRADE is a formal process designed to rate the quality of evidence in systematic reviews. A key emphasis of GRADE concerns the precision and certainty of results. It was decided *a priori* that factors related to study design and sampling procedures had the potential to impact precision, certainty, and bias. As a result, the following study-level characteristics were examined: 1) diagnostic process, 2) accounts of cognitive ability, 3) gender equality across clinical and control groups, 4) statistical methods accounting for gender equality, 5) processes for determining typicality for the control group, and 6) standardization procedures for test administration (see Appendix A).

After methodological quality was assessed for each study, a decision was made to retain studies resulting in low methodological quality. This was rationalized with the aim of capturing the current state of the literature. Inclusion of these studies was also helpful, given that methodological quality was examined as a moderator and thus, allowed for exploration of between-study variance (Cheung & Vijayakumar, 2016; Liberati et al., 2009).

Measures of Shifting

Search strategies were designed to retrieve a significant number of studies across a breadth of measurement methods, including a variety of PBMs and behavioral rating scales all purporting to measure set shifting. Table 3 organizes measures that were explored according to the nature of the task. The primary categories include card sorts, fluency switching (e.g., verbal fluency), intra-/extra-dimensional shifting, learning trials, trails, and rating scales. An important consideration regarded the inclusion of Stroop-like tasks as representative of set shifting. Initially, the Stroop Interference Test (Stroop, 1935) was developed to measure selective attention and cognitive flexibility (Homack & Riccio, 2004); however, it is typically theorized to measure cognitive inhibition of dominant/ automatic responses and interference control (Biederman et al., 2004; Bledsoe et al., 2013; Dimoska-Di Marco, McDonald, Kelly, Tate, & Johnstone, 2011). Given this rationale is consistent with its utility in several related studies (e.g., Boonstra et al., 2005; Miyake et al., 2000; Pennington & Ozonoff, 1996), Stroop-like tasks were not included in the current study.

Table 3 Measures of "Shift" in Studies Reviewed

Measure/ paradigm	Versions available	Example	Total studies	Scores utilized
Card Sorting Tasks	Global/ Local Task (Rinehart et al., 2001)	Gargaro et al. (2015)	2	Error rate; response time
	WCST (Berg, 1948; Grant & Berg, 1948; Heaton, 1981; Heaton et al., 1993; Nelson, 1976)	Antonini et al. (2015); Antshel et al. (2010); Aycicegi-Dinn et al. (2011); Chen et al. (2015)	51	Failure to maintain set; non-perseverative errors; learning to learning; number of categories; Milner/Nelson type errors; number of trials to completed the first category; percentage of conceptual level response; percentage of perseverative errors, percentage of perseverative responses; perseverative errors; perseverative responses; total correct; total errors; total responses; total time

Table 3 Continued

Measure/ paradigm	Versions available	Example	Total studies	Scores utilized
	WMST (Wilding et al., 2001)	Hobson et al. (2011)	1	Percentage of perseverative errors
Fluency Switching Tasks	Number-Color (Miyake et al., 2000)	Kofler et al. (2018)	1	Percent correct responses; response time
	RWT (Aschenbrenner et al., 2000)	Tucha et al. (2005)	1	Number of switches between clusters
	Verbal Fluency (D-KEFS; Delis et al., 2001)	Anderson et al. (2010)	3	Number of correct shifts
	Verbal Fluency (NEPSY; Korkman et al., 1998)	Loo et al. (2007)	2	Number of categories; number of correct responses; phonological score; total score
Intra/ Extra-Dimensional Shift	Creature Counting (TEA-Ch; Manly et al., 1998)	Anderson et al. (2011)	1	Number of correct trials (out of 7)
	I/ED (CANTAB; Cambridge Cognition Limited, 1996; 2006)	Chamberlain et al. (2007)	17	Adjusted total errors; adjusted total trials; ED level shift errors; number of stages completed; pre-ED level errors; total errors; total errors (ID); total reversal errors; total stage trials; total trials from levels 6 to 9; total trials to ED level; total trials to ID level

Table 3 Continued

Measure/ paradigm	Versions available	Example	Total studies	Scores utilized
	Shift Task (Rogers & Monsell, 1995)	Fuggetta (2006)	1	Alternating runs condition
	TAVIS-3 (Coutinho et al., 2007)	Coutinho et al. (2009)	1	Commission errors; hit reaction time; omission errors
Rating Scales	BDEFS (Barkley, 2011)	Mazursky-Horowitz et al. (2018)	1	BDEFS M
	BRIEF (Gioia et al., 2002)	Bodzy (2012)	22	Shift factor
Switch/ Learning Tasks	Attention Switching Paradigm (van Schouwenburg et al., 2010)	van Schouwenburg et al. (2014)	1	Novel switch trials
	Change Task (Oosterlaan & Sergeant, 1998)	Marzocchi et al. (2008)	1	Change mean reaction time; change number of errors; words produced
	Flexibility (KITAP; Zimmermann et al., 2002)	Drechsler et al. (2010)	1	Reaction time
	Modified Meiran Switch Task (Meiran, 1996)	Rubia et al. (2010)	1	Switch and repeat reaction times/ errors
	Navon Task (Miyake et al., 2000)	Holst & Thorell (2017)	2	Mean reaction time (3rd trial)
	Saccade-Antisaccade Mixed Task (O'Driscoll et al., 2011)	O'Driscoll et al. (2011)	1	Switch cost
	Shifting Attention Test (CNSVS; Gualtieri et al., 2004)	Gualtieri & Morgan (2006)	1	Average reaction time; correct responses; errors

Table 3 Continued

Measure/ paradigm	Versions available	Example	Total studies	Scores utilized
	Switch Task (Dibbets & Jolles, 2006)	Dibbets et al. (2010)	1	Response time and accuracy
	Task-Switching Test (e.g., Cepeda et al., 2000)	Gupta et al. (2011)	2	Error rate; reaction time; switch costs (difference in the overall reaction times of switch and nonswitch trials)
Trail Making Tests	CCTT (D'Elia et al., 1994; Llorente et al., 2003)	Corbett et al. (2009)	3	CTT_shift measure
	CST (Reitan, 1958)	Marchetta et al. (2008)	1	CST shifting score (CST shifting = CST_C – CST_AB)
	TMT (Lan et al., 2011; Reitan, 1955; 1958; 1969; 1971; 1979; Reitan & Davison, 1974; Reitan & Wolfson, 1985; 1993; Spreen & Gables, 1969; Spreen & Strauss, 1998)	Elosúa et al. (2017)	41	B – A (time difference in raw scores); B/A (time ratio); condition 4 – [condition 2 + condition 3]/ 2 (D-KEFS); part B time; part B total errors; percentage correct based on age norms; total time (A + B)
	TMT (D-KEFS; Delis et al., 2001)	Holmes et al. (2010)	6	Condition 4
Other	CWIT (D-KEFS; Delis et al., 2001)	Halleland et al. (2012)	1	Condition 4 – Condition 3

Table 3 Continued

Measure/ paradigm	Versions available	Example	Total studies	Scores utilized
	Computerized Stroop Test (Capovilla et al., 2005)	dos Santos Assef et al. (2007)	1	Interference reaction time and score
	Experimental task (De Sonneville, 1999)	Rommelse et al. (2007)	1	Mean reaction time; percentage of errors
	Go/ No-Go (Dodds et al., 2011; Linssen et al., 2012)	Linssen et al. (2012)	2	Commission errors; false alarm rate; hit rate; omission errors; response time
	Set-Shifting/ Attentional Shift Task (TAP; Zimmermann & Fimm, 1993)	Koschack et al. (2003)	2	Errors; false alarms; median and SD of hit reaction times; number of misses; reaction time; total hits
	Set Shifting Task (DeSonneville, 2001)	Greimel et al. (2011)	4	Number of errors; reaction time
	Shifting Sets Test (McKay et al., 1994)	Kovner (1998)	1	Mean reaction time; reaction time variability
	SSV (ANT)	Sinzig et al. (2014)	1	Mean hit reaction times; number of errors

Note. ANT = Amsterdam Neuropsychological Tasks; BDEFS = Barkley Deficits in Executive Functioning Scale; BRIEF = Behavior Rating Inventory of Executive Function; CANTAB = Cambridge Neuropsychological Test Automated Battery; CCTT – Children's Color Trails Test; CNSVS = CNS Vital Signs; CST = Concept Shifting Test; CWIT = Color-Word Interference Test; D-KEFS = Delis-Kaplan Executive Function System; I/ ED = Intra/ Extra-Dimensional; KITAP = Test of Attentional Performance in Children; NEPSY-II = A Developmental NEuroPSYchological Assessment, Second Edition; PRL = Probabilistic Reversal Learning Task; RWT = Der Regensburger Wortflüssigkeitstest; SSV = Shifting Attentional Set Visual; TAP = Test for Attentional Performance; TAVIS-3 = Test of Visual Attention; TEA-Ch = The Test of Everyday Attention for Children; TMT = Trail Making Test; WCST = Wisconsin Card Sorting Test; WMST = Wisconsin Monster Sorting Test.

Table 3 outlines six broad measurement categories and one "other" classification designated for more idiosyncratic tasks, oftentimes researcher-developed and specific to an individual study. It also informs the number of available versions and authorship for each version. Examples of studies using each measure also are highlighted, as well as the number of retrieved studies that used the specific task. Finally, variables and scores purporting to reflect a measure of shift also are provided. Results indicated that the most frequently used measures of set shifting include the Wisconsin Card Sort Test (e.g., Heaton, Chelune, Talley, Kay, & Curtiss, 1993), Trail Making Tests (e.g., Reitan and Wolfson, 1985), and the Shift index of the BRIEF (Gioia et al., 2000). Specific to the PBMs, several variables were reported across studies.

Decisions regarding which outcome variables to meta-analyze were based on a few considerations. First, it was important to extract variables that were the most prevalent across studies. Second, it was essential that each outcome demonstrated a clear linkage to at least one aspect of set shifting and cognitive flexibility as defined by Miyake et al. (2000). In other words, which variables were the most frequently reported in studies and do they purport to assess shifting and cognitive flexibility? Third, the present study aimed to provide results that were summative and less indicative of spurious variables; therefore, an *a priori* decision was made to forego calculating effect sizes of variables reported in less than 10 studies. This was an important consideration, given analyses such as funnel plots (Egger, 1997; Higgins & Green, 2017; Sterne et al., 2011) and meta-regression (Borenstein et al., 2009) are more optimal for study pools of 10 or more. As a result, studies that included the more idiosyncratic measures and potentially spurious variables (i.e., reported in fewer than 10 studies) were excluded. Between the WCST and TMT/D-KEFS measures, 13 variables were reported in fewer than 10 studies. For the BRIEF, Teacher and Self ratings were scores reported in fewer than 10 studies.

The Wisconsin Card Sort Test (WCST)

The WCST (Heaton, 1981) was originally designed to test general problem-solving and decision-making abilities (Berg, 1948; Grant & Berg, 1948). It is frequently used to measure neuropsychological functioning, specifically shifting and cognitive flexibility (Biederman et al., 2004; Puente, 1985; Romine et al., 2004). Since its development in 1948, the WCST has evolved across several variations and formats; however, the primary objective of the test has not changed in that examinees learn strategies for sorting cards based on color, form, and number. In the standard version, there are two identical decks each containing 64 cards (128 total). Cards contain up to four figures (plus sign, star, circle, or triangle), and figures are the same color for each card: red, green, yellow, or blue (Puente, 1985). The examiner presents cards one-at-a-time and the examinee must detect the underlying rule for each card sort (i.e., by color, form, or number) based on examiner feedback. The examiner presents cards continuously and criteria changes once the examinee achieves 10 correct responses for each category. The examinee is made aware of the shift solely by the examiner's feedback, either positive or negative.

The study pool reflects multiple versions of the WCST. While many studies included use of the standard 128-card version, other studies (Chen et al., 2015) utilized short (WCST-64; Haaland, Vranes, Goodwin, & Garry, 1987) and computerized versions (Heaton, Thompson, & Gomez, 1999; Tsuchiya, Oki, Yahara, & Fujieda, 2005). These variants have shown to be overall comparable measures to the standard version (Axelrod, Henry, & Woodard, 1992; Fortuny & Heaton, 1996; Heaton & Thompson, 1992; Sillanpaa et al., 1993), thus allowing for studies using different test formats to be coalesced for meta-analysis. Table 4 outlines the different variables that were extracted from the card sorting studies. Again, decisions surrounding which variables to include in the analyses were based highly on frequency and which ones were most frequently used across the literature.

The outcomes listed in Table 4 were the most prevalent among WCST variables reported across studies. It is well-documented that these variables are good indicators of one's ability to shift between attention and behavioral sets (e.g., Barceló & Knight, 2002; Greve et al., 2002; Nyhus & Barceló, 2009; Pasini, Paloscia, Alessandrelli, Porfirio, & Curatolo, 2007; Romine et al., 2004). The TE and #CAT variables have been shown to reflect one's ability to shift attentional sets (Brewer, Fletcher, Hiscock, & Davidson, 2001; Seidman et al., 2005). Barkley, DuPaul, and McMurray (1991) indicated that children with ADHD often make more PE compared to their typically developing peers. Moreover, high #CAT scores suggest a high level of general problem-solving ability (Holdnack et al., 1995). Further, lower numbers of errors indicated that the examinee effectively utilized the examiner's feedback to shift their sorting behaviors accordingly.

Table 4 Wisconsin Card Sorting Test Variables

Variable	Description	Number of studies
Perseverative Errors (PE)	The number of error responses considered correct for a previous category (Bowden et al., 1998).	33
Number of Categories Completed (#CAT)	The number of categories the examinee accurately sorted, based on their correct responses (Heaton et al., 1993).	32
Total Errors (TE)	The total number of error responses, perseverative and non-perseverative errors, across all categories (Kado et al., 2012).	19
Nonperseverative Errors (NPE)	The number of errors derived by subtracting the PEs from the TEs (Kado et al., 2012)	17
Failure to Maintain Set (FMS)	The number of sequences of two or more correct responses proceed one or more errors occurring before a single category is completed successfully (Kado et al., 2012).	17
Total Correct (TC)	The total number of correct sorts based on color, shape, and form (Heaton et al., 1993).	13
Perseverative Responses (PR)	The number of responses that were considered correct for a previous category (Bowden et al., 1998).	11

The Trail Making Tests (TMT and D-KEFS)

Partington and Leiter originally developed the TMT as a test of intelligence in 1938, referred to as the "Partington's Pathways Test." It was later renamed the "Test of Distributed Attention" (Partington, 1949). Following these iterations, it became part of the Army Individual Test Battery (1944) and has since been added to several neuropsychological test batteries including the Halstead-Reitan Battery (Reitan & Wolfson, 1992) and Delis-Kaplan Executive Function System (D-KEFS; Delis et al., 2001) as a measure of shifting (Allen, Haderlie, Dmitriy, & Mayfield, 2009; Reynolds, 2002; Skogli et al., 2014).

The TMT generally encompasses two parts: Trail Making, Part A (TMT-A) and Part B (TMT-B; Allen et al., 2009; Anderson, 1998; Reitan & Wolfson, 1992). The Trailmaking Test in the D-KEFS (Delis et al., 2001) contains five conditions in which case Condition 2: Number Sequencing and Condition 4: Number-Letter Sequencing are comparable to Part A and Part B, respectively (Delis et al., 2001). The TMT-A and D-KEFS Condition 2 involves drawing a line and connecting consecutive numbers from 1 to 25. For Part B or D-KEFS Condition 4, the examinee is asked to draw a line that connects alternating numbers and letters sequentially (i.e., 1-A-2-B, etc.; Arbuthnott & Frank, 2000; Hill, 2006; Reitan & Wolfson, 1993). These tasks are timed and have been standardized wherein the examiner identifies the errors as they occur (Lezak, 1995).

While the TMT-A yields information about motor and visual control as well as the speed of those mechanisms (Anderson 1998; Duff, Beglinger, Theriault, Allison, & Paulsen, 2010; Fisher, Garges, Yoon, Macguire, Zipay, & Gambino, 2014), studies suggest that the TMT-B loads on an attention factor (e.g., Fisher et al., 2014; O'Donnell, MacGregor, Dabrowski, Oestreicher, & Romero, 1994). For instance, a slower TMT-A performance suggests some form of cognitive impairment like general frontal lobe dysfunction (Ameiva et al., 1998). Combined, scores on the TMT-A and TMT-B are typically interpreted as the B-A difference and B/A ratio (Arbuthnott & Frank, 2000). Arbuthnott and Frank studied the B-A difference and B/A ratio in healthy adults. Their correlational analysis confirmed that scores on the TMT-B are more reflective of the attentional control needed to negotiate the immediate shift in tasks and requirements. Further, the attentional control tapped by the TMT-B is needed to adjust one's thinking and adherence to meet the requirements of TMT-B. Findings overall have confirmed that set shifting ability is one critical cognitive mechanism differentiating TMT-A and TMT-B.

While the relationship between A and B has been investigated broadly in both clinical and healthy individuals, fewer studies involving ADHD populations have utilized the difference and ratio variables as measures of set shifting (e.g., Elosúa, Del Olmo, & Contreras, 2017; Gonzalez-Gadea et al., 2015). In situations where studies reported raw completion times for both Part A and Part B, the difference and ratio were able to be calculated. To examine performance differences across several TMT variables, this was a post hoc decision that also involved manual calculation of standard deviations for the derived B-A and B/A variables. Based on strategies outlined in Higgins and Green (2011), calculations were performed using the group sample sizes and computed 95% upper- and lower-confidence intervals. In an effort to obtain B-A data, authors were contacted in cases where the raw completion times for either Part A or Part B, but not both, were reported. These requests were communicated via e-mail.

The TMT and D-KEFS Trail-Making use similar stimuli and employ the same task to measure mental flexibility (Delis et al., 2001). For this reason, studies using the D-KEFS version of the test were retained for this study. Further, variables used for this study include the raw score total time to complete TMT-A (or D-KEFS Condition 2) and TMT-B (or D-KEFS Condition 4), the raw score B-A index (or D-KEFS Condition 4-Condition 2), and the raw score B/A index (or D-KEFS Condition 4/ Condition 2). In some cases, scores were standardized or expressed as *T*-scores. These scores were also included in the analysis.

The Behavior Rating Inventory of Executive Function (BRIEF)

The BRIEF (Gioia et al., 2000) is a parent, teacher, and self-report rating scale that was developed to assess neuropsychological functioning in children and adolescents ages five through 18. Behavioral ratings aim to reflect the extent to which perceived neuropsychological deficits interfere with daily life. The BRIEF yields *T*-scores for eight individual scales, two indexes derived from factor analysis of the eight scales, and the Global Executive Composite (GEC). For the purposes of this study, *T*-scores for the Shift scale were extracted. The BRIEF defines Shift as the ability to transition between situations, activities, or aspects of a problem to another in accordance with demands of a given situation; the ability to adjust one's focus from one mindset or topic to another is also a key factor (Gioia et al., 2000).

The Shift scale contains 10 items targeting different aspects of cognitive flexibility, including one's propensity to ruminate over thoughts, negative feedback, and general disappointment, as well as problems accepting new ideas for solving problems, transition between activities, response to unexpected changes, adjustment in new situations, and getting stuck on topics and activities (Gioia et al., 2000). In terms of its construct validity, the Shift scale yielded low, moderate, and high correlations with scales measuring ADHD symptoms. For example, the Shift scale on the Parent Form moderately correlated with the Hyperactivity-Impulsivity scale of the ADHD-Rating Scale-IV ($r = .59$, $p < .01$; DuPaul, Power, Anastopoulos, & Reid, 1998). Moderate correlations also were noted with the Attention Problems scale on parent and teacher forms of the Child Behavior Checklist ($.49 \leq r \leq .59$, $p < .01$, Achenbach, 1991) and Hyperactivity and Attention Problems scales on the parent and teacher forms of the Behavior Assessment System for Children ($.46 \leq r \leq .53$, $p < .01$, Reynolds & Kamphaus, 1992).

Across the clinical and normative samples, internal consistency specific to the Shift scale ranged from .81 to .91. Interrater reliability between parent and teacher ratings was overall moderate ($r = .32$), while the correlation between parent and teacher ratings on the Shift scale ($r = .15$) was lower. McCandless and O'Laughlin (2007) similarly found that between parent and teacher ratings, the correlations were not significant on the Shift scale. Test-retest reliability correlations ranged from .72 to .92, with correlations for the Shift scale ranging from .72 to .83. The BRIEF is considered a valid measure, evidenced by its content and construct validity (Gioia et al., 2000; McCandless & O'Laughlin, 2007).

Following a comprehensive literature survey, it was discovered that no study utilized the second edition of the BRIEF (BRIEF-2, Gioia, Isquith, Guy, & Kenworthy, 2015), which is why its predecessor is discussed here. From records initially retrieved, 83 percent ($n = 19$) of BRIEF studies reported ratings for BRIEF Parent Form, followed by 22 percent ($n = 5$) of studies reporting ratings for the BRIEF Teacher Form and 13 percent ($n = 3$) using the self-report form. As a result, only studies including parent ratings were meta-analyzed.

CHAPTER IV

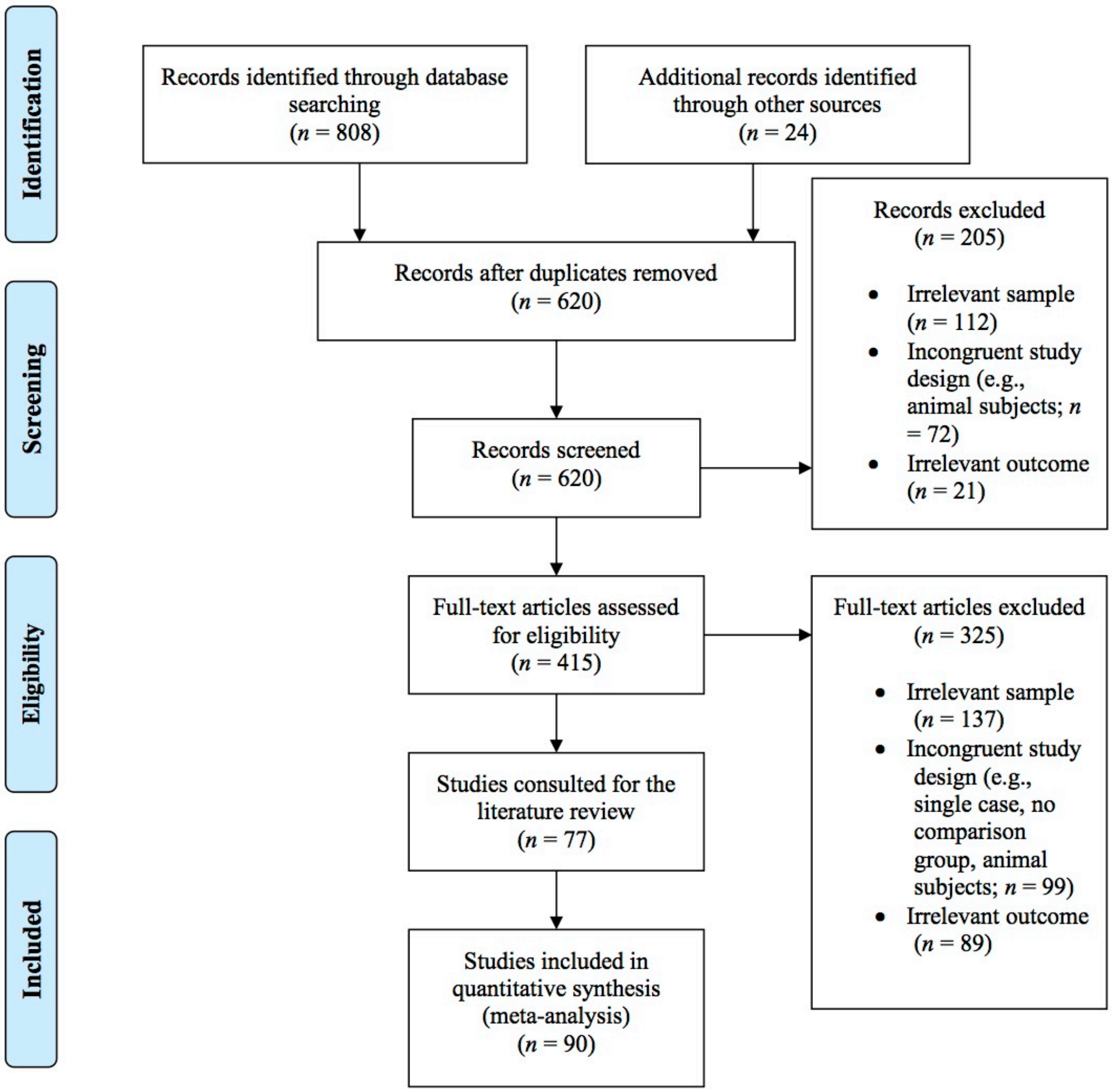
RESULTS

Across the two search databases, the search terms retrieved a total of 808 records (PsycINFO, *n* = 433; MEDLINE, *n* = 375). Figure 1 is an adapted version of the 2009 PRISMA Flow Diagram and shows the process by which the final study pool was developed (Liberati et al., 2009). Following the initial screening, duplicates were removed and the study pool was narrowed to 620 records. From that sample, 205 records were excluded and 415 records were screened to determine eligibility for full-text assessment. Of these, 325 records were excluded for one or more reasons indicated in Figure 1. In Figure 1, "Additional records identified through other sources" refers to studies that were retrieved through an ancestral search. An ancestral search refers to examination of reference lists within a given pool of related studies (Poirier & Behnen, 2014). The ancestral search was conducted across studies that met eligibility criteria following full-text assessment (*n* = 66) and yielded an additional 24 studies. In total, 90 studies met the eligibility criteria for the final analyses and 77 additional studies were consulted in authoring the literature review.

Description of the Studies

The publication timeline spanned from 1996 to 2018. Thirty studies were published prior to the meta-analysis by Willcutt et al. (2005) and 60 studies were published after Willcutt and colleagues' meta-analysis. Across studies, variables extracted differed. Table 5 provides information on the number of studies in the final sample for each variable.

Figure 1 Adapted from PRISMA 2009 Flow Diagram



Forty-three studies were conducted in regions within the United States and Canada, and 44 were studies conducted outside of North America. One study included a combined sample of participants from the United States and South Africa (Mattson et al., 2013) and two did not specify their locations (Kamradt, Ullsperger, & Nikolas, 2014; Lovejoy et al., 1999). In terms of setting, 34 studies were conducted in an academic medical setting and one study (Mataró, García-Sánchez, Junqué, Estévez-González, & Pujol, 1997) conducted their study in a juvenile detention setting. Additional settings included schools (*n* = 10), community centers (*n* = 16), private practices (*n* = 8), and universities (*n* = 13).

Table 5 Number Included Studies Based on Measure or Variable

Measure	Studies included for any variable*	Variable	# Studies included by variable
Wisconsin Card Sorting Test (WCST)	46	Perseverative Errors (PE)	33
		Number of Categories Completed (#CAT)	32
		Total Errors (TE)	19
		Nonperseverative Errors (NPE)	17
		Failure to Maintain Set (FMS)	17
		Total Correct (TC)	12
		Percent Correct (PR)	11
Trailmaking Test/Delis-Kaplan Executive Function System (TMT/D-KEFS)	47	Part B/Condition 4	42
		B-A	7
		B-A/Condition 4-2 (derived when A/Condition 2 and B/Condition 4 times were both provided)	21
		B/A and Condition 4/Condition 2 (derived when A/2 and B/4 times were both provided)	23
Behavior Rating Inventory of Executive Function (BRIEF)	18	Shift	18

Note. *21 studies contained multiple measures so this column will not add to the total studies (*N* = 90).

Study quality was defined according to assessment methods for ADHD, ways in which cognitive ability was accounted for, gender equality across clinical and control groups, statistical methods accounting for gender equality, processes for determining typicality of the control group, and standardization procedures for test administration. Ratings (-1, 0, or 1) were assigned to each factor, which summed into positive and negative scores, as well as scores of zero. To offset positive and negative scores, seven points were added to the total raw score for each study. Totals were transformed to scores ranging from 1 to 13 and were categorized as poor, moderate, or high quality. Scores ranging from 1 to 4 were categorized as poor quality; 5 to 8 were categorized as moderate quality; and 9 to 13 were categorized as high quality. These ranges were arbitrary and based on the author's judgement. The biggest challenge in assigning ratings to each domain was interpreting information included in the study, given that each one used different wording and language to describe their procedures (e.g., how IQ was handled).

Given these challenges (i.e., interpreting descriptions about methodology in studies), the overall quality for each study was evaluated using the intraclass correlation (ICC) statistic. Using SPSS (IBM Corp, 2017), the ICC estimate and its 95% confidence interval (CI) was computed based on a mean-rating ($k = 2$), absolute-agreement, and 2-way mixed-effects model (Koo & Li, 2016). Initially, the Average Measures variable (ICC = .77) 95% CI [.40, .91] indicated moderate reliability between both raters (Portney & Watkins, 2000).

Across the domains evaluating study quality, the two coders had varying levels of agreement, with more agreement in interpreting gender balance and whether standardization was maintained for test administration. In other areas, however, discrepancies were noted. Understanding descriptions of study designs and procedures was challenging, specifically the diagnostic processes for clinical and comparison groups. In other words, how did each study arrive at ADHD diagnoses and determine typicality of the control group? In addition, problems were noted in deciphering if and how IQ was controlled. Finally, the second coder noted difficulties in maintaining a clear understanding of how to assign scores for each factor. In some cases, the rationale for assigning a score was clear and recognizable, and in other cases, it was not. Disagreements regarding these ratings were resolved through discussion, thus resulting in 100% agreement between the coders (ICC = 1.00).

Study Sample

Effect sizes were derived by comparing the performance on measures of EF for clinical (ADHD) and healthy control (HC) groups. Seventy one percent of studies (n = 64) included clinical groups in which the subtype presentation of ADHD was either not specified or within a mixture of multiple subtype presentations. In other words, participants were collapsed into one group and labeled “ADHD.” Twenty six percent of studies (n = 23) specified the use of participants meeting diagnostic criteria for the combined presentation of ADHD (ADHD-C). A total of 17 studies (19% of the total sample) reported additional scores for groups presenting with the predominantly inattentive subtype of ADHD (ADHD-I), while two studies (Schmitz et al., 2002; Shuai, Chan, & Wang, 2011) included scores for groups presenting with the predominantly hyperactive/ impulsive subtype of ADHD (ADHD-H). In all cases, data was extracted for groups of participants identified as either ADHD-C or "ADHD." It is important to note that ADHD groups were defined according to the DSM nomenclature utilized for each study.

Using the mean age for each study, records were coded as "child," "adolescent," or "adult." The definitions for each age group were consistent with those outlined in prior studies (Antonini, Becker, Tamm, & Epstein, 2015; Kang, Han, Kim, Bae, & Renshaw, 2016; Keage et al., 2006; Stavro, Ettenhofer, & Nigg, 2007). For the total sample, mean ages for each group ranged from 7.95 to 37.35 years. Within child-focused studies, the mean ages ranged from 7.95 to 12.95. Studies targeting adolescents ranged from 13.18 to 16.00, and adult-focused studies ranged from 19.43 to 37.35. Table 6 outlines additional information regarding age of participants.

Table 6 Age of Participants

Range	*N*	%	ADHD mean age (*SD*)	HC mean age (*SD*)
Child	59	65.55	10.34 (1.89)	10.58 (1.97)
Adolescent	10	11.11	14.51 (1.34)	14.69 (1.33)
Adult	21	23.33	29.76 (8.33)	28.92 (7.58)

Note. *N* = number of studies; % = percent of studies from the total sample; ADHD = Attention-Deficit/Hyperactivity Disorder; HC = healthy controls; *SD* = standard deviation.

Table 7 shows the majority studies consisted primarily of male participants. Gender codes assigned to each study were based on the total percent of males between the ADHD and HC groups. The coding form was designed to extract race and ethnicity of participants; however, these characteristics were reported in only 31 studies (34.44% of the total sample). Within those studies, 25 described participants as Caucasian (27.78% of the total sample). Socioeconomic status of participants was not included in the initial coding form. Overall, informal observation indicated it was rarely specified.

Table 7 Gender Characteristics

Range	*N*	%
≤ 50% Male	13	14.44
> 50% Male	71	78.89
NS	6	6.67

Note. *N* = number of studies; % = percent of studies from the total sample; NS = not specified.

Table 8 indicates nearly half of studies contained relatively pure samples of ADHD. Alternatively, three studies (3.33%) included clinical samples in which more than 90% of participants met diagnostic criteria for a co-occurring condition in addition to ADHD. The most common co-occurring conditions were learning disorders, oppositional defiant disorder, anxiety, depression, and specific phobia. A small portion of the total sample included studies that did not specify information about comorbidity.

Table 8 Incidences of Comorbidity

Range	*N*	%
≤ 50% Comorbidity	53	58.89
> 50% Comorbidity	15	16.67
NS	22	24.44

Note. *N* = number of studies; % = percent of studies from the total sample; NS = not specified.

Participants' use of stimulant medication also was coded. Results indicated that 35 studies (38.89% of the total sample) included participants that took stimulant medication, but were not reportedly medicated during the experiment. Conversely, 10 studies allowed participants to continue their medication regimen throughout testing. Seventeen studies contained participants that were stimulant naïve, and 28 did not specify information regarding stimulant use.

Statistical Analyses

Effect Size Calculation

Using Microsoft Excel (2016), an effect size (ES) statistic was computed for each study. Effect size estimates reflect mean differences on measures of shifting between the ADHD and HC groups (Willcutt et al., 2005). Consistent with relevant meta-analyses (Alderson et al., 2007; Willcutt et al., 2005), Hedges' *g* (Hedges & Olkin, 1985) was calculated for all analyses involving effect size statistics. To obtain a mean effect for each outcome variable, studies were pooled together according to the variable being meta-analyzed. For most of the variables discussed herein, higher scores in the ADHD group indicated impairment and resulted in a positive effect size (Lai et al., 2017). Alternatively, variables such as the Number of Categories Completed (#CAT) and Total Correct (TC) of the WCST were expected to be lower for the ADHD group. Because effect size calculation involved subtracting the HC score from the ADHD score, the calculation of ES would be expected to yield a negative number and thus, indicative of impairment for these variables. Using standard interpretive metrics (Cohen, 1988), effect sizes are described as small ($g = \pm 0.20$), medium ($g = \pm 0.50$), or large ($g = \pm 0.80$).

Heterogeneity Analysis

Heterogeneity testing is a key step in determining the level of precision in the summary estimate and variability between the effects for each measure (Borenstein et al., 2009). Consistent with practices outlined in previous meta-analyses (e.g., Alderson et al., 2007; Kasper et al., 2012; Willcutt et al., 2005), the Q statistic was computed. Other options for reporting heterogeneity were considered, including use of the I^2 statistic. The I^2 statistic reflects the true proportion of dispersion that is not spurious and not dependent on the scale being used. It is not an indicator of the impact of dispersion (Borenstein et al., 2009). Alternatively, the Q statistic informs the extent to which the variance between study-level effects results from sampling error and the total variance. Q also informs whether the amount of heterogeneity is statistically significant (Higgins & Thompson, 2002; Higgins, Thompson, Deeks, & Altman, 2003).

The confidence interval (CI) is an index of precision and represents a 95% chance that the actual ES falls within a set of limits (Borenstein et al., 2009). CIs are specific to each study. Studies that resulted in statistically significant effects outside the CI were indicative of variability in samples across studies as well as grouping procedures (e.g., study groups based on age). Equation 1 shows the formula for the Q statistic:

Equation 1

$$Q = \sum (w \text{ x } ES^2) - \frac{[\sum (w \text{ x } ES)]^2}{\sum w}$$

Analysis of heterogeneity was performed using the macros program (Wilson, 2001) for SPSS (IBM Corp, 2013), which allowed for a better understanding of variability within and across moderator subgroups. A significant *Q* statistic suggests that the amount of variance across the effect sizes extends beyond what would be expected in the general population and cannot be explained by standard error alone. Alternatively, a nonsignificant *Q* statistic may suggest a small amount of dispersion across the effects, but it also may result from low power from an observed dispersion of imprecise studies (Borenstein et al., 2009).

Random Effects Model

Using the macros program (Wilson, 2001) for SPSS (IBM Corp, 2013), mean effects were computed and fitted within a random-effects model. Because ADHD is notably heterogeneous (e.g., Banaschewski et al., 2007; McAuley et al., 2014; Thapar et al., 2000), a random-effects model was useful in accounting for the expected variability between the effects. Further, a random-effects approach assumes that results are generalizable beyond the included studies (Borenstein, Hedges, Higgins, & Rothstein, 2010; Cheung & Vijayakumar, 2016). This assumption requires an estimate of the variance between studies; therefore, the method of moments was computed (Borenstein et al., 2009; DerSimonian & Laird, 1986; McKenzie et al., 2016). The method of moments assumes the true effect is confounded by multiple factors and thus, results in variability within a given distribution of effects (Del Re & Flückiger, 2016; DerSimonian & Laird, 1986). The between-study variance is reflected by the random effects variance component (*v*) and estimates the amount of heterogeneity within the general population, or the amount of *true* heterogeneity (Borenstein et al., 2009). Computing the variance component involves the overall heterogeneity and sum of the weights for each study as shown in Equation 2:

Equation 2

$$\hat{v}_\theta = \frac{Q_T - k - 1}{\sum w - \left(\frac{\sum w^2}{\sum w}\right)}$$

One key advantage to using Hedges' *g* is that it pools the results by applying weights to all the studies involved (Cheung & Vijayakumar, 2016). Weights function as a correction for bias that can result from with small sample sizes (McKenzie et al., 2016). Typically, studies more heavily weighted than others have more standard error and smaller variance, and are assumed to be more accurate estimates of the "true" population (Del Re & Flückiger, 2016; Hedges & Olkin, 1985). Consistent with Martinussen et al. (2005), the inverse-variance method was used to better account for all the information and affective variables from each study (Hedges & Vevea, 1998). With this method, the applied weights reflect the inverted variance of the study's ES (Lipsey & Wilson, 2001; McKenzie et al., 2016), while accounting for the between-study variance indicated by the variance component. Equation 3 shows the formula for computing the inverse-variance weight:

Equation 3

$$w = \frac{1}{se^2 + \hat{v}_\theta}$$

Publication Bias

Studies were restricted to those yielding empirical research with quantitative findings (Lipsey & Wilson, 2001). To avoid the "file drawer problem" and potential for inflated findings, a funnel plot was constructed for all outcome variables (Banks et al., 2012; Peters, Sutton, Jones, Abrams, & Rushton, 2008; Rothstein, Sutton, & Borenstein, 2005). Estimating variance between studies was addressed using the method of moments because it does not assume anything about the distribution of the random effects (Borenstein 2009).

The trim-and-fill method (Duval & Tweedie, 2000) also was incorporated into funnel plot analysis, which adjusts initial meta-analytic findings, thus allowing for examination of the potential impact of missing studies. Although the adjustment takes place over multiple iterations and theoretically adds studies to the sample, the intention is not to interpret imputed results. Rather, it is typically referred to as a "sensitivity" analysis (Rothstein et al., 2005, p. 127) wherein results inform the degree to which effects are sensitive to publication bias (i.e., missing studies). All funnel plot and trimming analyses were performed using STATA (StataCorp, 2017).

Meta-Regression Analysis

Meta-regression analysis was performed to explain possible changes in the relationship between ADHD and shifting across studies (Hedges & Pigott, 2004). It is useful for assessing the extent to which each moderator impacts outcomes. Prior research suggests variables such as cognitive ability (IQ), diagnostic procedures, sampling procedures, and age (Alderson et al., 2007; Boonstra et al., 2005; Denckla, 1996; Willcutt et al., 2005) moderate the strength and direction of the relationship between ADHD and EF. The present study utilized mean age of participants (e.g., children, adolescents, and adults), as well as the covariance of IQ, study quality, and test version (WCST only) as potential moderators of the effects.

The decision to include meta-regression in each analysis assumed that the selection of moderators was a mixture of categorical and continuous variables (Banks et al., 2012). Analyses were performed for variables that resulted in significant heterogeneity (Q) across their effects. To explore moderating effects of multiple covariates, it is essential to utilize the Q statistic over alternative options like the Z statistic (Borenstein et al., 2009). In addition, use of the Q statistic afforded consistency in reporting of the remaining analyses of heterogeneity.

For a fixed-effect model, study weights are based on within-study error and the total variance (Q or Q_{total}) can be partitioned into Q_{model} and Q_{resid}. These are synonymous to the between-study variance (Q_{bet}) and within-study variance (Q_{within}), respectively. Under a random effects model, however, study weights involve between-study variance that varies across each moderator analysis. In other words, the total variance is different for each moderator and thus, between-study variance and within-study variance are not additive components (Borenstein et al., 2009). For this study, Q_{model} and Q_{resid} are reported. Q_{model} reflects "the dispersion explained by the covariates" and Q_{resid} refers to "the distance of studies from the regression line" (Borenstein et al., 2009, "Assessing the Impact of the Slope," para. 4). In addition, the sum of variability within each subgroup (Q_{within}) is equal to Q_{resid}. In other words, Q_{within} for child, adolescent, and adult studies is equal to Q_{resid} for age.

The regression coefficient for each covariate (*B*) reflects a degree of latitude that corresponds to an increase or decrease of *B* units in effect size. For correlation coefficient, ESs are described as small ($B = \pm 0.1$), medium ($B = \pm 0.3$), and large ($B = \pm 0.5$; Del Re & Flückiger, 2016). When Q_{model} is statistically significant, it means the strength of the relationship between B, or latitude, and the score being meta-regressed exceeds what would be explained by chance alone. Statistical significance for Q_{resid} indicates that, although the moderator explains some of the observed variance in effects, a portion of the variance is unexplained (Borenstein et al., 2009). As part of this analysis, the method of moments random effects variance component also was computed (*v*). This informs the degree to which variability across the effects is explained by sampling error as well as the amount of absolute variation in a given scale. When $v > 0.00$, it suggests the variance cannot be explained by sampling error alone.

Research Question 1

Compared to healthy controls, do groups of participants identified as having ADHD yield lower outcomes on measures of shifting? It is hypothesized that shifting abilities are consistently deficient among individuals with ADHD (Pennington & Ozonoff 1996; Willcutt et al., 2005) regardless of measure. To test this hypothesis, the analyses described above were conducted for each of the three measures of interest.

The Wisconsin Card Sort Test (WCST)

Initially, 51 studies included some or all variables of the WCST. Variables reported across the studies were tallied to determine if their individual frequencies would yield a sufficient sample needed for effects size calculation (Lipsey & Wilson, 2001) and funnel plot analysis (Egger, 1997; Higgins & Green, 2011; Sterne et al., 2011). Based on the survey of literature, it was revealed that Perseverative Errors (PE) was the most reported variable across studies, totaling 35. Thirty-three of those studies met inclusion criteria for the current analysis. Conversely, Set Loss Maintenance was the least reported variable in studies ($n = 3$). Like Set Loss Maintenance, additional variables were idiosyncratic in their use across studies. Examples included the Percentage of Perseverative Errors ($n = 5$), Percentage of Conceptual Level Response ($n = 4$), Percentage of Perseverative Responses ($n = 4$), and Trials to Complete the First Category ($n = 8$). Because of their low frequency, these variables were removed from the analyses. In doing so, the sample was reduced to 46 records that used some version of the WCST (e.g., Heaton, 1981) and met full inclusion criteria for the present study.

Appendix B shows all the individual effect sizes specific to ADHD and HC groups for each variable. Once the individual study ES were calculated by variable, the mean ES and CI were calculated for each variable. Studies were weighted according to their sample size and ES. For the WCST, mean results are indicated by Hedges' *g* with the 95th upper and lower limits (CI) and presented in Table 9. The total sample sizes for each comparison group, according to the variables, as well as heterogeneity statistic and level of significance is also provided.

Table 9 Summary Table of Mean Effect Sizes for WCST Variables

Variable	*N* ADHD	*N* HC	Mean *g* (*SE*)	95% CI	*Q*	*p*-value
PE	1606	1674	0.44** (0.04)	0.25, 0.40	125.01**	< .001
#CAT	1846	1611	-0.54** (0.08)	-0.69, -0.38	126.58**	< .001
TE	794	643	0.45* (0.18)	0.10, 0.80	164.18**	< .001
NPE	975	929	0.27** (0.05)	0.17, 0.38	18.30	0.31
FMS	947	858	0.27** (0.09)	0.09, 0.44	41.72**	< .001
TC	763	553	-0.17 (0.14)	-0.45, 0.10	60.59**	< .001
PR	542	375	0.14 (0.26)	-0.37, 0.65	104.14**	< .001

Note. *WCST* = Wisconsin Card Sorting Test; *PE* = perseverate errors; *#CAT* = number of categories completed; *TE* = total errors; *NPE* = nonperseverative errors; *FMS* = failure to maintain set; *TC* = total correct; *PR* = perseverative responses; *N* ADHD = total sample for participants with attention-deficit/hyperactivity disorder; *N* HC = total sample for healthy control participants; *g* = Hedges' *g*; *SE* = standard error; *CI* = confidence interval; *Q* = heterogeneity statistic; *p* = significance for heterogeneity statistic. * $p < .05$. ** $p < .01$.

Perseverative Errors (PE)

The mean ES was small and statistically significant. Heterogeneity of studies using the PE score reached statistical significance ($Q = 125.01$, $df = 32$). The funnel plot in Figure 2 shows the dispersion of study effects specific to the PE score. To better understand reasons for outlier effects (i.e., effects that deviate largely from the population parameters), individual study characteristics were informally explored and considered. In the case of PE, one potential reason for outlying effects is the disproportionate groups based on gender. In the study by Tsuchiya et al. (2005), the male to female ratio in the ADHD sample was 20:2, whereas the male to female ratio in the HC group was 13:12. Additional characteristics potentially contributing to disparate effects include failure to report IQ scores (e.g., Silva et al., 2014), and imbalanced ADHD and HC group size (e.g., Stavro et al., 2007).

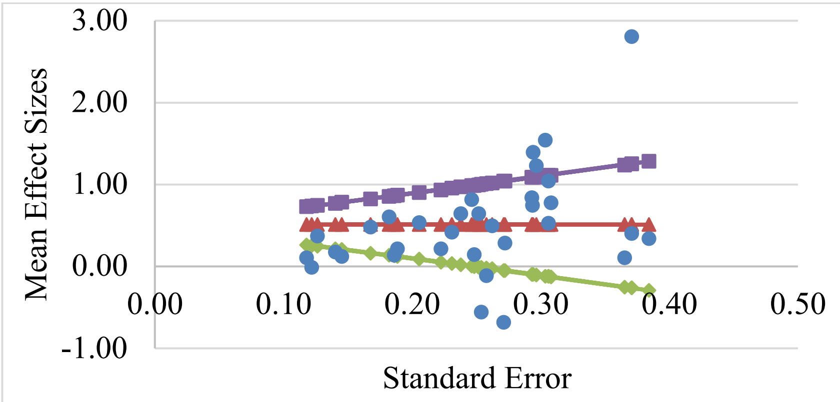
One study, Solanto et al. (2007), yielded a mean ES that substantially deviated from the remaining ESs and was consistently an outlier for all variables utilized from that study. This may have resulted from significant differences in IQ between the ADHD and HC groups ($p < .05$). In addition, sampling procedures for this study strictly defined the HC group as individuals without inattentive symptoms. This is important to note, given that results for the PE score, as well as other scores utilized by Solanto and colleagues may be a function of the sample characteristics.

Results did not reveal any substantial changes to the data following the trimming estimation and filling process. In other words, heterogeneity for PE neither gained nor lost statistical significance following the process and the data remained unchanged.

Heterogeneity statistics were computed for the moderators and subgroups (e.g., child versus adult studies), as shown in Table 10. Mean ESs were moderate and statistically significant for samples of child studies, as well as studies that did not covary IQ, those with poor and moderate study quality, and studies using the modified version of the WCST (64 cards).

For the Perseverative Errors study pool, some of the effects were small. For example, studies that covaried for IQ yielded small effect sizes. This was also the case for studies with high quality, as well as for studies using the 128-card test version. Despite effects being small, all except the pool of adult studies were statistically significant. Four of the effects were significant at the .01-level and three of the effects were significant at the .05-level.

Figure 2 Funnel Plot for Perseverative Errors (PE)



Variability reached statistical significance for samples of child studies and studies with high study quality.

Table 10 Heterogeneity Analysis for Potential Moderators of PE

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
Age	Child (<18 years)	0.51** (0.09)	23	0.33, 0.68	38.31*	.02
	Adult (>18 years)	0.26 (0.14)	8	-0.01, 0.53	9.80	.28
IQ	Covaried	0.35* (0.14)	8	0.06, 0.64	10.99	.20
	Not covaried	0.48** (0.09)	23	0.30, 0.66	34.78	.05
Study Quality	Poor	0.63* (0.24)	2	0.16, 1.11	5.17	.07
	Moderate	0.59* (0.13)	10	0.33, 0.84	4.96	.89
	High	0.32* (0.10)	18	0.13, 0.51	37.07**	<.01
Test version	64 cards	0.56** (0.14)	9	0.28, 0.85	16.92	.05
	128 cards	0.39** (0.09)	22	0.21, 0.57	29.30	.14

Note. *PE* = perseverative errors; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic. * $p < .05$. ** $p < .01$.

Table 11 shows results from the meta-regression analysis, indicating heterogeneity within the study pool could not be explained by any of the proposed moderators. That said, the method of moments random effects variance component (*v*) was greater than zero, meaning variability cannot be explained by sampling error alone. Further, within-group variability was statistically significant for each age, IQ, study quality, and test version, suggesting observed variance across the effects is likely explained by alternative considerations.

Table 11 Meta-Regression Results for PE

Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
Age	-0.12	0.08	-0.29, 0.04	48.11*	.03	.12
IQ	0.13	0.17	-0.21, 0.48	45.77*	.04	.14
Study quality	-0.20	0.11	-0.42, 0.02	47.77*	.03	.12
Test version	0.17	0.17	-0.17, 0.52	45.25*	.04	.14

Note. PE= perseverative errors; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval; Q_{resid} = within-group heterogeneity; *p* = significance for heterogeneity statistic; *v* = method of moments random effects variance component. * $p < .05$.

Number of Categories Completed (#CAT)

The mean ES was medium and statistically significant. Heterogeneity for studies using the #CAT score reached statistical significance. The funnel plot in Figure 3 shows the magnitude of dispersion across study effects. Again, the mean ES for Solanto et al. (2007), as well as Shin, Choi, Kim, Hwang, Kim, and Cho (2008), emerged as outliers in relation to the population parameters.

The latter study consisted of participant groups with statistically significant differences in IQ and age ($p < .01$), which may have magnified performance differences between the ADHD and HC groups. Following trim-and-fill analysis, six iterations were performed and yielded similar results. In other words, overall heterogeneity remained significant.

Figure 3 Funnel Plot for Number of Categories Completed (#CAT)

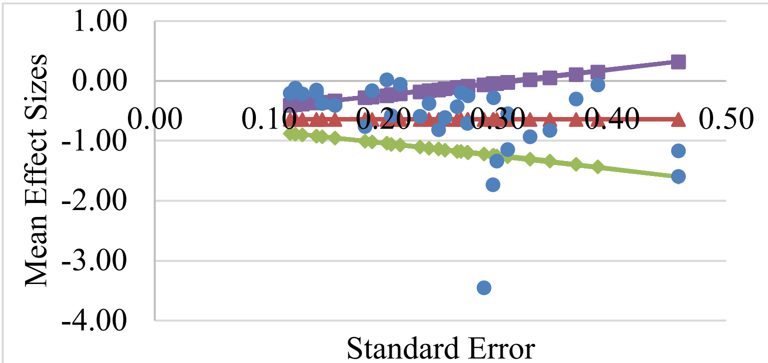


Table 12 shows results of heterogeneity testing. Medium effects for child and adolescent studies as well as for studies that covaried IQ were found. Studies within both subgroups for test version resulted in medium ESs. These effects were statistically significant. Mean effects for studies that did not covary IQ, as well as studies with high quality were small, but statistically significant. Variability was statistically significant for subgroups of child studies, studies that covaried IQ, studies with high quality, and studies that used the 128-card test version.

Table 12 Heterogeneity Analysis for Potential Moderators of #CAT

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
Age	Child	-0.74** (0.09)	18	-0.92, -0.56	41.93**	<.01
	Adolescent	-0.55* (0.24)	3	-1.01, -0.08	5.16	.16
	Adult	-0.17 (0.12)	8	-0.40, 0.07	0.86	.10
IQ	Covaried	-0.72** (0.16)	7	-1.04, -0.40	32.23**	<.001
	Not covaried	-0.49** (0.09)	23	-0.67, -0.30	17.78	.77
Study quality	Poor	-0.49 (0.27)	1	-1.01, 0.04	0.10	.75
	Moderate	-0.81** (0.13)	11	-1.07, -0.55	12.71	.31
	High	-0.39** (0.09)	17	-0.58, -0.21	38.84**	<.01
Test version	64 cards	-0.50** (0.17)	6	-0.84, -0.17	2.65	.85
	128 cards	-0.55** (0.09)	24	-0.73, -0.37	48.79**	<.01

Table 12 Continued

Note. *#CAT* = number of categories completed; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic. * $p < .05$. ** $p < .01$.

Results from the meta-regression, shown in Table 13, indicate the relationship between latitude of age and the #CAT score is stronger that what would be "expected by chance" (Borenstein et al., 2009, "Assessing the Impact of the Slope," para. 4). The is indicated by between-study variance (Q_{model} = 0.03, *df* = 1). The variance component also indicated, despite age explaining a portion of the variance, there is more within the sample than would be expected based on sampling error ($p < .01$).

Table 13 Meta-Regression Results for #CAT

Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
Age	0.28**	0.07	0.14, 0.43	48.70*	.02	0.09
IQ	0.23	0.19	-0.13, 0.60	50.00*	.01	0.14
Study quality	0.20	0.12	-0.04, 0.44	52.37**	< .01	0.12
Test version	0.05	0.19	-0.33, 0.43	51.44**	< .01	0.14

Note. *#CAT* = number of categories completed; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval; Q_{resid} = within-group heterogeneity; *p* = significance for heterogeneity statistic; *v* = method of moments random effects variance component. * $p < .05$. ** $p < .01$.

Total Errors (TE)

The main effect for TE was small and statistically significant. Heterogeneity of effects was statistically significant. The funnel plot in Figure 4 shows the magnitude of dispersion across study effects. Again, the mean ES for Solanto et al. (2007) emerged as an outlier in relation to the remaining effects. Trim-and-fill analysis performed five iterations and yielded statistically significant heterogeneity for the sample. Thus, results did not differ substantially from the initial calculations.

Figure 4 Funnel Plot for Total Errors (TE)

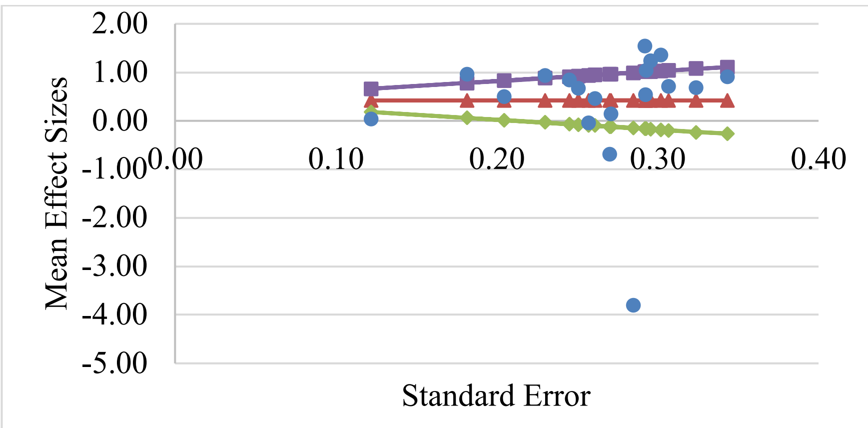


Table 14 shows across subgroups of moderators, the mean ES for child studies, as well as those that did not covary IQ and those that utilized the 64-card test version yielded medium effects that were statistically significant. The pool of studies with moderate quality resulted in a large and statistically significant mean effect. Heterogeneity was statistically significant for child studies and studies utilizing IQ as a covariate. Additionally, studies of high quality and those using the 128-card test were significantly heterogeneous.

Table 15 shows results from the meta-regression, suggesting the covariance of IQ is significantly related to the TE score. The strength of the relationship extends beyond what would be expected by chance. This is evidenced by the between-study variance (Q_{model} = 5.28, df = 1).

Table 14 Heterogeneity Analysis for Potential Moderators of TE

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
Age	Child (<18 years)	0.45** (0.21)	14	0.04, 0.86	30.35*	.01
	Adult (>18 years)	0.45 (0.40)	3	-0.33, 1.23	1.90	.59
IQ	Covaried	-0.24 (0.35)	4	-0.93, 0.44	26.21**	<.001
	Not covaried	0.69* (0.20)	13	0.29, 1.09	3.50	.10
Study quality	Poor	0.90 (0.55)	1	-0.18, 1.98	0.61	.43
	Moderate	0.83* (0.32)	5	0.20, 1.46	2.14	.83
	High	0.16 (0.24)	10	-0.30, 0.63	27.40**	<.01
Test version	64 cards	0.65** (0.32)	5	0.02, 1.29	.13	.10
	128 cards	0.36 (0.22)	12	-0.07, 0.79	32.25**	<.01

Note. *TE* = total errors; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic. * $p < .05$. ** $p < .01$.

Given within-study variability was significant for all potential moderators, between-study variability is explained by alternative factors ($p < .05$).

Table 15 Meta-Regression Results for TE

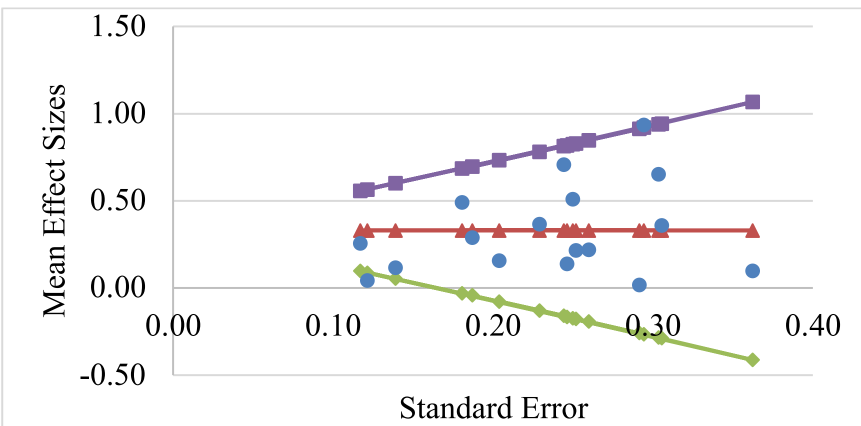
Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
Age	< -0.01	0.22	-0.44, 0.44	32.25*	.01	0.56
IQ	0.93*	0.40	0.14, 1.73	29.72*	.03	0.51
Study quality	-0.46	0.26	-0.97, 0.05	31.73*	.02	0.51
Test version	0.30	0.39	-0.47, 1.06	32.38*	.01	0.55

Note. *TE* = total errors; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval; Q_{resid} = within-group heterogeneity; *p* = significance for heterogeneity statistic; *v* = method of moments random effects variance component.
* $p < .05$.

Nonperseverative Errors (NPE)

The mean ES for NPE was small and statistically significant. The distribution of effects was relatively homogeneous, indicated by the non-significant *Q*-statistic. The funnel plot in Figure 5 shows the effects lie within the population parameters. As such, neither heterogeneity analysis nor meta-regression was performed on studies using the NPE. Similarly, trim-and-fill analysis was not performed.

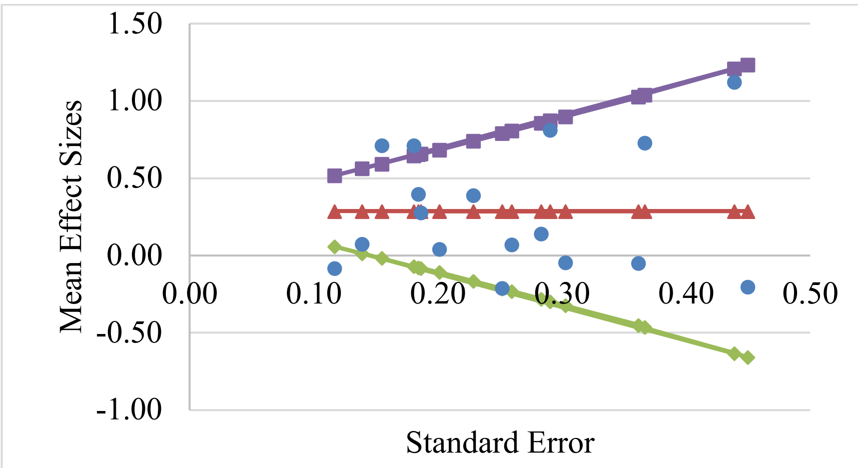
Figure 5 Funnel Plot for Nonperseverative Errors (NPE)



Failure to Maintain Set (FMS)

The mean ES was small and statistically significant. Total heterogeneity was also significant. Using the funnel plot in Figure 6, further investigation of studies with effects outside the population parameters revealed factors such as age (e.g., Beadle, 2010; Semrud-Clikeman, Steingard, Filipek, Biederman, Bekken, & Renshaw, 2000) and IQ (e.g., Shin et al., 2008) differences across participant groups may have contributed to the wide dispersion in effects. Trim-and-fill analysis performed two iterations and resulted in significant heterogeneity ($p < .001$), consistent with the initial calculation.

Figure 6 Funnel Plot for Failure to Maintain Set (FMS)



Results in Table 16 show the mean ES for studies with poor study quality was medium and statistically significant. Child studies, as well as those that did not covary IQ, those with moderate quality, and those using the 128-card test resulted in small, nonetheless, significant effects. Subgroups of studies were relatively homogeneous, indicated by their non-significant Q_w.

Table 16 Heterogeneity Analysis for Potential Moderators of FMS

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
Age	Child	0.39** (0.10)	9	0.18, 0.59	10.23	.33
	Adolescent	0.20 (0.21)	2	-0.22, 0.62	3.98	.14
	Adult	0.05 (0.15)	3	-0.25, 0.35	0.93	.82
IQ	Covaried	0.37 (0.20)	3	-0.02, 0.75	2.82	.42
	Not covaried	0.24** (0.10)	12	0.04, 0.45	11.23	.51

Table 16 Continued

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
Study quality	Poor	0.57** (0.19)	1	0.18, 0.95	0.65	.42
	Moderate	0.30* (0.15)	5	0.01, 0.60	7.02	.22
	High	0.16 (0.10)	8	-0.04, 0.37	8.35	.40
Test version	64 cards	0.24 (0.27)	14	-0.30, 0.78	0.61	.43
	128 cards	0.27** (0.10)	1	0.08, 0.47	14.06	.45

Note. *FMS* = failure to maintain set; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic. * $p < .05$. ** $p < .01$.

Results in Table 17 show that none of the selected moderators yielded significant effects. Within-group variability for each moderator was relatively homogeneous as indicated by their non-significant Q_{resid}.

Table 17 Meta-Regression Results for FMS

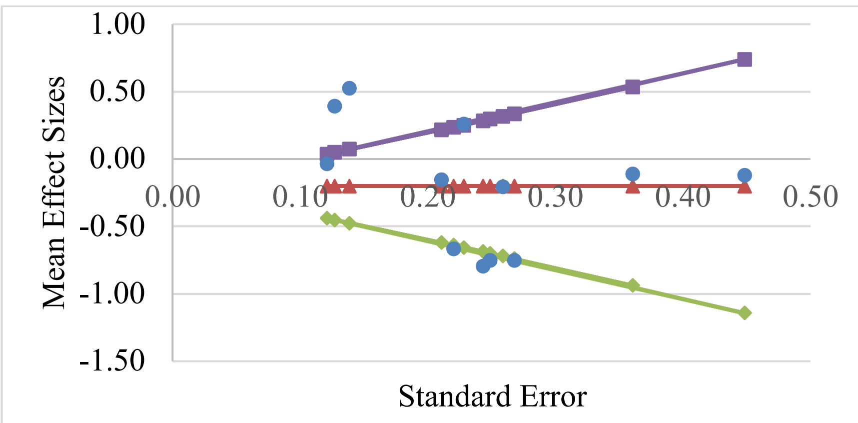
Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
Age	-0.17	0.09	-0.34, -0.003	16.35	.36	0.04
IQ	-0.12	0.22	-0.56, 0.32	14.05	.52	0.09
Study quality	-0.19	0.10	-0.39, 0.001	16.98	.32	0.04
Test version	-0.03	0.29	-0.60, 0.53	14.67	.47	0.08

Note. *FMS* = failure to maintain set; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval; Q_{resid} = within-group heterogeneity; *p* = significance for heterogeneity statistic; *v* = method of moments random effects variance component. * $p < .05$.

Total Correct (TC)

Results yielded a small effect for TC and significantly heterogeneous dispersion of study effects. Like the #CAT score, impairment was indicated by a negative ES. Using the funnel plot in Figure 7, two outlying effects were explored. In both studies (e.g., Faraone et al., 2006; Surman, Biederman, Spencer, Miller, Petty, & Faraone, 2015), the ADHD group performed significantly better than the HC group ($p < .05$). Trim-and-fill analysis indicated no trimming was performed and the data remained unchanged.

Figure 7 Funnel Plot for Total Correct (TC)



Regarding potential moderators, only one study using the TC score covaried IQ (e.g., Tripp, Ryan, & Peace, 2002); therefore, this analysis was not completed. For study quality, raw scores resulted in either moderate or high quality, meaning there were only two subgroups for this variable. Table 18 shows that child studies as well as those using the modified test version yielded medium heterogeneity effects that were statistically significant. Overall, subgroups were relatively homogenous as indicated by their non-significant Q_w.

Results in Table 19 highlight that both age and test version have statistically significant relationships with the TC score, evidenced by between-study variance that is statistically significant. For age (Q_{model} = 7.77, *df* = 1), this suggests that as people get older, they are likely to obtain more items correct.

Table 18 Heterogeneity Analysis for Potential Moderators of TC

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
Age	Child (<18 years)	-0.45* (0.15)	6	-0.74, -0.15	6.62	.36
	Adult (>18 years)	0.16 (0.16)	4	-0.15, 0.47	3.19	.53
Study quality	Moderate	-0.15 (0.33)	2	-0.79, 0.49	0.02	.99
	High	-0.18 (0.16)	8	-0.50, 0.13	9.66	.29
Test version	64 cards	-0.51* (0.18)	4	-0.86, -0.16	2.70	.61
	128 cards	0.07 (0.14)	6	-0.21, 0.35	7.71	.26

Note. *TC* = total correct; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic. * $p < .05$.

For test version, the slope is negative only for the smaller card deck and suggests going from a smaller card deck to a larger card deck increases the likelihood that ADHD have less correct (lower TC) than HC participants (Q_{model} = 6.56, *df* = 1; *p*'s < .05). It could indicate that sensitivity to impairment increases with the smaller deck (e.g., learning takes place with more opportunities to have correct responses decreasing the sensitivity to impairment).

Table 19 Meta-Regression Results for TC

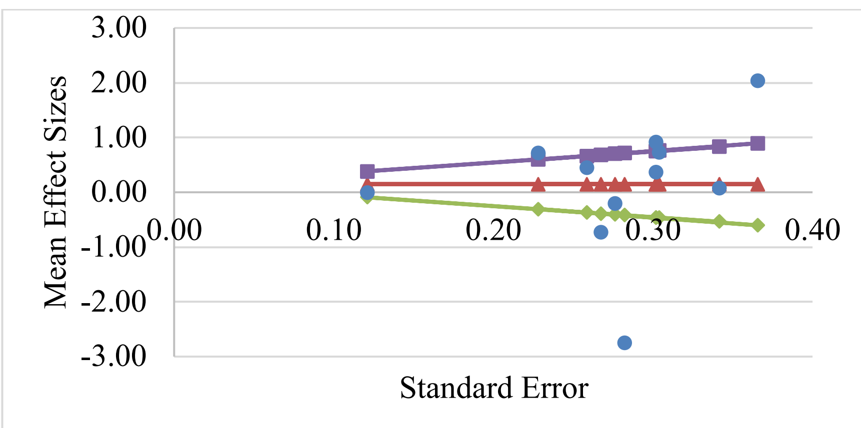
Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
Age	0.30**	0.11	0.09, 0.52	9.81	.46	0.09
IQ	0.63	0.51	-0.38, 1.63	9.41	.49	0.17
Study quality	-0.03	0.36	-0.74, 0.68	9.68	.47	0.19
Test version	-0.58*	0.23	-1.03, -0.17	10.40	.40	0.09

Note. *TC* = total correct; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval; Q_{resid} = within-group heterogeneity; *p* = significance for heterogeneity statistic; *v* = method of moments random effects variance component.
* $p < .05$. ** $p < .01$.

Perseverative Responses (PR)

The mean effect was small and statistically significant. Total heterogeneity was also significant. The funnel plot in Figure 8 shows two effects that deviated substantially from the population parameters (Kang et al., 2016; Solanto et al., 2007). Possible contributing factors in Solanto et al. have been discussed previously; however, considerations specific to Kang et al. have not. This study used a modified version of the test (64-cards). Results from the meta-regression analysis are discussed below and may help to explain why this moderator may have contributed to outlying effects. Trim-and-fill analysis performed five iterations and heterogeneity remained significant.

Figure 8 Funnel Plot for Perseverative Responses (PR)



The majority of studies using the PR variable were limited to participants ages six through 12. One study (e.g., Schmitz et al., 2002) used adolescent-age participants and another study (e.g., Silva et al., 2014) included adult participants. As a result, subgroup analysis for age is not reported. Like the TC score, these studies only yielded raw scores for study quality in the moderate or high range. Table 20 shows that, across the subgroups explored, studies using the 128-card version and those that did not covary IQ resulted in statistically significant small and medium effects, respectively. Although the subgroup of studies that covaried IQ yielded a significant and large effect, results for mean effects and variability are not meaningful because of the sample size ($k = 2$). Variability for studies with high quality was statistically significant.

Table 20 Heterogeneity Analysis for Potential Moderators of PR

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
IQ	Covaried	-1.65* (0.47)	1	-2.58, -0.73	4.58*	.03
	Not covaried	0.52** (0.21)	8	0.09, 0.94	7.48	.49
Study quality	Moderate	0.38 (0.45)	3	-0.51, 1.26	0.77	.86
	High	<0.01 (0.34)	6	-0.67, 0.68	14.97*	.02
Test version	64 cards	1.34 (0.60)	1	0.16, 2.51	1.18	.28
	128 cards	-0.10** (0.27)	8	-0.63, 0.43	13.76	.09

Note. *PR* = perseverative responses; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic. * $p < .05$. ** $p < .01$.

Table 21 shows that IQ ($Q_{model} = 17.65$, $df = 1$) and test version ($Q_{model} = 4.78$, $df = 1$) were significantly related to the PR score. With regards to each moderator, groups were relatively homogeneous as indicated by non-significant within-group variability (Q_{resid}).

Table 21 Meta-Regression Results for PR

Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
IQ	2.17**	0.52	1.16, 3.19	12.05	.21	0.33

Table 21 Continued

Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
Study quality	-0.37	0.57	-1.49, 0.74	15.73	.07	0.73
Test version	1.44*	0.66	0.15, 2.73	14.94	.09	0.57

Note. *PR* = perseverative responses; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval; Q_{resid} = within-group heterogeneity; *p* = significance for heterogeneity statistic; *v* = method of moments random effects variance component. * $p < .05$. ** $p < .01$

The Trailmaking Test/ Delis-Kaplan Executive Functioning Scale (TMT/D-KEFS)

Forty-seven studies using the TMT or D-KEFS were identified, and 39 of those studies used raw scores to inform set shifting. Thirty-six studies used the performance on Trails B and six studies used D-KEFS Condition 4 as a variable of set shifting. Within the TMT subsample, the time it took participants to complete Trails B (in raw score format) was expressed in seconds in 37 studies. One study used *z*-scores (Shanahan et al., 2006), one study reported *T*-scores (Tripp, Ryan, & Peace, 2002), one study reported standard scores (Davidson et al., 2016), and one study included log-transformed data (Tripp, Luk, Schaughency, & Singh, 1999). Within the D-KEFS subsample, three studies reported scaled scores (e.g., Holmes et al., 2010), one study used standard scores (Tafazoli et al., 2013), and one study reported raw scores (Kamradt et al., 2014). For this reason, using the same effect size statistic across a pool of studies was helpful in meta-analyzing a combination of scores with different scaling properties (Lipsey & Wilson, 2001). In doing so, this study remained consistent with previous meta-analyses (Lai et al. 2017; Willcutt et al., 2005).

Initially, the search strategy resulted in six studies that reported the TMT B-A difference. Qian et al. (2010) also reported the raw score time difference, which was a study obtained through additional searches in those initially identified for inclusion. For studies in which the individual raw score time (in seconds) for TMT-A/D-KEFS Condition 2 and TMT B/D-KEFS 4 were both reported, the B-A/4-2 difference, B/A ratio, Condition 4/Condition 2 ratio, and respective *SD*s were manually calculated. Using Microsoft Excel (2016), the time difference and ratio were manually calculated for 21 and 23 studies, respectively. *SD*s were derived from the CI of each group's sample size as well as its upper and lower limits. Using these variables, the square root of the group's sample size was multiplied by the ratio of the difference between the upper and lower limits and the confidence interval (Higgins & Green, 2011).

In cases where raw score times for TMT-A/D-KEFS Condition 2 were not reported, authors were contacted through e-mail to obtain additional data. Efforts to contact authors from 15 studies yielded additional data for one study (Perugini, Harvey, Lovejoy, Sanstrom, & Webb, 2000). Three authors indicated they did not have access to the original data and 11 did not reply. Appendix C shows a sample e-mail message that was sent to authors. For the TMT/D-KEFS variables, mean results are indicated by Hedges' *g* with the 95^{th} upper and lower limits (CI) and presented in Table 22. The total sample sizes for each comparison group, according to the variables, as well as heterogeneity statistic and level of significance for the *Q*-statistic also are provided.

Table 22 Summary Table of Mean Effect Sizes for TMT/D-KEFS Variables

Variable	*N* ADHD	*N* HC	Mean *g* (*SE*)	95% CI	*Q*	*p*-value
B/4	2160	2306	0.37** (0.07)	0.23, 0.52	200.70**	< .0001
B-A/4-2	1526	1186	1.07** (0.15)	0.77, 1.36	322.37**	< .0001
B/A and 4/2	989	899	0.04 (0.05)	-0.05, 0.13	4.11	1.00

Note. *TMT/D-KEFS* = Trailmaking Test/Delis-Kaplan Executive Function System; B/4 = Part B/Condition 4; B-A/4-2 = Part B-Part A/Condition 4-Condition 2; B/A and 4/2 = Part B/Part A and Condition 4/Condition 2; *N* ADHD = total sample for participants with attention-deficit/hyperactivity disorder; *N* HC = total sample for healthy control participants; *g* = Hedges' *g*; *SE* = standard error; *CI* = confidence interval; *Q* = heterogeneity statistic; *p* = significance for heterogeneity statistic. * $p < .05$. ** $p < .01$.

Trailmaking Test Part B/D-KEFS Condition 4 (TMT-B/D-KEFS 4)

The mean effect was small and statistically significant. The distribution of effects was significantly heterogeneous. To determine possible reasons for outlying effects, as shown in Figure 9, studies with effects outside the population parameters were explored further. Possible explanatory factors include significant differences in IQ between the comparison groups (e.g., Sartory, Heine, Müller, & Elvermann-Hallner, 2002) and disproportionate samples according to gender (e.g., Mattson et al., 2013). In several studies, mean scores yielded large *SD*s, suggesting noteworthy dispersion of scores at the participant level. Because of the role *SD* plays in ES calculation (Hedges & Olkin, 1985) and because raw score times can numerically appear large, which is how most TMT/D-KEFS scores are reported, it is possible these factors can result in larger, more disparate effects. Trim-and-fill analysis performed seven iterations and the heterogeneity of effects remained significant.

Figure 9 Funnel Plot for Trailmaking Test Part B/D-KEFS Condition 4 (TMT-B/D-KEFS 4)

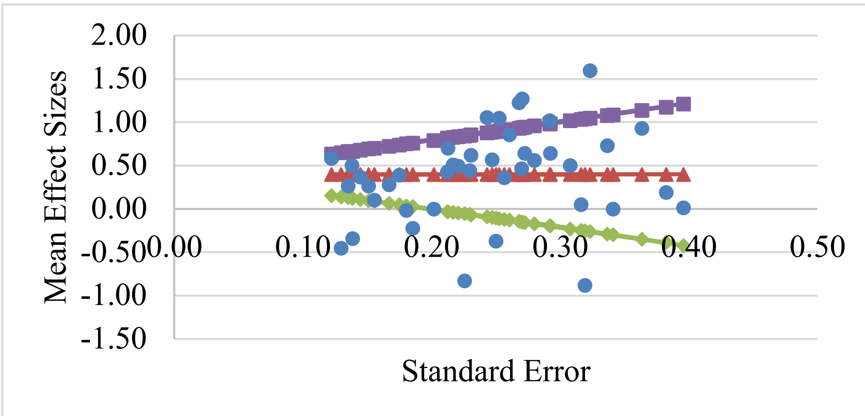


Table 23 shows that child and adult studies yielded small and medium effects, respectively, and at the level of statistical significance. Studies that did not covary IQ resulted in a statistically significant medium ES for heterogeneity ($p < .01$). Within-group variability (Q_w) was statistically significant for studies of moderate quality.

Table 23 Heterogeneity Analysis for Potential Moderators of TMT-B/D-KEFS Condition 4

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
Age	Child	0.27** (0.09)	24	0.08, 0.45	36.31	.05
	Adolescent	0.45 (0.28)	2	-0.09, 0.10	0.52	.77
	Adult	0.55** (0.13)	13	0.30, 0.80	7.30	.89
IQ	Covaried	0.23 (0.13)	13	-0.02, 0.48	13.29	.43
	Not covaried	0.45** (0.09)	27	0.27, 0.63	30.05	.31
Study quality	Poor	0.27 (0.21)	4	-0.15, 0.68	5.40	.25
	Moderate	0.32 (0.14)	11	0.03, 0.60	23.36*	.01
	High	0.43 (0.10)	24	0.23, 0.62	14.82	.93

Note. *TMT/D-KEFS* = Trailmaking Test/Delis-Kaplan Executive Function System; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic.
* $p < .05$. ** $p < .01$.

Results from the meta-regression, shown in Table 24, indicate the proposed moderators were not significantly related to the TMT-B/D-KEFS-4 scores. This finding, in conjunction with the variance component for each proposed moderator, suggests variability exceeds what can be explained by sampling error and that amount of excess cannot be explained by age, IQ covariance, or study quality.

Table 24 Meta-Regression Results for TMT B/D-KEFS 4

Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
Age	0.14	0.08	-0.01, 0.30	45.04	.27	0.16
IQ	0.22	0.16	-0.09, 0.53	43.35	.33	0.18
Study quality	0.09	0.10	-0.12, 0.30	45.21	.26	0.17

Note. *TMT B/D-KEFS 4* = Trailmaking Test Part B/Delis-Kaplan Executive Function System Condition 4; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval; Q_{resid} = within-group heterogeneity; *p* = significance for heterogeneity statistic; *v* = method of moments random effects variance component.
* $p < .05$.

Trailmaking Test Part B – Part A/D-KEFS Condition 4 – Condition 2 (TMT B-A/D-KEFS 4-2)

The mean ES was large and statistically significant. The distribution of effects was significantly heterogeneous. The funnel plot in Figure 10 shows the distribution extended well beyond the upper limit of the population parameters. Further exploration of outlying effects revealed several studies reported very large SDs (e.g., Gonzalez-Gadea et al., 2015). It was reported in one study (Taylor & Miller, 1997) that the IQ of ADHD participants was in the average range, while the IQ of HCs was in the above average range. In addition, high comorbidity was reported in the ADHD group.

Together, these characteristics may contribute to heterogeneity. Trim-and-fill analysis resulted in no trimming performed as the data remained unchanged.

Figure 10 Funnel Plot for Trailmaking Test Part B-A/D-KEFS Condition 4-2 (TMT B-A/D-KEFS 4-2)

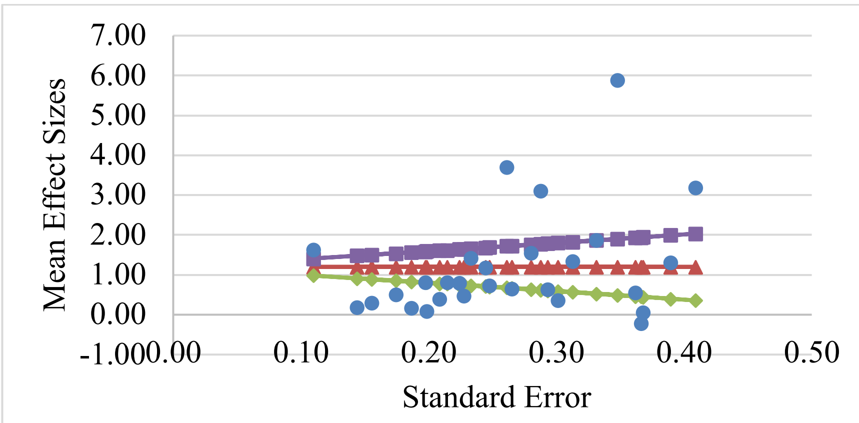


Table 25 shows that, specific to age, child and adolescent studies resulted in large effects, while adult studies yielded a small effect, all of which were statistically significant. Regarding IQ and study quality, studies that covaried IQ resulted in a medium and statistically significant effect. Studies that did not covary IQ, as well as those with moderate and high quality yielded large effects that were statistically significant. Subgroups of child and adult studies, as well as those that did not covary IQ and studies with moderate and high quality were significantly heterogeneous.

Table 25 Heterogeneity Analysis for Potential Moderators of TMT B-A/D-KEFS 4-2

Moderator	Group	Mean *g* (*SE*)	*df*	95% CI	Q_w	*p*-value
Age	Child	1.19** (0.24)	10	0.71, 1.66	26.21**	<.01
	Adolescent	0.99* (0.40)	3	0.20, 1.78	7.10	.07
	Adult	0.10** (0.22)	12	0.56, 1.44	23.41*	.02
IQ	Covaried	0.62* (0.26)	8	0.11, 1.13	4.86	.77
	Not covaried	1.31** (0.19)	18	0.94, 1.68	46.73**	<.001
Study quality	Poor	0.12 (0.51)	1	-0.88, 1.13	0.01	.94
	Moderate	1.54** (0.36)	4	0.83, 2.24	23.41**	<.001
	High	1.06** (0.16)	20	0.73, 1.38	33.98*	.03

Note. *TMT/D-KEFS* = Trailmaking Test/Delis-Kaplan Executive Function System; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic.
* $p < .05$. ** $p < .01$.

Results from the meta-regression, listed in Table 26, indicate a statistically significant relationship between the covariance of IQ and the TMT B-A/D-KEFS 4-2 score. Between-study variance (Q_{model} = 4.55, df = 1) also indicates the strength of this relationship extends beyond what would be expected by chance ($p < .05$). Given within-group variability (Q_{resid}) was significant for each of the proposed moderators, there is a degree of variance explained by factors other than IQ, age, and study quality. The variance component also suggests variability is not fully accounted for by sampling error alone.

Table 26 Meta-Regression Results for TMT B-A/D-KEFS 4-2

Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
Age	-0.09	0.16	-0.41, 0.23	58.53**	< .001	0.54
IQ	0.69*	0.32	0.05, 1.32	51.59**	< .01	0.58
Study quality	0.18	0.25	-0.32, 0.68	54.86**	< .001	0.59

Note. *TMT/D-KEFS* = Trailmaking Test/Delis-Kaplan Executive Function System; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval;

Table 26 Continued

Q_{resid} = within-group heterogeneity; p = significance for heterogeneity statistic; v = method of moments random effects variance component.
* $p < .05$.

Trailmaking Test Part B/Part A and D-KEFS Condition 4/Condition 2 (Trails Ratio)

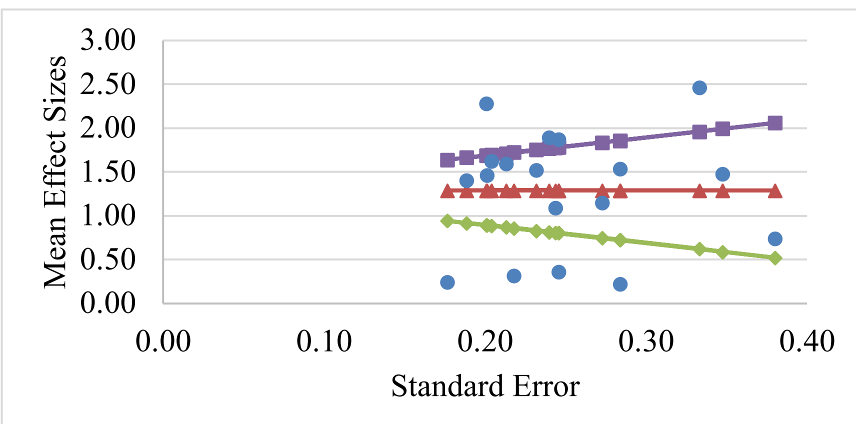
The mean ES was small and non-significant (p = .37). The funnel plot in Figure 11 shows the distribution of effects centered closely around the grand mean. Trim-and-fill analysis resulted in no trimming performed as the data remained unchanged. Because the sample of studies was well within the population parameters, neither heterogeneity analysis nor meta-regression was performed for the Trails ratio.

Figure 11 Funnel Plot for Trailmaking Test/Delis-Kaplan Executive Function System Ratio (TMT/D-KEFS Ratio)

Behavior Rating Inventory of Executive Function (BRIEF)

Across studies, participants with ADHD totaled 734 and HCs totaled 1105. The mean ES was large (g = 1.28, SE = 0.16) 95% CI [0.96, 1.60] and statistically significant. The distribution of effects was significantly heterogeneous (Q = 142.08, df = 17, p's < .0001). The funnel plot in Figure 12 was used to explore study characteristics potentially explaining outlying effects. Of note, two studies (Bodzy, 2011; Northington, 2009) yielding outlying effects were dissertations. Like other study pools described herein, possible contributing factors include significant gender differences across the comparison groups (e.g., Hovik et al., 2017) and noteworthy differences in IQ (e.g., Burmeister, Hannay, Copeland, Fletcher, Boudousquie, & Dennis, 2005). Trim-and-fill analysis resulted in no trimming performed as the data remained unchanged.

Figure 12 Funnel Plot for Behavior Rating Inventory of Executive Function (BRIEF)



Regarding age, one study (Miranda, Mercader, Fernández, & Colomer, 2013) included participants outside of the child and adolescent age range (6-17). Further, only two studies referenced participants between the ages of 13 and 17; therefore, differing subgroups were not feasible for moderator analysis or meta-regression. For study quality, one study yielded a raw score in the poor range (Northington, 2009) and it was omitted from this analysis. Only studies with moderate and high scores were included in the heterogeneity and meta-regression analyses. Table 28 shows that subgroups of studies that did and did not covary IQ, as well as studies with moderate and high quality yielded large effects that were statistically significant.

Table 27 Heterogeneity Analysis for Potential Moderators of the BRIEF

Moderator	Group	Mean *g* (*SE*)	*Df*	95% CI	Q_w	*p*-value
IQ	Covaried	1.33** (0.29)	5	0.75, 1.90	3.87	.57
	Not covaried	1.25** (0.21)	11	0.85, 1.67	11.18	.43
Study quality	Moderate	1.39** (0.22)	9	0.96, 1.82	8.86	.45
	High	1.27** (0.26)	6	0.75, 1.78	4.63	.59

Note. *BRIEF* = Behavior Rating Inventory of Executive Function; *g* = Hedges' *g* effect size; *SE* = standard error; *df* = degrees of freedom; *CI* = confidence interval; Q_w = within-group heterogeneity; *p* = significance for heterogeneity statistic.
* $p < .05$. ** $p < .01$.

Results shown in Table 29 indicate that neither IQ co-variance nor study quality were significantly related to BRIEF parent ratings. Within-group variability, indicated by Q_{resid}, suggests the subgroups are relatively homogeneous.

Table 28 Meta-Regression Results for the BRIEF

Moderator	*B*	*SE*	95% CI	Q_{resid}	*p*-value	*v*
IQ	-0.08	0.36	-0.78, 0.63	15.05	.52	0.45
Study quality	-0.12	0.34	-0.80, 0.55	13.49	.56	0.42

Note. *BRIEF* = Behavior Rating Inventory of Executive Function; *B* = coefficient of latitude for the covariate; *SE* = standard error; *CI* = confidence interval; Q_{resid} = within-group heterogeneity; *p* = significance for heterogeneity statistic; *v* = method of moments random effects variance component.

Summary

Results indicate mean effects range from small to large and evidence a strong relationship between ADHD and deficits in set shifting. In most instances, study effects are significantly heterogeneous. Although informal exploration of study characteristics identified unmatched samples according to age, gender, and IQ as possible reasons for several outlying effects, age and IQ explained portions of variablity for only a fraction of the varaibles investigated. This suggests that alternative factors not accounted for in the current study may contribute to observed variability in these effects.

Specific to the WCST studies, as noted previously, the study by Solanto et al. (2007) yielded substantially outlying effects across all of the variables, which raises the question whether the results are a function of the sample. Underlying this study was an aim to capture the current state of the literature and for this reason, this study, along with other studies yielding outlying effects, were retained for analyses. Although the results are relatively consistent with prior research, this observation is important to note when interpreting results of the current study.

Research Question 2

Does the pattern of impaired function, as evidenced by mean effect size for shifting, differ based on the method of EF assessment (i.e., PBMs versus rating scales)? Because PBMs tap a narrower aspect of shifting and rating scales reflect a broader application of shifting such as problem solving and goal-directed behaviors, it is hypothesized that rating scales will yield smaller effects compared to PBM (Ten Eycke & Dewey, 2016; Toplak et al., 2013).

To compare effects across the different variables, individual subgroups of studies across each moderator with statistically significant effects were explored. In order to compare effects between the PBMs and the rating scale, only subgroups yielding significant effects for the BRIEF were retained for the analysis. To stay consistent with expectations for meta-anlaysis (Egger, 1997; Sterne et al., 2011), subgroups with fewer than 10 studies were excluded from the anlaysis. Comparison of effects was achieved through boxplot anlaysis, which affords a visual aid in examining dispersion of effects (Thompson, 2006). Table 29 shows subgroups within each variable that were included in the boxplot analyses. Results are expressed in two groups: child studies and studies with moderate quality. It is important to note that, because calculation of the #CAT score yielded negative effects, the ES is presented here in absoulte terms. Transforming negative scores to positive scores allows for a clearer comparison of ES distribution for the #CAT score to the distributions of other variables.

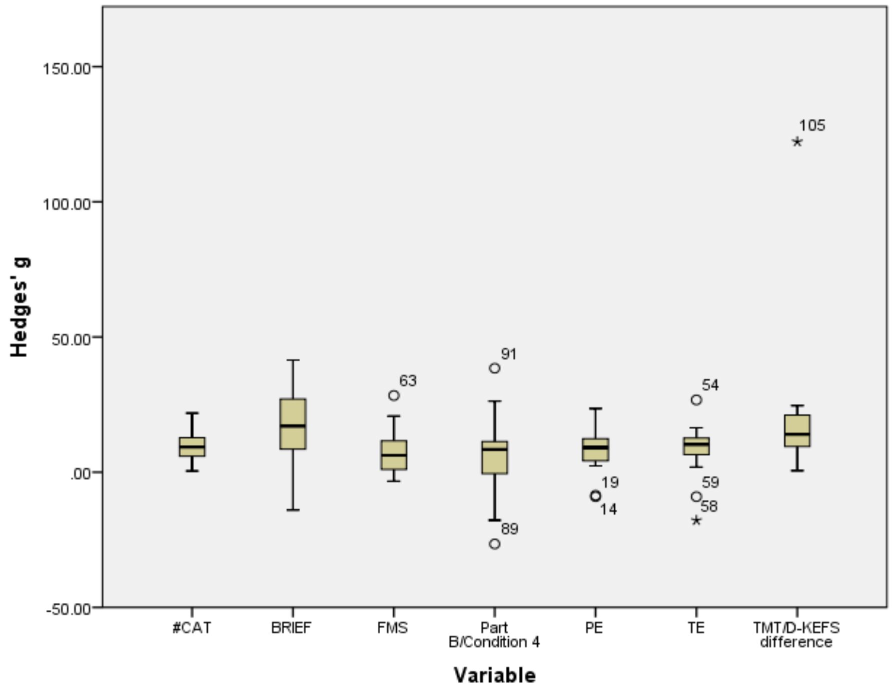
Table 29 Comparison of ES Across PBM and BRIEF

Test – Variable	Moderator subgroup	Total ADHD	Total HC	Mean Hedges' *g* (*SE*)	*df*
WCST – PE	Child	941	1052	0.51** (0.09)	23
WCST – #CAT[a]	Child	745	675	0.74** (0.09)	18
	Moderate Study Quality	299	308	0.81** (0.13)	11
WCST – TE	Child	474	496	0.45** (0.21)	14
WCST – FMS	Child	432	385	0.39** (0.10)	9
TMT B/D-KEFS 4	Child	1260	1636	0.27** (0.09)	24
TMT B-A/D-KEFS 4-2	Child	611	576	1.19** (0.24)	10
BRIEF	Child	642	1017	1.28** (0.16)	16
	Moderate Study Quality	408	686	1.39** (0.22)	9

Note. *WCST* = Wisconsin Card Sort Test; *PE* = perseverative errors; *#CAT* = number of categories completed; *TE* = total errors; *FMS* = failure to maintain set; *TMT/D-KEFS* = Trailmaking Test/Delis-Kaplan Executive Function System; *BRIEF* = Behavior Rating Inventory of Executive Function; *ADHD* = attention-deficit/hyperactivity disorder; *HC* = healthy control; *SE* = standard error; *df* = degrees of freedom.
[a]Effect sizes for the #CAT variable are presented as absolute values.
* $p < .05$. ** $p < .01$.

Figure 13 shows a comparison of the distribution of effects for children across the following variables: PE, #CAT, TE, FMS, TMT B/D-KEFS 4, TMT/D-KEFS difference, and BRIEF. Variables such as the PE, TE, FMS, TMT-B/D-KEFS-4, TMT/D-KEFS difference, and BRIEF all yielded outlying effects. The BRIEF scores had a higher median than the other variables. Given the size of the box, in comparison to that of the other variables, 50 percent of BRIEF effects were larger than effects for the other variables. Of the distributions not containing outliers, the BRIEF appears to have the largest distribution of effects.

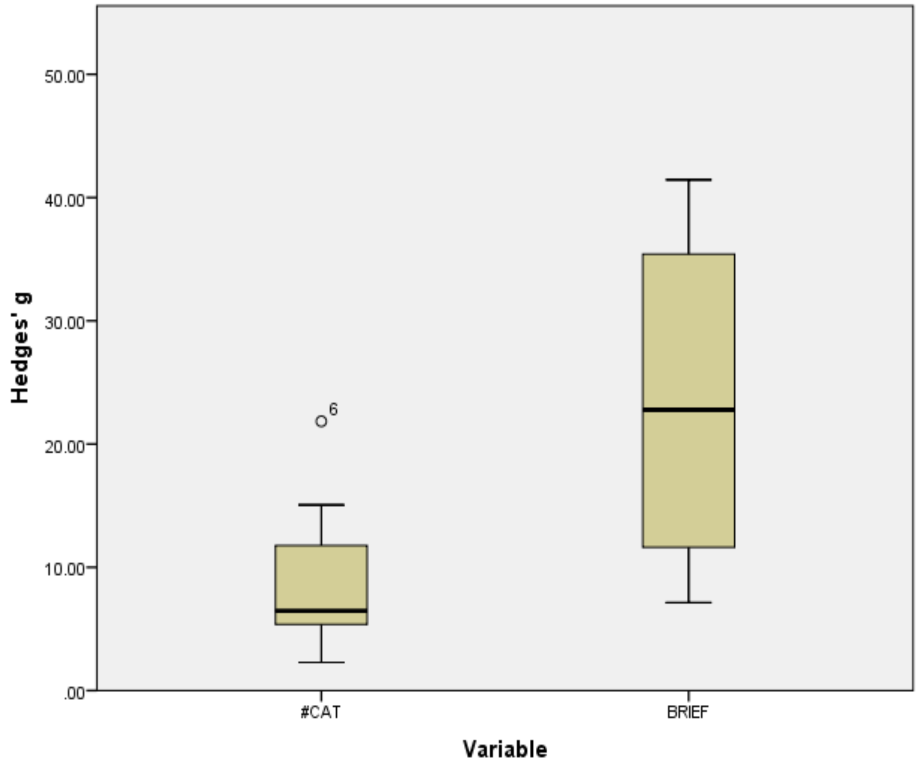
Figure 13 Comparison of Effects for Child Studies



Note. *#CAT* = number of categories completed; *BRIEF* = Behavior Rating Inventory of Executive Function Shift scale; *FMS* = failure to maintain set; *PE* = perseverative errors; *TE* = total errors; *TMT/D-KEFS* = Trailmaking Test/Delis-Kaplan Executive Function System.

Figure 14 shows the distribution of effects for two subgroups of greater than 10 studies of moderate quality. It shows that #CAT resulted in an outlying effect. Between the two subgroups, the BRIEF yielded a larger distribution of effects. In addition, the sample size for the BRIEF was larger compared to the #CAT subgroup. The difference in results across the two subgroups provides evidence that performance varies across PBMs and rating scales, depending on the performance variable to which the rating scale is being compared.

Figure 14 Comparison of Effects for Studies with Moderate Quality



Note. *#CAT* = number of categories completed; *BRIEF* = Behavior Rating Inventory of Executive Function Shift scale.

CHAPTER V

DISCUSSION AND CONCLUSION

This study aimed to answer two questions. The first question regards the relationship between ADHD and shifting deficits, as well as potential moderators that explain a pattern of deficits as indicated by measures of set shifting. The second question regards effect size differences between performance-based measures (PBM) and rating scales of set shifting. To answer these questions, studies utilizing clinical groups (ADHD) and healthy controls (HC) were examined through meta-analysis, specifically heterogeneity testing and meta-regression within a random effects model. Results indicated across two widely used PBMs of set shifting, the Wisconsin Card Sorting Test (WCST) and the Trailmaking Test/Delis-Kaplan Executive Function System Trail Making Test (TMT/D-KEFS), scores including Perseverative Errors (PE), Number of Categories Completed (#CAT), Total Errors (TE), Nonperseverative Errors (NPE), Failure to Maintain Set (FMS), Total Correct (TC), Perseverative Responses (PR), and TMT-B/D-KEFS Condition 4 are typically reported in ADHD research. Additionally, search methods used in this study evidenced wide use of the Behavior Rating Inventory of Executive Functioning (BRIEF) to assess shifting in individuals with ADHD.

Consistent with previous research, individuals with ADHD tended to perform lower than HCs on these measures of shifting, but at varying levels (Boonstra et al., 2005; Pennington & Ozonoff, 1996; Willcutt et al., 2005). Statistically significant effects ranged from small to large, with the PE, TE, NPE, FMS, PR (Romine et al., 2004), and TMT-B/D-KEFS-4 scores evidencing small overall effects. The #CAT, TMT/D-KEFS difference, and BRIEF Shift scale yielded medium and large effects. Findings suggest that, compared to HC, these measures are sensitive to set shifting deficits in people with ADHD. Within a given measure, however, the degree of impairment may vary based on the score and other factors, including age and study quality.

This study was different from previous meta-analyses in a few ways (e.g., Boonstra et al., 2005; Romine et al., 2004; Willcutt et al., 2005). First, the sample size for each variable was generally larger than samples in studies by Romine et al., and Willcutt et al. Second, statistical significance was included to describe the strength of the relationship between ADHD and shifting impairments. Because the presentation of ADHD is highly variable (Kofler et al., 2017; Polanczyk, Willcutt, Salum, Kieling, & Rohde, 2014; Shang & Gau, 2011), it is expected individuals would yield very different performances across measures of shifting. Variability was also expected given that, to date, studies have yet to definitively associate shifting impairment to ADHD. The use of statistical significance aimed to address this question. For example, some scores yielded small effects (e.g., PE); however, they were significant. This suggests that, although effects may reflect minor impairment, the area of performance deviates substantially from what is expected in typical development. It is important, therefore, to understand how and to what extent does ADHD relate to the different aspects of shifting.

It is also important to note that study pools were significantly heterogeneous for the PE, #CAT, TE, FMS, TC, PR, TMT-B/D-KEFS-4, TMT/D-KEFS difference, and BRIEF Shift scores. Again, this may be explained by the notion that ADHD is generally considered a heterogeneous condition (Kofler et al., 2017; Polanczyk et al., 2014; Shang & Gau, 2011). Kofler and colleagues found that differences in working memory were related to heterogeneity in interpersonal functioning. Heterogeneity according to parent ratings on the BRIEF may also result from assumptions that household expectations vary across families. For instance, varying levels at which caregivers impose demands on their children may help to explain how individuals are evaluated on things like attentional and behavior shifting. In families where children have less structure and fewer expectations, it may be more difficult to detect problems with shifting and flexible thinking, particularly if these skills are not being taxed by things like household chores, studying, and homework (Amorim & Marques, 2018). As such, these differences may be contributing to some of the observed variability within parent ratings.

Significant effects ranged from small to large across different subgroups of the proposed moderators. Specific to age, small effects were noted in children across the #CAT, TE, FMS, and TMT-B/D-KEFS-4 scores. Specific to the TMT/D-KEFS scores, Part B/Condition 4 yielded a small and medium effect for children and adults, respectively. The TMT/D-KEFS time difference resulted in large effects for child and adolescent studies, as well as an overall small effect for adult studies. Due to the limited number of adolescent studies for the PE, TE, and TC variables, child and adolescent studies were collapsed into one group (< 18 years). In these instances, small and large effects were indicated for the PE and TC scores, respectively. Regarding the #CAT score, results yielded a medium effect for adolescent studies and small effect for adult studies. Interpretation of results specific to adolescent studies warrant caution, given the sample size ($df = 3$). Results suggest that deficits are largely found in children compared to older age groups, which is consistent with prior research (Corbett et al., 2009; Oades & Christiansen, 2008; Romine et al., 2004; Willcutt et al., 2005).

From a developmental perspective, this study reiterates several considerations. First, the range of ESs specific to children is wide and consistent with assumptions that their developmental trajectory is highly variable. In other words, children change rapidly and make gains in different areas and at different rates. Although the range is wide, children with ADHD, nonetheless, do exhibit impairments on measures of shifting. Similarly, results specific to the BRIEF suggest impairments are indicated when compared to typical children. With effects being generally higher for the BRIEF, it suggests these individuals consistently display deficits when functioning in the real world. Although the pool of adolescent studies was small, there is some evidence that the performance gap narrows as skills and abilities develop more evenly.

Albeit small, the study pool suggests that abilities continue to develop more evenly in adulthood. Results from this study also evidence some of the inconsistencies in research related to adults with ADHD. Specific to the WCST, scores particularly like PE and TE have been shown to yield inconsistent effects across adult samples (Holdnack et al., 1995; Riccio et al., 2005; Seidman et al., 1997). A recent study suggested that, as people age, they are likely to achieve more categories and make fewer non-perseverating errors (Milioni et al., 2017). One exception noted regards the TMT-B/D-KEFS-4, as it evidenced a medium and significant effect. Specific to this measure, previous meta-analytic findings suggest the TMT-B/D-KEFS-4 score is sensitive to shifting deficits in adults (Boonstra et al., 2005). Despite these observations, more studies are needed to further investigate shifting in adults. This will help in making more informed inferences about developmental trends for shifting, particularly in adolescence and adulthood. These findings, nonetheless, are overall consistent with what is known about brain development, as well as developmental models of ADHD (Castellanos & Tannock, 2002) and executive function (Miyake et al., 2000).

Meta-regression provided evidence that age moderates a portion of variability yielded by scores such as the #CAT, TC, and TMT/D-KEFS difference. This is commensurate with findings in a recent meta-analysis by Lai et al. (2017). It is important to note, however, that their analysis did not include studies with adult participants and the target population was autism spectrum disorder (ASD). Given the number of shared traits between ADHD and ASD, however, it is expected that findings across the groups are somewhat consistent (Happé et al., 2006).

This study revealed there is a paucity of ADHD research utilizing intelligence (IQ) as a covariate. One reason might be because several studies incorporated other ways to control for IQ, one being the use of cutoff scores as eligibility criteria (e.g., IQ < 75). Studies using the covariance of IQ yielded significant effects that ranged from small to large for the PE, #CAT, PR, TMT/D-KEFS difference, and BRIEF scores. Given these pools of studies were relatively small ($2 \leq k \leq 9$), however, findings may not be as meaningful (Thompson, 2006). Specific to the #CAT, FMS, PR, PE, TE, TMT-B/D-KEFS-4, TMT/D-KEFS difference, and BRIEF variables, significant effects ranged from small to large for studies that did not covary IQ. Meta-regression indicated variables such as the TE and TMT/D-KEFS difference yielded variability that was explained in part by IQ. Interestingly, the PE, #CAT, TMT/D-KEFS difference, and BRIEF scores yielded significant effects for both subgroups, suggesting that, even under highly controlled circumstances, deficits in EF can emerge regardless (Pennington & Ozonoff, 1996). Together, these findings support the understanding that IQ and EF are interrelated, but at varying degrees.

Significant effects ranging from small to large emerged across studies grouped by study quality. Of note, the PE, #CAT, TE, FMS, TMT/D-KEFS difference, and BRIEF scores yielded significant effects across groups of studies with moderate and high quality. Studies rated as "poor" yielded a medium effect for the PE score. Informal exploration of outlying study effects revealed the impact of certain study characteristics that were addressed in the rubric for study quality. Concerns related to unmatched groups according to gender, age, IQ, and comorbidity potentially impacted the dispersion of effects. Prior research has indicated deficits in set shifting can result when IQ is not accounted for (Murphy et al., 2002; Schweitzer et al., 2006). In other instances, comorbidity has shown to influence performance on measures of EF (e.g., Marchetta et al., 2008; Walker et al., 2000). Meta-regression analysis indicated that variability specific to the TMT/D-KEFS difference is explained in part by study quality.

In sum, confounding factors such as IQ and study quality, in addition to age and gender may help to explain why information about ADHD and set shifting is inconclusive.

Unique to this study is the use of meta-analysis to explore effects of the test version of the WCST. For the modified version, effects were significant and medium for the #CAT, TE, and TC scores. The PE score resulted in a small effect. For the full test version, small effects were noted for the FMS and PR scores, while PE and #CAT resulted in medium effects. Further, meta-regression indicated the variability specific to TC was explained by the test version. Although research has shown both test versions to be comparable to one another (Axelrod et al., 1992; Fortuny & Heaton, 1996; Heaton & Thompson, 1992; Sillanpaa et al., 1993), results of the current study suggest that, between both decks, the #CAT is more consistently sensitive to shifting impairments compared to the other scores of the WCST. Given test length varies substantially and is determined by each of the decks, it may impact the degree of variability in performance, as well as how discrepant the performance is between clinical and healthy samples.

The second research question addressed the comparative effects of PBM versus the BRIEF. The results indicated that, in children, effects may be larger for PBMs when the PE score is the comparison. Within the moderate study quality group, the BRIEF was shown to yield a larger and perhaps more meaningful ES when compared to the #CAT score. This suggests that, under more ideal controls related to age, gender, and IQ balance, as well as appropriate diagnostic procedures, rating scales can tap deficits in shifting more so than PBMs. This was based on visual analysis using a boxplot.

Implications

This study affords both research and clinical practice key considerations. It is the first using meta-analysis to explore the relationship between ADHD and shifting across three measures of EF. Characteristics including study design (i.e., having studies coded separately by two coders) and statistical analyses of this study were theory driven and afforded rigor in key ways. To date, findings related to set shifting and ADHD have been inconclusive. Compared to other functions often explored in ADHD (e.g., inhibition and working memory), less was known definitively about set shifting. Results from this study, therefore, allow for a better conceptualization of ADHD in the context of three deficit areas (inhibition, working memory, and set shifting). Impairments across these areas of functioning are likely to indicate the presence of ADHD, meaning the three-factor model of EF by Miyake et al. (2000) may serve as another tool to conceptualize ADHD across the lifespan.

From a clinical perspective, results suggest the WCST is an effective assessment tool for use with children and adolescents, with certain variables yielding larger and more significant effects. The TMT/D-KEFS time difference has shown capability to detect shifting impairments not only in young people with ADHD, but in adults as well. These findings indicate that, although these assessments are equally useful in children and adolescents, the TMT/D-KEFS is currently one of a few measures with utility for assessing shift in adults. Further investigation, therefore, may help to identify additional assessments that can be used in the assessment of adult ADHD.

Regarding the use of rating scales, additional studies can help to evidence their contribution to the overall conceptualization of ADHD, namely, the extent to which shifting impairments interfere with daily functioning. Further, additional studies are needed to determine their utility in evaluating ADHD in adults. Combined, these findings reiterate the importance of using multiple methods, as well as multiple informants in assessment.

Having this knowledge can aid schools and families in their support of young people with ADHD. It can help them anticipate difficulties with transition, acquiring new problem-solving strategies, understanding novel tasks, and adjusting response styles (Gioia, Isquith, Guy, & Kenworthy, 2015). This study also provides more insight about ADHD in adolescence and adulthood. In older individuals, deficits in shifting can impact how individuals respond to routine/ schedule changes, navigate conversations across rapidly changing topics, and transition in their employment (e.g., learning new job skills and adapting to new work environments). As people begin to better understand what difficulties they may encounter, it may alleviate some of the stigma and negative feelings associated with having these inefficiencies.

Limitations

This study was met with some limitations. The study sample was derived from two major databases. It may have been more effective to include a third database as records appropriate for this study were likely missed. The use of an additional database would have been consistent with recommendations for conducting systematic reviews outlined in the Cochrane Handbook (Higgins & Green, 2017). Specific to coding considerations, inclusion criteria were not applied to the 20 re-coded studies, which raises the question whether the second evaluator would have even selected those studies for the analysis. The sample size of studies was relatively small for several variables. Having samples of less than 20 studies presented challenges when attempting to explore effects across different age groups, especially adolescents. Relatedly, all studies were coded according to the mean age of participants, which was not always an accurate representation of the overall sample. In other words, study samples may have contained participants with ages differentially categorized from the mean age. It was difficult, therefore, to generalize age-related results to people in the real world. The coding sequence also revealed challenges to evaluating and interpreting essential information, including mean IQ scores and details about the diagnostic process (i.e., how studies assigned participants to the ADHD and HC groups). In several instances, studies did not specify demographic information related to socioeconomic status, ethnicity, and primary language of the sample. Having this information may have helped to understand key characteristics about the sample and may have shed light on confounding factors that potentially impacted the results.

An aim of this study was to use meta-analysis of test variables with pools of studies greater than 10 and the search methods resulted in a limited pool of measures that met this criterion. As a result, tests like the CANTAB (Cambridge Cognition Limited, 2000) and others in development were excluded from this study. Findings related to these measures, therefore, remain inconclusive and do not afford a clear understanding of their utility in the assessment of ADHD. Ongoing investigation into these measures across the lifespan may help to identify those that are effective in evaluating ADHD in adolescents and adults.

Specific to the BRIEF, the search method retrieved a limited number of studies using the Teacher and Self-Report, as well as Adult and Other Informant forms. This restricted meta-analysis to studies using the Parent Rating form as well as participants aged 18 years and younger. Having additional forms of the BRIEF would have been helpful for looking at the manifestation of deficits in alternative settings and utility of rating scales in adults. For young people with ADHD, school often presents several challenges that can result from poor EF, in addition to symptoms of inattention and hyperactivity/ poor impulse control (Reid & Johnson, 2012). Having the ability to examine educator perspectives of shifting may have yielded different results (e.g., Davidson et al., 2016). Further, additional studies reporting teacher ratings may help to answer this question. In addition, the publication timeline of retrieved studies precluded the use of the updated (and current) version of the BRIEF (BRIEF-2; Gioia et al., 2015), which restricted findings to an outdated version of this measure.

Regarding moderators, the sample of studies using IQ as a covariate was small, thus limiting the ability to better understand the impact of IQ on scores of PBMs as well as ways in which IQ influences EF in everyday life (e.g., BRIEF scores). Having additional studies that include IQ as a covariate can afford more power for comparison of the effects. To better understand outcomes for which unexplained heterogeneity was detected, further exploration of possible factors, including the proportion of males to females in studies, subtypes of ADHD, and medication status of participants (naïve, medicated during the testing, effects of psychotropic medication) may have been helpful. An additional moderator of interest is gender. Because it is assumed ADHD is more prevalent in males, it was expected the majority of participants across studies would be male. Although this is consistent with prevalence rates of ADHD (APA, 2013; Rucklidge, 2010), results may not be generalizable beyond males. It further perpetuates a limited understanding of ADHD in females and hinders early identification of the disorder in females; thus, it can result in females being diagnosed later in life. Finally, of important note regards changes in DSM nomenclature. Given the amount of change in symptom criteria for adults, this is an especially salient consideration in furthering a clearer understanding of ADHD beyond childhood and adolescence. Informal observation from the present study revealed that these characteristics are not always specified and when they are specified, the descriptions are sometimes unclear. As such, future meta-analysis will benefit from replicating and adding to the rigor of analytic techniques used in this study to aid in defining these and other salient variables.

Future Directions

To better understand the trajectory of ADHD, it will be important for research to continue its investigation of the diagnosis in adolescence and adulthood. This will aid in the development of best practices related to assessment of EF in older individuals. As individual studies pursue this endeavor, meta-analysis can utilize this data to evidence a more conclusive and consistent conceptualization of ADHD symptomatology, as well as the role of executive impairment in the disorder. Given the knowledge that has been afforded by this study, it will be important for individual studies to account for the limitations addressed herein and future meta-analyses to consider the use of three databases instead of two.

Along the lines of diagnostic clarity, one possibility would be to examine more closely the contribution of rating scales to the identification of ADHD. In other words, does the use of either a rating scale or performance task (e.g., Stroop test) or clinical interview, mask or highlight differences in EF between ADHD and HC groups? Studies show that the exclusive use of rating scales in diagnosing ADHD leads to higher prevalence rates (Graetz et al., 2001; Nolan, Gadow, & Sprafkin, 2001; Vasconcelos et al., 2003). Knowing how this consideration relates to set shifting may enhance the overall understanding of set shifting in ADHD.

To better understand the role intelligence plays in the assessment of EF, it will be helpful if more studies pursue the covariance of IQ when exploring set shifting in people with ADHD. For example, a study by Milioni et al. (2017) found that in high-functioning adults with ADHD (IQ < 110), there was a negative correlation between IQ and completion time for TMT-B, and that they did not perform differently from healthy controls on most neuropsychological measures. Conversely, their second group of participants with ADHD (and of average intelligence) yielded significant performance deficits on measures of set shifting. Of important note, Milioni and colleagues did not covary IQ. Together, these findings highlight that in some cases, there is an interrelatedness between intelligence and EF (Frazier et al., 2004; Pennington & Ozonoff, 1996). With inconsistency across measures of shifting, however, this may be symptomatic of the heterogeneity in ADHD presentations.

Future studies may also want to consider how deficits in set shifting impairment impact long-term outcomes. As research on adult ADHD continues to emerge, it will be important to understand ways in which this can be harnessed to promote healthier and more positive outcomes. Along those lines, it may be helpful for future studies to investigate appropriate interventions for set shifting deficits. Given results herein suggest this is an area of impairment for many individuals with ADHD, it will be helpful to explore the effects of intervention and what outcomes can result when early versus late (i.e., adulthood) intervention is considered.

Conclusion

In conclusion, results of this study indicate that set shifting can be an area of deficit for people with ADHD. The relationship between ADHD and shifting impairment is impacted by several factors such as age and intelligence, although more research is needed to better understand additional factors moderating this relationship. This study also highlights the importance of using multi-methods in assessment of children, adolescents, and adults. Different methods help to provide multi-dimensional information about deficits. Measures explored in this study have shown to be good indicators of set shifting ability and are helpful when garnering evidence for a diagnosis of ADHD. It is the author's hope that researchers, clinicians, and patients alike will gain some benefit from information and knowledge yielded by this study.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

Achenbach, T. M., Howell, C. T., McConaughy, S. H., & Stanger, C. (1995). Six-year predictors of problems in a national sample of children and youth: I. Cross-informant syndromes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*(3), 336-347.

Adams, Z. W., Milich, R., & Fillmore, M. T. (2010). A case for the return of attention-deficit disorder in DSM-5. *The ADHD Report, 18*, 1-6. doi: 10.1521/adhd.2010.18.3.1

Adler, L. A., Barkley, R. A., & Newcorn, J. H. (2011). Performance improvement CME: Adult ADHD. *Journal of Clinical Psychiatry, 72*(4). doi:10.4088/JCP.9066pi4c

Agosti, V., Chen, Y., & Levin, F. R. (2011). Does Attention Deficit Hyperactivity Disorder increase the risk of suicide attempts? *Journal of Affective Disorders*, *133*, 595-599. doi: 10.1016/j.jad.2011.05.008

*Ahmadi, N., Mohammadi, M. R., Araghi, S. M., & Zarafshan, H. (2014). Neurocognitive profile of children with attention deficit hyperactivity disorders (ADHD): A comparison between subtypes. *Iranian Journal of Psychiatry*, *9*(4), 197.

Alderson, R. M., Rapport, M. D., & Kofler, M. J. (2007). Attention-deficit/hyperactivity disorder and behavioral inhibition: A meta-analytic review of the stop-signal paradigm. *Journal of Abnormal Child Psychology*, *35*, 745-758. doi: 10.1007/s10802-007-9131-6

Allen, D. N., Haderlie, M., Kazakov, D., & Mayfield, J. (2009). Construct and criterion validity of the Comprehensive Trail Making Test in children and adolescents with traumatic brain injury. *Child Neuropsychology*, *15*, 543-553. doi: 10.1080/09297040902748234

Alvarez, J., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*, 17-42. doi:10.1007/s11065-006-9002-x

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

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.

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

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., text rev.). Washington, DC: American Psychiatric Association.

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

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

Amorim, W. N., & Marques, S. C. (2018). Inhibitory control and cognitive flexibility in children with attention-deficit/hyperactivity disorder. *Psychology & Neuroscience*, *11*, 364-374. http://dx.doi.org/10.1037/pne0000156

Anderson, V., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships between cognitive and behavioral measures of executive functions in children with brain disease. *Child Neuropsychology, 8*(4), 231-240.

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology, 20*(1), 385-406.

Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, 40*(1), 57-87.

*Antonini, T. N., Becker, S. P., Tamm, L., & Epstein, J. N. (2015). Hot and cool executive functions in children with attention-deficit/hyperactivity disorder and comorbid oppositional defiant disorder. *Journal of the International Neuropsychological Society*, *21*, 584-595. doi:10.1017/S1355617715000752

*Antshel, K. M., Faraone, S. V., Maglione, K., Doyle, A. E., Fried, R., Seidman, L. J., & Biederman, J. (2010). Executive functioning in high-IQ adults with ADHD. *Psychological Medicine*, *40*, 1909-1918. doi: 10.1017/S0033291709992273

Arcia, E., & Conners, K. C. (1998). Gender Differences in ADHD? *Journal of Developmental & Behavioral Pediatrics, 19*(2), 77-83.

Awh, E., Anllo-Vento, L., & Hillyard, S. A. (2000). The role of spatial selective attention in working memory for locations: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *12*(5), 840-847.

Azouvi, P., Vallat-Azouvi, C., Millox, V., Darnoux, E., Ghout, I., Ruet, A.,… Jourdan, C. (2016). Ecological validity of the Dysexecutive Questionnaire: Results from the PariS-TBI study. *Neuropsychological Rehabilitation, 25*, 864-878, doi: 10.1080/09602011.2014.990907

*Aycicegi-Dinn, A., Dervent-Ozbek, S., Yazgan, Y., Bicer, D., & Dinn, W. M. (2011). Neurocognitive correlates of adult attention-deficit/hyperactivity disorder in a Turkish sample. *ADHD Attention Deficit and Hyperactivity Disorders*, *3*, 41-52. doi: 10.1007/s12402-010-0050-y

Babinski, D. E., Pelham Jr, W. E., Molina, B. S., Waschbusch, D. A., Gnagy, E. M., Yu, J.,…Biswas, A. (2011). Women with childhood ADHD: comparisons by diagnostic group and gender. *Journal of Psychopathology and Behavioral Assessment*, *33*, 420-429. doi: 10.1007/s10862-011-9247-4

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29. doi: 10.1146/annurev-psych-120710-100422

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49*, 5-28. doi: 10.1080/713755608

Baddeley, A. D. (1986). *Working Memory*. New York, NY: Oxford University Press.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, *18*(4), 362-365.

*Bahcivan Saydam, R., Ayvasik, H. B., & Alyanak, B. (2015). Executive functioning in subtypes of attention deficit hyperactivity disorder. *Noro Psikiyatri Arsivi, 52*, 386-392. doi: 10.5152/npa.2015.8712

Banaschewski, T., Neale, B. M., Rothenberger, A., & Roessner, V. (2007). Comorbidity of tic disorders & ADHD. *European Child & Adolescent Psychiatry*, *16*, 5-14. doi: 10.1007/s00787-007-1002-8

Banaschewski, T., Hollis, C., Oosterlaan, J., Roeyers, H., Rubia, K., Willcutt, E., & Taylor, E. (2005). Towards an understanding of unique and shared pathways in the psychopathophysiology of ADHD. *Developmental Science, 8*, 132–140. doi: 10.1111/j.1467-7687.2005.00400.x

Banks, G. C., Kepes, S., & Banks, K. P. (2012). Publication bias: The antagonist of meta-analytic reviews and effective policymaking. *Educational Evaluation and Policy Analysis, 34*, 259-277. doi: 10.3102/0162373712446144

Barkley, R. (2015). Health problems and related impairments in children and adults with ADHD. In R. Barkley (Ed.), *Attention- deficit hyperactivity disorder: A handbook for diagnosis and treatment* (4th ed., pp. 267-313). New York, NY: Guilford.

Barkley, R. A. (2011). *Barkley Adult ADHD Rating Scale-IV (BAARS-IV)*. Guilford Press. Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65-94.

Barkley, R. A., & Fischer, M. (2011). Predicting impairment in occupational functioning in hyperactive children as adults: self- reported executive function (EF) deficits vs. EF tests. *Developmental Neuropsychology, 36*, 137-161. doi: 10.1080/87565641.2010.549877

Barkley, R. A., & Murphy, K. R. (2011). The nature of executive function (EF) deficits in daily life activities in adults with ADHD and their relationship to performance on EF tests. *Journal of Psychopathology and Behavioral Assessment*, *33*(2), 137-158. doi:10.1007/s10862-011-9217-x

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

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2006). Young adult outcome of hyperactive children: adaptive functioning in major life activities. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*, 192-202. doi: 10.1097/01.chi.0000189134.97436.e2

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1991). Attention deficit disorder with and without hyperactivity: Clinical response to three dose levels of methylphenidate. *Pediatrics, 87*, 519-531.

*Beadle, C. L. (2009). *Attentional impairment and processing speed in children diagnosed with attention deficit hyperactivity disorder* (Unpublished doctoral dissertation). Texas Woman's University, Denton, TX.

Becker, S. P., & Langberg, J. M. (2014). Attention-deficit/ hyperactivity disorder and sluggish cognitive tempo dimensions in relation to executive functioning in adolescents with ADHD. *Child Psychiatry & Human Development*, *45*, 1-11. doi: 10.1007/s10578-013-0372-z

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *The Journal of General Psychology*, *39*(1), 15-22.

Biederman, J., Petty, C. R., Fried, R., Doyle, A. E., Spencer, T., Seidman, L. J., Gross, L., … Faraone, S. V. (2007). Stability of executive function deficits into young adult years: A prospective longitudinal follow-up study of grown up males with ADHD. *Acta Psychiatrica Scandinavica,116*, 129-136, doi: 10.1111/j.1600-0447.2007.01008.x

Biederman, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F.,… Faraone, S. V. (2004). Impact of executive function deficits and attention-deficit/ hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of Consulting and Clinical Psychology*, *72*(5), 757-766.

Biederman, J., Faraone, S. V., Keenan, K., Benjamin, J., Krifcher, B., Moore, C., ... Spencer, T. (1992). Further evidence for family-genetic risk factors in attention deficit hyperactivity disorder: Patterns of comorbidity in probands and relatives in psychiatrically and pediatrically referred samples. *Archives of General Psychiatry*, *49*(9), 728-738.

Blakemore, S., Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal Child Psychology and Psychiatry, 47*, 296-312. doi: 10.1111/j.1469-7610.2006.01611.x

Bledsoe, J. C., Semrud-Clikeman, M., & Pliszka, S. R. (2013). Anterior cingulate cortex and symptom severity in attention-deficit/ hyperactivity disorder. *Journal of Abnormal Psychology, 122*, 558-565. doi: 10.1037/a0032390

Blijd-Hoogewys, E. M., Bezemer, M. L., & van Geert, P. L. (2014). Executive functioning in children with ASD: An analysis of the BRIEF. *Journal of Autism and Developmental Disorders, 44*, 3089-3100. doi: 10.1007/s10803-014-2176-9

*Bodzy, M. E. (2011). *Examining group differences in executive functions in children with MTBI and ADHD* (Unpublished doctoral dissertation). Illinois Institute of Technology, Chicago, IL.

Bolfer, C., Pacheco, S. P., Tsunemi, M. H., Carreira, W. S., Casella, B. B., & Casella, E. B. (2017). Attention-deficit/hyperactivity disorder: The impact of methylphenidate on working memory, inhibition capacity and mental flexibility. *Arquivos De Neuro-Psiquiatria*, *75*(4), 204-208. https://doi.org/10.1590/0004-282X20170030

*Boswell, A. K. (2013). *Multidimensional assessment of attention-deficit/hyperactivity disorder (ADHD)* (Unpublished doctoral dissertation). Rutgers University, New Brunswick, NJ.

Boonstra, A. M., Oosterlaan, J., Sergeant, J. A., & Buitelaar, J. K. (2005). Executive functioning in adult ADHD: A meta-analytic review. *Psychological Medicine, 35*, 1097-1108. doi: 10.1017/S003329170500499X

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*, 97-111. doi: 10.1002/jrsm.12

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. doi: 10.1002/9780470743386

*Brewer, V. R., Fletcher, J. M., Hiscock, M., & Davidson, K. C. (2001). Attention processes in children with shunted hydrocephalus versus attention deficit-hyperactivity disorder. *Neuropsychology*, *15*, 185-198. doi: 10.1037//0894-4105.15.2.185

Bruchmuller, K., Margraf, J., & Schneider, S. (2012). Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. *Journal of Consulting and Clinical Psychology, 80*,128-138. doi: 10.1037/a0026582

Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society, 4*(6), 547-558.

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 81–116). Hove, UK: Psychology Press.

*Burmeister, R., Hannay, H. J., Copeland1, K., Fletcher, J. M., Boudousquie, A., & Dennis, M. (2005). Attention problems and executive functions in children with spina bifida and hydrocephalus. *Child Neuropsychology*, *11*, 265-283. doi: 10.1080/092970490911324

Busch, R. M., McBride, A., Curtiss, G., & Vanderploeg, R. D. (2005). The components of executive functioning in traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 27*,1022-1032. doi: 10.1080/13803390490919263

Carlin, D., Bonerba, J., Phipps, M., Alexander, G., Shapiro, M., & Grafman, J. (2000). Planning impairments in frontal lobe dementia and frontal lobe lesion patients. *Neuropsychologia*, *38*(5), 655-665.

Castellanos, F. X., Sonuga-Barke, E. J., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: beyond executive dysfunction. *Trends in Cognitive Sciences*, *10*(3), 117-123. doi: 10.1016/j.tics.2006.01.011

Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nature Reviews Neuroscience*, *3*, 617-628. doi: 10.1038/nrn896

*Chen, L., Huang, X., Lei, D., He, N., Hu, X., Chen, Y.,… Gong, Q. (2015). Microstructural abnormalities of the brain white matter in attention-deficit/hyperactivity disorder. *Journal of Psychiatry & Neuroscience: JPN, 40*, 280-287. doi: 10.1503/jpn.140199

Cheung, M. W. –L., & Vijayakumar, R. (2016). A guide to conducting a meta-analysis. *Neuropsychology Review, 26*, 121-128. doi: 10.1007/s11065-016-9319-z

Chiang, H. L., & Gau, S. S. F. (2014). Impact of executive functions on school and peer functions in youths with ADHD. *Research in Developmental Disabilities*, *35*, 963-972. doi: 10.1016/j.ridd.2014.02.010

Claesdotter, E., Cervin, M., Åkerlund, S., Råstam, M., & Lindvall, M. (2018). The effects of ADHD on cognitive performance. *Nordic Journal of Psychiatry*, *72*, 158-163. doi: 10.1080/08039488.2017.1402951

Clements, S. D., & Peters, J. E. (1962). Minimal brain dysfunctions in the school-age child: diagnosis and treatment. *Archives of General Psychiatry*, *6*(3), 185-197.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conners, C. K. (2000). Attention-deficit/hyperactivity disorder – Historical development and overview. *Journal of Attention Disorders, 3*(4), 173-191.

Cooper, H. M., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis (2nd ed.)*. New York, NY: Russell Sage Foundation.

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research*, *166*, 210-222. doi: 10.1016/j.psychres.2008.02.005

Corkum, P. V., & Siegel, L. S. (1993). Is the continuous performance task a valuable research tool for use with children with attention-deficit-hyperactivity disorder? *Journal of Child Psychology and Psychiatry*, *34*(7), 1217-1239.

Crichton. A. S. (n. d.). *An inquiry into the nature and origin of mental derangement: Comprehending a concise system of the physiology and pathology of the human mind and a history of the passions and their effects*. Retrieved from https://babel.hathitrust.org/cgi/pt?id=nyp.33433011465048;view=1up;seq=18

Cubillo, A., Halari, R., Ecker, C., Giampietro, V., Taylor, E., & Rubia, K. (2010). Reduced activation and inter-regional functional connectivity of fronto-striatal networks in adults with childhood attention-deficit hyperactivity disorder (ADHD) and persisting symptoms during tasks of motor inhibition and cognitive switching. *Journal of Psychiatric Research*, *44*, 629-639. doi: 10.1016/j.jpsychires.2009.11.016

D' Cruz, A. M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology*, *27*, 152-160. doi: 10.1037/a0031721

*Davidson, F., Cherry, K., & Corkum, P. (2016). Validating the Behavior Rating Inventory of Executive Functioning for children with ADHD and their typically developing peers. *Applied Neuropsychology: Child*, *5*, 127-137. doi: 10.1080/21622965.2015.1021957

Davis, N. O., & Kollins, S. H. (2012). Treatment for co-occurring attention deficit/hyperactivity disorder and autism spectrum disorder. *Neurotherapeutics*, *9*, 518-530. doi: 10.1007/s13311-012-0126-9

*De la Torre, G. G., Martin, A., Cervantes, E., Guil, R., & Mestre, J. M. (2017). Attention lapses in children with spina bifida and hydrocephalus and children with attention-deficit/hyperactivity disorder. *Journal of Clinical and Experimental Neuropsychology*, *39*, 563-573. doi: 10.1080/13803395.2016.1249828

de Sonneville, L. M. J. (2000). *ANT 2.1-Amsterdam Neuropsychological Tasks*.

Amstelveen, Netherlands: Sonar

Del Re, A. C., & Flückiger, C. (2016). Meta-analysis. In J. C. Norcross, G. R. VandenBos, & D. K. Freedheim (Eds.), *APA handbook of clinical psychology: Theory and research* (Vol. 2) (pp. 479-491). Washington, DC, US: American Psychological Association. doi: 10.1037/14773-022

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

Demakis, G. J. (2003). A meta-analytic review of the sensitivity of the Wisconsin Card Sorting Test to frontal and lateralized frontal brain damage. *Neuropsychology*, *17*(2), 255-264.

Demurie, E., Roeyers, H., Baeyens, D., & Sonuga-Barke, E. (2013). Domain-general and domain-specific aspects of temporal discounting in children with ADHD and autism spectrum disorders (ASD): A proof of concept study. *Research in Developmental Disabilities*, *34*, 1870-1880. doi: 10.1016/j.ridd.2013.03.011

Denckla, M. B. (1996). Biological correlates of learning and attention: What is relevant to learning disability and attention-deficit hyperactivity disorder? *Journal of Developmental & Behavioral Pediatrics*, *17*(2), 114-119.

Derekinko, K., & Pelham, W. (2016). ADHD and substance use. In (Ed.), *The Oxford Handbook of Substance Use and Substance Use Disorders: Volume 2*. doi: 10.1093/oxfordhb/9780199381708.013.18

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177-188.

Diamond, A. (2005). Attention-deficit disorder (attention-deficit/ hyperactivity disorder without hyperactivity): A neurobiologically and behaviorally distinct disorder from attention-deficit/ hyperactivity disorder (with hyperactivity). *Development and Psychopathology, 17*, 807-825. doi: 10.1017/S0954579405050388

Dijkers, M. (2013). Introducing GRADE: A systematic approach to rating evidence in systematic reviews and to guideline development. *KT Update*, *1*(5), 1-9.

Dimoska-Di Marco, A., McDonald, S., Kelly, M., Tate, R., & Johnstone, S. (2011). A meta-analysis of response inhibition and Stroop interference control deficits in adults with traumatic brain injury (TBI). *Journal of Clinical and Experimental Neuropsychology*, *33*, 471-485. doi: 10.1080/13803395.2010.533158

Dolan, M., & Lennox, C. (2013). Cool and hot executive function in conduct-disordered adolescents with and without co-morbid attention deficit hyperactivity disorder: Relationships with externalizing behaviours. *Psychological Medicine*, *43*, 2427-2436. doi: 10.1017/S0033291712003078

Donders, J., DenBraber, D., & Vos, L. (2010). Construct and criterion validity of the Behaviour Rating Inventory of Executive Function (BRIEF) in children referred for neuropsychological assessment after pediatric traumatic brain injury. *Journal of Neuropsychology, 4*, 197-209. doi: 10.1348/174866409X478970

Döpfner, M., Frölich, J., & Lehmkuhl, G. (2013). *Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS)* (Vol. 1). Hogrefe Verlag.

Douglas, V. I. (1972). Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioural Science*, *4*(4), 259-282.

*Doyle, A. E., Biederman, J., Seidman, L. J., Reske-Nielsen, J. J., & Faraone, S. V. (2005). Neuropsychological functioning in relatives of girls with and without ADHD. *Psychological Medicine*, *35*, 1121-1132. doi:10.1017/S0033291705004496

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.). *The cognitive neurosciences* (pp. 721-733). Cambridge, MA: The MIT Press.

DuPaul, G. J., Gormley, M. J., & Laracy, S. D. (2014). School-based interventions for elementary school students with ADHD. *Child and Adolescent Psychiatric Clinics of North America*, *23*, 687-697. doi: 10.1016/j.chc.2014.05.003

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

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale-IV*. New York: Guilford Press.

Durston, S., Tottenham, N. T., Thomas, K. M., Davidson, M. C., Eigsti, I. M., Yang, Y.,…Casey, B. J. (2003). Differential patterns of striatal activation in young children with and without ADHD. *Biological Psychiatry, 53*, 871-878. doi: 10.1016/S0006-3223(02)01904-2

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 455-463. doi: 10.1111/j.0006-341X.2000.00455.x

Egeland, J. (2010). Frequency of attention deficit in first-episode schizophrenia compared to ADHD. *Applied Neuropsychology*, *17*, 125-134. doi: 10.1080/09084280903297859

Egeland, J., & Fallmyr, Ø. (2010). Confirmatory factor analysis of BRIEF: Support for a distinction between emotional and behavioural regulation. *Child Neuropsychology, 16*, 326-337. doi: 10.1080/09297041003601462

Edvinsson, D. K., Bingefors, K., Lindstrom, E., & Lewander, T. (2010). ADHD-related symptoms among adults in out-patient psychiatry and female prison inmates as compared with the general population. *Upsala Journal of Medical Sciences, 115*, 30-40. doi: 10.3109/03009730903532333

Egger, M., Smith G. D., Schneider M., Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997 (315), 629-634.

*Elosúa, M. R., Del Olmo, S., & Contreras, M. J. (2017). Differences in executive functioning in children with attention deficit and hyperactivity disorder (ADHD). *Frontiers in Psychology*, *8*, 976. doi: 10.3389/fpsyg.2017.00976

Eme, R. (2016). ADHD and Risky Substance Use in Male Adolescents. *The ADHD Report*, *24*, 1-8. doi: 10.1521/adhd.2016.24.3.1

Eslinger, P., & Damasio, A. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation. *Neurology, 35*, 1731-1741. doi: 10.1212/WNL.35.12.1731

*Faraone, S. V., Biederman, J., Doyle, A., Murray, K., Petty, C., Adamson, J. J., & Seidman, L. (2006). Neuropsychological studies of late onset and subthreshold diagnoses of adult attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *60*, 1081-1087. doi: 10.1016/j.biopsych.2006.03.060

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology, 26*, 874-890. doi: 10.1080/13803390490510680

Fletcher, P. C., & Henson, R. N. A. (2001). Frontal lobes and human memory. *Brain, 124*(5), 849-881.

Fray, P. J., Robbins, T. W., & Sahakian, B. J. (1996). Neuorpsychiatyric applications of CANTAB. *International Journal of Geriatric Psychiatry, 11*(4), 329-336.

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. doi: 10.1037/0894-4105.18.3.543

Frazier-Wood, A. C., Bralten, J., Arias-Vasquez, A., Luman, M., Ooterlaan, J., Sergeant, J.,… Rommelse, N. N. (2012). Neuropsychological intra-individual variability explains unique genetic variance of ADHD and shows suggestive linkage to chromosomes 12, 13, and 17. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *159*, 131-140. doi: 10.1002/ajmg.b.32018

Fried, R., Hirshfield-Becker, D., Petty, C., Batchelder, H., & Biederman, J. (2015). How informative is the CANTAB to assess executive functioning in children with ADHD? A controlled study. *Journal of Attention Disorders, 19*, 468-475. doi: 10.1177/1087054712457038

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*, 172-179. doi: 10.1111/j.1467-9280.2006.01681.x

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference cognitive functions: A latent variable analysis. *Journal of Experimental Psychology: General, 133*, 101-135. doi: 10.1037/0096-3445.133.1.101

Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at head start. *Developmental Psychology*, *47*, 404. doi: 10.1037/a0021065

Fuster, J. M. (1989). *The prefrontal cortex: Anatomy, physiology and neuropsychology of the frontal lobe* (2nd edition). New York: Raven.

Gau, S. S., Chiu, C. D., Shang, C. Y., Cheng, A. T., & Soong, W. T. (2009). Executive function in adolescence among children with attention-deficit/ hyperactivity disorder in Taiwan. *Journal of Developmental and Behavioral Pediatrics 30*, 525-534. doi: 10.1097/DBP.0b013e3181c21c97

Gau, S., & Shang, C. (2010). Executive functions as phenotypes in ADHD: Evidence from the Cambridge Neuropsychological Test Battery (CANTAB). *Journal of Child Psychology and Psychiatry, 71*, 838-849. doi: 10.1111/j.1469-7610.2010.02215.x

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: a meta-analysis and critical review. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(8), 1036-1045.

Geurts, H. M., Verté, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology and Psychiatry*, *45*, 836-854. doi: 10.1111/j.1469-7610.2004.00276.x

Ghanizadeh, A., Mohammadi, M. R., Akhondzadeh, S., & Sanaei-Zadeh, H. (2011). Attention deficit hyperactivity disorder in imprisoned individuals: A review. *Psychiatria Danubina, 23*(2), 139-144.

Giacobini, M., Medin, E., Ahnemark, E., Russo, L. J., & Carlqvist, P. (2018). Prevalence, patient characteristics, and pharmacological treatment of children, adolescents, and adults diagnosed with ADHD in Sweden. *Journal of Attention Disorders*, *22*, 3-13. doi: 10.1177/1087054714554617

Giedd, J. N., Blumenthal, J., Molloy, E., & Castellanos, F. X. (2001). Brain imaging of attention deficit/ hyperactivity disorder. *Annals of the New York Academy of Sciences*, *931*(1), 33-49.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). *BRIEF-2: Behavior Rating Inventory of Executive Function: Professional manual*. Psychological Assessment Resources.

Gioia, G., Espy, K. A., & Isquith, P. K. (2003). *Behavior Rating Inventory of Executive*

Function Preschool version. Norwegian manual supplement 2007. Lutz, FL: Psychological Assessment Resources.

Gioia, G. A., & Isquith, P. K. (2002). Two faces of monitor: They self and thy task. *Journal of the International Neuropsychological Society, 8*, 229.

Gioia, G. A., Isquith, P. K., Kenworthy, L., & Barton, R. M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*, *8*(2), 121-137.

*Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & Espy, K. A. (2002). Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. *Child Neuropsychology, 8*(4), 249-257.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavioral Rating Inventory of Executive Function: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Glass, L., Ware, A. L., Crocker, N., Deweese, B. N., Coles, C. D., Kable, J. A.,… Riley, E. P. (2013). Neuropsychological deficits associated with heavy prenatal alcohol exposure are not exacerbated by ADHD. *Neuropsychology*, *27*, 713-724. doi: 10.1037/a0033994

Goltz, F. (1888). Uber die Verrichtungen des Grosshirns. Pfuger's archive fur die gesamte. *Physiologie, 42*, 419-467.

Golubchik, P., & Weizman, A. (2018). The effect of methylphenidate treatment on suspiciousness in children with ADHD alone or comorbid with ODD. *International Journal of Psychiatry in Clinical Practice*, *22*, 109-114. doi: 10.1080/13651501.2017.1383436

*Gonzalez-Gadea, M. L., Chennu, S., Bekinschtein, T. A., Rattazzi, A., Beraudi, A., Trippichio, P.,... Sigman, M. (2015). Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. *American Journal of Physiology-Heart and Circulatory Physiology*. doi:10.1152/jn.00543.2015

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment, 68*, 532-560. doi: 10.1207/s15327752jpa6803_5

Graetz, B. W., Sawyer, M. G., & Baghurst, P. (2005). Gender differences among children with DSM-IV ADHD in Australia. *Journal of American Academy of Child and Adolescent Psychiatry, 44*,159-168. doi: 10.1097/00004583-200502000-00008

Graetz, B. W., Sawyer, M. G., Hazell, P. L., Arney, F., & Baghurst, P. (2001). Validity of DSM-IV ADHD subtypes in a nationally representative sample of Australian children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(12), 1410-1417.

Grafman, J., & Litvan, I. (1999). Importance of deficits in executive functions. *The Lancet*, *354*(9194), 1921-1923.

Gray, J. A. (1991). Neural systems, emotion and personality. In J. Madden (Ed), *Neurobiology of learning, emotion and affect*. (pp. 273-306). New York: Raven Press.

Gray, J. A. (1982). *The neuropsychology of anxiety*. New York: Oxford University Press.

*Hale, J. B., Reddy, L. A., Decker, S. L., Thompson, R., Henzel, J., Teodori, A.,... Denckla, M. B. (2009). Development and validation of an attention-deficit/hyperactivity disorder (ADHD) executive function and behavior rating screening battery. *Journal of Clinical and Experimental Neuropsychology*, *31*, 897-912. doi: 10.1080/13803390802687423

Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/ hyperactivity disorder: examining profiles across domains and ages. *Brain and Cognition*, *61*, 25-39. doi: 10.1016/j.bandc.2006.03.004

Harstad, E., & Levy, S. (2014). Attention-deficit/ hyperactivity disorder and substance abuse. *American Academy of Pediatrics, 134*, 293-301. doi: 10.1542/peds.2014-0992

Hebb, D. O. (1939). Intelligence in man after large removals of cerebral tissue: Report of four left frontal lobe cases. *Journal of General Psychology, 21*(1), 73-87.

Hebb, D. O. (1945). Man's frontal lobes: A critical review. *Archives of Neurology and Psychology, 54*(1), 10-24.

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*, 426-445. doi: 10.1037/1082-989X.9.4.426

Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis.

Psychological Methods, 6, 203-217. doi: 10.1037/1082-989X.6.3.203

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*, 486-504. doi: 10.1037/1082-989X.3.4.486

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Herbert, M. (1964). The concept and testing of brain-damage in children: A review. *Journal of Child Psychology and Psychiatry*, *5*(3-4), 197-216.

Hibbs, E. D., & Jensen, P. S. (2004). *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice* (2nd ed.). Washington, DC: American Psychological Association.

Higgins, J. P., & Green, S. (2017). *Cochrane handbook for systematic reviews of interventions version 5.1.0 (Updated March 2011 ed)*. Retrieved from www.cochrane-handbook.org.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557-560. doi: 10.1136/bmj.327.7414.557

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*, 1539-1558. doi: 10.1002/sim.1186

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*, 189-233. doi: 10.1016/j.dr. 2004.01.001

Hinshaw, S. P., Owens, E. B., Sami, N., & Fargeon, S. (2006). Prospective follow-up of girls with attention-deficit/hyperactivity disorder into adolescence: Evidence for continuing cross-domain impairment. *Journal of Consulting and Clinical Psychology*, *74*, 489-499. doi: 10.1037/0022-006X.74.3.489

*Holmes, J., Gathercole, S. E., Place, M., Alloway, T. P., Elliott, J. G., & Hilton, K. A. (2010). The diagnostic utility of executive function assessments in the identification of ADHD in children. *Child and Adolescent Mental Health*, *15*, 37-43. doi: 10.1111/j.1475-3588.2009.00536.x_2009

Homack, S., & Riccio, C. A. (2004). A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. *Archives of Clinical Neuropsychology*, *19*, 725-743. doi: 10.1016/j.acn.2003.09.003

*Hooper, E. (1996). *The usefulness of the Wechsler Intelligence Scale for Children – Third Edition and the Wisconsin Card Sorting Test in the diagnosis of attention-deficit hyperactivity disorder* (Unpublished doctoral dissertation). Biola University, La Mirada, CA.

*Houghton, S., Douglas, G., West, J., Whitting, K., Wall, M., Lagsford, S.,… Carroll, A. (1999). Differential patterns of executive function in children with attention-deficit hyperactivity disorder according to gender and subtype. *Journal of Child Neurology, 14*(12), 801-805.

*Hovik, K. T., Egeland, J., Isquith, P. K., Gioia, G., Skogli, E. W., Andersen, P. N., & Øie, M. (2017). Distinct patterns of everyday executive function problems distinguish children with Tourette syndrome from children with ADHD or autism spectrum disorders. *Journal of Attention Disorders*, *21*, 811-823. doi: 10.1177/1087054714550336

*Huang, F., Sun, L., Qian, Y., Liu, L., Ma, Q. G., Yang, L.,... Wu, Z. M. (2016). Cognitive function of children and adolescents with attention deficit hyperactivity disorder and learning difficulties: A developmental perspective. *Chinese Medical Journal*, *129*, 1922.

Huang-Pollock, C. L., Mikami, A. Y., Pfiffner, L., & McBurnett, K. (2009). Can executive functions explain the relationship between attention deficit hyperactivity disorder and social adjustment? *Journal of Abnormal Child Psychology*, *37*, 679-691. doi: 10.1007/s10802-009-9302-9308

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*, 2017-2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Hummer, T. A., Kronenberger, W. G., Wang, Y., Dunn, D. W., Mosier, K. M., Kalnin, A. J., & Mathews, V. P. (2011). Executive functioning characteristics associated with ADHD comorbidity in adolescents with disruptive behavior disorders. *Journal of Abnormal Child Psychology, 39*, 11-19. doi: 10.1007/s10802-010-9449-3

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

IBM SPSS Statistics for Windows (Version 22.0) [Computer software]. Armonk, NY: IBM Corp.

Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child, 2*, 125-132. doi: 10.1080/21622965.2013.748389

*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. doi: 10.4103/0366-6999.187861

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology, 14*(89), 81.

*Johnson, D. E., Epstein, J. N., Waid, L. R., Latham, P. K., Voronin, K. E., & Anton, R. F. (2001). Neuropsychological performance deficits in adults with attention deficit/hyperactivity disorder. *Archives of Clinical Neuropsychology*, *16*(6), 587-604.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology Review*, *17*, 213-233. doi: 10.1007/s11065-007-9040-z

*Kado, Y., Sanada, S., Yanagihara, M., Ogino, T., Ohno, S., Watanabe, K.,… Ohtsuka, Y. (2012). Executive function in children with pervasive developmental disorder and attention-deficit/ hyperactivity disorder assessed by the Keio version of the Wisconsin Card Sorting Test. *Brain and Development, 34*, 354-359. doi: 10.1016/j.braindev.2011.08.008

*Kamradt, J. M., Ullsperger, J. M., & Nikolas, M. A. (2014). Executive function assessment and adult attention-deficit/hyperactivity disorder: Tasks versus ratings on the Barkley Deficits in Executive Functioning Scale. *Psychological Assessment*, *26*, 1095-1105. doi: 10.1037/pas0000006

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637-671.

*Kang, K. D., Han, D. H., Kim, S. M., Bae, S., & Renshaw, P. F. (2018). The correlation between cognitive and movement shifting and brain activity in children with ADHD. *Journal of Attention Disorders*, *22*, 661-670. doi: 10.1177/1087054716677816

Kasper, L. J., Alderson, R. M., & Hudec, K. L. (2012). Moderators of working memory deficits in children with attention-deficit/hyperactivity disorder (ADHD): a meta-analytic review. *Clinical Psychology Review*, *32*, 605-617. doi: 10.1016/j.cpr.2012.07.001

Kawabata, Y., Tseng, W. L., & Gau, S. S. F. (2012). Symptoms of attention-deficit/hyperactivity disorder and social and school adjustment: the moderating roles of age and parenting. *Journal of Abnormal Child Psychology*, *40*, 177-188. doi: 10.1007/s10802-011-9556-9

*Keage, H. A., Clark, C. R., Hermens, D. F., Kohn, M. R., Clarke, S., Williams, L. M.,... & Gordon, E. (2006). Distractibility in AD/HD predominantly inattentive and combined subtypes: The P3a ERP component, heart rate and performance. *Journal of Integrative Neuroscience*, *5*(01), 139-158.

Kebir, O., & Tabbane, K. (2009). Candidate genes and neuropsychological phenotypes in children with ADHD: Review of association studies. *Journal of Psychiatry & Neuroscience: JPN*, *34*(2), 88-101.

Kessler, R. C., Adler, L. A., Berglund, P., Green, J. G., McLaughlin, K. A., Fayyad, J.,… Zaslavsky, A. M. (2014). The effects of temporally secondary co-morbid mental disorders on the associations of DSM-IV ADHD with adverse outcomes in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A). *Psychological Medicine*, *44*, 1779-1792. doi: 10.1017/S0033291713002419

Kiefer, M., Marzinzik, F., Weisbrod, M., Scherg, M., & Spitzer, M. (1998). The time course of brain activations during response inhibition: Evidence from event-related potentials in a go/ no go task. *NeuroReport, 9*(4), 765-770.

Kingdon, D., Cardoso, C., & McGrath, J. J. (2016). Research review: Executive function deficits in fetal alcohol spectrum disorders and attention-deficit/ hyperactivity disorder – A meta-analysis. *Journal of Child Psychology and Psychiatry, 57*, 116-131. doi: 10.1111/jcpp.12451

Klein, R. G., Mannuzza, S., Olazagasti, M. A. R., Roizen, E., Hutchison, J. A., Lashua, E. C., & Castellanos, F. X. (2012). Clinical and functional outcome of childhood attention-deficit/ hyperactivity disorder 33 years later. *Archives of General Psychiatry*, *69*, 1295-1303. doi: 10.1001/archgenpsychiatry.2012.271

Kofler, M. J., Rapport, M. D., Bolden, J., Sarver, D. E., Raiker, J. S., & Alderson, R. M.

(2011). Working memory deficits and social problems in children with ADHD. *Journal of Abnormal Child Psychology*, *39*, 805-817. doi:10.1007/s10802-011-9492-8

Kok, A. (1999). Varieties of inhibition: manifestations in cognition, event-related potentials and aging. *Acta Psychologica*, *101*, 129-158. doi: 10.1016/S0001-6918(99)00003-7

Kolb, B., & Wishaw, I. Q. (1990). *Fundamentals of human neuropsychology*. New York: WH Freemont and Company.

Kofler, M. J., Sarver, D. E., Spiegel, J. A., Day, T. N., Harmon, S. L., & Wells, E. L. (2017). Heterogeneity in ADHD: Neurocognitive predictors of peer, family, and academic functioning. *Child Neuropsychology*, *23*, 733-759. doi: 10.1080/09297049.2016.1205010

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*, 155-163. doi: 10.1016/j.jcm.2016.02.012

Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY-II: A Developmental Neuropsychological Assessment*. San Antonio, TX: The Psychological Corporation.

Koziol, K., & Budding, D. (2009). *Subcortical structures and cognition*. New York, NY: Springer.

Kwon, H., Reiss, A. L., & Menon, V. (2002). Neural basis of protracted developmental changes in visuo-spatial working memory. *Proceedings of the National Academy of Sciences*, *99*(20), 13336-13341.

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., Garfinkel, B.,… Shaffer, D. (1994). DMS-IV field trials for attention deficit hyperactivity disorder in children and adolescents. *The American Journal of Psychiatry, 151*(11), 1673-1685

Lahey, B. B., Pelham, W. E., Schaughency, E. A., Atkins, M. S., Murphy, H. A., Hynd, G.,… Lorys-Vernon, A. (1988). Dimensions and types of attention deficit disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *27*(3), 330-335.

Lai, C. L. E., Lau, Z., Lui, S. S., Lok, E., Tam, V., Chan, Q.,... Cheung, E. F. (2017). Meta-analysis of neuropsychological measures of executive functioning in children and adolescents with high-functioning autism spectrum disorder. *Autism Research*, *10*, 911-939. doi: 10.1002/aur.1723

Lange, K. W., Reichl, S., Lange, K. M. Tucha, L., & Tucha, O. (2010). The history of attention deficit hyperactivity disorder. *Attention Deficit and Hyperactivity Disorders, 2*. doi: 10.1007/s12402-010-0045-8

Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence, 12*, 131-147. doi: 10.1016/0160-2896(88)90012-8

Laufer, M. W., Denhoff, E., & Solomons, G. (1957). Hyperkinetic impulse disorder in children's behavior problems. *Psychosomatic Medicine*, *19*(1), 38-49.

*Lawrence, V., Houghton, S., Douglas, G., Durkin, K., Whiting, K., & Tannock, R. (2004). Executive function and ADHD: A comparison of children's performance during neuropsychological testing and real-world activities. *Journal of Attention Disorders*, *7*(3), 137-149.

Lee, K., Bull, K., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development, 84*, 1933-1953. doi: 10.1111/cdev.12096

Lehto, J. E., Juujaervi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology,21*, 59-80. doi: 10.1348/026151003321164627

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Quarterly Journal of Experimental Psychology, 49*, 29–50. doi: 10.1080/713755616

Leung, R. C., & Zakzanis, K. K. (2014). Brief report: Cognitive flexibility in autism spectrum disorders: A quantitative review. *Journal of Autism and Developmental Disorders*, *44*, 2628-2645. doi: 10.1007/s10803-014-2136-4

Levin, H. S., Song, J., Ewing-Cobbs, L., & Roberson, G. (2001). Porteus maze performance following traumatic brain injury in children. *Neuropsychology*, *15*(4), 557-567.

Levine, B., Stuss, D., Milberg, W., Alexander, M., Schwartz, M., & MacDonald, R. (1998). The effects of focal and diffuse brain damage on strategy application: Evidence from focal lesions, traumatic brain injury and normal aging. *Journal of the International Neuropsychological Society, 4*(3), 247-264.

Levy, R., Friedman, H. R., Davachi, L., & Goldman-Rakic, P. S. (1997). Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *The Journal of Neuroscience*, *17*(10), 3870-3882.

Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (2003). *The Sage encyclopedia of social science research methods*. Sage Publications.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.

*Li, F., He, N., Li, Y., Chen, L., Huang, X., Lui, S.,... Gong, Q. (2014). Intrinsic brain abnormalities in attention deficit hyperactivity disorder: A resting-state functional MR imaging study. *Radiology*, *272*, 514-523. doi: 10.1148/radiol.14131622

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A.,… Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analysis of studies that evaluate healthcare interventions: Explanation and elaboration. *British Medical Journal, 339*, 37-64. doi: 10.1136/bmj.b2700

Lijffijt, M., Kenemans, J. L., Verbaten, M. N., & van Engeland, H. (2005). A meta-analytic review of stopping performance in attention-deficit/ hyperactivity disorder: Deficient inhibitory motor control? *Journal of Abnormal Psychology*, *114*, 216-222. doi: 10.1037/0021-843X.114.2.216

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Loe, I. M., & Feldman, H. M. (2007). Academic and educational outcomes of children with ADHD. *Journal of Pediatric Psychology*, *32*, 643-654. doi: 10.1093/jpepsy/jsl054

*Loftis, C. W. (2004). *An ecological validity study of executive function measures in children with and without attention deficit hyperactivity disorder* (Unpublished doctoral dissertation). University of Florida, Gainesville, FL.

Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189-239). San Diego, CA: Academic Press.

Loo, S. K., Humphrey, L. A., Tapio, T., Moilanen, I. K., McGough, J. J., McCracken, J. T.,… Järvelin, M. R. (2007). Executive functioning among Finnish adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 1594-1604. doi: 10.1097/chi.0b013e3181575014

*Lovejoy, D. W., Ball, J. D., Keats, M., Stutts, M. L., Spain, E. H., Janda, L., & Janusz, J. (1999). Neuropsychological performance of adults with attention deficit hyperactivity disorder (ADHD): Diagnostic classification estimates for measures of frontal lobe/executive functioning. *Journal of the International Neuropsychological Society*, *5*(3), 222-233.

Llorente, A. M., Williams, L., Satz, P., & D'Elia, L. (1998). *Children's Color Trails Test: Professional manual*. Psychological Assessment Resources.

Lundervold, A. J., Adolfsdottir, S., Halleland, H., Halmøy, A., Plessen, K., & Haavik, J. (2011). Attention Network Test in adults with ADHD - The impact of affective fluctuations. *Behavioral and Brain Functions*, *7*, 1-27. doi: 10.1186/1744-9081-7-27

Macdonald, G., & McCartan, C. J. (2014). Centre-based early education interventions for improving school readiness. *Cochrane Database of Systematic Reviews,* (1). doi: 10.1002/14651858.CD010913

Mahone, E. M., Hagelthorn, K. M., Cutting, L. E., Schuerholz, L. J., Pelletier, S. F., Rawlins, C.,… Denckla, M. B. (2002). Effects of IQ on executive function measures in children with ADHD. *Child Neuropsychology*, *8*(1), 52-65.

Manchester, D., Priestley, N., & Jackson, H. (2004). The assessment of executive functions: Coming out of the office. *Brain Injury*, *18*, 1067-1081. doi: 10.1080/02699050410001672387

Maniadaki, K., Kakouros, E., & Karaba, R. (2010). *Psychopathology in juvenile delinquents*. New York: Nova Science Publishers.

Mannuzza, S., Klein, R. G., Bessler, A., Malloy, P., & Hynes, M. E. (1997). Educational and occupational outcome of hyperactive boys grown up. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(9), 1222-1227.

Marchetta, N. D., Hurks, P. P., Krabbendam, L., & Jolles, J. (2008). Interference control, working memory, concept shifting, and verbal fluency in adults with attention-deficit/hyperactivity disorder (ADHD). *Neuropsychology*, *22*, 74-84. doi: 10.1037/0894-4105.22.1.74

*Martel, M. M., Nigg, J. T., & Von Eye, A. (2009). How do trait dimensions map onto ADHD symptom domains? *Journal of Abnormal Child Psychology*, *37*, 337-348. doi: 10.1007/s10802-008-9255-3

Martel, M., Nikolas, M., & Nigg, J. T. (2007). Executive function in adolescents with ADHD. *Journal of American Academy of Child and Adolescent Psychiatry, 46*, 1437-1444. doi: 10.1097/chi.0b013e31814cf953

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*, 377-384. doi: 10.1097/01.chi.0000153228.72591.73

*Mataro, M., Garcia-Sanchez, C., Junque, C., Estevez-Gonzalez, A., & Pujol, J. (1997). Magnetic resonance imaging measurement of the caudate nucleus in adolescents with attention-deficit hyperactivity disorder and its relationship with neuropsychological and behavioral measures. *Archives of Neurology*, *54*(8), 963-968.

*Mattson, S. N., Roesch, S. C., Glass, L., Deweese, B. N., Coles, C. D., Kable, J. A.,... Jones, K. L. (2013). Further development of a neurobehavioral profile of fetal alcohol spectrum disorders. *Alcoholism: Clinical and Experimental Research*, *37*, 517-528. doi: 10.1111/j.1530-0277.2012.01952.x

McAuley, T., Crosbie, J., Charach, A., & Schachar, R. (2014). The persistence of cognitive deficits in remitted and unremitted ADHD: A case for the state-independence of response inhibition. *Journal of Child Psychology and*

Psychiatry, *55*, 292-300. doi:10.1111/jcpp.12160

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the 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. doi: 10.1017/S1355617710000093

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*, 381-389. doi: 10.1177/1087054706292115

McInnes, A., Humphries, T., Hogg-Johnson, S., & Tannock, R. (2003). Listening comprehension and working memory are impaired in attention-deficit hyperactivity disorder irrespective of language impairment. *Journal of Abnormal Child Psychology*, *31*, 427-443. doi: 10.1023/A:1023895602957

McKenzie, J. E., Beller, E. M., & Forbes, A. B. (2016). Introduction to systematic reviews and meta-analysis. *Respirology, 21*, 626-637. doi: 10.1111/resp.12783

Mecklinger, A., Weber, K., Gunter, T. C., & Engle, R. W. (2003). Dissociable brain mechanisms for inhibitory control: Effects of interference content and working memory capacity. *Cognitive Brain Research*, *18*(1), 26-38.

Mesulam, M. (1986). Frontal cortex and behavior: Editorial. *Annals of Neurology, 19*, 320-325. doi: 10.1002/ana.410190403

Milioni, A. L. V., Chaim, T. M., Cavallet, M., de Oliveira, N. M., Annes, M., dos Santos, B.,... Zanetti, M. V. (2017). High IQ may “mask” the diagnosis of ADHD by compensating for deficits in executive functions in treatment-naïve adults with ADHD. *Journal of Attention Disorders*, *21*, 455-464. doi: 10.1177/1087054714554933

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, *13*, 395-423. doi: 10.1080/15248372.2011.585478

Miller, M., & Hinshaw, S. P. (2010). Does childhood executive function predict adolescent functional outcomes in girls with ADHD? *Journal of Abnormal Child Psychology*, *38*, 315-326. doi:10.1007/s10802-009-9369-2

*Miranda, A., Mercader, J., Fernández, M. I., & Colomer, C. (2017). Reading performance of young adults with ADHD diagnosed in childhood: Relations with executive functioning. *Journal of Attention Disorders*, *21*, 294-304. doi: 10.1177/1087054713507977

Mischel, W., Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., ... Shoda, Y. (2011). 'Willpower' over the life span: decomposing self-regulation. *Social Cognitive and Affective Neuroscience, 81*, 1-5. doi: 10.1093/scan/nsq081

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Current Directions in Psychological Science*, *21*, 8-14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T.

D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100. doi: 10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H.,… Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 2693-2698. doi: 10.1073/pnas.1010076108

Molina, B. S., & Pelham Jr, W. E. (2014). Attention-deficit/hyperactivity disorder and risk of substance use disorder: Developmental considerations, potential pathways, and opportunities for research. *Annual Review of Clinical Psychology*, *10*, 607-639. doi: 10.1146%2Fannurev-clinpsy-032813-153722

Monette, S., Bigras, M., & Lafrenière, M. (2015). Structure of executive functions in typically developing kindergarteners. *Journal of Experimental Child Psychology, 140*, 123-139. doi: 10.1016/j.jecp.2015.07.005

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp. 93-148). Hove, UK: Erlbaum.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*(2), 111-121.

Mueller, K. L., & Tomblin, J. B. (2012). Examining the comorbidity of language disorders and ADHD. *Topics in Language Disorders*, *32*, 228-246. doi: 10.1097/TLD.0b013e318262010d

*Mullane, J. C., & Corkum, P. V. (2007). The relationship between working memory, inhibition, and performance on the Wisconsin Card Sorting Test in children with and without ADHD. *Journal of Psychoeducational Assessment, 25*, 211-221. doi: 10.1177/0734282906297627

*Müller, B. W., Gimbel, K., Keller-Pließnig, A., Sartory, G., Gastpar, M., & Davids, E. (2007). Neuropsychological assessment of adult patients with attention-deficit/hyperactivity disorder. *European Archives of Psychiatry and Clinical Neuroscience*, *257*, 112-119. doi: 10.1007/s00406-006-0688-9

Munk, H. (1890). *Uber die functionen der grosshimrinde*. Berlin: Hirschwald.

Murphy, K. R., Barkley, R. A., & Bush, T. (2002). Young adults with attention deficit hyperactivity disorder: subtype differences in comorbidity, educational, and clinical history. *The Journal of Nervous and Mental Disease*, *190*(3), 147-157.

*Murphy, P. (2001). Cognitive functioning in adults with attention-deficit/hyperactivity disorder. *Journal of Attention Disorders*, *5*(4), 203-209.

Naglieri, J. A., & Goldstein, S. (2013). *Comprehensive Executive Function Inventory*. Multi-Health Systems.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353-383. doi: 10.1016/0010-0285(77)90012-3

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex*, *12*(4), 313-324.

Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, *57*, 1224-1230. doi: 10.1016/j.biopsych.2004.08.025

*Nigg, J. T., Blaskey, L. G., Huang-Pollock, C. L., & Rappley, M. D. (2002). Neuropsychological executive functions and DSM-IV ADHD subtypes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(1), 59-66.

Nijmeijer, J. S., Minderaa, R. B., Buitelaar, J. K., Mulligan, A., Hartman, C. A., & Hoekstra, P. J. (2008). Attention-deficit/ hyperactivity disorder and social dysfunctioning. *Clinical Psychology Review, 28*, 692-708. doi: 10.1016/j.cpr.2007.10.003

Nolan, E. E., Gadow, K. D., & Sprafkin, J. (2001). Teacher reports of DSM-IV ADHD, ODD, and CD symptoms in schoolchildren. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(2), 241-249.

Norman, L. J., Carlisi, C. O., Christakou, A., Murphy, C. M., Chantiluke, K., Giampietro, V.,... Rubia, K. (2018). Frontostriatal dysfunction during decision making in attention-deficit/hyperactivity disorder and obsessive-compulsive disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*, 694-703. doi: 10.1016/j.bpsc.2018.03.009

*Northington, S. A. (2009). *Executive function deficit variability within attention deficit hyperactivity disorder* (Unpublished doctoral dissertation). Illinois Institute of Technology, Chicago, IL.

*Nydén, A., Gillberg, C., Hjelmquist, E., & Heiman, M. (1999). Executive function/attention deficits in boys with Asperger syndrome, attention disorder and reading/writing disorder. *Autism*, *3*(3), 213-228.

Oades, R. D., & Christiansen, H. (2008). Cognitive switching processes in young people with attention-deficit/hyperactivity disorder. *Archives of Clinical Neuropsychology*, *23*, 21-32. doi: 10.1016/j.acn.2007.09.002

Oddo, L. E., Knouse, L. E., Surman, C. B., & Safren, S. A. (2018). Investigating resilience to depression in adults with ADHD. *Journal of Attention Disorders*, *22*, 497-505. doi: 10.1177/1087054716636937

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan — A web and mobile app for systematic reviews. *Systematic Reviews,* 5, 1-10. doi: 10.1186/s13643-016-0384-4

Owen, A. M., Morris, R. G., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1996). Double dissociations of memory and executive functions in working memory tasks following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain*, *119*(5), 1597-1615.

*Ozonoff, S., & Jensen, J. (1999). Brief report: Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders*, *29*(2), 171-177.

Papageorgiou, V., Kalyva, E., Dafoulis, V., & Vostanis, P. (2008). Differences in parents' and teachers' ratings of ADHD symptoms and other mental health problems. *The European Journal of Psychiatry*, *22*(4), 200-210.

*Pasini, A., Paloscia, C., Alessandrelli, R., Porfirio, M. C., & Curatolo, P. (2007).

Attention and executive functions profile in drug naive ADHD subtypes. *Brain and Development*, *29*, 400-408. doi: 10.1016/j.braindev.2006.11.010

Partington, J. (1949). Detailed instructions for administering Partington's pathways test. *Psychological Service Center Journal*, *1*, 46-48.

*Pazvantoğlu, O., Aker, A. A., Karabekiroğlu, K., Akbaş, S., Sarısoy, G., Baykal, S.,... Şahin, A. R. (2012). Neuropsychological weaknesses in adult ADHD; cognitive functions as core deficit and roles of them in persistence to adulthood. *Journal of the International Neuropsychological Society*, *18*, 819-826. doi: 10.1017/S1355617712000574

*Pedersen, J. A. (2002). *Executive functions in children with attention-deficit hyperactivity disorder, oppositional defiant disorder, and comorbid diagnosed attention-deficit hyperactivity disorder, and oppositional defiant disorder* (Unpublished doctoral dissertation). Pennsylvania State University, University Park, PA.

Pelham, W. E., Foster, E. M., & Robb, J. A. (2007). The economic impact of attention deficit/hyperactivity disorder in children and adolescents. *Ambulatory Pediatrics, 7*, 121-131. doi: 10.1016/j.ambp.2006.08.002

Pennington, B., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*(1), 51–87.

*Perugini, E. M., Harvey, E. A., Lovejoy, D. W., Sandstrom, K., & Webb, A. H. (2000). The predictive power of combined neuropsychological measures for attention-deficit/hyperactivity disorder in children. *Child Neuropsychology*, *6*(2), 101-114.

Peters, J., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*, 991-996. doi: 10.1016/j.jclinepi.2007.11.010

Piek, J. P., Dyck, M. J., Francis, M., & Conwell, A. (2007). Working memory, processing speed, and set-shifting in children with developmental coordination disorder and attention-deficit–hyperactivity disorder. *Developmental Medicine & Child Neurology*, *49*, 678-683. doi: 10.1111/j.1469-8749.2007.00678.x

Pierce, C. A. (2007). [Review of the software Comprehensive Meta-Analysis (Version 2.2.027) by M. Borenstein, L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein]. *Organizational Research Methods, 11*, 188-191. doi: 10.1177/1094428106296641

*Pineda, D., Ardila, A., Rosselli, M., Cadavid, C., Mancheno, S., & Mejia, S. (1998). Executive dysfunctions in children with attention deficit hyperactivity disorder. *International Journal of Neuroscience, 96*, 177-196. doi: 10.3109/00207459808986466

Poirier, T., & Behnen, E. (2014). Where and how to search for evidence in the education literature: The WHEEL. *American Journal of Pharmaceutical Education*, *78*(4), 1-8.

Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C., & Rohde, L. A. (2014).

ADHD prevalence estimates across three decades: An updated systematic review and meta-regression analysis. *International Journal of Epidemiology*, *43*, 434-442. doi: 10.1093/ije/dyt261

Poon, K., & Ho, C. S. (2014). Contrasting deficits on executive functions in Chinese delinquent adolescents with attention deficit and hyperactivity disorder symptoms and/ or reading disability. *Research in Developmental Disabilities, 35*, 3046-3056. doi: 10.1016/j.ridd.2014.07.046

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice* (Vol. 892). Upper Saddle River, NJ: Pearson/Prentice Hall.

Posner, M. I., & Rothbart, M. K. (2009). Toward a physical basis of attention and self-regulation. *Physics of Life Reviews, 6*, 103-120. doi: 10.1016/j.plrev.2009.02.001

Posner, M. I., & Raichle, M. E. (1994). *Images of mind*. Scientific American Library/ Scientific American Books.

*Potter, B. S. (2006). *Evaluating everyday executive functions and psychosocial behavior in children with neurofibromatosis type 1* (Unpublished doctoral dissertation). Antioch New England Graduate School, Keene, NH.

*Pratt, B. M. (1999). *The comparative development of executive functions in elementary school children with reading disorder and attention-deficit/hyperactivity disorder* (Unpublished doctoral dissertation). Alliant International University, Los Angeles, CA.

*Qian, Y., Shuai, L., Cao, Q., Chan, R. C., & Wang, Y. (2010). Do executive function deficits differentiate between children with attention deficit hyperactivity disorder (ADHD) and ADHD comorbid with oppositional defiant disorder? A cross-cultural study using performance-based tests and the Behavior Rating Inventory of Executive Function. *The Clinical Neuropsychologist*, *24*, 793-810. doi: 10.1080/13854041003749342

Quay, H. C. (1997). Inhibition and attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology*, *25*(1), 7-13.

Rajendran, K., O'Neill, S., Marks, D. J., & Halperin, J. M. (2015). Latent profile analysis of neuropsychological measures to determine preschoolers' risk for ADHD. *Journal of Child Psychology and Psychiatry*, *56*, 958-965. doi: 10.1111/jcpp.12434

Ramtekkar, U. P., Reiersen, A. M., Todorov, A. A., & Todd, R. D. (2010). Sex and age differences in attention-deficit/ hyperactivity disorder symptoms and diagnoses: Implications for DSM-V and ICD-11. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*, 217-228. doi: 10.1016/j.jaac.2009.11.011

*Rapport, L. J., Van Voorhis, A., Tzelepis, A., & Friedman, S. R. (2001). Executive functioning in adult attention-deficit hyperactivity disorder. *The Clinical Neuropsychologist*, *15*(4), 479-491.

*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*, 45. doi: 10.1037/a0022585

*Reeve, W. V., & Schandler, S. L. (2001). Frontal lobe functioning in adolescents with attention deficit hyperactivity disorder. *Adolescence*, *36*(144), 749-765.

Reid, R., & Johnson, J. (2011). *Teacher's guide to ADHD*. New York, NY: Guilford Press.

Reitan, R. M. (1967). Problems and prospects' in identifying the effects of brain lesions with psychological tests. *Sinai Hospital Journal*, *14*, 37-55.

Reitan, R. M. (1966). A research program on the psychological effects of brain lesions in human beings. *International review of research in mental retardation*, *1*, 153-218.

Reitan, R. M. (1955). Certain differential effects of left and right cerebral lesions in human adults. *Journal of Comparative and Physiological Psychology*, *48*(6), 474-477.

Reynolds, C. R. (2002). *Comprehensive Trail Making Test (CTMT)*. Austin, TX: PRO-ED, Inc.

Rhodes, S. M., Coghill, D. R., & Matthews, K. (2004). Methylphenidate restores visual memory, but not working memory function in attention deficit-hyperkinetic disorder. *Psychopharmacology*, *175*, 319-330. doi:10.1007/s00213-004-1833-7

Riccio, C. A., Blakely, A., Yoon, M., & Reynolds, C. R. (2013). Two-factor structure of the Comprehensive Trail-Making Test in adults. *Applied Neuropsychology: Adult*, *20*, 155-158. doi: 10.1080/09084282.2012.670169

Riccio, C. A., Homack, S., Jarratt, K. P., & Wolfe, M. E. (2006). Differences in academic and executive function domains among children with ADHD predominantly inattentive and combined types. *Archives of Clinical Neuropsychology, 21,* 657-667. doi: 10.1016/j.acn.2006.05.010

*Riccio, C. A., Wolfe, M., Davis, B., Romine, C., George, C., & Lee, D. (2005). Attention deficit hyperactivity disorder: Manifestation in adulthood. *Archives of Clinical Neuropsychology*, *20*, 249-269. doi: 10.1016/j.acn.2004.07.005

Rie, H. E. (1980). Definitional problems. *Handbook of Minimal Brain Dysfunctions*, 3-18.

Rigoli, D., Piek, J. P., Kane, R., & Oosterlaan, J. (2012). An examination of the relationship between motor coordination and executive functions in adolescents. *Developmental Medicine & Child Neurology*, *54*, 1025-1031. doi: 10.1111/j.1469-8749.2012.04403.x

*Riordan, H. J., Flashman, L. A., Saykin, A. J., Frutiger, S. A., Carroll, K. E., & Huey, L. (1999). Neuropsychological correlates of methylphenidate treatment in adult ADHD with and without depression. *Archives of Clinical Neuropsychology*, *14*(2), 217-233.

Roberts, B. A., Martel, M. M., & Nigg, J. T. (2017). Are there executive dysfunction subtypes within attention- deficit/hyperactivity disorder? *Journal of Attention Disorders, 21*, 284-293. doi: 10.1177/1087054713510349

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, *123*(4), 374-393.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207-231. doi: 10.1037/0096-3445.124.2.207

*Rohlf, H., Jucksch, V., Gawrilow, C., Huss, M., Hein, J., Lehmkuhl, U., & Salbach-Andrae, H. (2012). Set shifting and working memory in adults with attention-deficit/ hyperactivity disorder. *Journal of Neural Transmission, 119*, 95-106. doi: 10.1007/s00702-011-0660-3

Roman-Urrestarazu, A., Lindholm, P., Moilanen, I., Kiviniemi, V., Miettunen, J., Jääskeläinen, E.,… Nikkinen, J. (2016). Brain structural deficits and working memory fMRI dysfunction in young adults who were diagnosed with ADHD in adolescence. *European Child & Adolescent Psychiatry*, *25*, 529-538. doi:10.1007/s00787-015-0755-8

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/ full-term differences in academic achievement. *Developmental Science*, *14*, 1161-1175. doi: 10.1111/j.1467-7687.2011.01068.x

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

Roth, R. M., Lance, C. E., Isquith, P. K., Fischer, A. S., & Giancola, P. R. (2013). Confirmatory factor analysis of the behavior rating inventory of executive function – Adult version in healthy adults and application to attention-deficit/ hyperactivity disorder. *Archives of Clinical Neuropsychology, 28*, 425-434. doi: 10.1093/arclin/act031

Roth, R. M., Isquith, P. K., & Gioia, G. A. (2005). *Behavior Rating Inventory of Executive Function - Adult version (BRIEF-A)*. Lutz, FL: Psychological Assessment Resources.

Rothstein, H., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Chichester, England: Wiley. doi: 10.1002/0470870168

Rowe, J. B., Owen, A. M., Johnsrude, I. S., & Passingham, R. E. (2001). Imaging the mental components of a planning task. *Neuropsychologia*, *39*(3), 315-327.

Rucklidge, J. J. (2010). Gender differences in attention-deficit/ hyperactivity disorder. *Psychiatric Clinics of North America, 33*, 357-373. doi: 10.1016/j.psc.2010.01.006

*Rucklidge, J. J. (2006). Gender differences in neuropsychological functioning of New Zealand adolescents with and without attention deficit hyperactivity disorder. *International Journal of Disability, Development and Education*, *53*, 47-66. doi: 10.1080/10349120600577402

Rucklidge, J. J., & Tannock, R. (2002). Neuropsychological profiles of adolescents with ADHD: Effects of reading difficulties and gender. *Journal of Child Psychology and Psychiatry*, *43*(8), 988-1003.

Rucklidge, J. J., & Tannock, R. (2001). Psychiatric, psychosocial, and cognitive functioning of female adolescents with ADHD. *Journal of American Academy of Child and Adolescent Psychiatry, 40*, 530-540. doi: 10.1097/00004583-200105000-00012

Sadeh, S. S., Burns, M. K., Sullivan, A. L. (2012). Examining an executive function rating scale as a predictor of achievement in children at risk for behavior problems. *School Psychology Quarterly, 27*, 236-246. doi: 10.1037/spq0000012

Salthouse, T. A., & Tucker-Drob, E. M. (2008). Implications of short-term retest effects for the interpretation of longitudinal change. *Neuropsychology*, *22*, 800-811. doi: 10.1037/a0013091

Salthouse, T. A. (2007). Implications of within-person variability in cognitive and neuropsychological functioning for the interpretation of change. *Neuropsychology*, *21*, 401-411. doi: 10.1037/0894-4105.21.4.401

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*, 566-594. doi: 10.1037/0096-3445.132.4.566

*Sartory, G., Heine, A., Müller, B. W., & Elvermann-Hallner, A. (2002). Event-and motor-related potentials during the continuous performance task in attention-deficit/hyperactivity disorder. *Journal of Psychophysiology*, *16*, 97-106. doi: 10.1027//0269-8803.16.2.97

Schachar, R., Mota, V. L., Logan, G. D., Tannock, R., & Klim, P. (2000). Confirmation of an inhibitory control deficit in attention-deficit/hyperactivity disorder. *Journal* *of Abnormal Child Psychology*, *28*(3), 227-235.

Schachar, R. J., Tannock, R., & Logan, G. (1993). Inhibitory control, impulsiveness, and attention deficit hyperactivity disorder. *Clinical Psychology Review, 13*, 721-739. doi: 10.1016/S0272-7358(05)80003-0

*Schmitz, M., Cadore, L., Paczko, M., Kipper, L., Chaves, M., Rohde, L. A.,... Knijnik, M. (2002). Neuropsychological performance in DSM-IV ADHD subtypes: An exploratory study with untreated adolescents. *The Canadian Journal of Psychiatry*, *47*(9), 863-869.

Schmitz, N., Rubia, K., Daly, E., Smith, A., Williams, S., & Murphy, D. G. (2006). Neural correlates of executive function in autistic spectrum disorders. *Biological Psychiatry*, *59*, 7-16. doi: 10.1016/j.biopsych.2005.06.007

Schweitzer, J. B., Hanford, R. B., & Medoff, D. R. (2006). Working memory deficits in adults with ADHD: is there evidence for subtype differences? *Behavioral and Brain Functions*, *2*. doi: 10.1186/1744-9081-2-43

*Seidman, L. J., Biederman, J., Monuteaux, M. C., Valera, E., Doyle, A. E., & Faraone, S. V. (2005). Impact of gender and age on executive functioning: Do girls and boys with and without attention deficit hyperactivity disorder differ neuropsychologically in preteen and teenage years? *Developmental Neuropsychology, 27*, 79-105. doi: 10.1207/s15326942dn2701_4

Seidman, L. J., Valera, E. M., & Makris, N. (2005). Structural brain imaging of attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *57*(11), 1263-1272.

Seidman, L. J., Biederman, J., Faraone, S. V., Weber, W., & Ouellette, C. (1997). Toward defining a neuropsychology of attention deficit-hyperactivity disorder: Performance of children and adolescents from a large clinically referred sample. *Journal of Consulting and Clinical Psychology*, *65*(1), 150-160.

Semrud-Clikeman, M., Walkowiak, J., Wilkinson, A., & Butcher, B. (2010). Executive functioning in children with asperger syndrome, ADHD-combined type, ADHD-predominately inattentive type, and controls. *Journal of Autism and Developmental Disorders*, *40*, 1017-1027. doi: 10.1007/s10803-010-0951-9

*Semrud-Clikeman, M., Steingard, R. J., Filipek, P., Biederman, J., Bekken, K., & Renshaw, P. F. (2000). Using MRI to examine brain-behavior relationships in males with attention deficit disorder with hyperactivity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(4), 477-484.

Sergeant, J. A., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder? *Behavioural Brain Research*, *130*(1), 3-28.

Shallice, T. (1988). *From neuropsychology to mental structure*. New York, NY: Cambridge University Press.

*Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G.,... DeFries, J. C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology*, *34*, 584-602. doi: 10.1007/s10802-006-9037-8

Shang, C. Y., & Gau, S. S. (2011). Visual memory as a potential cognitive endophenotype of attention deficit hyperactivity disorder. *Psychological Medicine*, *41*, 2603-2614. doi:10.1017/S0033291711000857

Sharp, S. I., McQuillin, A., & Gurling, H. M. (2009). Genetics of attention-deficit/ hyperactivity disorder (ADHD). *Neuropharmacology, 57*, 590–600. doi: 10.1016/j.neuropharm.2009.08.011

Shaw, P., Lalonde, F., Lepage, C., Rabin, C., Eckstrand, K., Sharp, W.,… Rapoport, J. (2009). Development of cortical asymmetry in typically developing children and its disruption in attention-deficit/ hyperactivity disorder. *Archives of General Psychiatry, 66*, 888–896. doi:10.1001/archgenpsychiatry.2009.103

Shen, I. H., Lee, D. S., & Chen, C. L. (2014). The role of trait impulsivity in response inhibition: Event-related potentials in a stop-signal task. *International Journal of Psychophysiology*, *91*, 80-87. doi: 10.1016/j.ijpsycho.2013.11.004

*Shin, M. S., Choi, H., Kim, H., Hwang, J. W., Kim, B. N., & Cho, S. C. (2008). A study of neuropsychological deficit in children with obsessive-compulsive disorder. *European Psychiatry*, *23*, 512-520. doi: 10.1016/j.eurpsy.2008.03.010

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S. C., & Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Developmental Neuropsychology*, *35*, 679-697. doi: 10.1080/87565641.2010.508546

*Shuai, L., Daley, D., Wang, Y. F., Zhang, J. S., Kong, Y. T., Tan, X., & Ji, N. (2017).

Executive function training for children with attention deficit hyperactivity disorder. *Chinese Medical Journal*, *130*, 549-558. doi: 10.4103/0366-6999.200541

*Shuai, L., Chan, R. C., & Wang, Y. (2010). Executive function profile of Chinese boys with attention-deficit hyperactivity disorder: Different subtypes and comorbidity. *Archives of Clinical Neuropsychology*, *26*, 120-132. doi: 10.1093/arclin/acq101

*Silva, K. L., Rovaris, D. L., Guimarães-da-Silva, P. O., Victor, M. M., Salgado, C. A., Vitola, E. S.,… Bau, C. H. (2014). Could comorbid bipolar disorder account for a significant share of executive function deficits in adults with attention-deficit hyperactivity disorder? *Bipolar Disorders, 16*, 270-276. doi: 10.1111/bdi.12158

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. doi: 10.1016/j.ecresq.2011.08.004

Sinzig, J., Morsch, D., Bruning, N., Schmidt, M. H., & Lehmkuhl, G. (2008). Inhibition, flexibility, working memory and planning in autism spectrum disorders with and without comorbid ADHD-symptoms. *Child and Adolescent Psychiatry and Mental Health*, *2*, 1-12. doi: 10.1186/1753-2000-2-4

Skogan, A. H., Egeland, J., Zeiner, P., Øvergaard, K. R., Oerbeck, B., Reichborn-Kjennerud, T., & Aase, H. (2016). Factor structure of the Behavior Rating Inventory of Executive Functions (BRIEF-P) at age three years. *Child Neuropsychology, 22*, 472-492. doi: 10.1080/09297049.2014.992401

*Skogli, E. W., Egeland, J., Andersen, P. N., Hovik, K. T., & Merete, Ø. (2014). Few differences in hot and cold executive functions in children and adolescents with combined and inattentive subtypes of ADHD. *Child Neuropsychology, 20*, 162-181. doi: 10.1080/09297049.2012.753998

Skogli, E. W., Teicher, M. H., Andersen, P. N., Hovik, K. T., & Merete, Ø. (2013). ADHD in girls and boys: Gender differences in co-existing symptoms and executive function measures. *BioMed Central Psychiatry, 13*, 298-325. doi: 10.1186/1471-244X-13-298

Silver, C. H. (2000). Ecological validity of neuropsychological assessment in childhood traumatic brain injury. *Journal of Head Trauma Rehabilitation, 15*(4), 973–988.

Slick, D. J., Lautzenhiser, A., Sherman, E. M. S., & Eyrl, K. (2006). Frequency of scale elevations and factor structure of the Behavior Rating Inventory of Executive Function (BRIEF) in children and adolescents with intractable epilepsy. *Child Neuropsychology, 12*, 181-189. doi: 10.1080/09297040600611320

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283*(5408), 1657-1661.

*Solanto, M., Gilbert, S., Raj, A., Zhu, J., Pope-Boyd, S., Stepak, B.,… Newcorn, J. (2007). Neurocognitive functioning in AD/HD, predominantly inattentive and combined subtypes. *Journal of Abnormal Child Psychology, 35*, 729-744. doi: 10.1007/s10802-007-9123-6

*Sørensen, L., Plessen, K. J., & Lundervold, A. J. (2012). The impact of inattention and emotional problems on cognitive control in primary school children. *Journal of Attention Disorders*, *16*, 589-599. doi: 10.1177/1087054711417394

Stanovich, K. (2011). *Rationality and the reflective mind*. Oxford University Press.

*Stavro, G. M., Ettenhofer, M. L., & Nigg, J. T. (2007). Executive functions and adaptive functioning in young adult attention-deficit/hyperactivity disorder. *Journal of the International Neuropsychological Society*, *13*, 324-334. doi: 10.10170S1355617707070348

Steenbergen-Hu, S., & Olszewski-Kubilius, P. (2016). How to conduct a good meta-analysis in gifted education. *Gifted Child Quarterly, 60*, 134-154. doi: 10.1177/0016986216629545

*Stern, A., & Maeir, A. (2014). Validating the measurement of executive functions in an occupational context for adults with attention deficit hyperactivity disorder. *American Journal of Occupational Therapy*, *68*, 719-728. doi: 10.5014/ajot.2014.012419

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J.,... Tetzlaff, J. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *343*, d4002.

Strauss A. A., Lehtinen L. E. (1947) *Psychopathology and education of the brain-injured child*. New York, NY: Grune & Stratton.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662. doi: 10.1037/h0054651

Stuss, D. T., Bisschop, S. M., Alexander, M. P., Levine, B., Katz, D., & Izukawa, D.

(2001). The Trail Making Test: a study in focal lesion patients. *Psychological Assessment*, *13*(2), 230-239.

Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: A Conceptual view. *Psychological Research*, *63*(3-4), 289-298.

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.

Sullivan, J. R., & Riccio, C. A. (2006). An empirical analysis of the BASC Frontal Lobe/ Executive Control scale with a clinical sample. *Archives of Clinical Neuropsychology*, *21*, 495-501. doi: 10.1016/j.acn.2006.05.008

*Surman, C. B., Biederman, J., Spencer, T., Miller, C. A., Petty, C. R., & Faraone, S. V. (2015). Neuropsychological deficits are not predictive of deficient emotional self-regulation in adults with ADHD. *Journal of Attention Disorders*, *19*, 1046-1053. doi: 10.1177/1087054713476548

*Tafazoli, S., O'Neill, J., Bejjani, A., Ly, R., Salamon, N., McCracken, J. T.,... Levitt, J. G. (2013). 1H MRSI of middle frontal gyrus in pediatric ADHD. *Journal of Psychiatric Research*, *47*, 505-512. doi: 10.1016/j.jpsychires.2012.11.011

Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(10), 1231-1238.

Tamnes, C. K., Ostby, Y., Walhovd, K. B., Westlye, L. T., Due-Tonnesen, P., & Fjell, A. M. (2010). Neuroanatomical correlates of executive functions in children and adolescents: A magnetic resonance imaging (MRI) study of cortical thickness. *Neuropsychologia, 48*, 2496–2508. doi: 10.1016/j.neuropsychologia.2010.04.024

*Taylor, C. J., & Miller, D. C. (1997). Neuropsychological assessment of attention in ADHD adults. *Journal of Attention Disorders*, *2*(2), 77-88.

Ten Eycke, K. D., & Dewey, D. (2015). Parent-report and performance-based measures of executive function assess different constructs. *Child Neuropsychology*, 1-18. doi: 10.1080/09297049.2015.1065961

Thapar, A., Harrington, R., Ross, M., & McGuffin, P. (2000). Does the definition of ADHD affect heritability? *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 1528-1536. doi: 10.1097/00004583-200012000-00015

Thissen, A. J. A. M., Rommelse, N. N. J., Hoekstra, P. J., Hartman, C., Heslenfeld, D., Luman, M.,... Buitelaar, J. K. (2014). Attention deficit hyperactivity disorder (ADHD) and executive functioning in affected and unaffected adolescents and their parents: challenging the endophenotype construct. *Psychological Medicine*, *44*, 881-892. doi: 10.1017/S0033291713001153

Thomas, S., Sciberras, E., Lycett, K., Papadopoulos, N., & Rinehart, N. (2018). Physical functioning, emotional, and behavioral problems in children with ADHD and comorbid ASD: a cross-sectional study. *Journal of Attention Disorders*, *22*, 1002-1007. doi: 10.1177/1087054715587096

Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach*. Guilford Press.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2017). The assessment of executive functions in Attention-Deficit/Hyperactivity Disorder (ADHD): Performance-based measures versus ratings of behavior. In M. Hoskyn, G. Iarocci, & A. Young (Eds.), *Executive functions in children's everyday lives: A handbook for professionals in applied psychology*. New York: Oxford University Press.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, 5*, 131-143. doi: 10.1111/jcpp.12001

*Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2008). 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*, 53-72. doi: 10.1080/09297040802070929

Tricco, A. C., Chit, A., Hallett, D., Soobiah, C., Meier, G., Chen, M.,... Loeb, M. (2012). Effect of influenza vaccines against mismatched strains: A systematic review protocol. *Systematic Reviews*, *1*, 1-5. doi: 10.1186/2046-4053-1-35

*Tripp, G., Luk, S. L., Schaughency, E. A., & Singh, R. (1999). DSM-IV and ICD-10: A comparison of the correlates of ADHD and hyperkinetic disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*(2), 156-164.

*Tripp, G., Ryan, J., & Peace, K. (2002). Neuropsychological functioning in children with DSM-IV combined type attention deficit hyperactivity disorder. *Australian and New Zealand Journal of Psychiatry*, *36*(6), 771-779.

Tseng, W., & Gau, S. S. (2013). Executive function as a mediator in the link between attention-deficit/ hyperactivity disorder and social problems. *Journal of Child Psychology and Psychiatry, 54*, 996-1004. doi: 10.1111/jcpp.12072

*Tsuchiya, E., Oki, J., Yahara, N., & Fujieda, K. (2005). Computerized version of the Wisconsin Card Sorting Test in children with high-functioning autistic disorder or attention-deficit/ hyperactivity disorder. *Brain and Development, 27*, 233-236. doi: 10.1016/j.braindev.2004.06.008

Turgay, A., Goodman, D. W., Asherson, P., Lasser, R. A., Babcock, T. F., Pucci, M. L., & Barkley, R. (2012). Lifespan persistence of ADHD: The life transition model and its application. *The Journal of Clinical Psychiatry*, *73*, 192-201. doi: 10.4088/JCP.10m06628

Uebel, H., Albrecht, B., Asherson, P., Börger, N. A., Butler, L., Chen, W.,... Andreou, P. (2010). Performance variability, impulsivity errors and the impact of incentives as gender-independent endophenotypes for ADHD. *Journal of Child Psychology and Psychiatry*, *51*(2), 210-218.

Usai, M. C., Viterbori, P., Traverso, L., & De Franchis, V. (2014). Latent structure of executive function in five-and six-year-old children: A longitudinal study. *European Journal of Developmental Psychology*, *11*, 447-462. doi: 10.1080/17405629.2013.840578

*Van Goozen, S. H., Cohen-Kettenis, P. T., Snoek, H., Matthys, W., Swaab-Barneveld, H., & Van Engeland, H. (2004). Executive functioning in children: A comparison of hospitalised ODD and ODD/ADHD children and normal controls. *Journal of Child Psychology and Psychiatry*, *45*(2), 284-292.

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.

van Rooij, D., Hoekstra, P. J., Maarten, M., von Rhein, D., Thissen, A., Heslenfeld, D.,… Hartman, C. A. (2015). Distinguishing adolescents with ADHD from their unaffected siblings and healthy comparison subjects by neural activation patterns during response inhibition. *American Journal of Psychiatry, 172*, 674-683. doi: 10.1176/appi.ajp.2014.13121635

Vasconcelos, M. M., Werner Jr, J., Malheiros, A. F. D. A., Lima, D. F. N., Santos, Í. S. O., & Barbosa, J. B. (2003). Attention deficit/hyperactivity disorder prevalence in an inner city elementary school. *Arquivos de Neuro-Psiquiatria*, *61*(1), 67-73.

Visser, S. N., Danielson, M. L., Bitsko, R. H., Holbrook, J. R., Kogan, M. D., Ghandour, R. M.,... Blumberg, S. J. (2014). Trends in the parent-report of health care provider-diagnosed and medicated attention-deficit/ hyperactivity disorder: United States, 2003-2011. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*, 34-46. doi: 10.1016/j.jaac.2013.09.001

Vriezen, E. R., & Pigott, S. E. (2002). The relationship between parental report on the BRIEF and performance-based measures of executive function in children with moderate to severe traumatic brain injury. *Child Neuropsychology, 8*(4), 296-303.

*Walker, A. Y., Shores, A. E., Trollor, J. N., Lee, T., & Sachdev, P. S. (2000).

Neuropsychological functioning of adults with attention deficit hyperactivity disorder. *Journal of Clinical and Experimental Neuropsychology*, *22*(1), 115-124.

Wasserman, T., & Wasserman, L. D. (2012). The sensitivity and specificity of neuropsychological tests in the diagnosis of attention deficit hyperactivity disorder. *Applied Neuropsychology: Child, 1*, 90-99. doi: 10.1080/21622965.2012.702025

Weare, K., & Nind, M. (2011). Mental health promotion and problem prevention in schools: What does the evidence say? *Health Promotion International*, *26*(1), 29-69. doi: 10.1093/heapro/dar075

Weintraub, M. J., Axelson, D. A., Kowatch, R. A., Schneck, C. D., & Miklowitz, D. J. (2019). Comorbid disorders as moderators of response to family interventions among adolescents with bipolar disorder. *Journal of Affective Disorders*, *246*, 754-762. doi: 10.1016/j.jad.2018.12.125

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7*, 131-149. doi: 10.1080/87565649109540483

Welsh, M. C., & Pennington, B. F. (1988). Assessing frontal lobe functioning in children: Views from developmental psychology. *Developmental Neuropsychology*, *4*(3), 199-230.

Weyandt, L., Swentosky, A., & Gudmundsdottir, B. G. (2013). Neuroimaging and ADHD: fMRI, PET, DTI findings, and methodological limitations.

Developmental Neuropsychology, 38, 211-225. doi: 10.1080/87565641.2013.783833

Weyandt, L. (2006). *The physiological bases of cognitive and behavioral disorders*. Mahwah, NJ: Lawrence Erlbaum and Associates.

*Weyandt, L. L., Rice, J. A., Linterman, I., Mitzlaff, L., & Emert, E. (1998). Neuropsychological performance of a sample of adults with ADHD, developmental reading disorder, and controls. *Developmental Neuropsychology*, *14*, 643-656. doi: 10.1080/87565649809540734

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, *108*, 436-452. doi: 10.1016/j.jecp.2010.08.008

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*, 575-587. doi: 10.1037/0012-1649.44.2.575

Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R.,… Lahey, B. B. (2012). Validity of DSM–IV attention deficit/ hyperactivity disorder symptom dimensions and subtypes. *Journal of Abnormal Psychology, 121*, 991-1010. doi: 10.1037/a0027347

Willcutt, E., 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. doi:10.1016/j.biopsych.2005.02.006

*Willcutt, E. G., Pennington, B. F., Olson, R. K., Chhabildas, N., & Hulslander, J. (2005). Neuropsychological analyses of comorbidity between reading disability and attention deficit hyperactivity disorder: In search of the common deficit. *Developmental Neuropsychology*, *27*(1), 35-78.

*Willcutt, E. G., Pennington, B. F., Boada, R., Ogline, J. S., Tunick, R. A., Chhabildas, N. A., & Olson, R. K. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/ hyperactivity disorder. *Journal of Abnormal Psychology*, *110*(1), 157-172.

Williams, D., Stott, C. M., Goodyer, I. M., & Sahakian, B. J. (2000). Specific language impairment with or without hyperactivity: Neuropsychological evidence for frontostriatal dysfunction. *Developmental Medicine & Child Neurology*, *42*(6), 368-375.

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, *24*, 418. doi: 10.1037/a0025779

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment*, *22*, 306-317. doi: 10.1037/a0018708

Wojczynski, M. K., & Tiwari, H. K. (2008). Definition of phenotype. *Advances in Genetics*, *60*, 75-105. doi: 10.1016/S0065-2660(07)00404-X

Wolraich, M. L., Hannah, J. N., Pinnock, T. Y., Baumgaertel, A., & Brown, J. (1996). Comparison of diagnostic criteria for attention-deficit hyperactivity disorder in a county-wide sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*, 319-324. doi: 10.1002/hbm.22047

*Yang, J., Zhou, S., Yao, S., Su, L., & McWhinnie, C. (2009). The relationship between theory of mind and executive function in a sample of children from mainland China. *Child Psychiatry and Human Development*, *40*, 169-182. doi: 10.1007/s10578-008-0119-4

*Yang, L., Cao, Q., Shuai, L., Li, H., Chan, R. C., & Wang, Y. (2012). Comparative study of OROS-MPH and atomoxetine on executive function improvement in ADHD: A randomized controlled trial. *International Journal of Neuropsychopharmacology*, *15*, 15-26. doi: 10.1017/S1461145711001490

*Yilmaz, A., Gokcen, C., Fettahoglu, E., & Ozatalay, E. (2013). The effect of methylphenidate on executive functions in children with attention-deficit hyperactivity disorder. *Klinik Psikofarmakoloji Bulteni-Bulletin of Clinical Psychopharmacology, 23*, 162-170. doi: 10.5455/bcp.20121130090251

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 5*, 7–17. doi: 10.1177/001872086300500102

Zandbelt, B. B., Bloemendaal, M., Neggers, S. F., Kahn, R. S., & Vink, M. (2013).

Expectations and violations: delineating the neural network of proactive inhibitory control. *Human Brain Mapping*, *34*(9), 2015-2024.

*Zarrabi, M., Shahrivar, Z., Doost, M. T., Khademi, M., & Nejad, G. Z. (2015). Concurrent validity of the Behavior Rating Inventory of Executive Function in children with attention deficit hyperactivity disorder. *Iranian Journal of Psychiatry and Behavioral Sciences*, *9*, 1-6. doi: 10.17795/ijpbs213

Zelazo, P. D., & Cunningham, W. A. (2007). Executive function: Mechanisms underlying emotion regulation. In J. Gross (Ed.), *Handbook of emotion regulation* (pp. 135–158). New York, NY: Guilford.

Zulauf, C. A., Sprich, S. E., Safren, S. A., & Wilens, T. E. (2014). The complicated relationship between attention deficit/hyperactivity disorder and substance use disorders. *Current Psychiatry Reports*, *16*, 1-11. doi: 10.1007/s11920-013-0436-6

APPENDIX A

Final Coding Form	
	1. Bibliographic reference [BR]
	2. Study ID [STUDYID]
	3. Type of publication [PUBTYPE] 1 journal article 2 dissertation 3 technical report 4 conference paper 5 other (specify)
	4. Publication year (9999 if unknown) [PUBYEAR]
Sample Descriptors	
	5. Age range [AGERANGE] 1 child (mean age ≤ 12.11) 2 adolescent (mean age 13-17) 3 adult (mean age 18-35) 4 child and adolescent (6-17) 5 child, adolescent, and adult (6-35+)
	6. Mean age - ADHD [ADHDAGE]
	7. Mean age - HC [HCAGE]

	8. Predominant race [RACE] 1 >60% White 2 >60% Black 3 >60% Hispanic 4 >60% other minority 5 mixed, none more than 60% 6 mixed, cannot estimate proportion 7 ethnic minority not specified 9 cannot tell
	9. Predominant sex [SEX] 1 <5% male 2 5%-50% male 3 50% male 4 50%-95% male 5 >95% male 9 cannot tell
	10. Sample size [SSIZE] ADHD sample size HC sample size
	11. Comorbidity [COMORBID] 0 none 1 other neurodevelopmental disorder

	2 mood disorder 3 anxiety 4 serious mental illness 5 medical comorbidity 6 multiple comorbidity 9 unspecified
	12 Percentage of Comorbidity [%COMORBID] 1 <5% 2 5%-50% 3 50% 4 50%-95% 5 >95% 9 cannot tell
	13. Medication status [MEDSTATUS] 1 takes medication, but not medicated during testing 2 takes medication, and was medicated during testing 3 naïve 9 unspecified
	14. Recruitment [SETTING] 1 academic medical center 2 community 3 private practice

	4 school-based (public, charter, private) 5 juvenile/ incarceration 6 academia/ university 7 hospital/ medical center not tied to academic institution 9 unspecified
	15. Location [LOC] Name of city/ country
Research Design Procedures	
	16. Measures used Name of assessments used to diagnose ADHD
	17. Group assignment [GROUP] 1 ADHD 2 ADHD-C 3 ADHD-I 4 ADHD-H 5 ADHD-comorbid
	18. DSM nomenclature [DSM] 3 DSM-III 3.5 DSM-III-R 4 DSM-IV 4.5 DSM-IV-TR 5 DSM-V

	9 unspecified
	19. Mean IQ for the groups Mean IQ for the ADHD group Mean IQ for the HC group
	20. Control for IQ [IQ] 1 yes 2 no
	21. Measurement method [MEASUREMETHOD] 1 PBM 2 rating scale
	22. PBM measure [PBM] 1 TMT-B 2 WCST 3 DKEFS
	23. Study quality [SQ] 1 Diagnostic process for the ADHD and HC groups 2 Statistical methods accounting for IQ 3 Balanced of gender between the groups 4 Statistical methods accounting for gender balance between the groups 5 Standardization of test administration

APPENDIX B

	Card sorting							Trails			BRIEF
Number of studies	33	32	19	17	17	12	11	42	28	23	18
Study referenced	PE	#CAT	TE	NPE	FMS	TC	PR	B	B-A	B/A	Parent
Ahmadi et al. (2014)	0.78		0.71	0.36			0.73				
Antonini et al. (2015)	0.22					-0.67					
Antshel et al. (2010)	0.22	-0.16		0.29	0.28						
Aycicegi-Dinn et al. (2011)	0.10	-0.30		0.10	-0.05	-0.11		0.94	1.32	0.02	
Bahçivan Saydam et al. (2015)	0.64	-0.37									
Beadle (2010)		-0.41			0.71			0.11			
Bodzy (2011)											0.32
Boswell (2013)								-0.82			0.36
Brewer et al. (2001)	1.39	-0.28	1.05	0.02							
Burmeister et al. (2005)											2.47
Chen et al. (2015)	0.82	-0.81	0.84	0.71		-0.79					
Davidson et al. (2016)								-0.88			1.48

	Card sorting							Trails			BRIEF
Number of studies	33	32	19	17	17	12	11	42	28	23	18
Study referenced	PE	#CAT	TE	NPE	FMS	TC	PR	B	B-A	B/A	Parent
De la Torre et al. (2017)								1.27			
Doyle et al. (2005)	0.11	-0.20		0.26	-0.08						
Elosúa et al. (2017)								0.64	0.64	0.55	
Faraone et al. (2006)		-0.21				0.39					
Gioia et al. (2002)											-0.60
Gonzalez-Gadea et al (2015)									0.06		
Hale et al. (2009)								0.51			
Holmes et al. (2010)								-0.02			
Hooper (1996)	0.53	-1.14		0.65	-0.04		0.37				
Houghton et al. (1999)	0.42	-0.58	0.93	0.37	0.39	0.26	0.72	0.45	1.18	0.05	
Hovik et al. (2017)											1.87
Huang et al. (2016)									0.51		0.24
Jarratt et al. (2005)											1.15

	Card sorting							Trails			BRIEF
Number of studies	33	32	19	17	17	12	11	42	28	23	18
Study referenced	PE	#CAT	TE	NPE	FMS	TC	PR	B	B-A	B/A	Parent
Johnson et al. (2001)		-0.05				-0.15		0.43	0.48	0.01	
Kado et al. (2012)	0.54	-0.58	0.49	0.15							
Kamradt et al. (2014)								0.27	0.18	0.00	
Kang et al. (2016)	2.80						2.04				
Keage et al. (2006)								0.39	0.17	0.00	
Lawrence et al. (2004)	1.04	-0.54					0.92				
Li et al. (2014)	0.64	-0.61	0.67	0.51		-0.75					
Loftis (2004)								-0.37			
Lovejoy et al. (1999)								0.56	0.36	-0.02	
Martel et al. (2009)								0.27			
Mataró et al. (1997)	0.34	-0.06						0.19	3.18	0.01	
Mattson et al. (2013)								-0.34			
Miranda et al. (2013)											0.61

	Card sorting							Trails			BRIEF
Number of studies	33	32	19	17	17	12	11	42	28	23	18
Study referenced	PE	#CAT	TE	NPE	FMS	TC	PR	B	B-A	B/A	Parent
Mullane & Corkum (2007)	0.40				0.73						
Müller et al. (2007)								1.23	3.11	0.06	
Murphy (2002)								0.73	0.56	-0.01	
Nigg et al. (2002)								0.51	1.43	0.09	
Northington (2009)											0.22
Nydén et al. (1999)		-1.17									
Ozonoff & Jensen (1999)							-0.20				
Pasini et al. (2007)	-0.56			0.21	-0.21			1.05	0.79	0.03	
Pazvantoglu et al. (2012)								1.06	3.69	0.08	
Pedersen (2002)	0.13				0.40			-0.22			
Perugini et al. (2000)								0.50	1.87	0.06	
Pineda et al. (1998)	0.61	-0.75	0.96	0.49	0.71						
Potter (2006)											2.28
Pratt (1999)											1.09
Qian et al. (2010)									0.80		

	Card sorting							Trails			BRIEF
Number of studies	33	32	19	17	17	12	11	42	28	23	18
Study referenced	PE	#CAT	TE	NPE	FMS	TC	PR	B	B-A	B/A	Parent
Rapport et al. (2001)	0.14			0.14				0.57	0.66	0.01	
Reddy et al. (2011)											1.62
Reeve & Schandler (2001)		-1.59			-0.20	-0.12					
Riccio et al. (2005)	-0.11		-0.04					0.36			
Riordan et al. (1999)								0.00	-0.22	-0.02	
Rohlf et al. (2012)	0.74		0.54					0.64	1.34	0.04	
Rucklidge (2006)		-0.19									
Sartory et al. (2002)		-0.93	0.68					1.60	5.89	0.18	
Schmitz et al. (2002)		-0.82	0.91				0.07				
Seidman et al. (2005)	0.17	-0.19		0.11	0.08						
Semrud-Clikeman et al. (2000)					1.12						
Shanahan et al. (2006)								-0.45			
Shin et al. (2008)	0.84	-1.72	1.55		0.81			1.02			

	Card sorting							Trails			BRIEF
Number of studies	33	32	19	17	17	12	11	42	28	23	18
Study referenced	PE	#CAT	TE	NPE	FMS	TC	PR	B	B-A	B/A	Parent
Shuai et al. (2011)								0.59	1.63	0.02	
Shuai et al. (2017)									0.08		
Silva et al. (2014)	-0.01	-0.11	0.04	0.04		-0.04	-0.01				
Skogli et al. (2014)								0.49			1.89
Solanto et al. (2007)		-3.45	-3.81		0.14		-2.75				
Sørensen et al. (2012)											1.46
Stavro et al. (2007)	0.12	-0.37						0.37	0.29	0.01	
Stern & Maeir (2014)											1.40
Surman et al. (2015)		-0.15				0.52					
Tafazoli et al. (2013)								0.02			
Taylor & Miller (1997)		0.03			0.04			0.00	0.39	0.01	
Toplack et al. (2009)								0.70			1.52
Tripp et al. (1999)								0.05			

	Card sorting							Trails			BRIEF
Number of studies	33	32	19	17	17	12	11	42	28	23	18
Study referenced	PE	#CAT	TE	NPE	FMS	TC	PR	B	B-A	B/A	Parent
Tripp et al. (2002)	-0.69	-0.70	-0.68			-0.75	-0.73	0.47			
Tsuchiya et al. (2005)	1.23	-1.33	1.23	0.93							
van Goozen et al. (2004)								0.62	0.73	0.05	
Walker et al. (2000)								0.86	1.54	0.05	
Weyandt et al. (1998)	1.54		1.35								
Willcutt et al. (2001)	0.48							0.28			
Willcutt et al. (2005)	0.37										
Yang et al. (2009)	0.28	-0.24	0.14								
Yang et al. (2012)									0.81		
Yilmaz et al. (2013)	0.50	-0.42	0.46	0.22	0.07	-0.21	0.44				
Zarrabi et al. (2015)											1.53

APPENDIX C

To Whom it May Concern,

My name is Ryan Hinojosa and I am entering my final year in the school psychology doctoral program at Texas A&M University in College Station, TX. For my dissertation, I am meta-analyzing data specific to differences in executive functioning between people with ADHD and healthy controls, namely set-shifting and cognitive flexibility. Primarily, I am looking to include studies indicating the means and standard deviations of various scores from different performance-based measures of shifting (e.g., Trails B, WCST, etc.) as well as behavior checklists (e.g., BRIEF). As such, I would like to include data from your publication, titled **[STUDY TITLE]**.

In your study, you reference the **[VARIABLE]**. I would like to inquire about the **[VARIABLE]** (and standard deviation) and whether that score was captured in your study. If so, would you be willing to share that data with me? With your help, I would be able to compute other referenced variables like **B-A** and **B/A** for additional analyses. I understand you are a very busy academician, but at your earliest convenience, I would sincerely appreciate your input. I can also be reached by phone at (###) ###-#### if it would be easier to talk through our ideas. Thanks for your consideration and I look forward to hearing back from you.

Respectfully,

Ryan Hinojosa